AQA

AQA Approved

Geography

For A-level

Fifth Edition

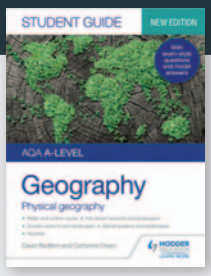
Malcolm Skinner
Paul Abbiss
Philip Banks
Helen Fyfe
Ian Whittaker

Boost

HODDER EDUCATION
LEARN MORE

Support throughout your course

Hodder Education publishes a range of study and revision guides for AQA Geography, so you can be confident in the knowledge and skills you need for success.

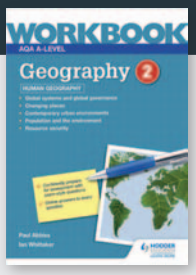


Student Guides

Reinforce your geographical understanding throughout your course with these helpful student guides. Combining clear topic summaries and sample questions and answers, you can improve your exam technique, and work towards achieving your best.

Physical Geography (topics water and carbon cycles, hot desert systems and landscapes, glacial systems and landscapes, hazards) 9781398328181

Human Geography (global systems and global governance, changing places, contemporary urban environments, population and the environment, resource security) | 9781398328198

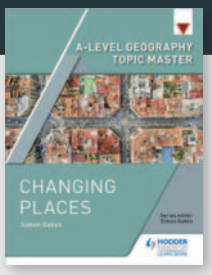


Student Workbooks

These new edition write-in workbooks have been fully updated to help you practise your skills and improve your subject knowledge both in school and at home.v

Physical Geography (topics water and carbon cycles, hot desert systems and landscapes, glacial systems and landscapes, hazards) 9781398332416

Human Geography (global systems and global governance, changing places, contemporary urban environments, population and the environment, resource security) | 9781398332423



A-level Topic Masters

Master the in-depth knowledge and higher-level skills you need to succeed in A-level Geography; these focused topic books help to extend your learning far beyond your course textbooks.

Global Systems | 9781510427938

Global Governance | 9781510427891

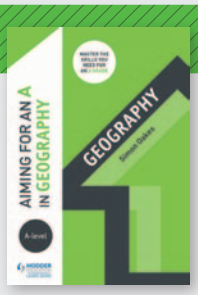
Glaciated Landscapes | 9781510427914

Coastal Landscapes | 9781510434622

The Water and Carbon Cycles | 9781510434615

Changing Places | 9781510427532

Suitable for all exam boards

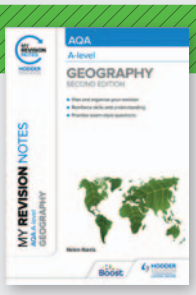


Aiming for an A

Master the skills you need to set yourself apart and hit the highest grades. This year-round course companion provides step-by-step guidance, examples and tips for getting an A grade.

Aiming for an A in A-level Geography 9781510424081

Suitable for all exam boards



My Revision Notes

Target success in AQA A-level Geography with this comprehensive revision guide; with key content, exam-style tasks and practical tips, you can rely on My Revision Notes to review, strengthen and test your knowledge.

My Revision Notes: AQA AS/A-level Geography 2nd edition | 9781398325500

Need to Know

The key facts at your fingertips: check content, review your understanding, use quick tips for success and improve your exam performance.

Need to Know: AQA A-level Geography | 9781510428485

For more information, visit
www.hoddereducation.co.uk/Geography
or contact Customer Services on 01235 827827

HODDER EDUCATION
LEARN MORE

AQA
A-level

Geography

For A-level

Fifth Edition

Malcolm Skinner
Paul Abbiss
Philip Banks
Helen Fyfe
Ian Whittaker

Approval message from AQA

This textbook has been approved by AQA for use with our qualification. This means that we have checked that it broadly covers the specification and we are satisfied with the overall quality. Full details of our approval process can be found on our website.

We approve textbooks because we know how important it is for teachers and students to have the right resources to support their teaching and learning. However, the publisher is ultimately responsible for the editorial control and quality of this book.

Please note that when teaching the ***AQA A-level Geography*** course, you must refer to AQA's specification as your definitive source of information. While this book has been written to match the specification, it cannot provide complete coverage of every aspect of the course.

A wide range of other useful resources can be found on the relevant subject pages of our website: www.aqa.org.uk.

Boost

HODDER EDUCATION
AN HACHETTE UK COMPANY

Acknowledgements

All acknowledgements can be found on p.690.

Maps included by kind permission of Ordnance Survey (OS) on **p.49, p.85, p.128, p.255, p.640** and **p.667**. Ordnance Survey (OS) is the national mapping agency for Great Britain, and a world-leading geospatial data and technology organisation. As a reliable partner to government, business and citizens across Britain and the world, OS helps its customers in virtually all sectors improve quality of life.

Every effort has been made to trace all copyright holders, but if any have been inadvertently overlooked, the Publishers will be pleased to make the necessary arrangements at the first opportunity.

Although every effort has been made to ensure that website addresses are correct at time of going to press, Hodder Education cannot be held responsible for the content of any website mentioned in this book. It is sometimes possible to find a relocated web page by typing in the address of the home page for a website in the URL window of your browser.

Hachette UK's policy is to use papers that are natural, renewable and recyclable products and made from wood grown in well-managed forests and other controlled sources. The logging and manufacturing processes are expected to conform to the environmental regulations of the country of origin.

Orders: please contact Hachette UK Distribution, Hely Hutchinson Centre, Milton Road, Didcot, Oxfordshire, OX11 7HH. Telephone: +44 (0)1235 827827. Email education@hachette.co.uk
Lines are open from 9 a.m. to 5 p.m., Monday to Friday. You can also order through our website: www.hoddereducation.co.uk

ISBN: 978 1 3983 1254 8

© Malcolm Skinner, Paul Abbiss, Phil Banks, Helen Fyfe and Ian Whittaker 2021

First published in 2016
This edition published in 2021 by
Hodder Education,
An Hachette UK Company
Carmelite House
50 Victoria Embankment
London EC4Y 0DZ

www.hoddereducation.co.uk

Impression number 10 9 8 7 6 5 4 3 2 1

Year 2025 2024 2023 2022 2021

All rights reserved. Apart from any use permitted under UK copyright law, no part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying and recording, or held within any information storage and retrieval system, without permission in writing from the publisher or under licence from the Copyright Licensing Agency Limited. Further details of such licences (for reprographic reproduction) may be obtained from the Copyright Licensing Agency Limited, www.cla.co.uk

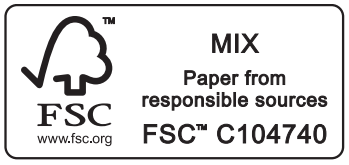
Cover photo © Cherries – stock.adobe.com

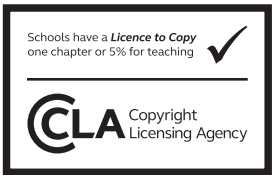
Illustrations by Barking Dog Art Design & Illustration and Aptara Inc.

Typeset in ITC Berkeley Oldstyle Std Book 11/14 pt by Aptara Inc.

Printed in Slovenia

A catalogue record for this title is available from the British Library.





Contents

Introduction v

Part 1 Physical Geography

Chapter 1 **Water and carbon cycles** 2
1.1 Systems frameworks and their application 2
1.2 The water cycle 4
1.3 The carbon cycle 24
1.4 Water, carbon, climate and life on Earth 41

Chapter 2 **Hot desert systems and landscapes** 53
2.1 Introduction to hot deserts and their margins 53
2.2 Systems and processes in hot deserts 61
2.3 Arid landscape development in contrasting parts of the world 70
2.4 Desertification 77
2.5 Quantitative and qualitative skills in hot deserts 82

Chapter 3 **Coastal systems and landscapes** 90
3.1 Introduction to coastal systems and landscapes 90
3.2 Systems and processes in coastal environments 92
3.3 Coastal landscape development 104
3.4 Coastal management 118
3.5 Quantitative and qualitative skills in coastal landscapes 126

Chapter 4 **Glacial systems and landscapes** 137
4.1 Introduction to glacial systems and landscapes 137
4.2 The nature and distribution of cold environments 138
4.3 Systems and processes in glaciers and glacial landscapes 143
4.4 Human occupation and development 166
4.5 Quantitative and qualitative skills in glaciated landscapes 177

Chapter 5 **Hazards** 185
5.1 The concept of a hazard in a geographical context 185
5.2 Plate tectonics 190
5.3 Volcanic hazards 197
5.4 Seismic hazards 204
5.5 Storm (typhoons, hurricanes) hazards 212
5.6 Fires in nature 219

Chapter 6 **Ecosystems under stress** 230
6.1 Ecosystems and processes 231
6.2 Ecosystems and sustainability 252
6.3 Marine ecosystems 259
6.4 Biomes 265

Part 2 Human Geography

Chapter 7 Global systems and global governance 283
7.1 Globalisation 283
7.2 Global systems 298
7.3 International trade and access to markets 306
7.4 Global governance 327
7.5 The 'global commons' 333
7.6 Globalisation critique 346

Chapter 8 Changing places 351
8.1 The nature and importance of places 351
8.2 Changing places: relationships and connections 361
8.3 Changing places: meaning and representation 364
8.4 Representations of place and the use of quantitative and qualitative sources 369
8.5 Place studies 377

Chapter 9 Contemporary urban environments 395
9.1 Patterns of urbanisation since 1945 395
9.2 Urban forms 407
9.3 Social and economic issues associated with urbanisation 415
9.4 Urban climate 422
9.5 Urban drainage 432
9.6 Urban waste and its disposal 437
9.7 Other contemporary urban environmental issues 443
9.8 Sustainable urban development 449

Chapter 10 Population and the environment 462
10.1 Introduction 462
10.2 Environment and population 466
10.3 Environment, health and well-being 485
10.4 Population change 506
10.5 Principles of population ecology applied to human populations 521
10.6 Global population futures 532

Chapter 11 Resource security 554
11.1 Resource development 554
11.2 Natural resource issues – Energy 564
11.3 Natural resource issues – Mineral ores 572
11.4 Natural resource issues – Water 578
11.5 Water security 585
11.6 Energy security 595
11.7 Mineral security 616
11.8 Resource futures 626

Part 3 Skills and Fieldwork Investigation

Chapter 12 Geographical skills and fieldwork 641
12.1 Qualitative skills and quantitative skills 641
12.2 Specific skills 645
12.3 Fieldwork investigation 664
Glossary 671
Index 681
Acknowledgements 690

Introduction

This textbook covers the subject content for AQA A-level Geography, as it is laid out in the specification document. This new edition brings the content up to date in order to keep pace with the changing world and to ensure it is even more in tune with students' and teachers' needs.

In response to teacher and student feedback:

- the content has been reviewed for quality and accuracy and updated accordingly
- case studies and examples have been updated and, where appropriate, replaced with more contemporary ones to reflect the dynamic nature of the subject
- shorter, more regular review questions have been included to prompt a check on learning
- practice questions have been added to reflect the style and structure of those set in A-level exams
- links between different topics have been identified
- suggestions for further reading have been updated to include new titles and articles.

The AQA A-level Geography curriculum encourages students to explore the interactions between the physical environment and human activities. Within this, you can investigate important contemporary issues at local, national and global scales in contexts specific to different parts of the world. The book forms the backbone for studies of AQA Geography, but it should also be supplemented with reference to current sources of information, including newspapers, television, periodicals aimed at post-16 geography students and, of course, the internet.

The following key features of the fourth edition have been retained:

- Concepts are clearly explained and related issues are explored and analysed.
- A variety of stimulus material is provided, including maps, graphs, diagrams and photographs.
- Key terms for each topic are placed in text boxes within the relevant chapter and are to be found in bold type within the text.
- Opportunities for fieldwork are suggested.
- The geographical skills required in the specification are covered in Chapter 12. In the other chapters, there will also be a focus on using particular skills in the context of the topic material.
- Review questions help reinforce your understanding of the information covered in the preceding section.

Understanding and preparing for assessment

Knowing and understanding the content of the core and your selected option topics is essential for the final assessment, and all the information that you need for the topics is covered in this textbook. In addition to knowing and understanding the content, to be truly successful in your final assessment, you need to know how to approach the examinations.

This introduction will help you understand and familiarise yourself with the structure of the exams, the types of questions that you will be asked and what examiners are seeking from your responses. If you are more aware of the **scheme of assessment**, **assessment objectives**, **command words** and how **links** between topics are used, it will improve your performance, as you will be able to apply your knowledge, understanding and skills more appropriately.

Scheme of assessment

The A-level course is designed to be studied over two years. The components of the assessment are as follows:

Component 1 (Physical geography)
- Written examination paper of 2 hours and 30 minutes
- Three sections:

Section A – Water and carbon cycles – core topic

All questions are **compulsory**. (36 marks)

Section B – either Hot desert systems and landscapes or Coastal systems and landscapes or Glacial systems and landscapes

You will have to answer questions on **one** of these topics. (36 marks)

Section C – either Hazards or Ecosystems under stress

You will have to answer questions on **one** of these topics. (48 marks)

Component 2 (Human geography)
- Written examination paper of 2 hours and 30 minutes
- Three sections:

Section A – Global systems and global governance – core topic

All questions are compulsory. (36 marks)

Section B – Changing places – core topic

All questions are compulsory. (36 marks)

Section C – either Contemporary urban environments or Population and the environment or Resource security

You will have to answer questions on **one** of these topics. (48 marks)

In the three sections of Components 1 and 2, question types will consist of some lower mark questions and some with a higher tariff requiring answers with more depth. Total marks for each paper are 120.

Component 3 - Geography fieldwork investigation
This is a non-examination assessment

- You must complete an individual investigation, which must include data collected in the field. The investigation must be based on a question or issue defined and developed by you and **must relate to any part of the specification content**.
- Your final report should be between 3,000 and 4,000 words in length and will be marked by your teachers; your whole centre will then be moderated by AQA.
- This component carries 60 marks in total.

Assessment objectives

The three assessment objectives (AOs) for A-level Geography are shown below. These AOs determine how your knowledge, understanding and skills (and application of these) are assessed for each question on the exam paper.
To get a clearer idea of how the AOs are allocated on the exam papers, see the typical assessment structure on page xi and also check mark schemes for past papers on the AQA website at www.aqa.org.uk. Here is a summary of each AO in practice.

Assessment Objective 1

- AO1 shows **knowledge** and **understanding**.
- You should learn concepts, theories, processes and case studies in detail to be able to reach higher AO1 marks.
- On lower mark (for example, 4-mark) questions, prefixed with command words such as 'explain' or 'outline', **all** of the marks are likely to be assessing AO1 knowledge and understanding, and your ability to recall concepts.
- There is no AO1 credit in the non-examination assessed fieldwork investigation.

Assessment Objective 2

- AO2 requires **application** of knowledge and understanding and is the key to achieving higher marks on most of the higher mark exam questions, prefixed with command words such as 'assess' or 'evaluate'.
- To access more AO2 marks, you are expected to apply your knowledge and respond to links identified by the question.
- You therefore need to consider a range of potential connections within and between different topics to be able to achieve the AO2 marks in higher mark questions. (This is covered in more detail in the 'Making connections and following links' section of this introduction on pages viii–x.)

Student advice on AO1 and AO2

For higher mark questions, the marks will usually be divided between AO1 and AO2. This means that the question is not only asking you to recall your knowledge and understanding but also to apply this to a new or 'novel' context posed by the question. It is important, therefore, to realise that on higher mark questions, you may only achieve around half the available marks for recalling learned material (AO1). To achieve the remaining marks, you must apply this recalled material to the question (AO2).

The use of case studies further illustrates the distinction between AO1 and AO2 marks. When answering higher mark questions, particularly those of 20 marks, recalling information based on your knowledge of a case study will only give AO1 credit. To gain AO2 marks you need to use case studies to support your answer in the context of the question. This is applying your knowledge and understanding and will gain AO2 marks.

Assessment Objective 3

- AO3 assesses your **use of skills** and is primarily used when assessing your fieldwork enquiry.
- You should gain competence in the full range of geographical skills listed on the checklist in the specification.
- AO3 marks are also available in Papers 1 and 2 where a question is linked to stimulus material, such as data, maps or graphs, to assess your statistical, cartographic or graphical skills.
- On Papers 1 and 2, usually **one** six-mark question for each topic will be devoted to a given stimulus resource and all the credit will be covered by AO3. In these questions, knowledge, understanding and application will not be credited as you will only be assessed on interpreting and analysing the resources.

Command words used in the examinations

One of the major challenges in any examination is interpreting the demands of the questions. Thorough revision is essential, but an awareness of what is expected from you is also required. Too often candidates attempt to answer the question which they think is there, rather than the one that has actually been set.

Correct interpretation of the **command words** is therefore vital to your success. In A-level Geography exam papers, a variety of command words are used. Some commands are more demanding than others; some require a relatively simple task to be performed; others are more complex, requiring greater thought and a more in-depth response. The information below offers advice on the command words used in the examinations:

Analyse ...

This is a commonly used command. It requires you to break down the given information on a topic or issue into its constituent elements, in order to provide an in-depth account and convey an understanding of it. You should look for connections in the information provided and causes and/or effects related to these connections. This command is often used for the 6-mark AO3 question in each topic where you are required to analyse resource material, such as data.

Annotate ...

You are required to add to a diagram, image or graphic material a number of words that describe and/or explain features, rather than just identify them (which is labelling).

Assess ... (see also Evaluate)

This is commonly used. It requires you to consider several options or arguments and weigh them up, in order to come to a conclusion about their effectiveness or validity. The additional commands of '*assess the extent*' or '*assess the relative importance*' require a similar judgement as to the merit or validity of a statement or view. All of these associated commands can often be used in questions that require you to make connections within or between topics and they usually require you to write extended prose.

Critically (assess or evaluate) ...

Sometimes occurs before 'assess' or 'evaluate'. It invites you to examine an issue from the point of view of a critic, with a particular focus on the strengths and weaknesses of the point of view being expressed. You should question not only the evidence presented but also its source and how it was collected.

Define ...

You have to state the precise meaning of an idea or concept. There is usually a low mark allocation for these questions, which mainly require a relatively short answer.

Describe ...

You are required to give a written account of something, which may be an entity, an event, a feature, a pattern, a distribution or a process. Explanations are **not** required.

Discuss ...

You are expected to set out both sides of an argument (for and against), making good use of evidence and appropriate examples, expressing an opinion on the merits of each side and finally coming to a conclusion. There should be some evidence of balance, though not necessarily of equal weighting. The command can be used in a variety of contexts such as: 'Discuss the extent to which...' where a judgement about the validity of the evidence is required (see 'To what extent…').

Evaluate ...

A similar command word to 'assess', it is used in questions with higher marks. To evaluate requires you to consider several options or arguments about an issue and to give an overall statement of value in your conclusion. As with 'assess' it can be used as '*Evaluate the extent…*' or '*Evaluate the importance…*' or '*Critically evaluate…*'.

Examine ...

You must consider carefully and provide a detailed account of the indicated topic. Similar to 'describe' and will usually be used on lower mark questions.

Explain ... /Why ... /Suggest reasons for ...

You are required to set out the causes of something and/or the factors which influence its form/nature. This usually requires an understanding of processes and as such, it will usually be assessing AO1 only and it will typically be found in the 4-mark questions.

Interpret ...

You are required to ascribe a meaning to something presented in the question. Usually you are required to show an understanding of the meaning of a set of data or other stimulus material.

Justify ...

You need to give reasons for the validity of a view or idea or state why some action should be undertaken. This will involve discussing and discounting alternative views or actions. Each of the views or optional actions will have positive and negative aspects. For the chosen outcome, the overall positives should outweigh the negatives.

Outline ... /Summarise ...

You are required to provide a brief account of relevant information. This command word will require a relatively short answer and will usually only be found in the 4-mark questions.

To what extent ... /How far do you agree ...

As with 'Assess' and 'Evaluate', you are required to use your judgement. It is sometimes used in conjunction with other command words, for example, '*Assess the extent…*'. The command is asking you to examine the available evidence, and/or different sides of an argument, and then form and express an opinion as to the merit or validity of a viewpoint or statement. 'To what extent' therefore often follows a given statement in the stem of the question as shown in this example:

'Strategies used to increase energy supply are more important for sustainability than those used to conserve energy consumption.'

To what extent **do you agree** with this view?

Making connections and following links

The specification states clearly that the A-level exams will include questions that: '**draw together skills, knowledge and understanding from the full course of study**'. When you are studying one aspect of geography (for example, the water cycle), it is important that you recognise that it has connections to other aspects of geography. You will be expected to demonstrate your knowledge and understanding of these connections and you will be assessed on your ability to do so by answering questions that have links identified within them. These links are **not signposted** in the specification, but we have helped you to identify some in this textbook through the 'links' feature box. You are expected to **apply** your knowledge and understanding to the link within the question. Linking questions, that assess your ability to make connections, are associated with AO2.

Links may be assessed in the following ways:

- **Stimulus material**, such as a map, photograph or data. You are required to apply your knowledge and understanding to answer the question about the stimulus material. Recalling your knowledge and understanding only gains AO1 credit. To gain AO2 marks you need to apply your knowledge and understanding to the given situation.
- Questions that make **links** within specification **topics**. For example, in the core topic 'Water and carbon cycles', questions about concepts or theories relating to the water cycle can be linked to climate change (Figure 0.1) or applied to the specified case study of a tropical rainforest within the same topic. For example: *'Evaluate the extent to which feedback systems in the water cycle will be affected by climate change.'*
- Questions that make **links across** specification **topics**. For example, in the option topic 'Hazards', questions about specific hazards, such as storm events, volcanic eruptions or wildfires, might be linked to an understanding of ideas, theories and concepts in the 'Water and carbon cycle' or to ideas in other core topics, such as 'Changing places' and 'Global systems and governance' (Figure 0.1). For example: *'Assess the importance of factors in globalisation in supporting the response to major seismic hazards.'*

Figure 0.2 shows the potential links **between core** and **option** topics. While this might look complex, linking questions will only ask you to make links across topics that **all** students have studied.

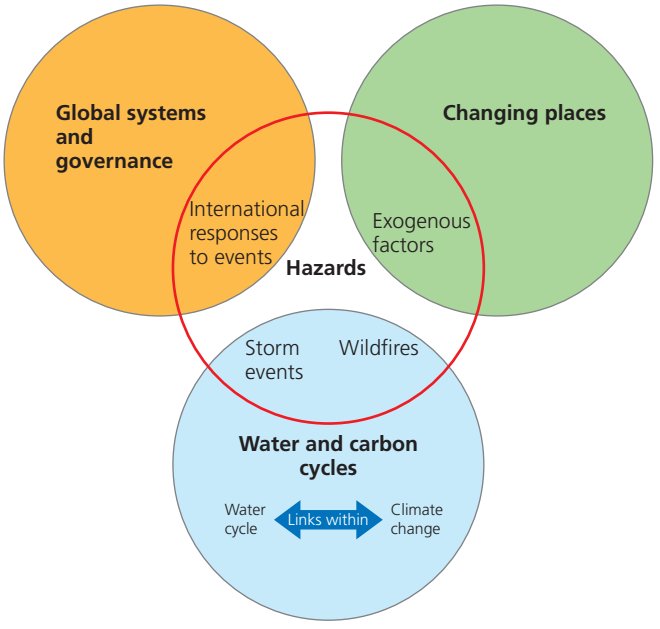


Figure 0.1 Venn diagram showing examples of links a) **within** the topic 'Water and carbon cycles' and b) **between** the **option topic** 'Hazards' and other **core topics** in the specification

- **Questions** within **core** topics on the exam paper may be linked to **other** core topics but will **not** be linked to option topics.
- Questions within **option** topics will **not** expect you to make links to other **option** topics but may require you to make a connection back to one of the **core** topics (Figure 0.2).

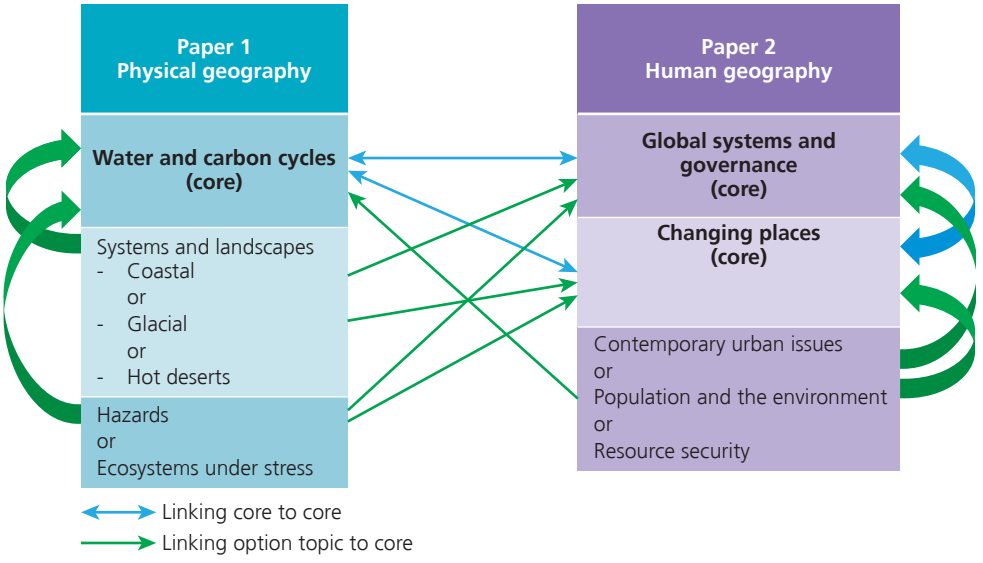


Figure 0.2 The possible links that could be assessed **across** core and option topics

Developing the ability, skills and confidence to address linking questions will come with practice and by taking a step-by-step approach. You can start to build the necessary skills by completing short exercises to identify links between different parts of a topic or across two different topics. For example, a good starting point is using the specification to draw your own mind map of all the different aspects within a topic. From this, you can then construct useful tools, such as Venn diagrams, (Figure 0.1, page ix), or a matrix (Table 0.1), to identify links both within a topic and across different topics.

For example, **within** the **option** topic 'Population and the environment' consider the links between 'natural population change' and 'improvements in healthcare'. They can both be linked to:

- economic and social development
- better food security
- demographic transition and epidemiological transition
- impacts on mortality, fertility and population structure.

Making connections **between** your **chosen option** topics and **core** topics may appear more difficult, but you can practise identifying these using a matrix, as shown in Table 0.1.

The partially completed matrix shows some links that can be made between the **option** topics 'Coastal systems and landscapes' and 'Contemporary urban environments' and each of the **core** study topics. You can construct a similar matrix to show links between your three selected option topics and the core topics. Remember that it is possible for connections to be made across two different core topics as well. An obvious example would be flows of people or flows of investment into an area ('Global systems'), both of which would be **exogenous** factors that would impact on a place ('Changing places').

Once you have more confidence in understanding links, start looking at past exam questions that require you to make connections and plan possible responses. For example, a question appears in the specimen assessment materials on the AQA website (www.aqa.org.uk) that links the destruction of ecosystems with globalisation. Look at the mark scheme for this question to understand more clearly how they are linked.

Questions in the Question practice sections of this textbook which require you to make connections are identified by this symbol: 🔗 This symbol also appears on the Links boxes throughout the text.

Essentially, making connections between different areas of your study, whether physical or human geography, requires you to '**think like a geographer**'.

Key term

Exogenous – External factors that affect the demographic, socio-economic and cultural characteristics of a place.

Table 0.1 A matrix to assist in making connections from option to core topics

Option/Core	Water and carbon cycles	Global systems and global governance	Changing places
Hot desert systems and landscapes			
Coastal systems and landscapes	Climate change > sea level change > impacts on coastal landforms > coastal management	Increased international trade > development of deep sea ports > impacts on coastal systems	Coastal landscape = endogenous factor Coastal management > built environment
Glacial systems and landscapes			
Hazards			
Ecosystems under stress			
Contemporary urban environments	Urban climate, heat islands etc. Urban run-off, drainage (SuDS) etc. River restoration and conservation. Flooding	Urbanisation as part of economic development. World cities and role in global economy. Flows of people and investment	Urban forms, pattern of land use etc. Social segregation, cultural diversity, gentrification, other urban processes
Population and the environment			
Resource security			

Review questions

1 Select one of the core topics (for example, Water and carbon cycle) and draw a mind map showing all the sub-topics that are included within it.

2 With reference to the mind map, draw a Venn diagram with one of the sub-topics (for example, The carbon cycle) as a central circle and overlap the other sub-topics, identifying links where they overlap.

3 Using Table 0.1 (page x), draw a matrix with your three option topics cross referenced with your three core topics. Complete the matrix by identifying links in each of the boxes. (Develop or use different links if you are using Coastal systems and/or Urban environments).

Specific concepts

In addition to the connections explained so far, you need to have a clear understanding of other specific concepts that are relevant to the core, and the non-core content of your studies, such as:

causality, **systems**, **equilibrium**, **feedback**, **inequality**, **representation**, **identity**, **globalisation**, **interdependence**, **mitigation**, **adaptation**, **sustainability**, **risk**, **resilience** and **thresholds**.

Other assessment aims are listed in Section 4.1 of the specification document, which can be found online.

Assessment structure

Understanding all the elements of the assessment and their interrelationship should give you a clearer picture of what to expect and how to approach your examinations. The box below gives you an overview of how the exam papers are typically structured in terms of mark allocation, command words and assessment objectives.

Please note: this is an example for general guidance. While it represents a typical structure, it will vary from year to year in terms of the use of command words and where the links appear.

PAPER 1			PAPER 2		
Marks	**Command word/ type of question**	**AO(s)**	**Marks**	**Command word/ type of question**	**AO(s)**
Section A – Water and carbon cycles			**Section A – Global systems and governance**		
4 marks	Explain ...	AO1	4 marks	Explain ...	AO1
6 marks	Analyse ... Resource to analyse	AO3	6 marks	Analyse ... Resource to analyse	AO3
6 marks	Assess ... Stimulus resource	AO1/ AO2	6 marks	To what extent ... Stimulus resource	AO1/ AO2
20 marks	To what extent ... (link within topic)	AO1/ AO2	20 marks	Assess the relative importance ... (link within topic)	AO1/ AO2
Section B – Landscape options			**Section B – Changing places**		
4 marks	Outline ...	AO1	4 marks	Explain ...	AO1
6 marks	Analyse ... Resource to analyse	AO3	6 marks	Analyse ... Resource to analyse	AO3
6 marks	Assess ... Stimulus resource	AO1/ AO2	6 marks	Suggest how ... Stimulus resource	AO1/ AO2
20 marks	To what extent ...	AO1/ AO2	20 marks	Critically assess ...	AO1/ AO2
Section C – Hazards or Ecosystem options			**Section C – Human options**		
4 marks	Summarise ...	AO1	4 marks	Outline ...	AO1
6 marks	Analyse ... Resource to analyse	AO3	6 marks	Analyse ... Resource to analyse	AO3
9 marks	Assess ... Stimulus resource	AO1/ AO2	9 marks	Evaluate ... Stimulus resource	AO1/ AO2
9 marks	To what extent ... (link within topic)	AO1/ AO2	9 marks	Assess the extent ... (link across topics)	AO1/ AO2
20 marks	Assess the view ...	AO1/ AO2	20 marks	To what extent do you agree ...	AO1/ AO2

Essay writing – producing extended prose

For many students, this represents the most challenging part of any examination, but it is an opportunity to demonstrate your strengths and gain marks. Before you start to write an essay, you must have a plan of what you will write, either in your mind or on paper. All essays must have a beginning (introduction), a middle section (arguments) and an ending (conclusion).

A written essay plan should be brief, taking only two or three minutes to write. The plan should outline a structure to your response, addressing the command word and key words within the question. Be logical and only write an outline plan. Make sure that you adhere to the plan. Keep the examples you have studied in mind to use at the most appropriate point in your answer.

The introduction

This does not have to be lengthy; a few sentences should be sufficient. It may define the terms set out in the question, set the context for the argument to follow and demonstrate any links that are to be made – initiating these at an early stage. It can also provide a brief statement of the idea or viewpoint to be developed in the main part of the answer.

The main part of the answer – arguments and judgements

Essay style questions will instruct you to 'assess' or 'evaluate', in some cases 'critically …' and in other cases 'to what extent …'. All these commands require you to consider options and/or arguments and to make judgements about whether they are valid, or the extent to which you agree or disagree with them.

The main section should consist of a series of paragraphs, each developing an argument or considering an option. Each argument that you make should be supported by evidence, from a case study or other example. Develop your answer by explaining how or why the evidence you have provided supports your argument. Remember to follow on logically by linking to the next point and/or back to the theme of the question. If your points can be pulled together and integrated clearly, it demonstrates a more sophisticated response that will gain more credit, particularly if the question requires you to make links.

The use of case studies is important and essential in those questions that require a case study prescribed by the specification. Make good use of exemplar material, naming real places and giving accurate details specific to those locations. Remember that using facts and information from your case study will achieve AO1 marks but using the information and facts as evidence to support your argument will gain AO2 marks.

Avoid paragraphs that just list information without expansion and avoid reciting all you know about a particular case study or topic without any link to the question. You should aim to demonstrate good, and often critical, knowledge and understanding of concepts.

The conclusion

You should reach a clear and definite conclusion or point of view that draws together your arguments and provides an overall statement. Be sure to address the command word/s given in the question, such as 'assess', or 'evaluate'. Reiterate the key points, evidence and facts provided in your essay to support this statement. The conclusion should contain concise detail rather than just a basic statement of agreement/disagreement with the question.

Other important points to consider:

- Always keep an eye on the time.
- Make sure that you write clearly and concisely.
- Avoid giving confused answers, producing endlessly long sentences or pages of prose with no paragraphs.
- Above all, read the question carefully and answer the question that has been asked.

Part 1

Physical Geography

Chapter 1 Water and carbon cycles
Chapter 2 Hot desert systems and landscapes
Chapter 3 Coastal systems and landscapes
Chapter 4 Glacial systems and landscapes
Chapter 5 Hazards
Chapter 6 Ecosystems under stress

Chapter

1 Water and carbon cycles

The cycling of water has obvious and significant implications for the health and prosperity of society. The availability and quantity of water is vital to life on Earth and helps to tie together the Earth's lands, oceans and atmosphere into an integrated physical system. Added to this is the fact that water vapour is the most important greenhouse gas and is a major driving factor in determining climate. The global water cycle is driven by many complex processes and interactions at a variety of scales; these are often poorly understood and badly represented in model predictions.

Carbon is everywhere: in the oceans, in rocks and soils, in all forms of life and in our atmosphere. Without carbon, life would not exist as we know it. The well-being and functioning of our planet depends on carbon and how it cycles through the Earth's system. The carbon cycle plays a key role in regulating the Earth's global temperature and climate by controlling the amount of another greenhouse gas, carbon dioxide, in the atmosphere.

Both the water and carbon cycles (and other aspects of geography) can be studied by considering them as systems.

In this chapter you will study:

- systems frameworks and their application
- the water cycle
- the carbon cycle
- water, carbon, climate and life on Earth

1.1 Systems framework and their application

Key terms

Energy – The ability to do work. In physical geography, much of this energy ultimately comes from the Sun.

Flow/transfer – A form of linkage between one store/component and another that involves movement of energy or mass.

Input – The addition of matter and/or energy into a system.

Output – The results of the processes within a system.

Store/component – A part of the system where energy/mass is stored or transformed.

System – A set of interrelated components working together towards some kind of process.

Because the Earth is highly complex, geographers have attempted to simplify aspects of it so that relationships between components can be better understood. These simplifications are called models (for example, the water cycle or the demographic transition model). One type of model that is widely used, particularly in physical geography, is the **system**.

A system is an assemblage of interrelated parts that work together by way of some driving process. They are a series of **stores** or **components** that have **flows** or **connections** between them. There are three types of property:

- **elements:** things that make up the system of interest
- **attributes:** the perceived characteristics of the elements
- **relationships:** descriptions of how the various elements (and their attributes) work together to carry out some kind of process.

Most systems:

- have a structure that lies within a **boundary**.
- are generalisations of reality, removing incidental detail that obscures fundamental relationships.

- function by having **inputs** and **outputs** of material (**energy** and/or matter) that is processed within the components causing it to change in some way.
- involve the flow of material between components.

Systems can be classified as:

- **Isolated systems**: these have no interactions with anything outside the system boundary. There is no input or output of energy or matter. Many controlled laboratory experiments are this type of system and they are rare in nature.
- **Closed systems**: these have transfers of energy both into and beyond the system boundary but not transfer of matter (Figure 1.1).
- **Open systems**: these are where matter and energy can be transferred from the system across the boundary into the surrounding environment. Most ecosystems are examples of open systems (Figure 1.2).

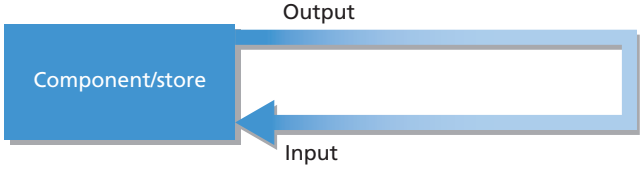


Figure 1.1 A closed system

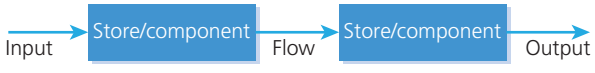


Figure 1.2 An open system (e.g. an ecosystem)

When there is a balance between the inputs and outputs then the system is said to be in a state of **dynamic equilibrium.** This means that the stores stay the same. If, however, one of the elements of the system changes, for example, one of the inputs increases without any corresponding change in the outputs, then the stores change and the equilibrium is upset. This is called feedback**.** There are two types of feedback:

- **positive feedback** where the effects of an action (for example, an increase in carbon dioxide) are amplified or multiplied by subsequent knock-on or secondary effects (Figure 1.3)
- **negative feedback** where the effects of an action (for example, the increased use of fossil fuels) are nullified by its subsequent knock-on effects (Figure 1.4).

The Earth can be studied using a systems approach; indeed the Earth as a whole could be considered a closed system. Energy comes into the system in the form of solar energy. This is balanced by radiant energy lost by the Earth. It could be related to the

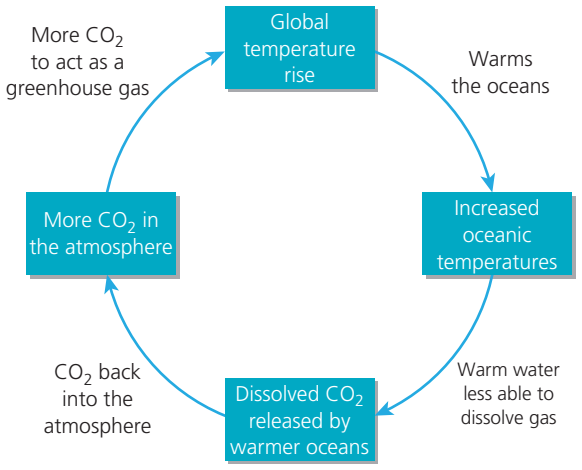


Figure 1.3 Example of positive feedback in a system

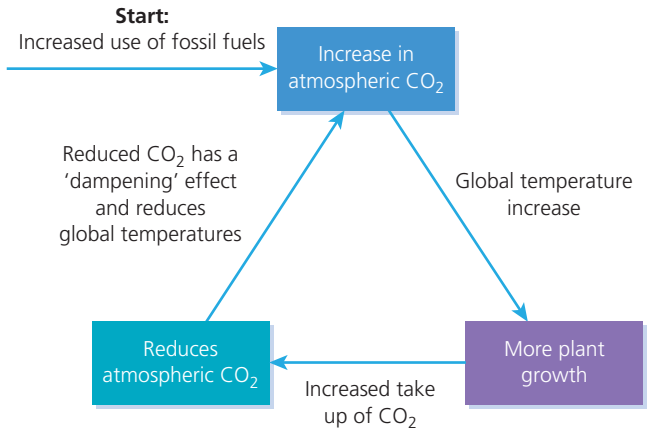


Figure 1.4 Example of negative feedback in a system

concept of 'spaceship Earth' which is a term usually expressing concern over the use of limited resources available on Earth and encouraging everyone on it to act as a harmonious crew working towards the greater good.

At the global level the Earth has four major subsystems, including the atmosphere, lithosphere, hydrosphere and biosphere. Each of these can be considered to be an open system that forms part of a chain: a **cascading system**.

Interlocking relationships among the atmosphere, lithosphere, hydrosphere and biosphere have a profound effect on the Earth's climate and climate change.

Review questions

1 Give two examples of:
   - physical geography systems
   - human geography systems.

2 For each example, consider the nature of any feedback that may apply to that system.

1.2 The water cycle

Key terms

Atmospheric water – Water found in the atmosphere; mainly water vapour with some liquid water (cloud and rain droplets) and ice crystals.

Cryospheric water – The water locked up on the Earth's surface as ice.

Discharge – The amount of water in a river flowing past a particular point expressed as m^3s^{-1} (cumecs).

Greenhouse gas – Any gaseous compound in the atmosphere that allows short wave ultraviolet radiation from the Sun to pass through the atmosphere, but then prevents outgoing terrestrial infrared radiation from escaping to space.

Hydrosphere – A discontinuous layer of water at or near the Earth's surface. It includes all liquid and frozen surface waters, groundwater held in soil and rock and atmospheric water vapour.

Oceanic water – The water contained in the Earth's oceans and seas but not including such inland seas as the Caspian Sea.

Terrestrial water – This consists of groundwater, soil moisture, lakes, wetlands and rivers.

'Water is life's matter, mother and medium.'

Albert Szent-Gyorgyi, 1937 Nobel Prize acceptance speech

The major stores of water

Water on or close to the Earth's surface is called the **hydrosphere**. Scientists have attempted many times to estimate the total amount of water in the hydrosphere. There is general agreement that it amounts to some 1.338×10^9 km^3. It is thought that approximately 97 per cent of this is **oceanic water** (Figure 1.5). Fresh water, which makes up the remaining 3 per cent, is locked up in land ice, glaciers and permafrost (**cryospheric water**), groundwater, lakes, soil, wetland, rivers, biomass (**terrestrial water**) and **atmospheric water**.

12,900 km^3 of water vapour are found in the atmosphere. This amounts to a global average of 26 kg/m^2 of water for each column of air on the surface of the Earth. There are large variations in this, however. Although atmospheric water only makes up 0.4 per cent of all water, it has a profound effect on our lives at present.

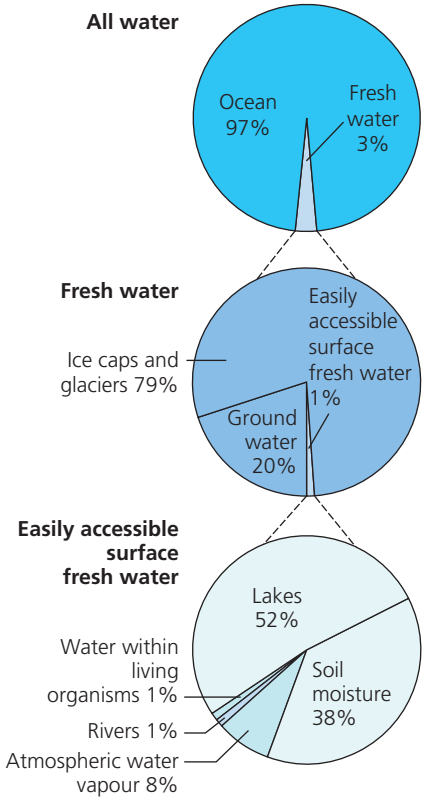


Figure 1.5 The distribution of the world's water

The amount of water in these stores is in a state of dynamic equilibrium with changes at a range of timescales from diurnal to geological. Changing amounts of atmospheric water in the future could be a major cause and/or important effect of climate change.

Oceanic water

The oceans dominate the amount of available water. Its exact amount is unknown with figures varying from 1,320,000,000 to 1,370,000,000 km^3 with an average depth of 3,682 m. That difference is greater than the sum of all the rest of the water put together. They cover approximately 72 per cent of the planet's surface (3.6×10^8 km^2). They are customarily divided into several principal oceans and smaller seas. Although the ocean contains 97 per cent of the Earth's water, oceanographers have stated that only 5 per cent has been explored.

Oceanic water tastes salty because it contains dissolved salts. These salts allow it to stay as liquid water below 0°C. They are alkaline with an average pH of about 8.14. The pH has fallen from about 8.25 in the last 250 years and it seems destined to continue falling. This change in the pH is linked to the increase in atmospheric carbon and may have a profound influence on marine ecosystems.

The cryosphere

The cryosphere is those portions of the Earth's surface where water is in solid form. Figure 1.6 shows the five locations of cryospheric water.

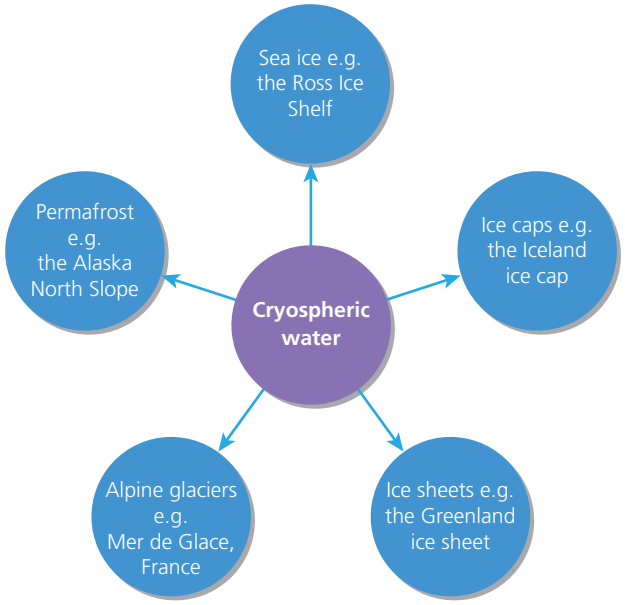


Figure 1.6 The locations of cryospheric water

Sea ice

Much of the Arctic Ocean is frozen; the amount of which grows in winter and shrinks in summer. The same is true of the waters surrounding Antarctica. Sea ice forms when water in the oceans is cooled to temperatures below freezing. Sea ice does not raise sea level when it melts, because it forms from ocean water. It is closely linked with our planet's climate, so scientists are concerned about its recent decline.

Figure 1.7 Chunks of broken sea ice in Yelverton Bay, Ellesmere Island, Canada

Ice shelves are platforms of ice that form where ice sheets and glaciers move out into the oceans. Ice shelves exist mostly in Antarctica and Greenland, as well as in the Arctic near Canada and Alaska. Icebergs are chunks of ice that break off glaciers and ice shelves and drift in the oceans. They raise sea level only when they first leave land and push into the water, but not when they melt in the water.

Ice sheets

An ice sheet is a mass of glacial land ice extending more than 50,000 km^2. The two major ice sheets on Earth today cover most of Greenland and Antarctica. During the last ice advance, ice sheets also covered much of North America, northern Europe and Argentina.

Figure 1.8 Mountains rising out of part of the Greenland ice sheet

Together, the Antarctic and Greenland ice sheets contain more than 99 per cent of the freshwater ice on Earth. The Antarctic Ice Sheet extends almost 14 million km^2, roughly the area of the United States and Mexico combined. It contains 30 million km^3 of ice. The Greenland Ice Sheet extends about 1.7 million km^2, covering most of the island of Greenland.

Ice sheets form in areas where snow that falls in winter does not melt entirely over the summer. Over thousands of years, the layers of snow pile up into thick masses of ice, growing thicker and denser as the weight of new snow and ice layers compresses the older layers. Ice sheets are constantly in motion, slowly flowing downhill under their own weight. Near the coast, most of the ice moves through relatively fast-moving outlets called ice streams. This type of glacier is significant in the Antarctic

where they can be up to 50 km wide, 2 km thick and hundreds of kilometres long. As long as an ice sheet accumulates the same mass of snow as it loses to the sea, it remains stable.

Ice sheets contain enormous quantities of frozen water. If the Greenland Ice Sheet melted, scientists estimate that sea level would rise about six metres. If the Antarctic Ice Sheet melted, sea level would rise by about 60 m.

Ice caps

Ice caps are thick layers of ice on land that are smaller than 50,000 km². They are usually found in mountainous areas. Ice caps tend to be dome-shaped and are centred over the highest point of an upland area. They flow outwards, covering almost everything in their path and becoming the major source for many glaciers.

Ice caps occur all over the world, from the polar regions to mountainous areas such as the Himalayas, the Rockies, the Andes and the Southern Alps of New Zealand. The Furtwangler Glacier on Kilimanjaro, at 60,000 m², is Africa's only remaining ice cap. It is melting rapidly and may soon disappear.

Figure 1.9 The Furtwangler ice cap. The last ice cap in Africa

Alpine glaciers

Alpine glaciers are thick masses of ice found in deep valleys or in upland hollows. Most valley glaciers are fed by ice from ice caps or smaller corrie glaciers. These glaciers are particularly important in the Himalayas where about 15,000 Himalayan glaciers form a unique reservoir which supports perennial rivers such as the Indus, Ganges and Brahmaputra which, in turn, are the lifeline of millions of people in South Asian countries (Pakistan, Nepal, Bhutan, India and Bangladesh).

Permafrost

Permafrost is defined as ground (soil or rock and included ice or organic material) that remains at or below 0°C for at least two consecutive years. The thickness of permafrost varies from less than one metre to more than 1,500 m. Most of the permafrost existing today formed during cold glacial periods and has persisted through warmer interglacial periods, including the Holocene (the last 10,000 years). Some relatively shallow permafrost (30 to 70 m) formed during the second part of the Holocene (the last 6,000 years) and some during the Little Ice Age (from 400 to 150 years ago). Subsea permafrost occurs at close to 0°C over large areas of the Arctic continental shelf, where it formed during the last glacial period on the exposed shelf landscapes when sea levels were lower. Permafrost is found beneath the ice-free regions of the Antarctic continent and also occurs beneath areas in which the ice sheet is frozen to its bed.

The permafrost has begun to melt as climate warms. This melting is releasing large amounts of carbon dioxide and methane, potentially affecting global climates.

Terrestrial water

Terrestrial water falls into four broad classes:

- surface water
- groundwater
- soil water
- biological water.

Surface water

Surface water is the free-flowing water of rivers as well as the water of ponds and lakes.

- **Rivers** act as both a store and a transfer of water; they are streams of water within a defined channel. They transfer water from the ground, from soils and from the atmosphere to a store. That store may be wetlands, lakes or the oceans. Rivers make up only a small percentage (0.0002 per cent) of all water, covering just 1,000,000 km² with a volume of 2,120 km³. One river alone, the Amazon in South America, is the largest river in the world by **discharge** of water, averaging a discharge of

about 209,000 m³/s, greater than the next seven largest independent rivers combined. It drains an area of about 7,050,000 km² and accounts for approximately one-fifth of the world's total river flow. The portion of the river's drainage basin in Brazil alone is larger than any other river's basin. The Amazon enters Brazil with only one-fifth of the flow it finally discharges into the Atlantic Ocean, yet already has a greater flow at this point than the discharge of any other river.

- **Lakes** are collections of fresh water found in hollows on the land surface. They are generally deemed a lake if they are greater than two hectares in area. Any standing body of water smaller than this is termed a pond.

  The majority of lakes on Earth are freshwater, and most lie in the Northern Hemisphere at higher latitudes. Canada has an estimated 31,752 lakes larger than 3 km² and an estimated total number of at least 2 million. Finland has 187,888 lakes 500 m² or larger, of which 56,000 are large (10,000 m²).

  The largest lake is the Caspian Sea at 78,200 km³. It is a remnant of an ancient ocean and is about 5.5 million years old. It is generally fresh water, though becomes more saline in the south where there are few rivers flowing into it. The deepest lake in the world is Lake Baikal in Siberia with a mean depth of 749 m and a deepest point at 1,637 m.
- **Wetlands:** The **Ramsar Convention** defines wetlands as 'areas of marsh, fen, peatland or water, whether natural or artificial, permanent or temporary, with water that is static or flowing where there is a dominance by vegetation'.

  They are areas where water covers the soil, or is present either at or near the surface of the soil all year or for varying periods of time during the year, including during the growing season. Water saturation determines how the soil develops and the types of plant and animal communities living in and on the soil. Wetlands may support both aquatic and terrestrial species. The prolonged presence of water creates conditions that favour the growth of specially adapted plants and promotes the development of characteristic wetland soils.

  Wetlands vary widely because of regional and local differences in soils, topography, climate, hydrology, water chemistry, vegetation and other factors, including human disturbance. They are found from the polar regions to the tropics and on every continent except Antarctica.

The Pantanal wetlands

The Pantanal of South America is often referred to as the world's largest freshwater wetland system. It extends through millions of hectares of central western Brazil, eastern Bolivia and eastern Paraguay (Figure 1.10).

It is a complex system of marshlands, flood plains, lagoons and interconnected drainage lines. It also provides economic benefits by being a huge area for water purification and groundwater discharge and recharge, climate stabilisation, water supply, flood abatement, and an extensive, transport system, among numerous other important functions.

Figure 1.10 The location of the Pantanal wetlands (shown in dark green)

Wetlands are the main ecosystem in the Arctic. These peatlands, rivers, lakes, and shallow bays cover nearly 60 per cent of the total surface area. Arctic wetlands store enormous amounts of greenhouse gases and are critical for global biodiversity.

Groundwater (lithosphere)

Groundwater is water that collects underground in the pore spaces of rock (Figure 1.11, page 8). Scientists have set a lower level for groundwater at a depth of 4,000 m but it is known that there are large quantities of water below that. A very deep borehole in the Kola Peninsula in Northern Russia found huge quantities of hot mineralised water at a depth of 13 km.

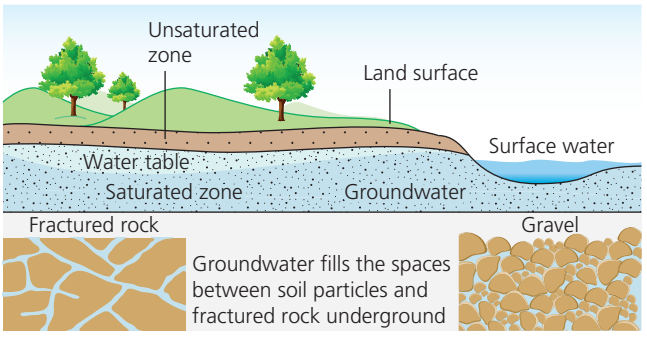


Figure 1.11 Groundwater

The depth at which soil pore spaces or fractures and voids in rock become completely **saturated** with water is called the water table. Groundwater is recharged from, and eventually flows to, the surface. Natural discharge often occurs at springs and seeps, and can form oases or wetlands. The amount of groundwater is reducing rapidly due to extensive extraction for use in irrigating agricultural land in dry areas.

Soil water

Soil water is that which is held, together with air, in unsaturated upper weathered layers of the Earth. It is of fundamental importance to many hydrological, biological and biogeochemical processes. It affects weather and climate, **run-off** potential and flood control, soil erosion and slope failure, reservoir management, geotechnical engineering and water quality. Soil moisture is a key variable in controlling the exchange of water and heat energy between the land surface and the atmosphere through **evaporation** and plant **transpiration**. As a result, soil moisture plays an important role in the development of weather patterns and the production of precipitation.

Biological water

Biological water constitutes the water stored in all the biomass. It varies widely around the globe depending on the vegetation cover and type. Areas of dense rainforest store much more water than deserts. The role of animals as a water store is minimal.

Trees take in water via their roots. This is either transported or stored in the trunk and branches of the tree. The water is lost by the process of transpiration through stomata in the leaves. This storage provides a reservoir of water that helps maintain some climatic environments. If the vegetation is destroyed, this store is lost to the atmosphere and the climate can become more desert-like. Many plants are adapted to store water in large quantities. Cacti are able to gather water via their extensive root system and then use it very slowly until the next rainstorm. The baobab tree stores water, but it is thought that this is to strengthen the structure of the tree rather than to be used in tree growth.

The atmosphere

Atmospheric water exists in three states. The most common atmospheric water exists as a gas: water vapour. This is clear, colourless and odourless and so we take its presence for granted. This atmospheric water vapour is important as it absorbs, reflects and scatters incoming solar radiation, keeping the atmosphere at a temperature that can maintain life. The amount of water vapour that can be held by air depends upon its temperature. Cold air cannot hold as much water vapour as warm air. This results in air over the poles being quite dry, whereas air over the tropics is very humid.

A small increase in water vapour will lead to an increase in atmospheric temperatures. This becomes positive feedback as a small increase in global temperature would lead to a rise in global water vapour levels, thus further enhancing the atmospheric warming.

Cloud is a visible mass of water droplets or ice crystals suspended in the atmosphere. Cloud formation is the result of air in the lower layers of Earth's atmosphere becoming saturated due to either or both of two processes: cooling of the air and an increase in water vapour. When the cloud droplets grow they can eventually fall as rain.

Factors driving the change in magnitude of water stores

Water exists on Earth in three forms: liquid water, solid ice and gaseous water vapour. Figure 1.12 shows the processes that occur as water changes from one state to another. Energy, in the form of latent heat, is either absorbed or released depending on the process. This is particularly important in atmospheric processes such as cloud or precipitation formation.

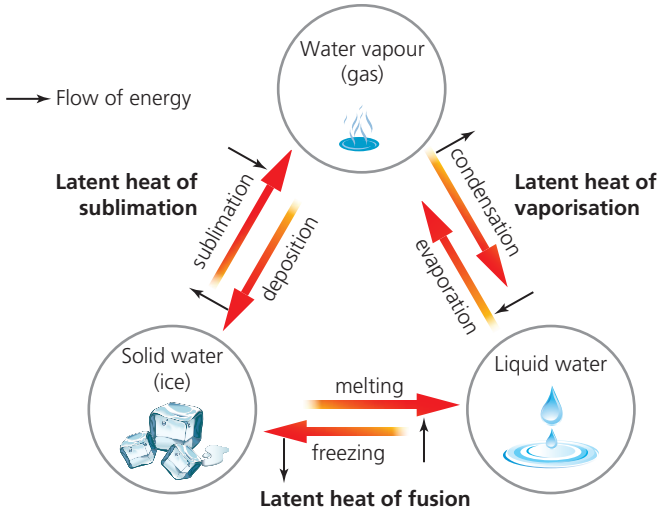


Figure 1.12 The phases of water and the phase changes

Evaporation

Evaporation occurs when energy from solar radiation hits the surface of water or land and causes liquid water to change state from a liquid to a gas (water vapour). The rate of evaporation depends upon:

- the amount of solar energy
- the availability of water (for example, there is more evaporation from a pond than from a grassy field).
- the humidity of the air; the closer the air is to saturation point, the slower the rate of evaporation
- the temperature of the air; warmer air can hold more water vapour than cold air.

Linked to this is the fact that all terrestrial plants lose water through **transpiration**. This is where water is transported from the roots of a plant to its leaves and then lost through pores on the leaf surface. Leaves also intercept rain as it falls, and this water can be evaporated before it reaches the soil.

As water evaporates it uses energy in the form of latent heat and so cools its surroundings.

Condensation

As air cools it is able to hold less water vapour. This means that if it is cooled sufficiently then it will get to a temperature at which it becomes saturated. This is known as the dew point temperature. Excess water in the air will then be converted to liquid water in the process of **condensation**. The water molecules need something to condense on. These can be tiny particles (smoke, salt, dust, etc.) that are collectively called condensation nuclei, or surfaces (leaves, grass stems, windows, etc.) that are below the dew point temperature. If the surface is below freezing point then water vapour sublimates, changing directly from gas to solid in the form of hoar frost.

Condensation is the direct cause of all forms of precipitation.

- It takes place when the temperature of air is reduced to dew point but its volume remains constant. This occurs when:
  - warm moist air passes over a cold surface
  - on a clear winter's night heat is radiated out to space and the ground gets colder, cooling the air directly in contact with it.
- It also occurs when the volume of air increases but there is no addition of heat (adiabatic cooling). This happens when air rises and expands in the lower pressure of the upper atmosphere. This can occur when:
  - air is forced to rise over hills (relief or orographic effect)
  - masses of air of different temperatures and densities meet. The less dense warm air rises over the denser cold air (frontal effect)
  - localised warm surfaces heat the air above. This expands, becomes less dense and rises (convectional effect).

Review questions

3 Rain, drizzle, snow, sleet and hail are all forms of precipitation. What are the key differences in their formation?

4 What are the similarities and differences in the formation of dew and hoar frost?

Cryospheric processes

Chapter 4 discusses in detail the two main **cryospheric processes** of **accumulation** and **ablation** (melting) on page 145, and slope processes on page 160.

It is thought that there have been five major glacial periods in Earth's history. The most recent started 2.58 million years ago and continues today; it is called the **Quaternary** glaciation. During this time there were:

- glacial periods when, due to the volume of ice on land, sea level was approximately 120 m lower than present and continental glaciers covered large parts of Europe, North America, and Siberia. This represents an interruption in the global hydrological cycle

- interglacial periods when global ablation exceeds accumulation and the hydrological cycle as we know it today returns.

Over the past 740,000 years there have been eight such glacial cycles.

Permafrost is formed when air temperatures are so low that they freeze any soil and groundwater present. It rarely occurs under ice because the temperatures are not low enough.

Drainage basin systems

Key terms

Condensation – The process by which water vapour changes to liquid water.

Cryospheric processes – Those processes that affect the total mass of ice at any scale from local patches of frozen ground to global ice amounts. They include accumulation (the build-up of ice mass) and ablation (the loss of ice mass).

Drainage basin – This is an area of land drained by a river and its tributaries. It includes water found on the surface, in the soil and in near-surface geology.

Evaporation – The process by which liquid water changes to a gas. This requires energy, which is provided by the Sun and aided by wind.

Evapotranspiration – The total output of water from the drainage basin directly back into the atmosphere.

Groundwater flow – The slow movement of water through underlying rocks.

Infiltration – The downward movement of water from the surface into soil.

Interception store – The precipitation that falls on the vegetation surfaces (canopy) or human-made cover and is temporarily stored on these surfaces. Intercepted water either can be evaporated directly to the atmosphere, absorbed by the canopy surfaces or ultimately transmitted to the ground surface.

Overland flow – The tendency of water to flow horizontally across land surfaces when rainfall has exceeded the infiltration capacity of the soil and all surface stores are full to overflowing.

Percolation – The downward movement of water within the rock under the soil surface. Rates vary depending on the nature of the rock.

Run-off – All the water that enters a river channel and eventually flows out of the drainage basin.

Saturated – This applies to any water store that has reached its maximum capacity.

Stemflow – The portion of precipitation intercepted by the canopy that reaches the ground by flowing down stems, stalks or tree boles.

Storm and rainfall event – An individual storm is defined as a rainfall period separated by dry intervals of at least 24 hours and an individual rainfall event is defined as a rainfall period separated by dry intervals of at least 4 hours (Hamilton and Rowe, 1949).

Throughfall – The portion of the precipitation that reaches the ground directly through gaps in the vegetation canopy and drips from leaves, twigs and stems. This occurs when the canopy-surface rainwater storage exceeds its storage capacity.

Throughflow – The movement of water downslope through the subsoil under the influence of gravity. It is particularly effective when underlying permeable rock prevents further downward movement.

Transpiration – The loss of water from vegetation through pores (stomata) on their surfaces.

Water balance – The balance between inputs (precipitation) and outputs (run-off, evapotranspiration, soil and groundwater storage) in a drainage basin.

Drainage basins

A drainage basin (or catchment area) is the area that supplies a river with its supply of water (see Figure 1.13). This includes water found below the water table as well as soil water and any surface flow. Drainage basins are separated from one another by high land called a watershed.

A useful way of looking at **drainage basins** is to consider them as cascading systems. They are a series of open systems that link together so that the output of one is the input of the next.

Table 1.1 shows the inputs, stores, transfers and outputs of a drainage basin system.

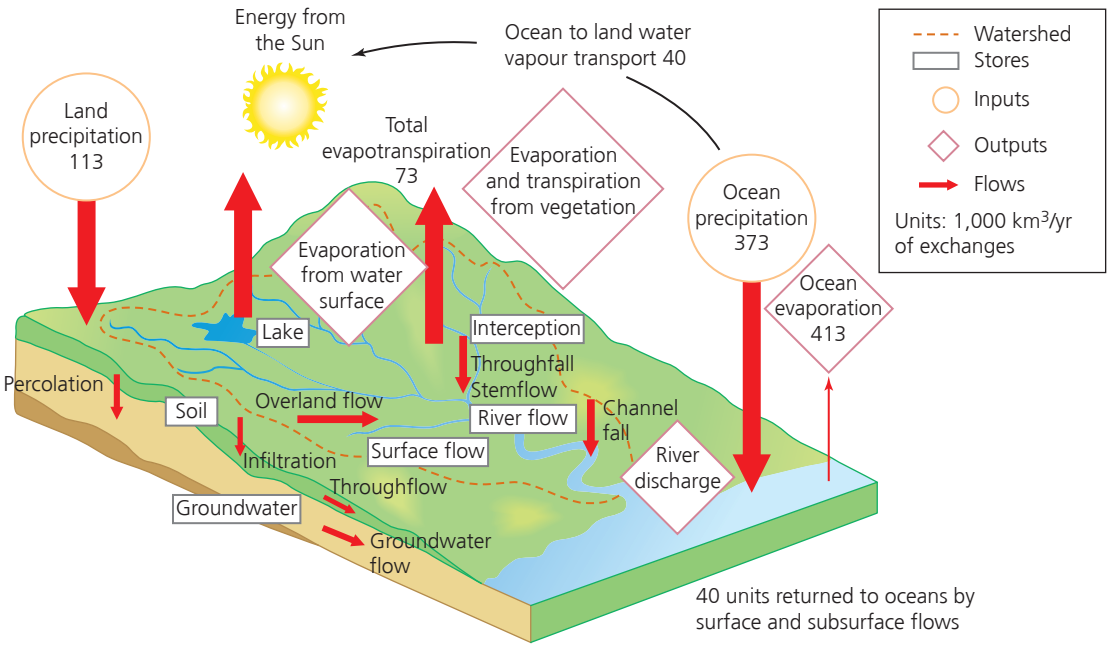


Figure 1.13 The drainage basin hydrological cycle

Table 1.1 Inputs, stores, transfers and outputs of a drainage basin system

Inputs	Stores	Transfers	Outputs
Precipitation onto land	Lakes and surface water	Overland flow	Evaporation and transpiration from vegetation
Precipitation onto the sea	River channels	Channel flow	Evaporation from water surfaces
	Interception from plants	Throughflow Stemflow	Run-off from rivers
	Soil water	Infiltration	Evaporation from the sea
	Ground water	Percolation	
		Throughflow	
		Ground water flow	

The input to the system is precipitation. The nature, intensity and longevity of precipitation have a direct bearing on what happens when the water hits the ground. In general, however, on a hill slope the following happens:

- Precipitation lands on the bare surface or, more likely, vegetation cover. This vegetation provides a store: **interception store**. This occurs on the leaves/branches of vegetation. Some vegetation is a better interception store than others. The density of the vegetation cover also plays a part. Studies show that forests made up of needle-leaf trees captured 22 per cent of the rainfall, while broad-leaf deciduous forests intercepted 19 per cent. This difference may be due to the density of the vegetation cover rather than the structure of the leaves. Some tropical rainforests intercept as much as 58 per cent of the rainfall.
- A lot of the water captured by vegetation surfaces is evaporated back into the atmosphere.
- Water making its way from the leaves of trees to the ground drips or falls by the process of **throughfall** from one interception store to another until it eventually reaches the ground. Water also flows down the stems of grasses, etc., and in very heavy storms it can flow straight down the trunks of trees. This is called **stemflow**.
- On reaching the ground, water soaks into the soil by the process of **infiltration**. The rate of this movement is called the **infiltration rate**. This movement of water into the **soil** is controlled by gravity, capillary action and soil porosity. Of these factors soil porosity is most important. A soil's porosity is controlled by its texture, structure and organic content. Coarse textured soils have larger pores and fissures than fine-grained soils

and therefore allow for more water flow. Pores and fissures found in soils can be made larger through a number of factors that enhance internal soil structure. For example, the burrowing of worms and other organisms and penetration of plant roots can increase the size and number of macro- and micro-channels within the soil.

Fieldwork opportunities

There is an opportunity here for an investigation based on the infiltration rates of different surfaces.

The rate of infiltration normally declines rapidly during the early part of a **rainstorm event** and reaches a constant value after several hours of rainfall. A number of factors are responsible for this phenomenon:

- The filling of small pores on the soil surface with water reduces the ability of capillary forces to actively move water into the soil.
- As the soil moistens, the clay particles absorb water causing them to expand. This expansion reduces the size of soil pores.
- Raindrop impact breaks large soil clumps into smaller particles. These particles then clog soil surface pores reducing the movement of water into the soil.

- **Soil storage** is the amount of water stored in the soil. Soils consist of solid particles with pore spaces between them. Those pore spaces can be filled with air but also with water. The amount of pores in a soil is different for different types of soil. The pores in a clay soil account for 40 to 60 per cent of the volume. In fine sand this can be 20 to 45 per cent.
- Vegetation requires water to survive. Most plants remove water from the soil and store it in the structure of the plant. This is called **vegetation storage**. Plants lose water back to the atmosphere through stomata on their leaves. This is called **transpiration.**
- If rainfall intensity is greater than the infiltration rate then the soil has reached **infiltration capacity** and the soil will be **saturated**. Water will build up on the surface as **surface storage** which is usually in the form of puddles. While common in a man-made environment their natural occurrence is quite rare. Water usually disappears into the ground at a greater rate than it rains (that is, the infiltration rate is greater than the rate of precipitation). Water can only build up on the surface after a long period of rain, an intense rainstorm or on an impermeable surface (either man-made, for example, an impacted footpath; or natural, for example, a bare rock or frozen surface).
- Much of this surface storage then **evaporates** back into the atmosphere and is lost to the drainage basin. It is difficult to separate evaporation from transpiration and so the total amount outputted from the system in this way is referred to as **evapotranspiration.** When the surface stores are full then **overland flow** or **sheet flow** will begin on slopes. This is very fast flow, rapidly reaching the nearest channel.
- Any lateral movement of soil water downslope is called **throughflow** and eventually reaches the nearest channel. This tends to be much slower than overland flow. Generally speaking, the more vegetated an area, the faster the rate of throughflow because it is aided by root channels in the soil.
- Following infiltration, water moves vertically down through the soil and unsaturated rock by the process of **percolation.** This can then be held in pore spaces in the rocks as **groundwater.** It then passes slowly into the zone of saturated rock where it can move vertically and laterally by the process of **groundwater flow.** This is very slow movement. It can feed rivers through long periods of drought. Some rocks are able to store a lot of water, especially if they are porous. These are called aquifers.

The sum of all these movements and stores adds up to form the drainage basin hydrological cycle. All of these flows lead water to the nearest river (or deep groundwater store). The river then transfers water by **channel flow.** The amount of water that leaves the drainage basin in this way is called **run-off.**

The rate of movement of all this water varies as shown in Figure 1.14. The fastest movement of water is along the surface. There are relatively few obstacles slowing it down. Urban surfaces have especially fast water movement because they are often designed to move the water quickly by having strategically placed slopes and very smooth surfaces. Water moves through soil at a slower rate. Again this varies greatly. Under

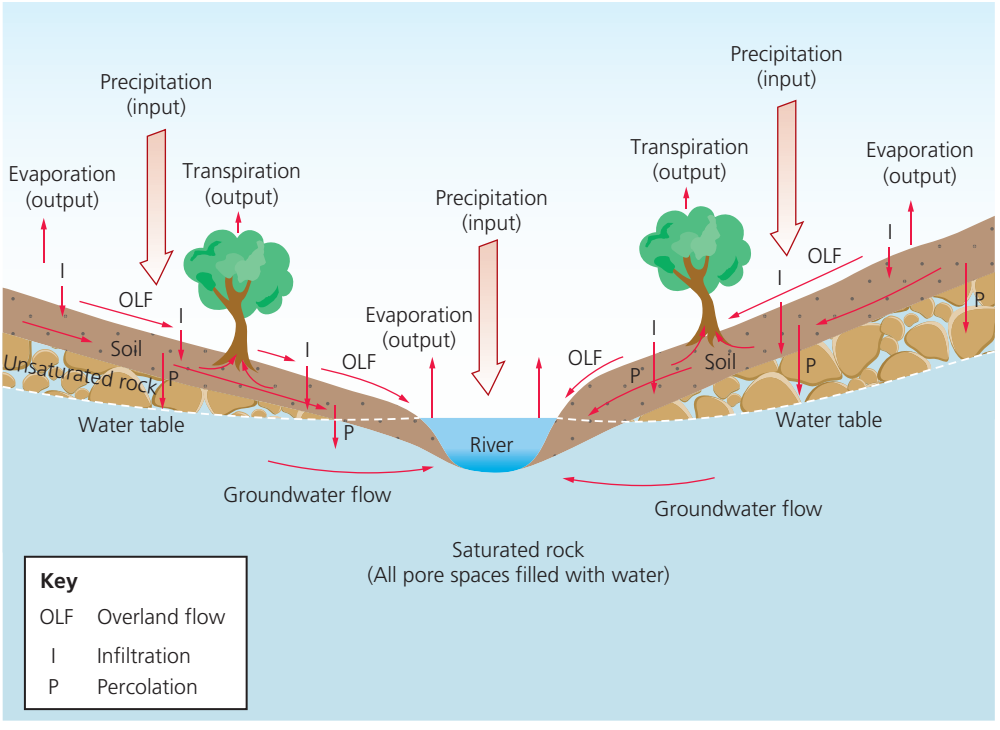


Figure 1.14 The hill slope drainage system

woodland there are many channels created by roots, as well as burrowing animals and these allow relatively free movement. Clay soils retain water, hindering any movement. They can dry out from the surface down before they allow horizontal transfer.

Once in the rocks below the soil the rate of transfer slows considerably. Groundwater can be held for millennia.

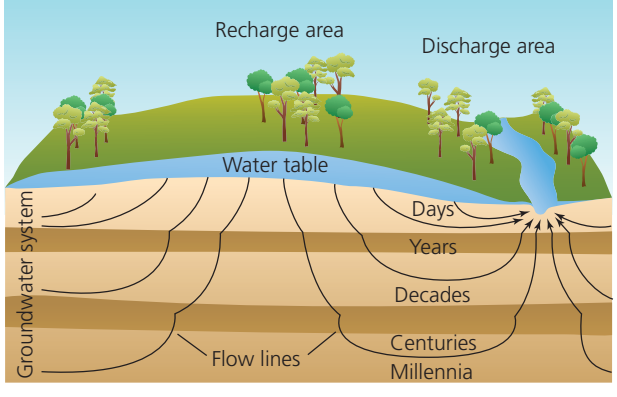


Figure 1.15 The varying timescales of water movements

Flows within the global water cycle

There is a net transport of 40 units from ocean to land with about the same amount returning by the rivers to the ocean (Figure 1.16, page 14). However, the amount of precipitation over the continents is almost three times as high as the net transport, indicating a considerable recirculation of water over land. This recirculation has a marked annual cycle as well as having large variations between continents. The recirculation is larger during the summer and for tropical land areas.

The hydrological cycle of the world's oceans interacts differently with that of the continents.

- Most of the water from the Pacific Ocean recirculates between different parts of the Pacific itself and there is little net transport towards land.
- The pattern of water exchange between ocean and land is different in the Atlantic and Indian Oceans.
- Two-thirds of the total net transport of water towards the continents comes from the Atlantic Ocean, with the rest from the Indian Ocean.
- Most of the continental water for North and South America, Europe and Africa emanates from the Atlantic and is also returned to the Atlantic by rivers.

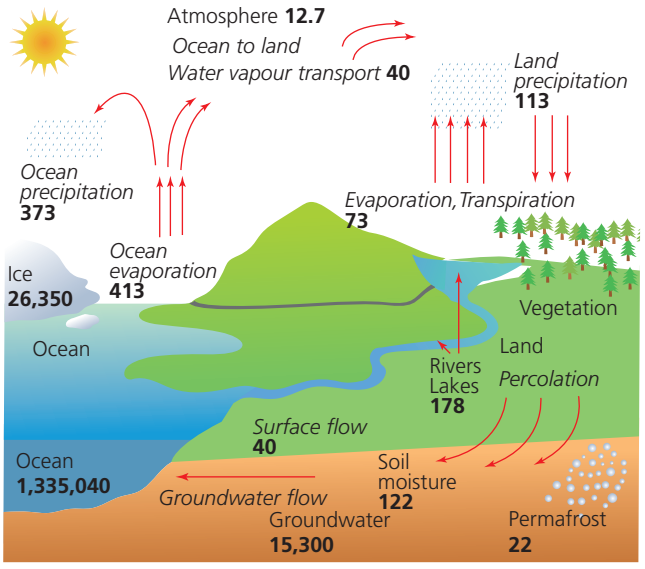


Figure 1.16 The varying amounts of water cycle movements (in thousand km^3/year for exchanges and thousand km^3 for storage)

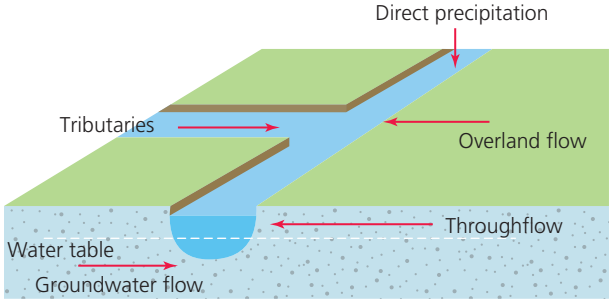


Figure 1.17 How water reaches a river

Water balance

Within a drainage basin, the balance between inputs (precipitation) and outputs (run-off, evapotranspiration, soil and groundwater storage) is known as the **water balance** or budget. Rivers only occur if the stores are able to release water, there is direct precipitation or there is overland flow into the river. As they move downstream, larger rivers are also fed by their tributary streams (Figure 1.17).

Discharge levels rise and fall, often showing an annual pattern (called the river's **regime**). They also vary in the short term following heavy rainfall. The water balance can be shown using the following formula:

Precipitation (P) = discharge (Q) + evapotranspiration (E) ± changes in storage (S)

The two most important parts of this relationship are precipitation and **'potential' evapotranspiration**. Evapotranspiration is closely related to the prevailing temperature. The warmer it is, the higher the evapotranspiration. It often happens that the temperature, and so the atmosphere's ability to hold water vapour, is greater than the amount of water available. Potential evapotranspiration is the amount of water that could be evaporated or transpired (or both) from an area if there was sufficient water available. The soil moisture graph (Figure 1.18) shows the relationship between precipitation and potential evapotranspiration over a year for an area of eastern England.

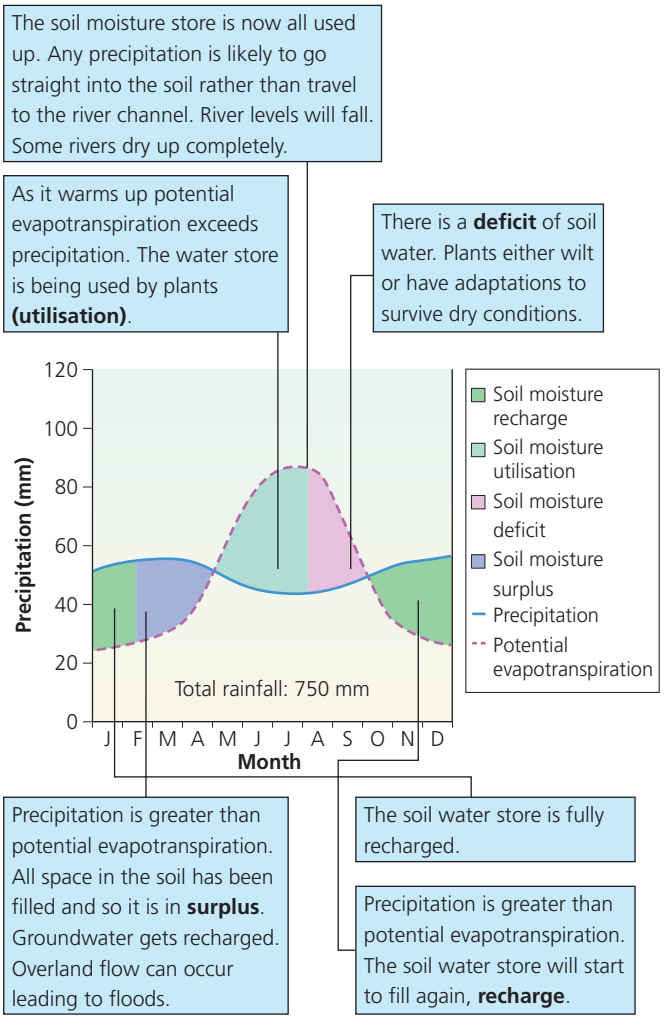


Figure 1.18 Soil water budget graph for eastern England

Soil moisture graphs vary considerably depending on the climate. Figure 1.19 shows the soil moisture budget for an equatorial area. In this case, the rainfall has two marked maxima. Because the temperatures vary very little throughout the year the potential evapotranspiration stays relatively constant. The high rainfall fills the soil stores rapidly. In the short times between the rainfall maxima, soil water does not go into deficit and so rivers

and plants have a source of water all year round. There is a high potential for flooding between February and June and again in August to November.

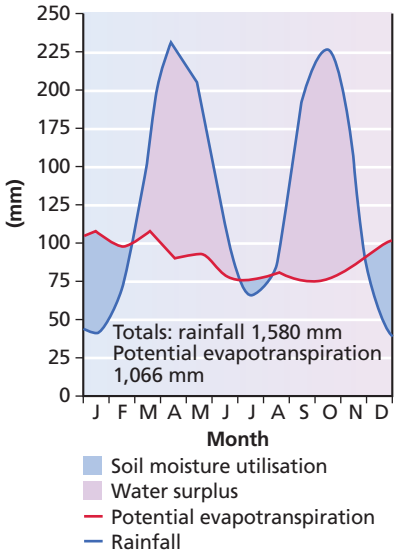


Figure 1.19 Soil moisture budget for Yaoundé, Cameroon

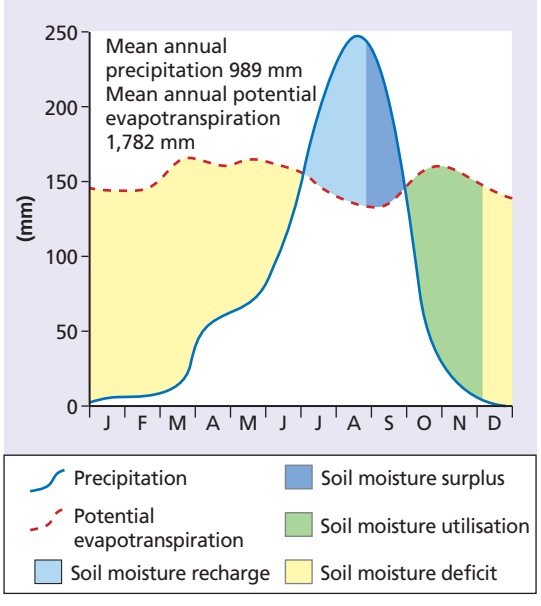


Figure 1.20 Soil moisture budget for Navrongo, northern Ghana (11°N)

Skills focus

Compare the soil moisture budgets for Yaoundé and Navrongo.

Figure 1.20, however, is for a tropical wet/dry climate. During the rainy season (July to October) the soils are recharged until there is surplus. This surplus does not last long when the rains stop. There is a short period of utilisation but from December to July the soil is dry. River levels fall with many drying up altogether. Vegetation has adapted to this seasonal water supply by evolving characteristics that enable it to survive drought. Humans have adapted traditionally by following migrating herds of animals that have themselves migrated following the rains north and south.

Review question

5 The Mediterranean area has hot dry summers and mild wet winters. Create a soil moisture budget graph for southern France.

Run-off variation

River flow is studied by measuring the discharge of a river. This is defined as the volume of water passing a measuring point in a given time. It is calculated by multiplying the cross-sectional area of the river by its velocity at the measuring point. It is measured in m^3/sec, or cumecs.

Key terms

Bankfull – the maximum discharge that a river channel is capable of carrying without flooding.

Base flow – This represents the normal day-to-day discharge of the river and is the consequence of slow moving soil throughflow and groundwater seeping into the river channel.

Lag time – the time between the peak rainfall and peak discharge.

Peak discharge – the point on a flood hydrograph when river discharge is at its greatest.

Storm flow – discharge resulting from storm precipitation involving both overland flow, throughflow and groundwater flow.

Storm hydrograph – A graph of discharge of a river over the time period when the normal flow of the river is affected by a storm event.

Discharge is the combined result of the many climatological and geographical factors which interact within a drainage basin. This means knowledge of discharge is very important in the assessment and management of water resources (including irrigation provision), the design of water-related structures

(reservoirs, bridges, flood banks, urban drainage schemes, sewage treatment works, etc.) and flood warning and alleviation schemes. The information can also help in developing hydroelectric power and protecting both the ecological health of watercourses and wetlands and their amenity and recreational value.

With climate change expected to impact very unevenly on river discharge patterns, keeping records of river flows is the key to identifying, quantifying and interpreting hydrological trends. This will help in the development of more effective ways of dealing with future flood and drought episodes.

The UK has a dense gauging station network of around 1,500 measuring stations supported by secondary and temporary monitoring sites (Figure 1.21). Such a large number is needed because the UK contains a multiplicity of mostly small river basins and is diverse in terms of its climate, topography, geology, land use and patterns of water usage.

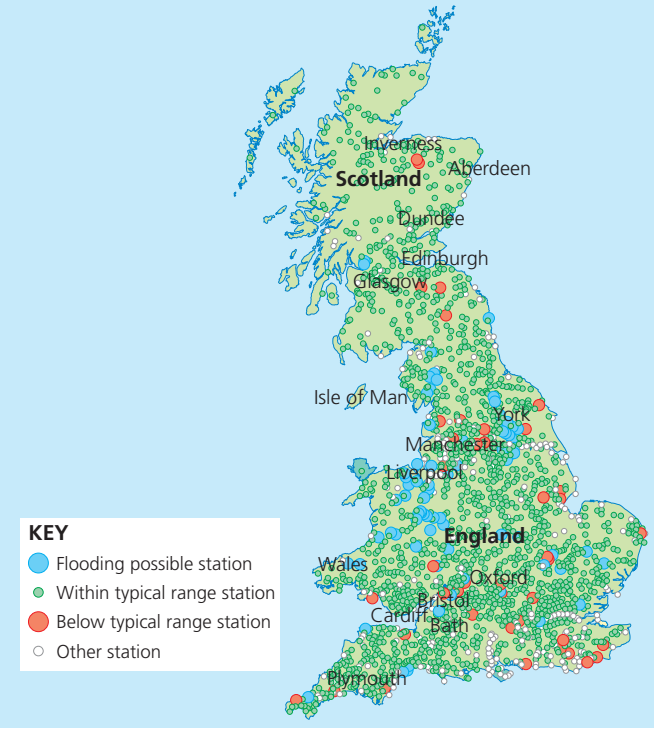


Figure 1.21 The UK gaug ing station network
Source: Shoothill

There are many ways to measure the discharge of a river. Figure 1.22 shows a weir of known shape. The amount of water flowing over the weir is proportional to the depth of water at the weir. The depth of the water flowing over the weir can be converted into discharge using a simple equation.

Figure 1.22 Brant Broughton Gauging Station on the River Brant in Lincolnshire, England

River regimes

A river regime can be defined as *'the variability in its discharge throughout the course of a year in response to precipitation, temperature, evapotranspiration and drainage basin characteristics.'*

There are almost 1,500 river systems within the UK and they vary greatly, being extremely sensitive to climatic variation, land use changes and water abstraction. They are also greatly influenced by the landscapes through which they flow. UK rivers range from mountain torrents draining areas receiving up to five metres of rain a year, to much more placid groundwater-fed streams supported by much lower rainfall in parts of southeast England.

UK rainfall is fairly evenly distributed throughout the year with a slight autumn/winter maximum, particularly in the west. However, seasonal variations in temperature and sunshine amounts mean that evapotranspiration losses are heavily concentrated in the summer. This means that UK rivers have a marked seasonality in their discharge with a late summer/early autumn minimum (See Figure 1.23).

In other parts of the world river regimes differ widely. The Mohawk River flows out of the mountains of upper New York State. Here the extremely snowy, cold and long winters and drier, hot summers lead to a different pattern of river discharge (Figure 1.24).

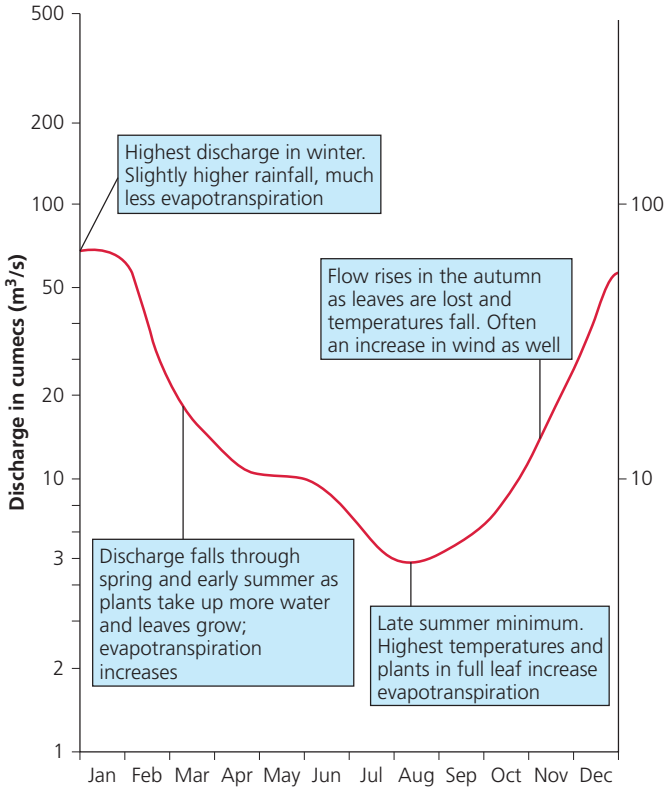


Figure 1.23 Sketch of the regime of the River Avon, Bath 1999

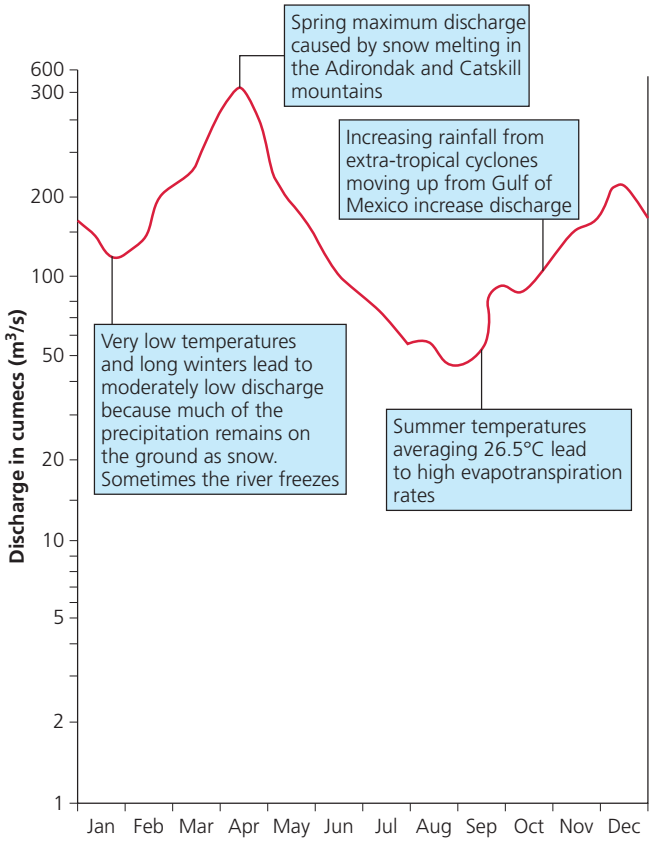


Figure 1.24 Sketch of the regime of the Mohawk River (New York State) USA

Skills focus

Research secondary data to find the regime of a river in your own locality. Explain the pattern you have discovered.

The geological characteristics of individual catchments have a major influence on flow patterns. Figure 1.25 shows daily flows for neighbouring catchments in the Thames basin. Although they experience almost identical climatic conditions, the flow regime for the Lambourn, which is supplied by springs from the underlying chalk, is much less variable than that for the Ock which drains an impermeable clay catchment.

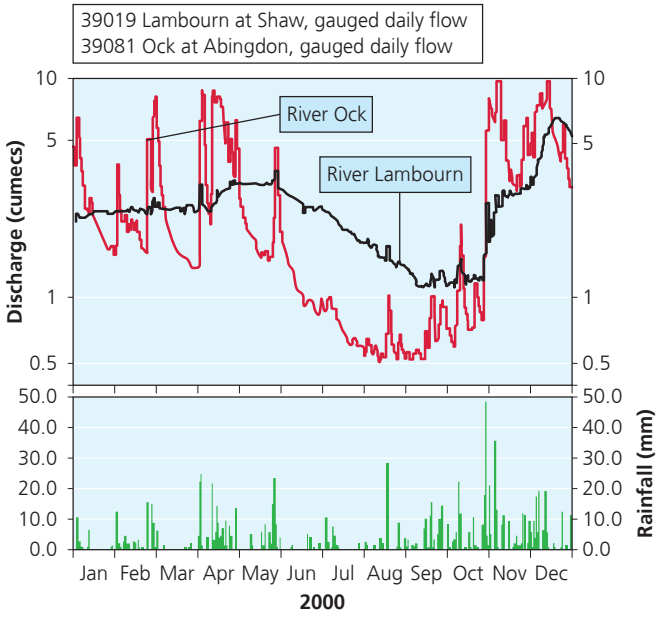


Figure 1.25 The variation in run-off for two rivers with the same rainfall pattern but different underlying geology

Source: Centre for Ecology & Hydrology

Skills focus

You should be able to describe trends from graphs. What is the general pattern of flow for the River Lambourn?

Rainfall events mean that the pattern of river flow is not smooth. This is illustrated by the hydrograph for the River Ock in Figure 1.25. Each time it rains the River Ock responds quickly, giving the hydrograph a jagged appearance. Each sudden rise and then subsequent fall in discharge can be closely

studied using a storm (or flood) hydrograph. Storm hydrographs are graphs of discharge over the time period when the normal flow of a river is affected by a storm event.

The storm (flood) hydrograph

We have already seen that a hydrograph is a graph of river discharge against time. A **storm hydrograph** is the graph of the discharge of a river leading up to and following a storm or rainfall event (**storm flow**) (Figure 1.26). They are important because they can predict how a river might respond to a rainstorm. This can help in managing the river.

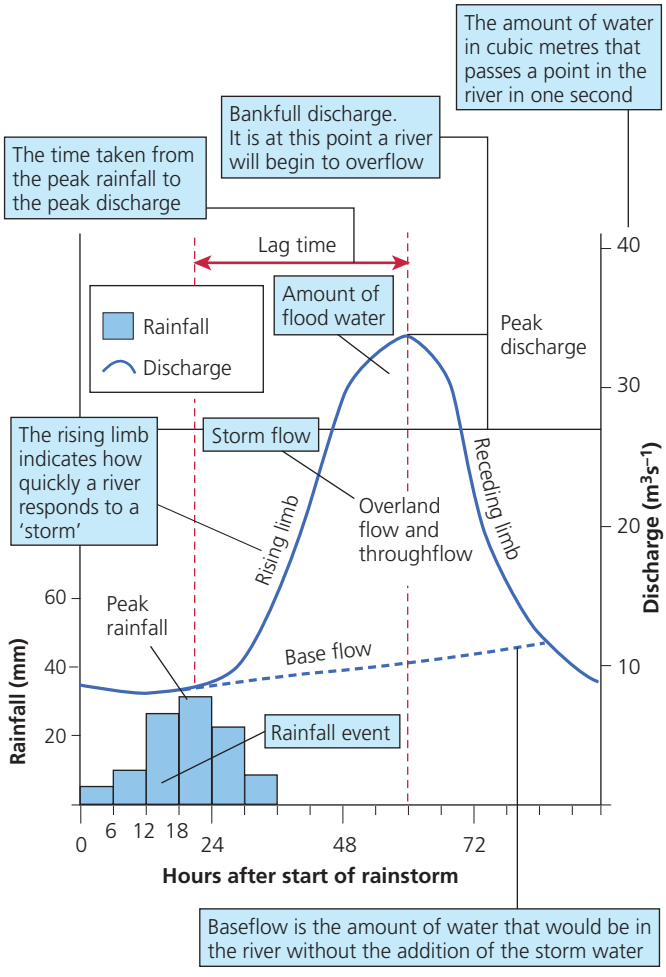


Figure 1.26 A storm hydrograph

- The storm hydrograph starts with the **base flow**. The river is fed by throughflow of soil water and groundwater. The slow movement of this water means that the changes in discharge are small.
- As storm water enters the drainage basin the river begins to be fed by much more fast-moving water. The discharge rises, as shown by the rising limb of the storm hydrograph.
- It eventually reaches a **peak discharge**, the highest flow in the channel for that event. The time taken from the peak rainfall to the peak discharge is called the **lag time**.
- The discharge begins to fall as shown by the **receding limb** of the graph.
- When all the storm water has passed through, the river returns to its base flow.

Although all storm hydrographs have the same common elements, they are not all the same shape. Hydrographs that have a short lag time, high peak discharge, steep rising and falling limbs are described as being '**flashy**'. Others are a lot more subdued with gentle rising and falling limbs, long lag times and low peak discharge. This shape is determined by both physical and human factors.

Physical factors

- Drainage basins that are more circular in shape (Figure 1.27) lead to more flashy hydrographs than those that are long and thin because each point in the drainage basin is roughly equidistant from the measuring point on a river.

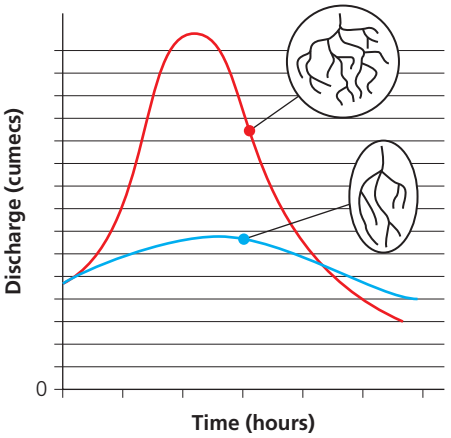


Figure 1.27 The influence of drainage basin shape on storm hydrographs: Water takes less time to reach the river in a round drainage basin than an elongated one

- Drainage basins with steep sides tend to have flashier hydrographs than gently sloped river basins. This is

because water flows more quickly on the steep slopes, whether as throughflow or overland flow and so gets to the river more quickly (Figure 1.28).

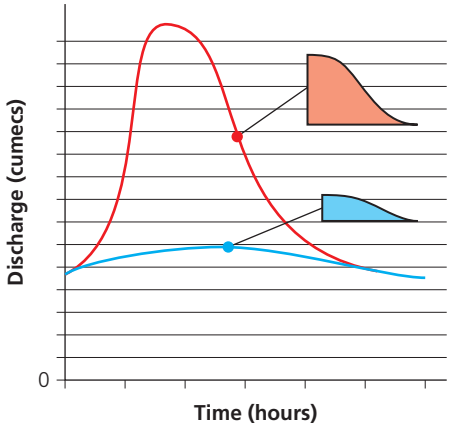


Figure 1.28 The influence of steepness on storm hydrographs: In steep sided drainage basins (or the upper reaches of a river) water gets to the river more quickly than in an area of gentle slopes

- Basins that have a high drainage density (that is, they have a lot of surface streams acting as tributaries to the main river) have flashy hydrographs (Figure 1.29). All the water arrives at the measuring station at the same time.

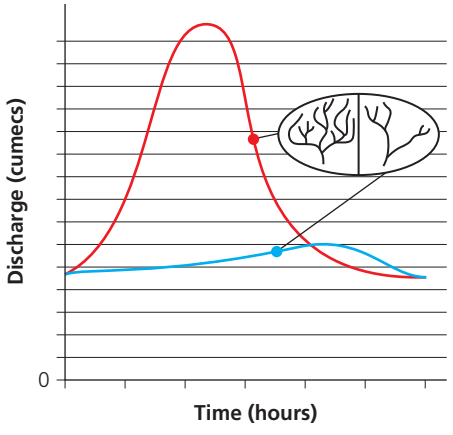


Figure 1.29 The influence of drainage density on flood hydrographs: Drainage basins with high drainage density have flashier river flows

- If the drainage basin is already saturated by antecedent rainfall then overland flow increases because infiltration capacity has been reached. Since overland flow is the fastest of the transfers the lag time is reduced. Again, peak discharge is higher resulting in a flashy hydrograph.
- If the soil or rock type within the river basin is impermeable (for example, clay soils or shale rocks) overland flow will be higher. Throughflow and infiltration will also be reduced, meaning a flashy hydrograph. The same can be said of surfaces baked hard by the Sun during a long period of dry weather or frozen surfaces resulting from cold weather.

  On the other hand, drainage basins underlain by sandstones have a subdued hydrograph because the water soaks into this porous rock.
- Thick vegetation cover in drainage basins will have a significant effect on a storm hydrograph. Vegetation intercepts the precipitation, holding the water on its leaves; this slows the movement of rainwater to the ground and so to river channels. Water is also lost due to evaporation and transpiration from vegetation surfaces reducing how much gets to the river. This subdues the storm hydrograph, increasing lag time and reducing peak discharge.
- The amount and intensity of precipitation can affect storm hydrographs. Heavy storms with a lot of water entering the drainage basin over a short time result in higher discharge. The type of precipitation can also have an impact. The lag time is likely to be greater if the precipitation is snow rather than rain. This is because snow takes time to melt before the water enters the river channel. When there is rapid melting of snow the peak discharge could be high.
- Large drainage basins catch more precipitation and so have a higher peak discharge compared to smaller basins. Smaller basins generally have shorter lag times because precipitation does not have as far to travel.

Human factors

- Deforestation reduces interception rates allowing rainwater to hit the surface directly. The lack of vegetation roots reduces the infiltration rate into the soil. These both result in rapid overland flow and flashy hydrographs. Deforestation also exposes the soil to greater rates of erosion, which leads to sedimentation of the channel. This reduces the **bankfull** capacity of a river and leads to a greater chance of flooding.
- Afforestation has the opposite effect making it a useful flood prevention measure.
- Agriculture has a variety of effects, among which are:
  - Ploughing breaks up the topsoil and allows greater infiltration, subduing hydrographs. This can be enhanced by contour ploughing where furrows

are created that run directly down slope, then they can act as small stream channels and lead to flashier hydrographs. Ploughing wet soils can cause impermeable smears in the subsoil called plough pans. These inhibit percolation producing greater surface flows.
  - Terracing on hillsides stops movement of water downhill and subdues hydrographs.
  - Grass crops increase infiltration and lead to subdued hydrographs.
  - Large numbers of animals on small areas can impact soils leading to overland flows.
- Growth of urban areas and other large impermeable surfaces such as roads produce flashy rivers (Figure 1.30) (see Chapter 9). This is exacerbated by the very fact that settlements have been built on floodplains. This urban growth leads to the expansion of built-up, impermeable surfaces such as roads, car parks, shopping centres, etc. Most settlements are designed to transfer water as quickly as possible away from human activity to the nearest river. This is achieved through road camber, building design and drainage systems. In many cities in the UK there has been a continued loss of front gardens in favour of paved drives. According to a Royal Horticultural Society report in 2016, London had half of all its front gardens paved over, an increase of 36 per cent over ten years.

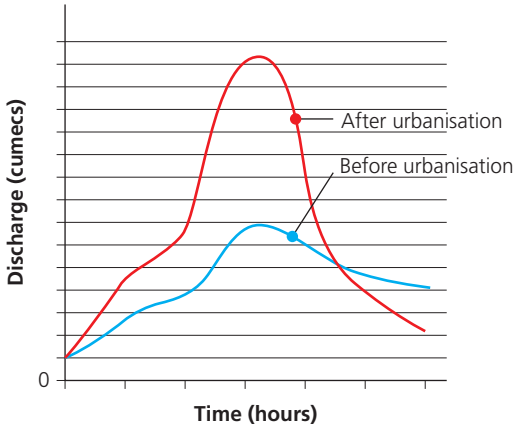


Figure 1.30 The influence of land use change on flood hydrographs: Impermeable surfaces shaped to get rid of water quickly, combined with a dense network of smooth drains, means that water gets to the river very quickly. The river itself can also be altered, for example, to move the water rapidly away from the urban area

Skills focus

You need to be able to draw storm hydrographs accurately and be able to label them with reasons for their pattern.

- Some soft engineering flood management schemes attempt to reduce flashiness in a river's hydrograph. Afforestation increases interception and infiltration. This slows down the progress of water to the river channel and subdues any changes in discharge.
- Water abstraction reduces the base flow and so more water must reach the channel before it reaches bankfull capacity.

Examples of how land use changes can affect the water cycle

Deforestation

Tropical South America contains the world's largest continuous tropical forest and savanna ecosystems. This region is environmentally important not only because of traditional ecological measures, such as its high biodiversity, but also because it generates more than a quarter of the world's river discharge. It has undergone explosive development and deforestation in the last 50 years as national and international demand for cattle feed (mostly soy), beef and sugar cane for ethanol has increased. Already over 10 per cent of the rainforest in this large region has been converted to cattle pasture and agriculture.

Deforestation and forest degradation result in a complex set of changes to streams of all sizes. When forests are removed, the new vegetation generally has fewer leaves and shallower roots. This means it uses less water than the forest it replaces. The result is that less water evaporates from the land surface to be returned to the atmosphere; more water runs off the land and stream flow is increased. The amount of change that occurs depends on local conditions including the amount of rainfall, how much of a watershed is deforested (Figure 1.31), topography, soils and the land use after deforestation. Studies have shown that there is little effect with less than 20 per cent of a basin deforested but a large increase with 50 to 100 per cent of a basin deforested. These changes occur at the local scale, but rivers of all sizes are affected when deforestation is extensive.

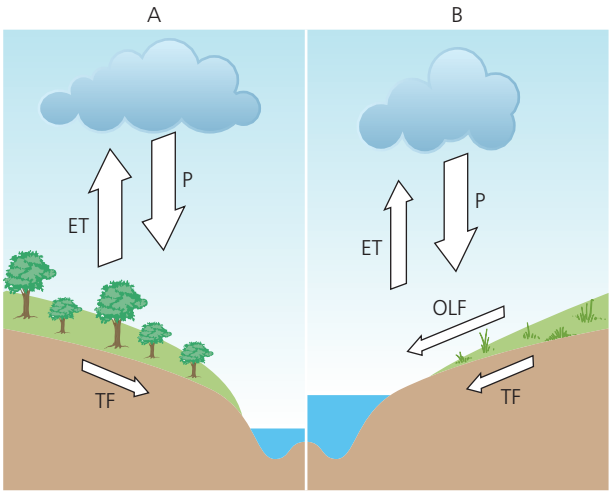


Figure 1.31 The effects of deforestation on the water cycle – localised deforestation

Before deforestation (A) much of the high precipitation (P) is returned to the atmosphere by evapotranspiration (ET). Overland flow is minimal. Most of the water that reaches the forest floor infiltrates into the soil and travels slowly to the river by throughflow, maintaining a steady flow in the river.

After deforestation (B), although the precipitation stays the same, the evapotranspiration is lower because the replacement vegetation has smaller leaves and roots and is less dense. Overland flow and throughflow occur because of the lack of vegetation. This leads to increased discharge and flashiness. This can cause localised flooding

This suggests several important points about the climate, land surface and water cycle:

- If deforestation does not cause decreased rainfall via atmospheric feedbacks, discharge will likely be significantly increased throughout the entire southern Amazon (Figure 1.31).
- If rainfall does decrease via atmospheric feedbacks, the resulting decrease in river discharge may be greater than the changes without feedbacks (Figure 1.32).
- Changes in water resources caused by atmospheric feedbacks will not be limited to those catchment areas where deforestation has occurred but will be spread unevenly throughout the whole Amazon basin by atmospheric circulation.

Soil drainage

Subsurface drainage removes excess water from the soil profile. It is carried out usually through a network of perforated tubes installed 60–120 cm below the soil surface (Figure 1.33). These tubes are commonly called 'tiles' because they were originally made from short lengths of clay pipes known as tiles. Water would seep into the small spaces between the tiles and drain away.

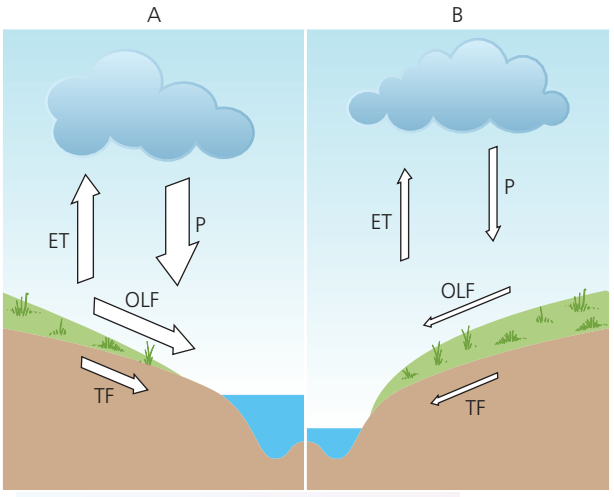


Figure 1.32 The effects of deforestation on the water cycle – extensive deforestation

Where deforestation is extensive, positive feedback can occur in the basin hydrological system. In A, because evapotranspiration is low, much of the water leaves the area in the river channel rather than being recycled continuously between the forest and the atmosphere. Once the water has left the area (B), there is less water vapour available in the atmosphere for precipitation and so precipitation levels fall. Less water gets to the river channel and the flow is reduced

The most common type of 'tile' now is corrugated plastic tubing with small perforations to allow water entry. When the water table in the soil is higher than the tile, water flows into the tubing, either through holes in the plastic tube or through the small cracks between adjacent tiles. This lowers the water table to the depth of the tile over the course of several days. Drain tiles allow excess water to leave the field, but once the water table has been lowered to the elevation of the tiles, no more water flows through them. In most years in the UK, drain tiles have no flow between June and October.

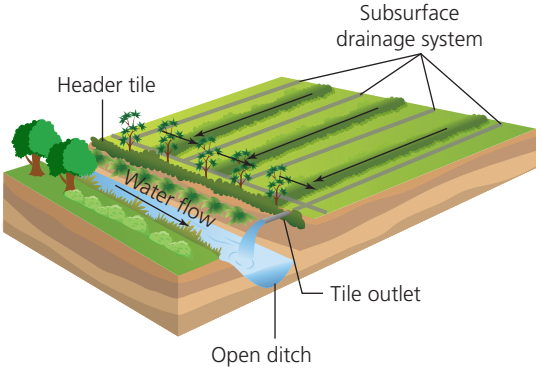


Figure 1.33 A typical tile drainage set-up
Source: Queen's Printer for Ontario

Although costly, agricultural drainage is very good for moderately to poorly drained soils. It increases the productivity of the field and helps to improve the efficiency of growers. A study by Ohio State University demonstrated that for every dollar spent on drainage the grower got back between $1.20 and $1.90 when growing corn and soybeans.

Advantages of soil drainage

- The build-up of an improved soil structure makes it more friable and easier to work. It also makes it easier to achieve greater root penetration, enabling roots to travel faster and further.
- Improved aeration makes conditions more favourable for microorganisms to thrive. This increases the rate at which organic matter is broken down into humus and plant nutrients are mineralised into an available form. It also provides the necessary supply of air for root cell respiration.
- The increased aeration increases the ease with which the soil can be warmed. This can make possible earlier sowing of seeds, with greater likelihood of improved germination.
- Heavy machinery can work on the land without danger of compaction (and so leading to increased overland flow).
- Larger numbers of animals can be allowed to graze the land, once again without compacting the soil.

Disadvantages of soil drainage

- The insertion of drains artificially increases the speed of throughflow in the soil. Much more water reaches watercourses more quickly than before drainage. This can increase the likelihood of flooding and increase the range of flows in rivers. It is interesting to note that before the drainage of many floodplains in the UK from the eighteenth century onwards, rivers were more navigable than today; the annual flow regime was much more even.
- The dry topsoil can be subject to wind erosion if not properly protected. Soil loss by wind erosion has mainly been documented for sandy and peaty soils in the eastern and middle counties of England, especially arable fields in the East Midlands and East Anglia. Generally the area of England and Wales subject to wind erosion is small, although those fields that are affected are more severely eroded than by water erosion. The proportion of a field subject to wind erosion is likely to be greater than that subject to water erosion. Available estimates suggest that the mean wind erosion rate is of the order of 0.1 to 2 tonnes/ha/year, although maximum values for fields can be one or two orders of magnitude higher.
- Another major concern with regard to land drainage is nitrate loss. It can lead to eutrophication. Water draining from fields finds its way into local watercourses. There it enriches ponds, etc., with nitrogen or phosphorus. It causes algae and higher forms of plant life to grow too fast. This disturbs the balance of organisms present in the water and the quality of the water concerned.

One way to overcome some of these problems is to use controlled drainage. This keeps the water table high during the off-season when crops are not growing. The high water table increases the rate of denitrification (a process that converts nitrate to harmless nitrogen gas (N_2) as soon as the saturated soil warms up in the spring) and reduces nitrate loss to the environment.

Water abstraction

Problems can occur when the demand for water exceeds the amount available during a certain period. They happen frequently in areas with low rainfall and high population density, and/or in areas with intensive agricultural or industrial activity.

In many areas of Europe, groundwater is the dominant source of fresh water. In a number of places water is being pumped from beneath the ground faster than

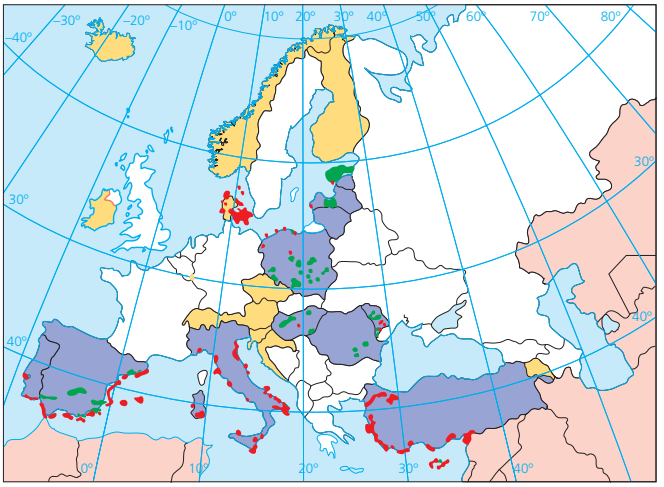


Figure 1.34 Salt water intrusion in Europe

Source: European Environment Agency

it is being replenished through rainfall. The result is sinking water tables, empty wells, higher pumping costs and, in coastal areas, the intrusion of saltwater from the sea which degrades the groundwater (Figure 1.34 page 22). This saline intrusion is widespread along the Mediterranean coastlines of Italy, Spain and Turkey, where the demands of tourist resorts are the major cause of over-abstraction. In Malta, most groundwater can no longer be used for domestic consumption or irrigation because of saline intrusion, and the country has resorted to expensive desalination plants.

Sinking water tables can also make rivers less reliable, since many river flows are maintained in the dry season by springs that dry up when water tables fall. Groundwater also helps sustain surface reservoirs of water such as lakes and wetlands that are often highly productive ecosystems and resources for tourism as well as leisure activities. These, too, are threatened by over-abstraction of groundwater.

Irrigation is the main cause of groundwater overexploitation in agricultural areas. Examples include the Greek Argolid plain of eastern Peloponnesus where it is common to find boreholes 400 m deep contaminated by seawater intrusion. In Italy, overexploitation of the Po River in the region of the Milan aquifer has led to a 25–40 m decrease in groundwater levels over the last 80 years. In Spain, more than half of the abstracted groundwater volume is obtained from areas facing overexploitation problems.

Water abstraction from the chalk of southern England

The water within the chalk aquifer of southern England is replenished by rainfall that lands on the exposed chalk hills of the North and South Downs and the Chilterns. Normally recharge takes place during the winter months when potential evapotranspiration is low and soil moisture deficits are negligible. Groundwater amounts vary seasonally, with levels rising from autumn through winter into spring. During the summer months, potential evapotranspiration generally exceeds rainfall, soil moisture deficits build up, and little, if any, percolation takes place.

In the summer, water still leaves the chalk from springs as well as by abstraction from boreholes. This pattern is not constant, since rainfall varies both over time and location. Rivers fed by groundwater from chalk aquifers can have intermittent sections. These streams, often referred to as 'bournes', are a natural characteristic of chalk downlands. The positions of the springs feeding these rivers differ throughout the year, being at a greater altitude in winter and spring. If there are one or more dry winters when the effective rainfall available for recharge is low then these rivers can dry up altogether.

Some of the most acute problems with over-abstraction have been found in chalk stream systems, where up to 95 per cent of the flow is derived from underground aquifers. The catchments of chalk streams provide underground reservoirs of generally high quality groundwater which can be abstracted for public supply. This abstraction has dramatically reduced the flow in many chalk streams and, in some cases, completely dried up sections of these important rivers, particularly during dry summers when public demand is at its highest. This also has an economic impact on local communities, resulting from the inability to fish, enjoy river views due to encroaching vegetation or undergo other recreational activities.

Water abstraction in the London Basin

Figure 1.35 (page 24) shows the subsurface geology of the London Basin. The chalk layers form a syncline beneath the London area with the uplands of the Chilterns to the NW and the North Downs to the SE. Precipitation on these exposed chalk hills soaks into the porous chalk where it is stored and released naturally at springs where it is in contact with either Greensand or Palaeogene rocks.

Throughout history, in London water has been abstracted from wells and boreholes that penetrate down to the chalk. During the nineteenth century and first part of the twentieth century, the chalk-basal sands aquifer had been increasingly exploited, as a result of increased industrialisation and the associated development of groundwater sources. At the peak of abstraction in the 1960s, groundwater levels beneath central London had dropped to 88 m below sea level, creating a large depression in the water table. A smaller cone of depression also developed to the east beneath the River Roding in Essex.

Since the mid-1960s, industries, such as brewing, in central London relocated or were closed down. Economic activity turned more to service industries and commerce than heavy industry. The subsequent

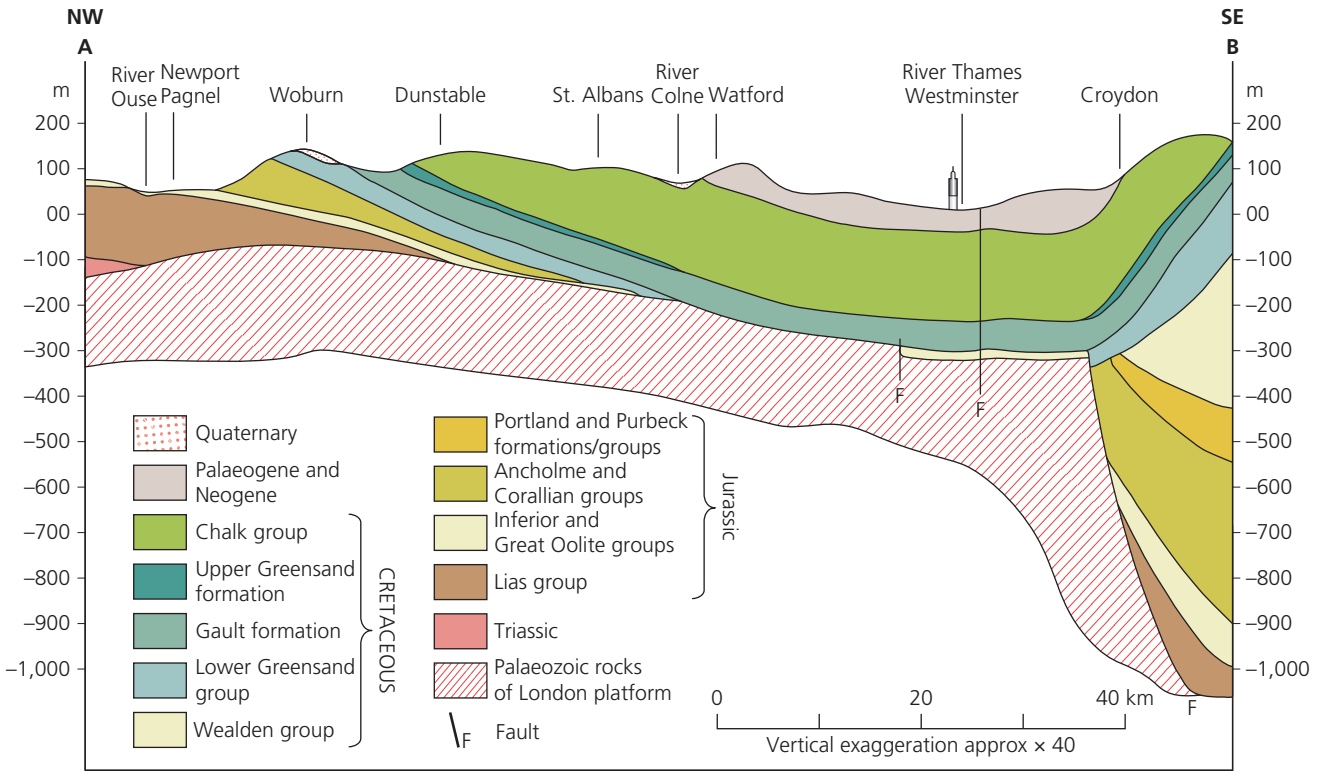


Figure 1.35 Cross section showing the confined chalk aquifer of the London Basin (BGS, 1996)
Source: British Geological Society

reduction in abstraction resulted in groundwater levels recovering by as much as 3 m/year in places by the early 1990s, leading to a gradual rebound of the water table. This then has posed the threat of rising groundwater to structures in the London Basin such as London Underground and building foundations. This led to the implementation of the General Aquifer Research, Development and Investigation Team (GARDIT) strategy to control water levels. As a result of careful management of both abstraction and artificial recharge, the rise in groundwater that the GARDIT strategy was designed to arrest had largely been achieved by 2000, so this year provides a useful baseline year for comparisons.

According to the Environment Agency, the differences in groundwater levels for January 2000 and January 2018 are:

- Groundwater levels have risen in the order of 8–15 m across north London, due to limited abstraction. They have levelled off in recent years.
- In east London, groundwater levels have fallen in the order of 10 m as a result of increased abstraction.
- In south-west and south-central London, groundwater levels have fallen in the order of 2–5 m, also as a result of increased abstraction.
- In east London, where there are chalk outcrops around the River Thames from Greenwich to Woolwich, there is a risk of saline intrusion. When groundwater levels near the river are lower than the water level in the River Thames, saline river water can enter the chalk aquifer.

Review question

6 Outline the benefits and problems caused by water abstraction in an environment that contrasts with that of south-east England.

1.3 The carbon cycle

Carbon (C) (from the Latin *carbo* meaning coal) is one of the most chemically versatile of all the elements. Carbon forms more compounds than any other element and scientists predict that there are more than ten million different carbon compounds in existence today on Earth. Carbon is found in all life forms in addition to sedimentary rocks, diamonds, graphite, coal and petroleum (oil and natural gas).

Carbon follows a certain route on Earth, called the carbon cycle. It is the complex processes carbon undergoes as it is transformed from organic carbon (the form found in living organisms such as plants and trees) to inorganic carbon and back again. Through following the carbon cycle we can also study energy flows on Earth, because most of the chemical energy needed for life is stored in organic compounds as bonds between carbon atoms and other atoms.

Carbon atoms move through the carbon cycle in many different forms. Some important examples of carbon compounds include:

- carbon dioxide (CO_2), a gas found in the atmosphere, soils and oceans
- methane (CH_4), a gas found in the atmosphere, soils and oceans and sedimentary rocks
- calcium carbonate ($CaCO_3$), a solid compound found in calcareous rocks, oceans and in the skeletons and shells of ocean creatures
- hydrocarbons – solids, liquids or gases usually found in sedimentary rocks
- bio-molecules – complex carbon compounds produced in living things. Proteins, carbohydrates, fats and oils, and DNA are examples of bio-molecules.

Of all these forms of carbon, we study CO_2 in most detail because it is thought that this has a profound effect on climate. It is also difficult to separate a *natural* carbon cycle from one that is affected by human activity. Human activity and associated emissions of carbon dioxide (**anthropogenic CO_2**) fundamentally affect the carbon cycle and so affect climate.

Origins of carbon on Earth

The primary source of carbon/CO_2 is the Earth's interior. It was stored in the mantle when the Earth formed. It escapes from the mantle at constructive and destructive plate boundaries as well as hot-spot volcanoes. Much of the CO_2 released at destructive margins is derived from the metamorphism of carbonate rocks subducting with the ocean crust. Some of the carbon remains as CO_2 in the atmosphere, some is dissolved in the oceans, some carbon is held as biomass in living or dead and decaying organisms, and some is bound in carbonate rocks. Carbon is removed into long-term storage by burial of sedimentary rock layers, especially coal and black shales (these store organic carbon from undecayed biomass) and carbonate rocks like limestone (calcium carbonate).

The major stores of carbon

The Intergovernmental Panel on Climate Change (IPCC) uses gigatonne of carbon equivalent (GtC) and gigatonne of carbon equivalent per year (GtCy-1) to measure the storage and transfer of global carbon. Estimates of both these amounts vary considerably. Those given in this chapter are the ones used by the IPCC.

Key terms

Anthropogenic CO_2 – Carbon dioxide generated by human activity.

Carbon sequestration – The capture of carbon dioxide (CO_2) from the atmosphere or capturing anthropogenic (human) CO_2 from large-scale stationary sources like power plants before it is released to the atmosphere. Once captured, the CO_2 gas (or the carbon portion of the CO_2) is put into long-term storage.

Carbon sink – A store of carbon that absorbs more carbon than it releases.

Weathering – The breakdown of rocks *in situ* by a combination of weather, plants and animals.

The lithosphere

The Earth's **lithosphere** includes the crust and the uppermost mantle; this constitutes the hard and rigid outer layer of the Earth. Carbon is stored in the lithosphere in both inorganic and organic forms:

- Inorganic deposits of carbon in the lithosphere include fossil fuels like coal, oil, and natural gas, oil shale (kerogens) and carbonate-based sedimentary deposits like limestone.
- Organic forms of carbon in the lithosphere include litter, organic matter and humic substances found in soils.

Carbon in the lithosphere is stored as:

- marine sediments and sedimentary rocks contain up to 100 million GtC
- soil organic matter contains between 1,500 and 1,600 GtC
- fossil fuel deposits of coal, oil and gas contain approximately 4,100 GtC
- peat, which is dead but undecayed organic matter found in boggy areas contains approximately 250 GtC.

The hydrosphere

The ocean plays an important part in the carbon cycle. Attempts to collate measurements of the amount of carbon in the oceans have been made by the Global Ocean Data Analysis project (GLODAP) using data from research ships, commercial ships and buoys. The measurements come from deep and shallow waters from all the oceans. There is some variation in the results and figures can only be an approximation.

The oceanic stores can be divided into three:

- the surface layer (euphotic zone), where sunlight penetrates so that photosynthesis can take place, contains approximately 900 GtC
- the intermediate (twilight zone) and the deep layer of water contain approximately 37,100 GtC
- living organic matter (fish, plankton, bacteria, etc.) amounts to approximately 30 GtC and dissolved organic matter 700 GtC.

This gives a total for oceanic carbon of between 37,000 GtC to 40,000 GtC.

When organisms die, their dead cells, shells and other parts sink into deep water. Decay releases carbon dioxide into this deep water. Some material sinks right to the bottom, where it forms layers of carbon-rich sediments. Over millions of years, chemical and physical processes may turn these sediments into rocks. This part of the carbon cycle can lock up carbon for millions of years. It is estimated that this sedimentary layer could store up to 100 million GtC.

The biosphere

This is defined as the total sum of all living matter. For our purposes we are going to consider the terrestrial biosphere as being separate from the oceanic biosphere. The total amount of carbon stored in the terrestrial biosphere has been estimated to be 3,170 GtC. The distribution of this carbon depends upon the ecosystem as shown in Figure 1.36.

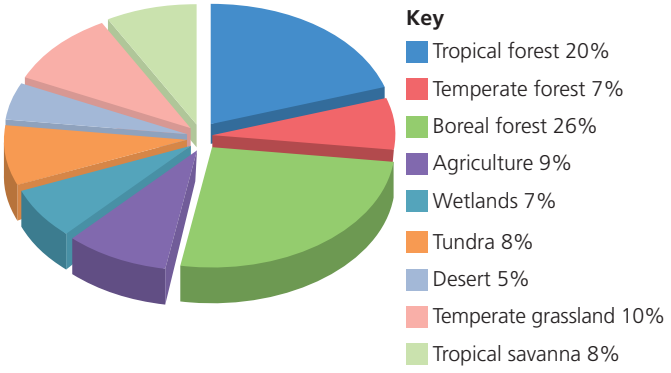


Figure 1.36 Terrestrial carbon stores by ecosystem

The main stores of carbon in the terrestrial biosphere are as follows:

- **Living vegetation:** at the global level, 19 per cent of the carbon in the Earth's biosphere is stored in plants. Unlike the oceans, much of this carbon is stored directly in the tissues of the plants. Although the exposed part of the plant is the most visible, the below-ground biomass (the root system) must also be considered. The amount of carbon in the biomass varies from between 35 and 65 per cent of the dry weight. The amount varies depending on the location and the vegetation type. It is estimated that half of the carbon in forests occurs in high-latitude forests, and a little more than one-third occurs in low-latitude forests. The two largest forest reservoirs of carbon are the vast expanses of boreal forest in Russia, which hold roughly 25 per cent of the world's forest carbon, and the Amazon basin, which contains about 20 per cent.
- **Plant litter:** this is defined as fresh, undecomposed, and easily recognisable (by species and type) plant debris. This can be anything from leaves, cones, needles, twigs, bark, seeds/nuts, etc. The type of litter is directly affected by the type of ecosystem. Leaf tissues account for about 70 per cent of litter in forests, but woody litter tends to increase with forest age. In grasslands there is very little above ground perennial tissue so the annual fall of litter is very low.
- **Soil humus:** this originates from litter decomposition. Humus is a thick brown or black substance that remains after most of the organic litter has decomposed. It gets dispersed throughout the soil by soil organisms such as earthworms. In

all forests, tropical, temperate and boreal together, approximately 31 per cent of the carbon is stored in the biomass and 69 per cent in the soil. In tropical forests, approximately 50 per cent of the carbon is stored in the biomass and 50 per cent in the soil. Altogether the world's soils hold more carbon (2,500 GtC) than the vegetation. Soil carbon can be either organic (1,550 GtC) or inorganic carbon (950 GtC). The inorganic carbon component consists of carbon itself as well as carbonate materials such as calcite, dolomite and gypsum. The amount of carbon found in living plants and animals is comparatively smaller than that found in soil (560 GtC). The soil carbon pool is approximately 3.1 times larger than the atmospheric pool of 800 GtC. Only the ocean has a larger carbon store.

- **Peat:** this is an accumulation of partially decayed vegetation or organic matter that is unique to natural areas called peatlands or mires. Peat forms in wetland conditions, where almost permanent water saturation obstructs flows of oxygen from the atmosphere into the ground. This creates low oxygen anaerobic conditions that slow down rates of plant litter decomposition. Peatlands cover over four million km² or three per cent of the land and freshwater surface of the planet; they occur on all continents, from the tropical to boreal and Arctic zones and from sea level to high alpine conditions. It is estimated that peat stores more than 250 GtC worldwide.
- **Animals:** these play a small role in the storage of carbon. They are, however, very important in the generation of movement of carbon through the carbon cycle.

The atmosphere

Carbon has been in the atmosphere from early in the Earth's history. Atmospheric carbon dioxide levels have reached very high values in the deep past, possibly topping over 7,000 ppm (parts per million) in the Cambrian period around 500 million years ago. Its lowest concentration has probably been over the last two million years during the Quaternary glaciation when it sank to about 180 ppm.

Today, carbon dioxide is a trace gas in the Earth's atmosphere. Estimates of the overall amount of carbon stored in the atmosphere vary from 720 GtC to 800 GtC. It makes up about 0.04 per cent (400 ppm) of the atmosphere; this low concentration belies its importance to the planet and all life on it. Due to human activities, the present concentration of CO_2 in the Earth's atmosphere is higher than it has been for at least 800,000 years, and, in all likelihood, the highest in the past 20 million years. Despite its relatively small concentration, CO_2 is a potent greenhouse gas and plays a vital role in regulating the Earth's surface temperature. The recent phenomenon of global warming has been attributed primarily to increasing industrial CO_2 emissions into the Earth's atmosphere.

Atmospheric carbon has been measured at the Mauna Loa Observatory (MLO) on Hawaii since 1958. The undisturbed air, remote location and minimal influences of vegetation and human activity at MLO are ideal for monitoring atmospheric constituents. The observatory is part of the American National Oceanic and Atmospheric Administration (NOAA). The measurements show that the global annual mean concentration of CO_2 in the atmosphere has increased markedly since the Industrial Revolution, from 280 ppm to 317.7 ppm in March 1958 to 416.2 ppm as of May 2020. This increase is largely attributed to human derived (anthropogenic) sources, particularly the burning of fossil fuels and deforestation.

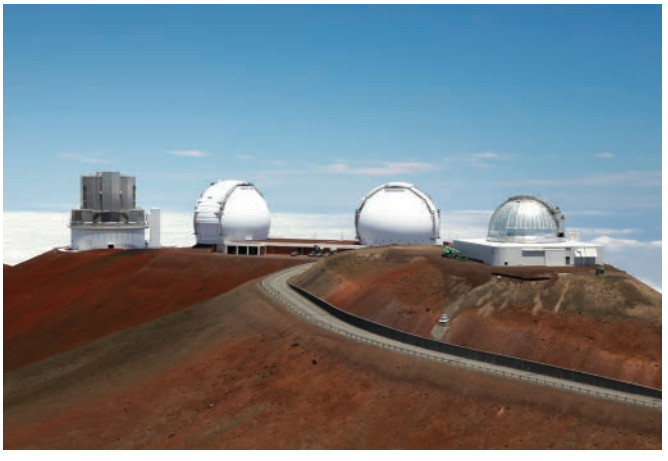


Figure 1.37 The NOAA atmospheric observatory station, Mauna Loa, Hawaii

A graph of this change has been named after the scientist who first started this research; it is called the Keeling Curve (Figure 1.38, page 28).

Keeling was one of the first scientists to gather evidence that linked fossil fuel emissions to rising levels of carbon dioxide. Keeling's research has been backed up by other readings from around the world;

for example, the CO_2 trapped in ice cores from Antarctica and Greenland can be used to give a 'proxy' measure of the CO_2 in the atmosphere at the time that snow was laid down.

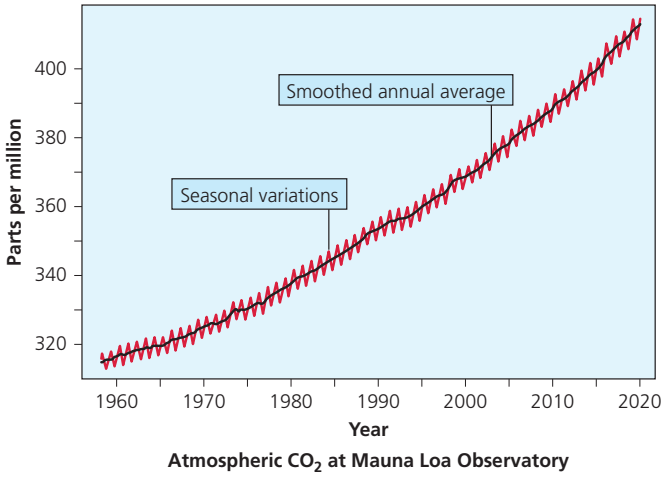


Figure 1.38 The Keeling curve of the changing concentration of atmospheric CO_2

Source: NOAA

The daily average at Mauna Loa first exceeded 400 ppm on 10 May 2013. It is currently rising at a rate of approximately 2 ppm/year and accelerating.

Movement of carbon

Carbon moves from one store to another in a continuous cycle (see Figure 1.39).

This cycle consists of several carbon stores as described above. The processes by which the carbon moves between these stores are known as transfers or fluxes. If more carbon enters a store than leaves it, that store is considered a **net carbon sink**. If more carbon leaves a store than enters it, that store is considered a **net carbon source**.

The geological component

The geological component of the carbon cycle is where it interacts with the rock cycle in the processes of **weathering**, burial, subduction and volcanic eruptions.

In the atmosphere, carbon dioxide is removed from the atmosphere by dissolving in water and forming carbonic acid:

$$CO_2 + H_2O \rightarrow H_2CO_3 \text{ (carbonic acid)}$$

As this weakly acidic water reaches the surface as rain, it reacts with minerals at the Earth's surface, slowly dissolving them into their component ions through the process of chemical weathering. These component ions are carried in surface waters like streams and rivers, eventually to the ocean, where they settle out as minerals like calcite, a form of calcium carbonate ($CaCO_3$).

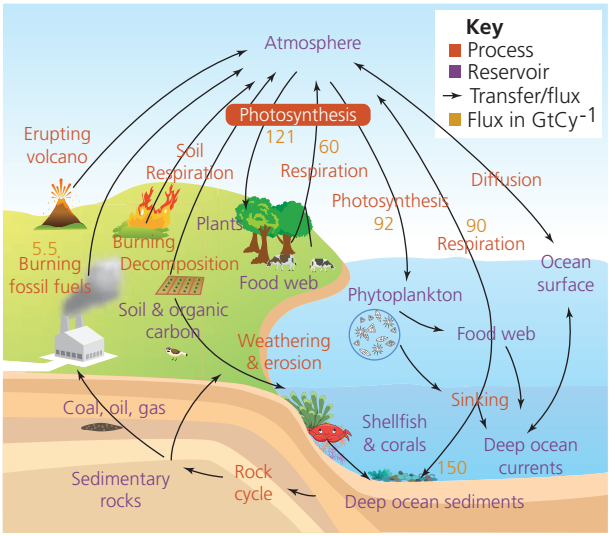


Figure 1.39 The carbon cycle

Calcium carbonate is also precipitated from calcium and bicarbonate ions in seawater by marine organisms like foraminifera, coccoliths or molluscs. When these creatures die, their skeletons sink to the bottom of the oceans where they collect as sediment. Burial by overlying layers of sediment can eventually turn these sediments into sedimentary limestone. Coral also extracts $CaCO_3$ from seawater. These creatures live and eventually die in the same location. Dead coral is built upon by later generations of live coral and so it too becomes buried. The carbon is now stored below the sea floor in layers of limestone. Tectonic uplift can then expose this buried limestone. One example of this occurs in the Himalayas where some of the world's highest peaks are formed of material that was once at the bottom of the ocean.

Tectonic forces cause plate movement to push the sea floor under continental margins in the process of subduction. The carbonaceous sea-floor deposits are pushed deep into the Earth where they heat up, eventually melt, and can rise back up to the surface through volcanic eruptions or in seeps, vents and CO_2-rich hot springs. In this way CO_2 returns to the atmosphere. Weathering, burial, subduction and volcanism control atmospheric carbon dioxide concentrations over time periods of hundreds of millions of years.

Photosynthesis

Tiny marine plants (phytoplankton) in the sunlit surface waters (the euphotic zone) of the oceans, as well as all terrestrial plants, photosynthetic algae and bacteria, turn the carbon into organic matter by the process of **photosynthesis**. They use energy from sunlight to combine carbon dioxide from the atmosphere with water to form carbohydrates. These carbohydrates store energy. Oxygen is a by-product that is released into the atmosphere.

carbon dioxide + water + sunlight ⟶ carbohydrate + oxygen

$$CO_2 + H_2O + \text{sunlight} \longrightarrow CH_2O + O_2$$

Respiration

Plants (and photosynthetic algae and bacteria) then use some of the stored carbohydrates as an energy source to carry out their life functions by the process of **respiration**. Some of the carbohydrates remain as biomass (the bulk of the plant, etc.). Consumers such as animals and bacteria get their energy from this excess biomass. Oxygen from the atmosphere is combined with carbohydrates to liberate the stored energy. Water and carbon dioxide are by-products.

oxygen + carbohydrate ⟶ energy + water + carbon dioxide

$$O_2 + CH_2O \longrightarrow \text{energy} + H_2O + CO_2$$

Note: Photosynthesis and respiration are essentially the opposite of one another. Photosynthesis removes CO_2 from the atmosphere and replaces it with O_2. Respiration takes O_2 from the atmosphere and replaces it with CO_2. However, these processes are not in balance. Not all organic matter is oxidised. Some is buried in sedimentary rocks. The result is that over geologic time, there has been more oxygen put into the atmosphere and carbon dioxide removed by photosynthesis than the reverse.

Decomposition

Decomposition is more difficult to define than photosynthesis. In broad terms, decomposition includes physical, chemical and biological mechanisms that transform organic matter into increasingly stable forms. This broad definition includes physical break-up of organic material by wet-dry, shrink-swell, hot-cold and other cycles.

- **Physical mechanisms**: Animals, wind and even other plants can also cause this fragmentation. Leaching and transport in water is another important physical mechanism.
- **Chemical mechanisms**: These include oxidation and condensation. Biological mechanisms involve feeding and digestion aided by the catalytic effect of enzymes.

The decomposition process is carried out by **decomposers** whose special role is to break down the cells and tissues in dead organisms into large biomolecules and then break those biomolecules down into smaller molecules and individual atoms.

Decomposition ensures that the important elements of life – carbon, hydrogen, oxygen, nitrogen, phosphorus, sulphur, magnesium – can be continually recycled into the soil and made available for life. For example, a plant cannot make its DNA molecules unless it has a supply of nitrogen, phosphorus and sulphur atoms from the soil in addition to the carbon, hydrogen and oxygen atoms it obtains through photosynthesis. For this reason, plant growth is limited by the availability of nitrogen, phosphorus, magnesium and sulphur atoms in addition to the availability of carbon dioxide, water and light energy.

Oceanic carbon pumps

Water is able to dissolve CO_2. There is a negative correlation between the temperature of the water and the amount of CO_2 that can be dissolved. This leads to **vertical deep mixing**, a term used to describe the most important movement of CO_2 in the oceans. It occurs when warm water in oceanic surface currents is carried from the warm tropics to the cold polar regions. Here the water is cooled, making it dense enough to sink below the surface layer, sometimes all the way to the ocean bed. When cold water returns to the surface and warms up again, it loses carbon dioxide to the atmosphere (Figure 1.40, page 30). In this fashion, vertical circulation ensures that carbon dioxide is constantly being exchanged between the ocean and the atmosphere. This vertical circulation also acts as an enormous **carbon pump**, giving the ocean a lot more carbon than it would have if this surface water was not being constantly replenished (Figure 1.41, page 30).

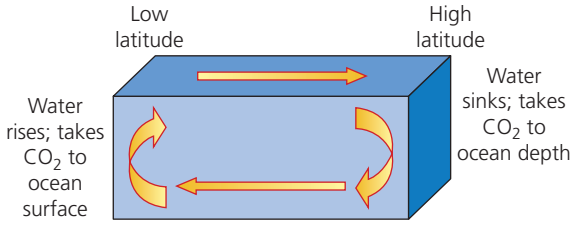


Figure 1.40 The oceanic carbon pump: The concept of vertical deep mixing, where carbon dioxide is transported from the ocean surface to the ocean depths by sinking cold water in the high latitudes. If brought to the surface (for instance, by upwelling) the cold water will warm up and release some of its carbon dioxide to the atmosphere

Living things in the ocean move carbon from the atmosphere into surface waters then down into the deeper ocean and eventually into rocks. This action of organisms moving carbon in one direction is often called a **biological pump**. Carbon gets incorporated into marine organisms as organic matter or structural calcium carbonate. When organisms die, their dead cells, shells and other parts sink into deep water. Decay releases carbon dioxide into this deep water. Some material sinks right to the bottom, where it forms layers of carbon-rich sediments. Over millions of years, chemical and physical processes may turn these sediments into rocks. This part of the carbon cycle can lock up carbon for millions of years.

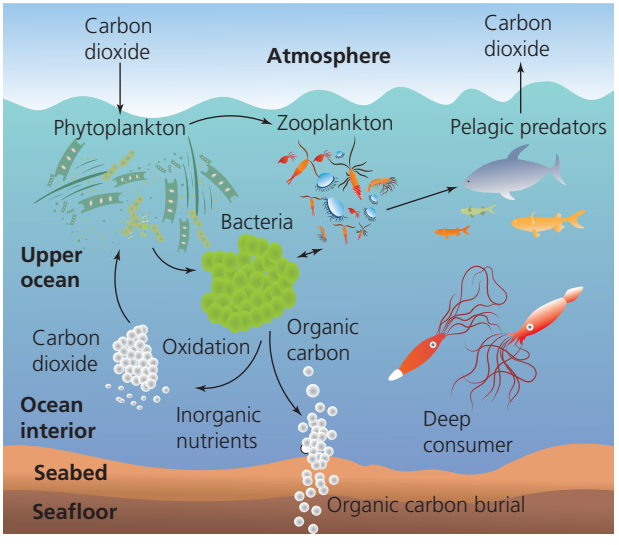


Figure 1.41 The biological carbon pump

Combustion

Combustion occurs when any organic material is reacted (burned) in the presence of oxygen to give off the products of carbon dioxide, water and energy. The organic material can be any vegetation or fossil fuel such as natural gas (methane), oil or coal. Organic materials contain at least carbon and hydrogen and may include oxygen. If other elements are present they also ultimately combine with oxygen to form a variety of pollutant molecules such as sulphur oxides and nitrogen oxides.

Biomass combustion

This is the burning of living and dead vegetation. It includes human-induced burning as well as naturally occurring fires (Figure 1.42). It happens most in:

- the boreal (northern) forests in Alaska, Canada, Russia, China and Scandinavia
- savanna grasslands and forests in Africa, Brazil and northern Australia
- tropical forests in Brazil, Indonesia, Colombia, Ivory Coast, Thailand, Laos, Nigeria, Philippines, Myanmar and Peru
- temperate forests in the US, western Europe and southern Australia
- agricultural waste after harvests in the US and western Europe.

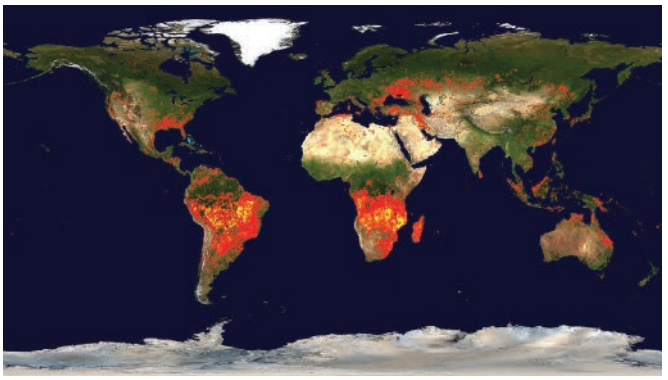


Figure 1.42 The Moderate Resolution Imaging Spectroradiometer (MODIS) on NASA's Terra satellite shows fires around the world

Source: NASA

The relationship between forests and carbon dioxide emissions is complex. Forests have a life cycle: trees die after severe fire, setting the stage for new growth to begin. If a forest fully replaces itself, there will be no net carbon change over that life cycle. The fire consumes only about 10 to 20 per cent of the carbon and immediately emits it into the atmosphere. It kills trees but doesn't consume them. So, new trees grow (storing carbon), old trees decompose (emitting carbon) and the organic layer of the soil accumulates (storing carbon). This balance between simultaneous production and decomposition determines whether

the forest is a net source or sink. Left alone, terrestrial and atmospheric carbon stays more or less in balance. However, increasingly large and/or more frequent fires, possibly made worse by warming temperatures and changing precipitation levels, can change that carbon balance.

Every year, fires burn 3 to 4 million km^2 of the Earth's land surface area, and release more than a billion tonnes of carbon into the atmosphere in the form of carbon dioxide. However, massive old-growth northern latitude forests are also considered a carbon 'sink' because older trees are repositories of decades or centuries of carbon; their heavy canopy blocks sunlight from reaching the forest floor, slowing decomposition of the forest litter.

Volcanic activity

According to the United States Geological Survey (USGS) '*the carbon dioxide released in recent volcanic eruptions has never caused detectable global warming of the atmosphere.*' This is probably because:

- the warming effect of emitted CO_2 is counterbalanced by the large amount of sulphur dioxide that is given out – **conversion of this sulphur dioxide to sulphuric acid, which forms fine droplets,** increases the reflection of radiation from the Sun back into space, cooling the Earth's lower atmosphere
- the amounts of carbon dioxide released have not been enough to produce detectable global warming. For example, all studies to date of global volcanic CO_2 emissions indicate that present-day sub-aerial and submarine volcanoes have released less than 1 per cent of the CO_2 released currently by human activities. It has been proposed that intense volcanic release of carbon dioxide in the deep geologic past did cause a large enough increase in atmospheric CO_2 to cause a rise in atmospheric temperatures and possibly some mass extinctions, though this is a topic of scientific debate.

Hydrocarbon extraction and burning – cement manufacture

Dead plants or animals turn into fossil fuels following burial. The pressure from multiple layers of sediment leads to an anoxic (oxygen free) environment that allows for decomposition to take place without oxygen. When this is combined with heat from the Earth, the carbon in organic molecules is rearranged to form other compounds. The type of material that is buried helps to determine what the final product will be. Animal remains tend to form petroleum (crude oil) while plant matter is more likely to form coal and natural gas. When these fossil fuels are extracted from the ground and then burnt, carbon dioxide and water are released into the atmosphere.

Cement manufacture contributes CO_2 to the atmosphere when calcium carbonate is heated, producing lime and carbon dioxide. CO_2 is also produced by burning the fossil fuels that provide the heat for the cement manufacture process.

It is estimated that the cement industry produces around 5 per cent of global anthropogenic CO_2 emissions, of which 50 per cent is produced from the chemical process itself, and 40 per cent from burning fuel to power that process. The amount of CO_2 emitted by the cement industry is more than 900 kg of CO_2 for every 1,000 kg of cement produced.

Figure 1.43 (page 32) shows the acceleration in the growth of CO_2 emissions over the last 260 years. Between 2017 and 2018, global energy-related CO_2 emissions rose by 1.7 per cent to a historic high of 33.1 GtC. Emissions from all fossil fuel burning increased, with electricity generation accounting for just over 60 per cent of this growth. Coal burning (mainly in China, India and the USA) accounted for 10 GtC of CO_2 and 85 per cent of the net increase in emissions. There was a decline in Germany, Japan, Mexico, France and the United Kingdom as a result of increased use of renewable energy.

Farming practices

When soil is ploughed, the soil layers invert, air mixes in, and soil microbial activity dramatically increases. It results in soil organic matter being broken down much more rapidly, and carbon being lost from the soil into the atmosphere. In addition to the effect on soil from ploughing, emissions from the farm tractors increases carbon dioxide levels in the atmosphere.

According to a 2019 report from the United Nations Food and Agriculture Organization (FAO):

- agriculture and forestry cause almost one-quarter of all human greenhouse gas emissions
- a third of all soils are degraded, releasing 78 GtC into the atmosphere
- the largest source of carbon emissions is enteric fermentation – where methane (CH_4) is released by livestock following digestion
- livestock account for 14.5 per cent of all anthropogenic emissions
- greenhouse gases from rice paddies make up ten per cent of total agricultural emissions, while the burning of tropical grasslands accounts for five per cent.

The FAO also reported that 44 per cent of agriculture-related greenhouse gas outputs occurred in Asia, followed by the Americas (25 per cent), Africa (15 per cent), Europe (12 per cent) and Oceania (4 per cent). This regional distribution has been constant over the last decade.

Land use change

CO_2 emissions that result from land use change (mainly **deforestation**) account for up to 30 per cent of anthropogenic CO_2 emissions.

Deforestation

Most deforestation is driven by the need for extra agricultural land. Often subsistence farmers will clear a few hectares to feed their families by cutting down trees and burning them in a process known as 'slash and burn' agriculture.

Logging operations also remove forest. Loggers, some of them acting illegally, also build roads to access more and more remote forests, which in turn leads to further deforestation. Forests are also cut as a result of growing urban sprawl. Not all deforestation is intentional. Some is caused by a combination of human and natural factors like wildfires and subsequent overgrazing, which may prevent the re-establishment of young trees.

According to Global Forest Watch, there was a total of 361 Mha of tree cover lost globally between 2001 and 2018 (roughly one and a half times the size of the UK). This has released an estimated 98.7 GtC of CO_2 emissions.

At the same time, planting of trees has resulted in forests being established or extended. One study by NASA and the University of Maryland in 2018, using modern satellite imagery, shows that tree cover is increasing. The loss caused by deforestation in the tropics is being more than compensated for by reforestation in temperate areas, particularly agricultural regions in Asia. The long-term outcome could be a complete loss of rainforest but an increase in temperate forest on former natural grasslands (Figure 1.44).

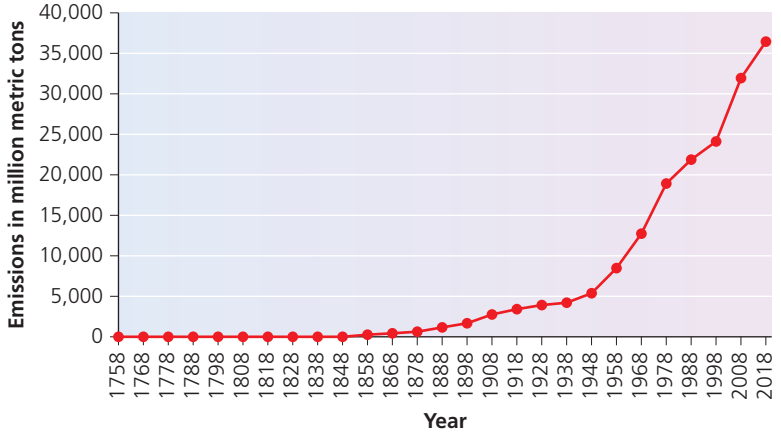


Figure 1.43 Global CO_2 emissions from fossil fuel combustion and industrial processes, 1758 to 2018

Source: Statista

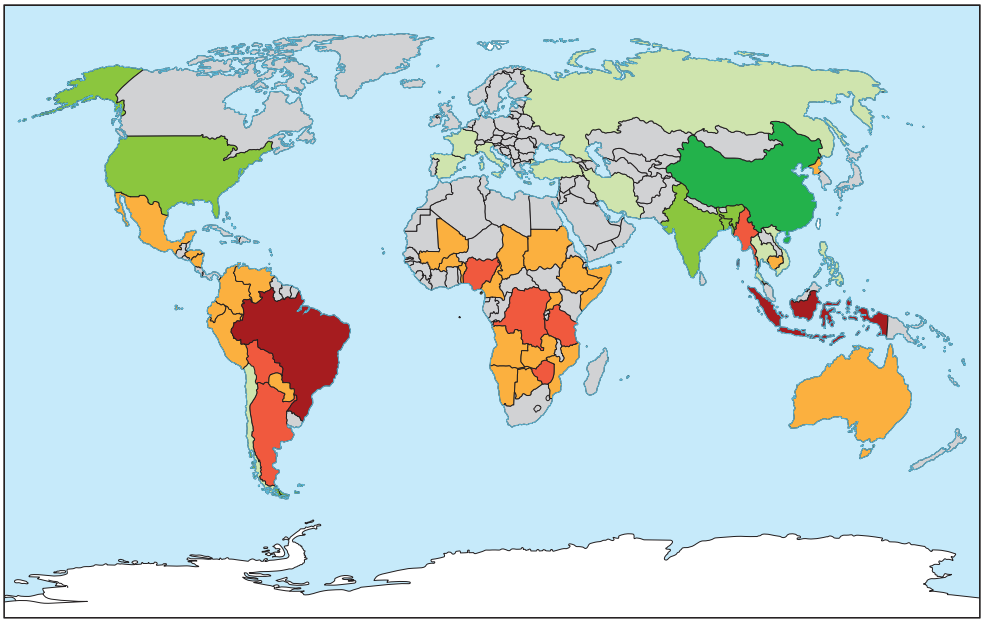


Figure 1.44 Countries with large net changes in forest area, 1990–2015

Source: FAO

Skills focus

You should be able to read and interpret atlas maps. Think about how you might describe the contrast in the distribution of those areas with a net gain in forest and those with a net loss.

When forests are cleared for conversion to agriculture or pasture, a large proportion of the above-ground biomass may be burned, rapidly releasing most of its carbon into the atmosphere. Some of the wood may be used as wood products and so preserved for a longer time. Forest clearing also accelerates the decay of dead wood, litter and below-ground organic carbon.

Figure 1.45 (page 34) shows how deforestation changes the carbon cycle but it is not the only way it affects climate.

Links

There are clear links between deforestation and how this affects both the water cycle and the carbon cycle.

Urban growth

For the first time in human history, over half the world's population now lives in urban areas. As a proportion of global population, the urban population is expected to reach 60 per cent by 2030, with urban areas growing at a rate of 1.3 million people every week. As cities grow, the land use changes from either natural vegetation or agriculture to one which is built up. The CO_2 emissions resulting from energy consumption for transport, industry and domestic use, added to the CO_2 emitted in the cement manufacture required for all the buildings and infrastructure, have increased.

In 2019, according to the United Nations Community Programme, urban areas were estimated to be responsible for around 75 per cent of global carbon emissions. If something does not change this will rise with increasing city populations.

The distribution of emissions is highly concentrated: with the 20 highest emitting cities contributing a large proportion of global energy-related carbon emissions (Figure 1.46, page 34). This concentration is likely to increase in the future.

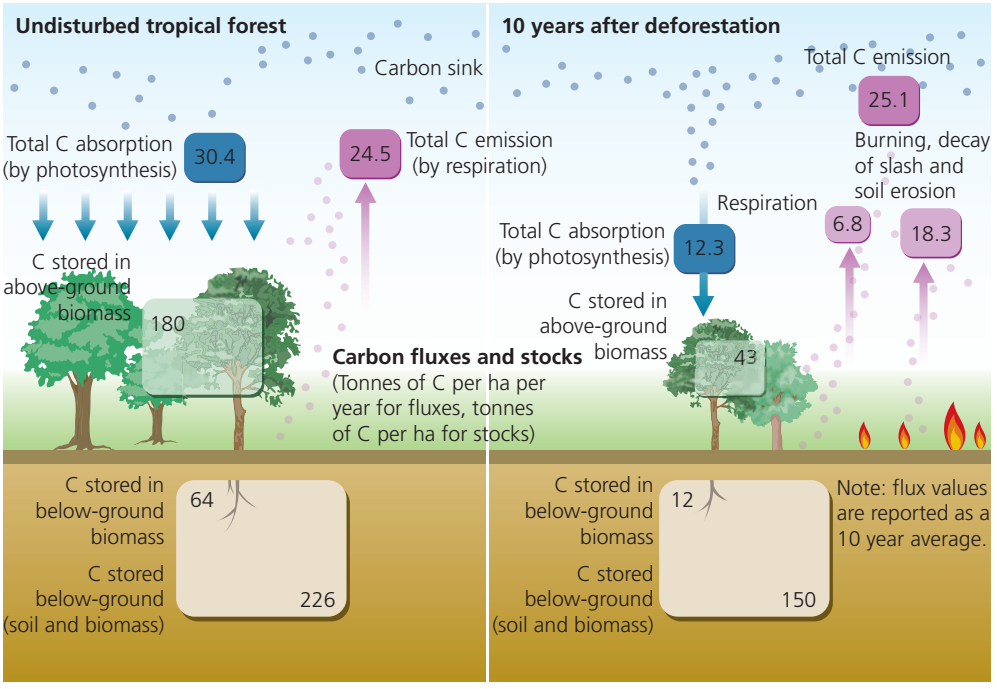


Figure 1.45 The effects of deforestation on the carbon cycle

Source: GRIDA

Carbon sequestration

The process of **carbon sequestration** involves capturing CO_2 from the atmosphere and putting it into long-term storage. There are two primary types of sequestration – geologic and terrestrial or biologic.

Geologic sequestration

CO_2 is captured at its source (for example, power plants or industrial processes) and then injected in liquid form into stores underground. These could be depleted oil and gas reservoirs, thin, uneconomic coal

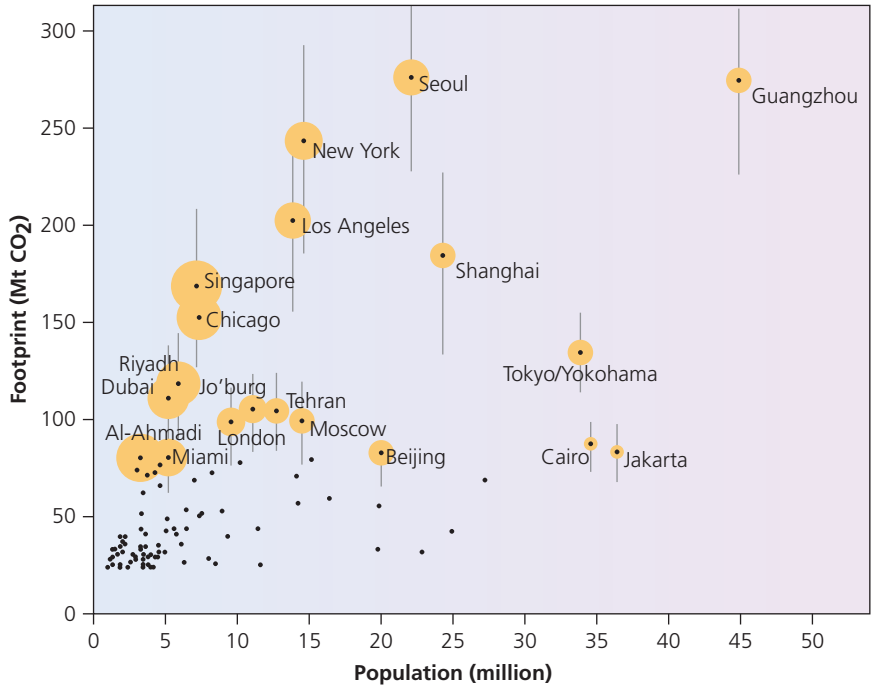


Figure 1.46 The top 20 urban areas with the largest carbon footprint (million tonnes carbon equivalent)

Source: Environment Research Letters, IOPscience

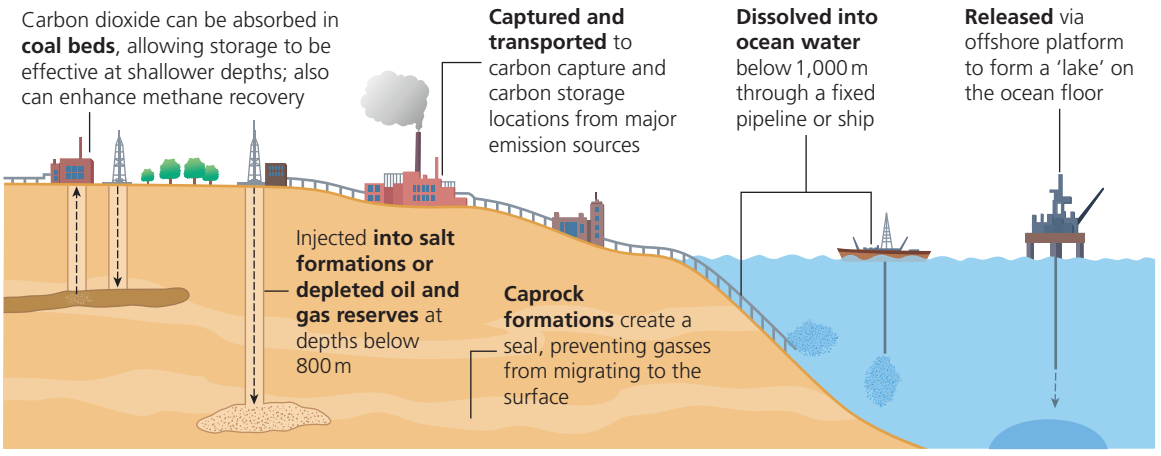


Figure 1.47 Geologic carbon sequestration

seams, deep salt formations and the deep ocean (Figure 1.47). This is still at the experimental stage.

The ocean is very capable of absorbing much more additional carbon than terrestrial systems simply because of its sheer size. An advantage of ocean carbon sequestration is that the carbon sequestered is quite literally 'sunk' within weeks or months of being captured from the air/water. Once in the deep ocean it is in a circulation system commonly measured in thousands of years. By the time this carbon reaches the seabed it has entered the Earth's geological cycle.

Terrestrial sequestration

This involves the use of plants to capture CO_2 from the atmosphere and then to store it as carbon in the stems and roots of the plants as well as in the soil (Figure 1.48 page 36). The aim is to develop a set of land management practices that maximises the amount of carbon that remains stored in the soil and in plant material for the long term. Most authorities also believe that the enrichment of plant ecosystems is a positive environmental action with many associated benefits, including the enrichment of wildlife.

There are disadvantages to terrestrial sequestration:

- For instance, a forest planted to capture carbon might lose that carbon back to the air in a catastrophic forest fire or if the forest suffers disease or infestation.
- Land-based sequestration plantations are slow growing and require active monitoring and management for the lifetime of the plantation, usually many decades.
- The carbon within those systems is never removed permanently from the atmospheric system.

Review question

7 What are the differences between changes to the carbon cycle brought about by natural variation and those brought about by human activity?

Effects of the changing carbon budget

It is clear that CO_2 levels have varied through time. It is certain that 100 million years ago carbon dioxide values were many times higher than now though the exact value is in doubt. Five hundred million years ago atmospheric CO_2 was some 20 times higher than present values. It dropped, then rose again 200 million years ago to a maximum of four to five times present levels. It then followed a slow decline until recent pre-industrial time.

The only way to calculate the effects of the changing levels of carbon is to build a computer model. A number of such models have been built over the years. They can have between 50 and 100 interacting equations describing all the different processes of the carbon cycle. The result of this is that models only predict possibilities, not probabilities.

Impact on the land

The impact of increasing atmospheric CO_2 on the land has been subject to intense research. Unfortunately the results are unclear because the study has, so far, been over a relatively short time period. This is coupled with the fact that there are so many other variables that

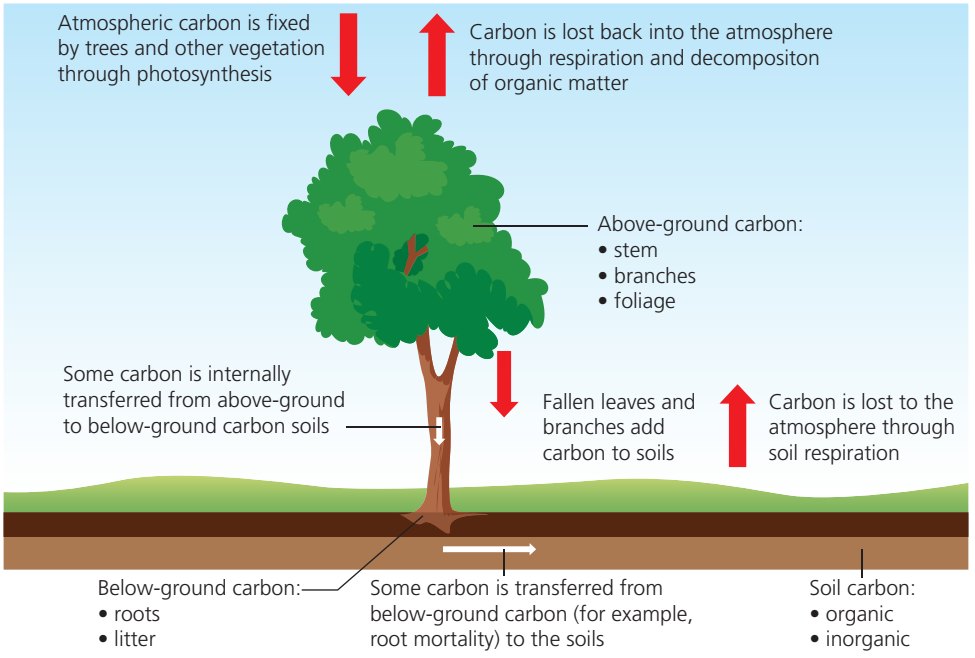


Figure 1.48 Biologic or terrestrial carbon sequestration

could have an impact on the land and the atmosphere. These are both human and physical variables. Some of the impacts are summed up in Figure 1.49.

Note: The National Ocean and Atmospheric Administration (NOAA) estimates that permafrost in the Northern Hemisphere holds between 1,460 and 1,600 GtC. If just 10 per cent of this permafrost were to thaw, it could release enough extra carbon dioxide to the atmosphere to raise temperatures an additional 0.7 °C by 2100.

Impact on the oceans

Many of the observed physical and chemical changes in the ocean are consistent with increasing atmospheric CO_2 and a warming climate. Unfortunately, because of the complex nature of the chemistry and biochemistry of the oceans and its inhabitants, many of the causative links to climate change are still not well understood. It is difficult to predict the precise rate, magnitude and direction of change of CO_2 uptake and how that affects acidity, salinity, storminess and nutrient enrichment. It is even more difficult to map these effects at a local scale.

Ocean acidification

About 30 per cent of the CO_2 that has been released into the atmosphere has diffused into the ocean through direct chemical exchange. Dissolving carbon dioxide in the ocean creates carbonic acid. This makes the slightly alkaline ocean become a little less alkaline. Since 1750, the pH of the ocean's surface has dropped by 0.1, a 30 per cent change in acidity.

The impact of ocean acidification on ecosystems is largely unknown, with the exception of coral reefs. It affects marine organisms in two ways:

- Carbonic acid reacts with carbonate ions in the water to form bicarbonate. However, those same carbonate ions are what animals like coral and many planktonic species need to create their calcium carbonate shells.
- With less carbonate available, the animals need to expend more energy to build their shells. As a result, the shells end up being thinner and more fragile.

Coral reefs provide food and livelihood security for some 500 million people worldwide. Significant reef loss, due to loss of carbonate ions, and the consequent fall in marine biodiversity threatens the survival of coastal communities through reduced food availability and a reduced capacity of coastlines to buffer the impact of sea level rise, including increased storm surges.

Polar and sub-polar marine ecosystems are projected to become so low in carbonate ions within this century that waters may actually become corrosive to unprotected shells and skeletons of organisms currently living there.

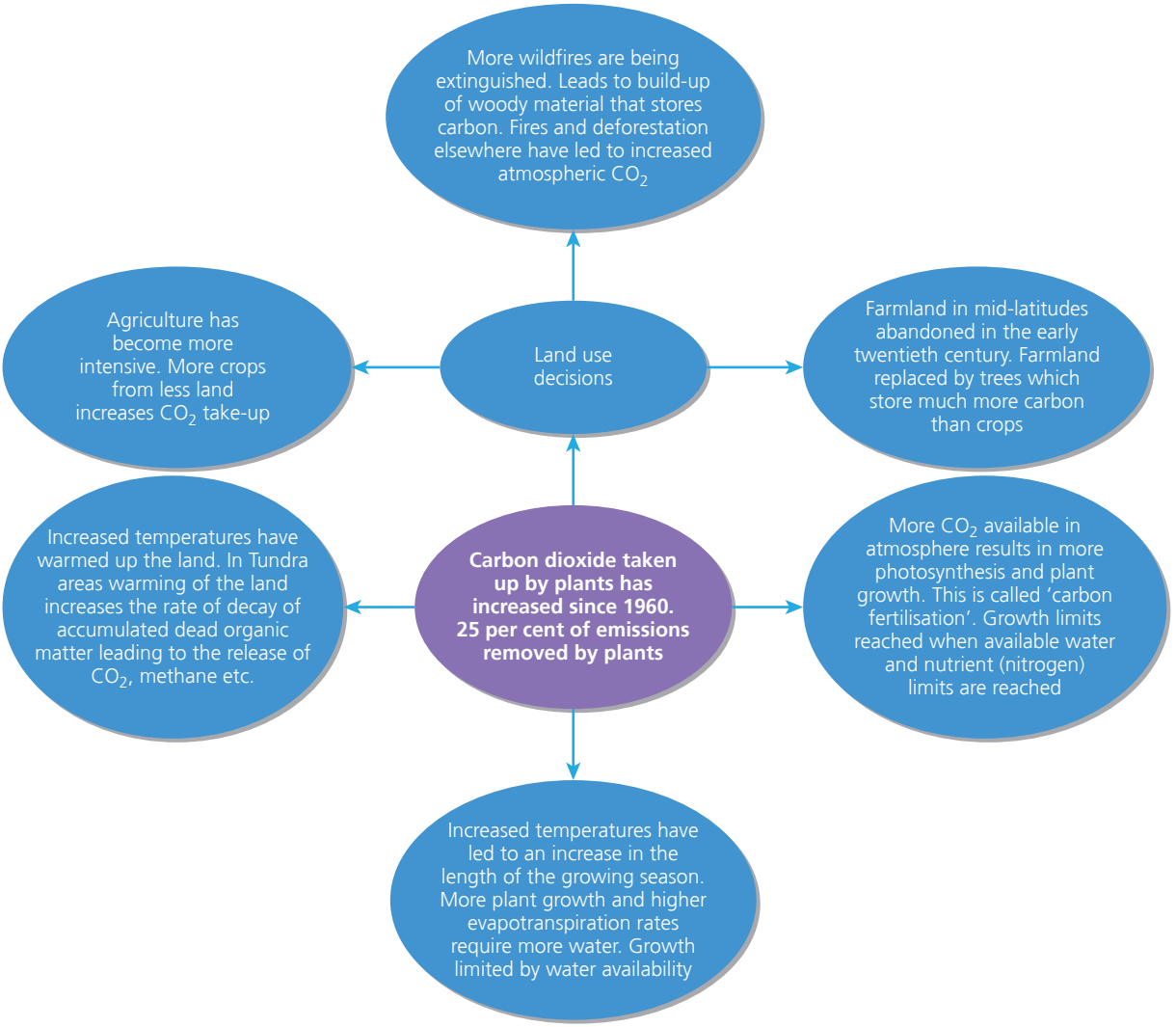


Figure 1.49 The effects of changes in the carbon cycle on the land.

In 2007, farmed oyster larvae off the coast of Oregon and Washington in the USA began dying by the millions. It was found that these losses were directly linked to ocean acidification. The rise in acidity and subsequent oyster crash took a significant toll on coastal communities – from 2005 to 2009 lost production cost millions of dollars in lost sales. Further research showed evidence of acidic seawater rising up from the ocean depths and that the water rising from the deep ocean today holds CO_2 that was absorbed approximately 30 to 50 years ago. Benoit Eudeline, chief hatchery scientist for Taylor Shellfish Farms, the largest US producer of farmed shellfish, likened the current situation to 'sitting on a ticking time bomb'.

A more optimistic viewpoint is that the more acidic seawater is, the better it dissolves calcium carbonate rocks (chalk and limestone). Over time this reaction will allow the ocean to soak up excess CO_2 because the more acidic water will dissolve more rock, release more carbonate ions and increase the ocean's capacity to absorb CO_2.

Ocean warming

Warmer oceans, a product of climate change, could decrease the abundance of phytoplankton, which grow better in cool, nutrient-rich waters. This could limit the ocean's ability to take carbon from the atmosphere through the biological carbon pump and lessen the effectiveness of the oceans as a carbon sink. On the other hand, CO_2 is essential for plant and phytoplankton growth. An increase in CO_2 could increase their growth by fertilising those few species of phytoplankton and ocean plants (like sea grasses) that take carbon dioxide directly from the water.

Ocean warming also kills off the symbiotic algae which coral needs in order to grow, leading to bleaching and eventual death of reefs.

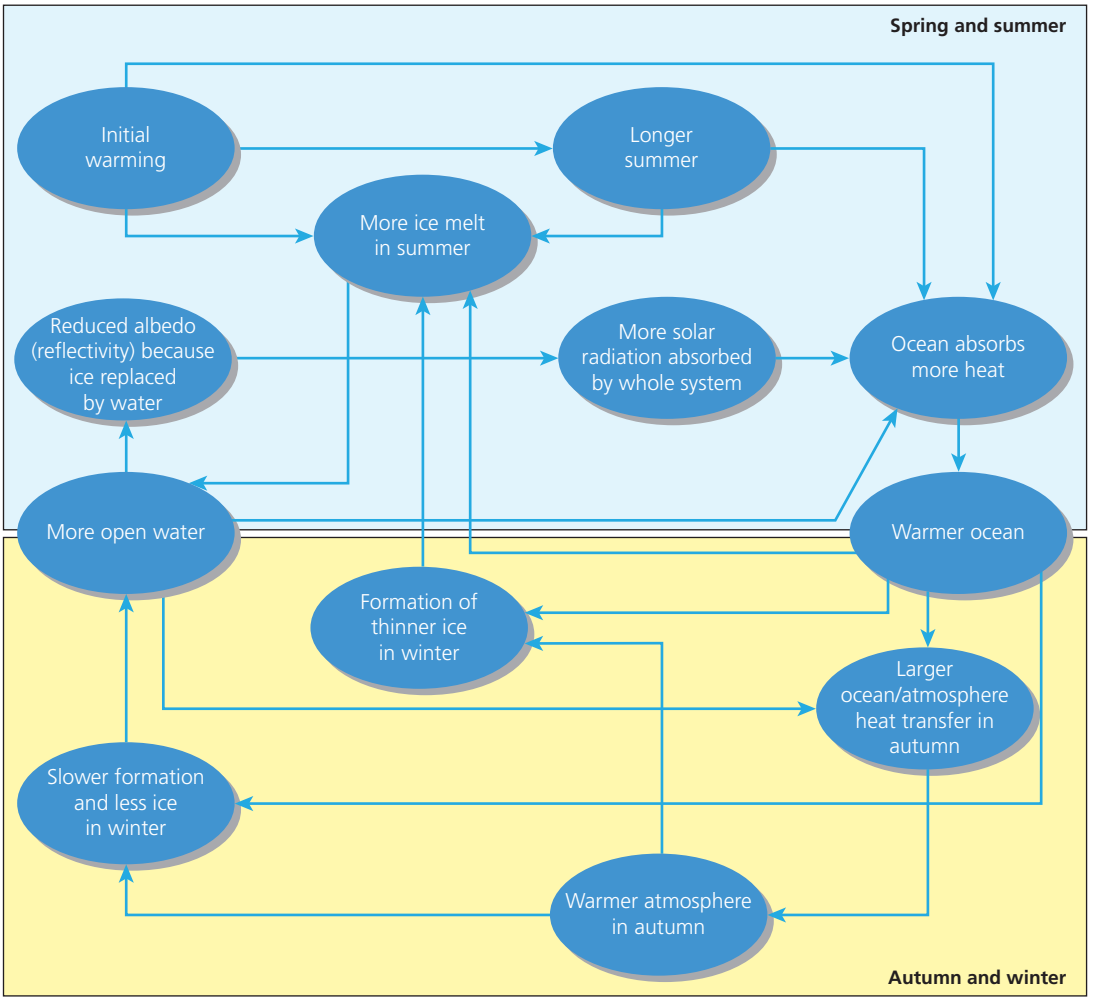


Figure 1.50 A systematic view of the oceanic carbon cycle showing the positive feedback effects of oceanic warming

Melting sea ice

In the last 35 years satellites monitoring sea ice in the Arctic have measured its retreat at 12.8 per cent per decade. When sea ice melts, it is not just an indicator of a warming climate but also part of a feedback mechanism because the highly reflective ice is replaced by more heat absorbent water. When it starts to melt the ocean is able to absorb more sunlight, which in turn amplifies the warming that caused it to melt in the first place (Figure 1.50).

Sea ice also provides a unique habitat for algae that appear in more concentrated forms and with more fat content in the ice. The loss of those ice-bound algae affects marine life all the way up the food chain, from krill and fish to seals, walruses and polar bears. Added to this is the fact that animals like the polar bear, that rely on sea ice to get to their main food source of seals, can no longer travel upon it.

Figure 1.51 A polar bear standing on an ice floe

Ocean salinity

A decrease in salinity in the deep North Atlantic has been observed. This is probably caused by higher levels of precipitation and higher temperatures, the result of increased atmospheric carbon. The precipitation leads to higher river run-offs that eventually reach the sea.

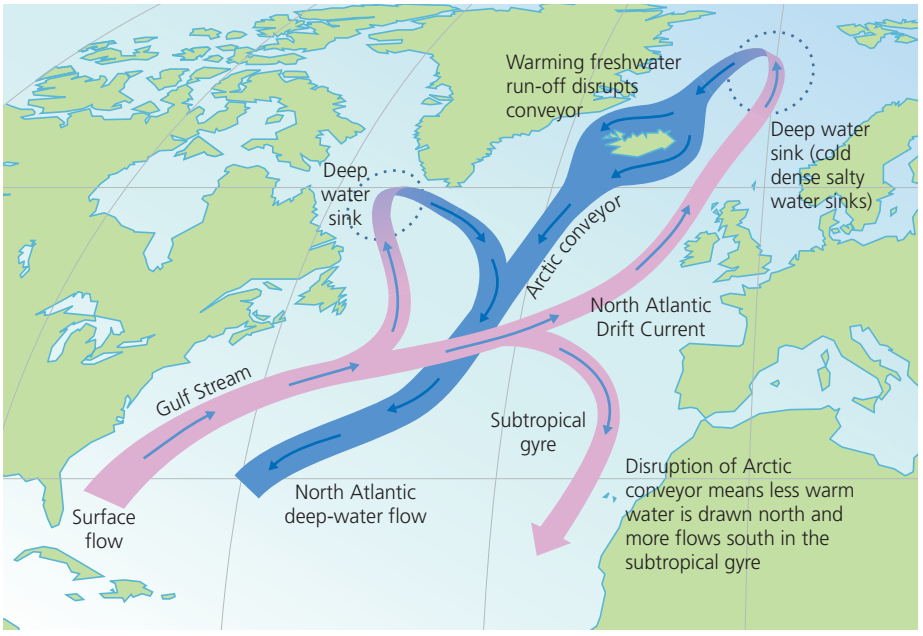


Figure 1.52 The effect of changing ocean salinity on the Atlantic Ocean currents

These higher temperatures are also causing melting of the Greenland ice sheet and many alpine glaciers. This too will lead to an increase in fresh water reaching the oceans. These changes have been linked to a possible slowing down of the large-scale oceanic circulation in the North-East Atlantic (Figure 1.52). This in turn will have an effect on the climate of North West Europe.

Sea level rise

Although global sea levels have been more or less constant for the last 5,000 years they are subject to change. Studies of coastal landforms show that they have been much lower in the past than they are today. The last glacial retreat led to a worldwide rise in sea levels about 10,000 years ago. This increase was caused by the melting of land-locked freshwater ice.

Research now indicates that sea levels worldwide have been rising at a rate of 3.1 mm/year since the early 1990s. Once again, higher atmospheric CO_2 and the resultant increase in atmospheric temperatures have led to:

- **Melting of terrestrial ice:** there has been an increased rate of summer melting as well as a drop in snowfall in the shorter winters. This imbalance results in a significant net gain in water entering the oceans from rivers against evaporation from the ocean. Added to this is the fact that the massive ice sheets of Antarctica and Greenland are moving more quickly towards the oceans due to the increased amount of meltwater lubricating their bases.
- **Thermal expansion:** when water heats up, it expands. About half of the past century's rise in sea level is thought to be attributable to warmer oceans having a greater volume and so occupying more space. Accurate measurements of this phenomenon have not yet been possible.

If the Earth continues to warm up then we can expect the oceans to rise between 0.6 and 2.5 m by 2100. This is not an exact science and Figure 1.53 (page 40) shows a range of predictions for rising sea levels.

Impact on the atmosphere

Key terms

Enhanced greenhouse effect – The impact on the climate from the additional heat retained due to the increased amounts of carbon dioxide and other greenhouse gases that humans have released into the Earth's atmosphere since the industrial revolution.

Geo-sequestration – The technology of capturing greenhouse gas emissions from power stations and pumping them into underground reservoirs.

Radiative forcing – The difference between the incoming solar energy absorbed by the Earth and energy radiated back to space.

Soil organic carbon (SOC) – The organic constituents in the soil: tissues from dead plants and animals, products produced as these decompose and the soil microbial biomass.

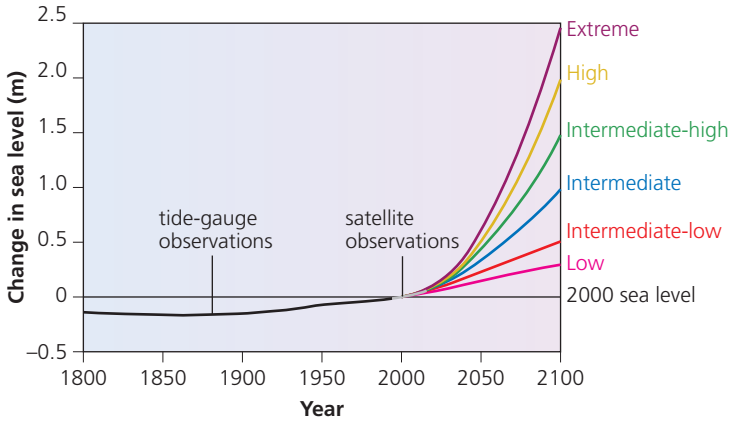


Figure 1.53 Observed and predicted sea level change between 1800 and 2100 – predictions based on different greenhouse gas pathways

Source: NOAA

Although the land and oceans will take up most of the extra CO_2, as much as 20 per cent may remain in the atmosphere for many thousands of years. This is significant because CO_2 is the most important gas for controlling the Earth's temperature. Carbon dioxide, methane and halocarbons are greenhouse gases that absorb a wide range of energy – including infrared energy (heat) emitted by the Earth – and then re-emit it. The re-emitted energy travels out in all directions but some returns to Earth where it heats the surface. Without greenhouse gases, the Earth would be a frozen –180°C. With too many greenhouse gases, the Earth would be like Venus, where the greenhouse atmosphere keeps temperatures around 400°C.

Clearly, the fact that there is a greenhouse effect at all is good for life on Earth. The problem that is facing us is that of an **enhanced greenhouse effect** (Figure 1.54). This is where the extra CO_2 and other greenhouse gases in the atmosphere are causing something called **radiative forcing**.

The concept of radiative forcing is that energy is constantly flowing into the atmosphere in the form of sunlight that always shines on half of the Earth's surface. Some of this sunlight (about 30 per cent) is reflected back to space and the rest is absorbed by the planet. Some of this absorbed energy is radiated back into the much colder surrounding space as

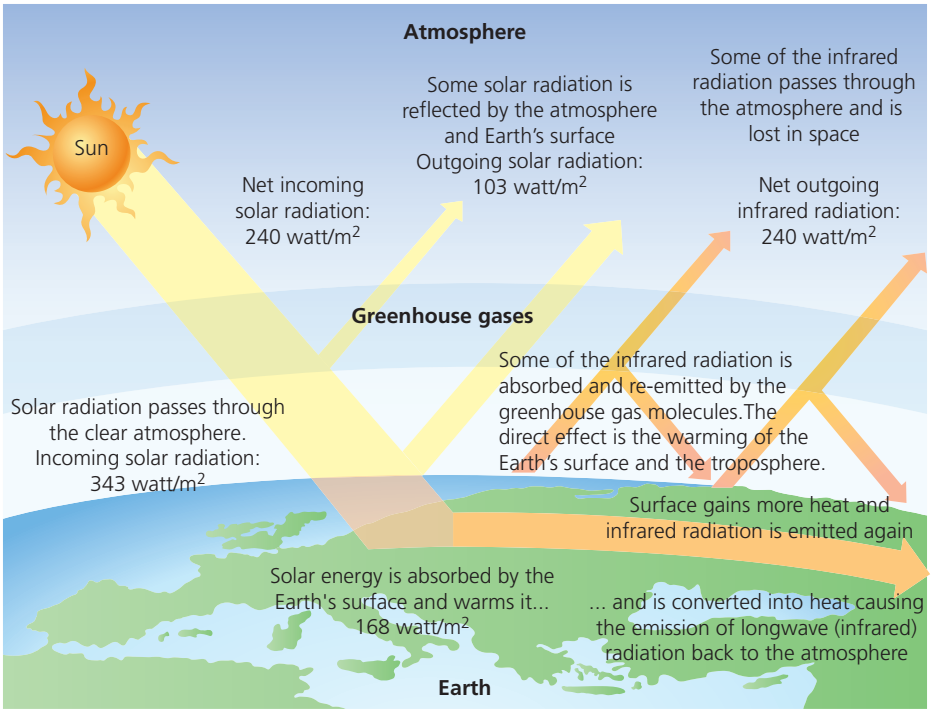


Figure 1.54 The greenhouse effect

infrared energy. If the balance between the incoming and the outgoing energy is anything other than zero there has to be some warming (or cooling, if the number is negative) going on. The amount that the Earth's energy budget is out of balance is called the radiative forcing and is a measure of recent human activities. It is measured in watts/m² of the Earth's surface.

Studies have shown that prior to 1750 radiative forcing was negligible. Since then it has increased, not only because of increased greenhouse gas emissions but also changing albedos because of land use changes. Measurement of the actual amount of radiative forcing is difficult because of many complicating factors including natural changes in solar radiation and the effects of aerosols such as carbon particles from diesel exhausts (which lead to warming).

The current level of radiative forcing, according to the IPCC AR5, is 2.8 watts/m^2 (with a range of uncertainty from 2.5 to 3.1). If CO_2 levels continue to rise at projected rates, experts predict that the Earth will become much hotter, possibly hot enough to melt much of the existing ice cover. Figure 1.55 depicts projected surface temperature changes up to 2060 as estimated by NASA's Global Climate Model.

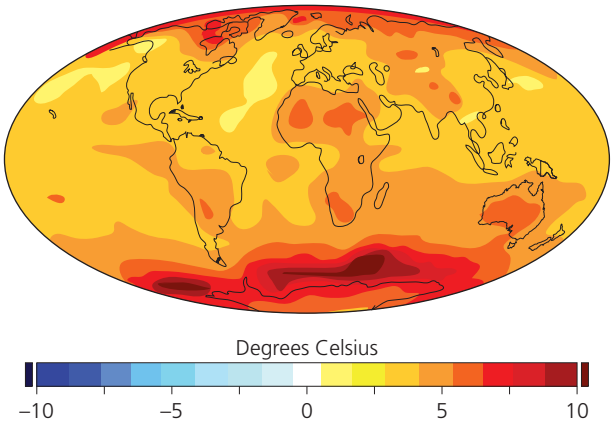


Figure 1.55 Predicted increases in surface air temperature 1960 to 2060 according to NASA's Global Climate Model

Review questions

8 What are the likely effects of the predicted sea level rises on coastal communities?

9 How might the response of those communities depend on their level of development?

1.4 Water, carbon, climate and life on Earth

Carbon forms the key component for all known life on Earth. It bonds with other elements such as oxygen, hydrogen and nitrogen to form complex molecules. Water too makes life as we know it possible. Every drop cycles continuously through air, land and sea to be used by someone (or something) else in the cycle. Although there is a lot of water on Earth, only a fraction of one per cent supports all life on land. Climate change and growing populations are increasing the pressures on that reserve.

Water and carbon cycles and the atmosphere

As stated above, the increased emission of CO_2 is warming the atmosphere. This increased temperature results in higher evaporation rates and a wetter atmosphere, which leads to a positive feedback situation of further warming. Carbon dioxide causes about 20 per cent of the Earth's greenhouse effect, water vapour accounts for about 50 per cent and clouds 25 per cent. The rest is caused by small particles (aerosols) and minor greenhouse gases like methane.

When carbon dioxide concentrations rise, air temperatures go up. The oceans warm up and more water vapour evaporates into the atmosphere, which then amplifies greenhouse heating. Although CO_2 contributes less to the overall greenhouse effect than water vapour, scientists have found that it is CO_2 that sets the temperature. It controls the amount of water vapour in the atmosphere and therefore the size of the enhanced greenhouse effect. This is summarised in Figure 1.56 (page 42).

This rise in temperature is not all the warming that we will see based on current CO_2 concentrations. There is a time lag between the increase in CO_2 and increased warming because the ocean soaks up heat. This means that the Earth's temperature will increase at least another 0.6 °C because of carbon dioxide already in the atmosphere.

Climate change mitigation

The various means of mitigation (efforts to reduce or prevent greenhouse gases) are summarised in Figure 1.57 (page 42).

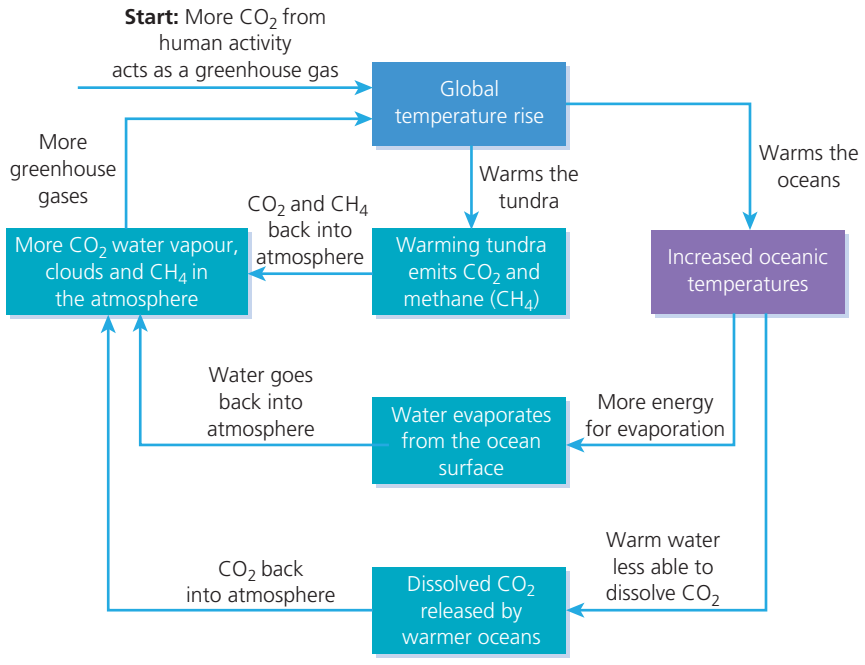


Figure 1.56 A system diagram for the impact of water and carbon on climate change showing positive feedback

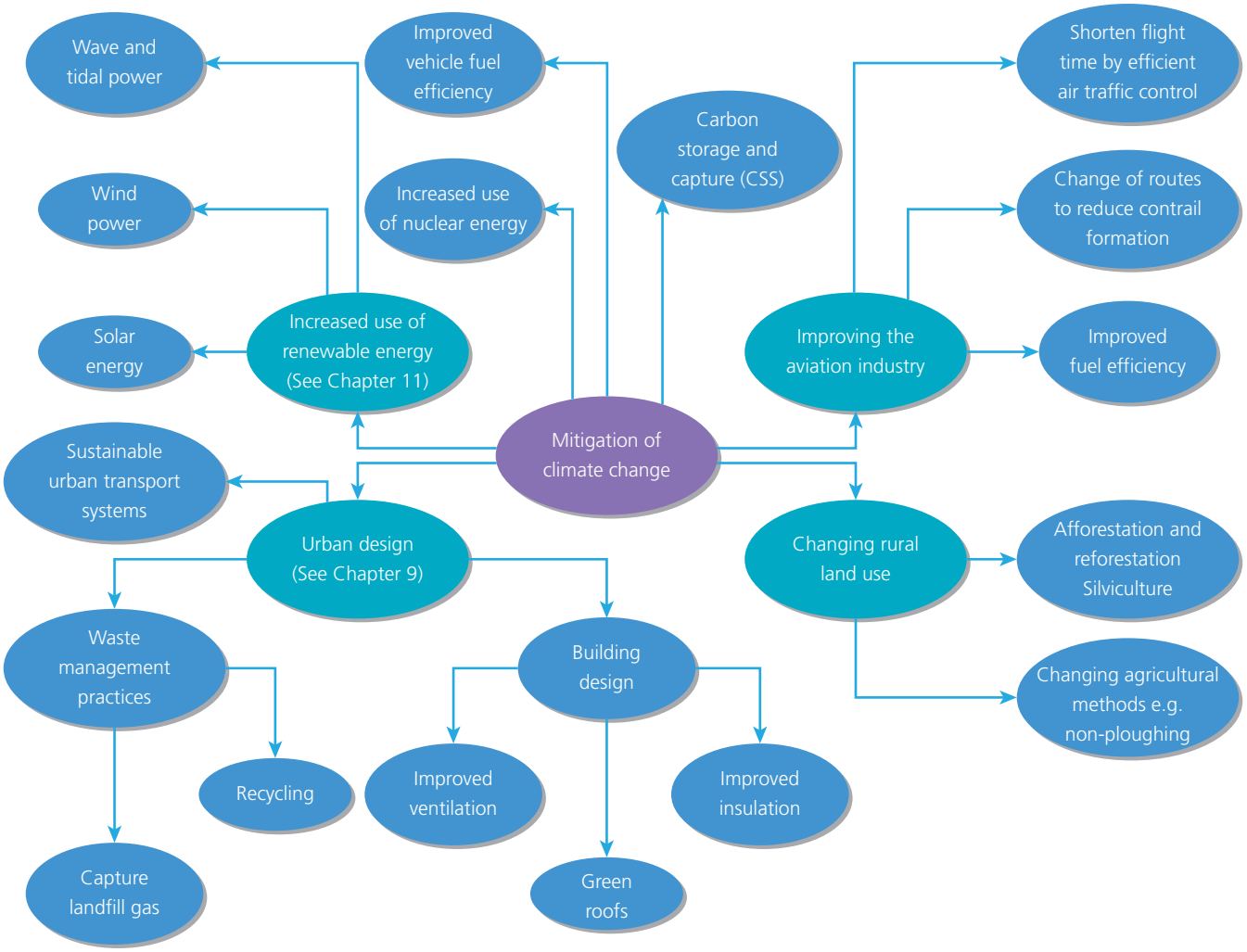


Figure 1.57 Management of climate change

Carbon capture and sequestration (CCS) technologies

Carbon capture and storage (CCS) is a technology that can capture up to 90 per cent of CO_2 emissions produced from the use of fossil fuels in electricity generation and industrial processes, preventing the carbon dioxide from entering the atmosphere.

The CCS chain consists of three parts (see Figure 1.58):

- **Capturing** the CO_2: Capture technologies allow the separation of CO_2 from gases produced in electricity generation and industrial processes by one of three methods: pre-combustion capture, post-combustion capture and oxy-fuel combustion.

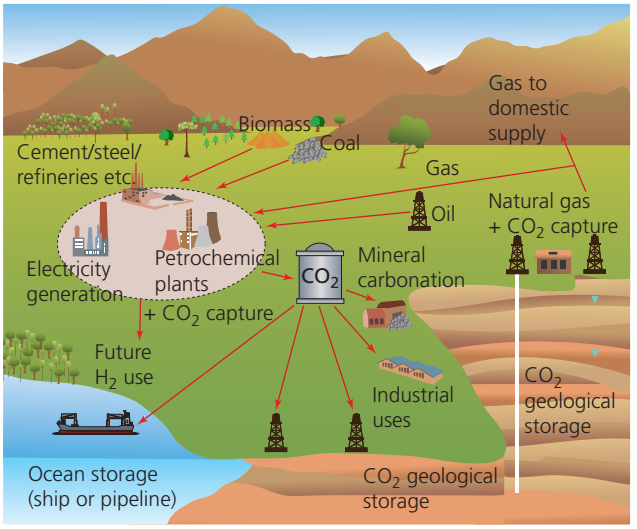


Figure 1.58 Possible CCS systems

Source: IPCC

Boundary Dam CCS plant

An example of CCS in action is at the 110-megawatt coal power and CCS plant in Saskatchewan, called Boundary Dam, built by the provincial utility SaskPower. It is a coal-fired power station complex that has been retrofitted to capture 90 per cent of its CO_2 output (approximately 1 million tonnes per year). The CO_2 will eventually be piped 66 km to the Weyburn Oil Unit and injected into an oil-bearing formation at 1,500 m depth. This will add pressure to the oil-bearing rock and so help push more oil out of the ground, a process called enhanced oil recovery (EOR). Until that is ready it will be injected into local salt formations. The capture process was started in October 2014 and CO_2 injection started in April 2015.

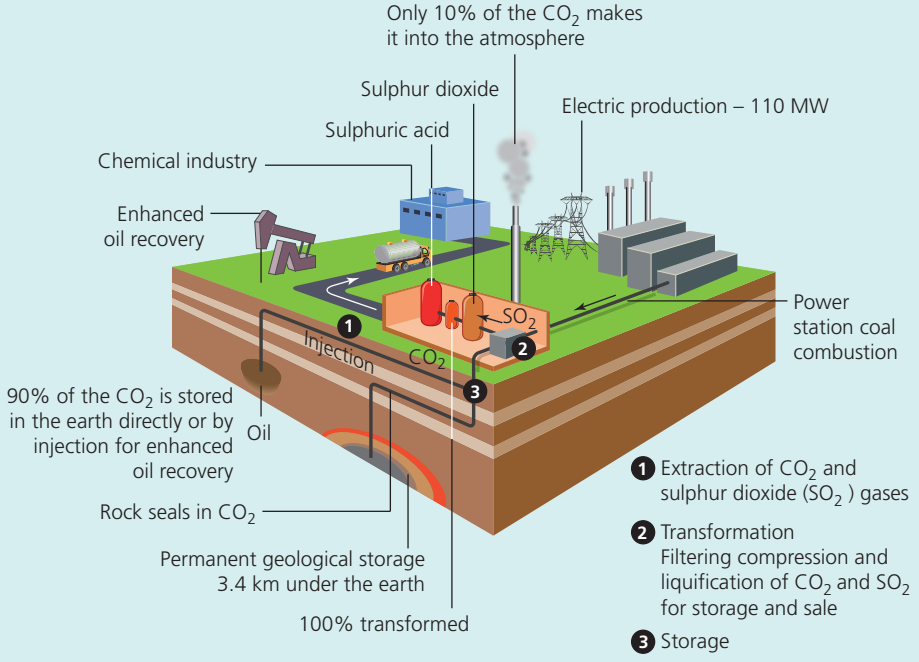


Figure 1.59 Carbon capture and storage at Boundary Dam power station, Canada

Source: SaskPower

CCS imposes big costs and energy penalties: the Boundary Dam plant's CCS unit cost $800 million to build and consumes 21 per cent of the coal plant's power output in order to scrub out the carbon dioxide and compress it into a liquid for burial. It is hoped that this extra cost will be offset by the extra oil recovered from the Weyburn oil field.

- **Transporting** the CO_2 by pipeline or by ship to the storage location: Millions of tonnes of CO_2 are already transported annually for commercial purposes by road tanker, ship and pipelines.
- **Storing** the carbon dioxide emissions securely underground in depleted oil and gas fields, and deep saline aquifer formations several kilometres below the surface or the deep ocean.

CCS systems could be used to extract a greater percentage of oil and gas out of existing reservoirs by the CO_2 being injected under such pressure as to force the oil or gas out. Although this would partly pay for the CCS technology, it would also enhance the original problem by producing more fossil fuel for burning. When CO_2 is stored in deep geological formations it is known as **geo-sequestration** (see Figure 1.60. CO_2 is converted into a high pressure liquid-like form known as 'supercritical CO_2' which behaves like a runny liquid. This supercritical CO_2 is injected directly into sedimentary rocks. The rocks may be in old oil fields, gas fields, saline formations or thin coal seams. Various physical (for example, impermeable 'cap rock') and geochemical trapping mechanisms prevent the CO_2 from escaping to the surface.

Captured CO_2 could also be stored in the ocean by a variety of means, as shown in Figure 1.61. This method has the main disadvantage of the CO_2 causing acidification of the oceans and all the problems that arise from that.

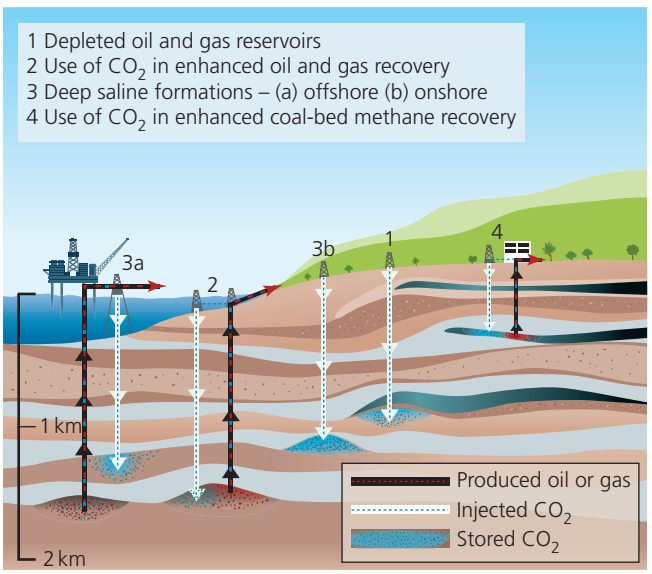


Figure 1.60 Methods for storing CO_2 in deep underground geological formations

Source: IPCC

Changing rural land use

Carbon stores can be improved by ensuring that carbon inputs to the soil are greater than carbon losses from it. There are a variety of different strategies depending on land use, soil properties, climate and land area.

Grasslands

These offer a global greenhouse gas mitigation potential of 810 million tonnes of CO_2 (in the period up to 2030), almost all of which would be sequestered in the soil. Soil carbon storage in grasslands can be improved by:

- avoidance of **overstocking** of grazing animals
- **adding manures and fertilisers** that have a direct impact on **soil organic carbon (SOC)** levels through the added organic material. There are also the indirect benefits of increasing plant productivity and stimulating soil biodiversity (for example, with earthworms that help degrade and mix the organic material)
- **revegetation**, especially using improved pasture species and legumes, can increase productivity, resulting in more plant litter and underground biomass, which can add to the SOC stock
- **irrigation** and water management can improve plant productivity and the production of SOM.

Croplands

Techniques for increasing SOC include the following:

- **mulching**, which can add organic matter. If crop residues are used, mulching also prevents carbon losses from the system

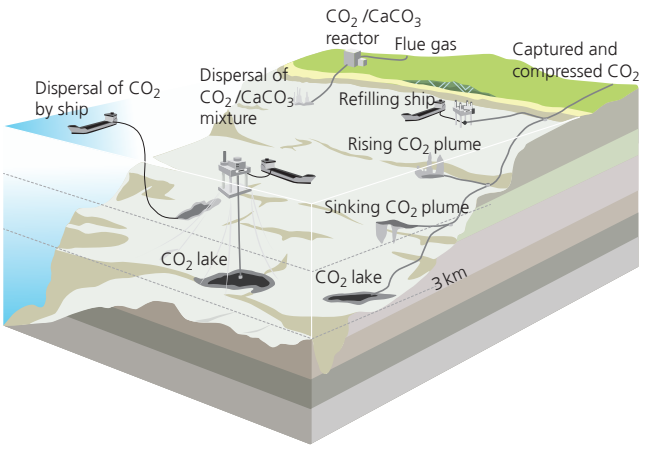


Figure 1.61 Methods of ocean storage

Source: IPCC

- **reduced or no tillage** (ploughing and harrowing) avoids the accelerated decomposition of organic matter and depletion of soil carbon that can otherwise occur. Reduced tillage also prevents the break-up of soil aggregates that protect carbon
- **some use of animal manure** or chemical fertilisers, which can increase plant productivity and thus SOC
- **rotations of cash crops** with pasture or the use of cover crops and green manures, which have the potential to increase biomass returned to the soil
- using **improved crop varieties** to increase productivity above and below ground, as well as increasing crop residues, thereby enhancing SOC.

Forests and tree crops

Forests are able to reduce CO_2 emissions to the atmosphere by storing large stocks of carbon both above and below ground.

- **Protection** of existing forests will preserve current soil carbon stocks.
- **Reforesting** degraded lands and increasing tree density in degraded forests increase biomass density and therefore carbon density, above and below ground.
- **Trees in croplands** (silviculture) and orchards can store carbon above and below ground. CO_2 emissions can be reduced if they are grown as a renewable source of fuel.

It is important to note that many of these mitigation schemes have different and unwanted side effects.

Improved aviation practices

According to the Air Transport Action Group, in 2019 the global aviation industry carried 4.5 billion passengers, producing 915 million tonnes of CO_2. Although the industry has made major strides in reducing its production of CO_2 (for example, the Airbus A380 and the Boeing 787 both use less than three litres of fuel per 100 passenger km), the EU Directorate General for Climate Action state that in 2020 the global emissions of CO_2 will be 70 per cent more than in 2005 and could be a further 300 per cent more by 2050.

The ways in which the aviation industry could reduce their emissions are shown in Figure 1.62. These must be treated with caution because many of them are still at the aspirational or theoretical stage.

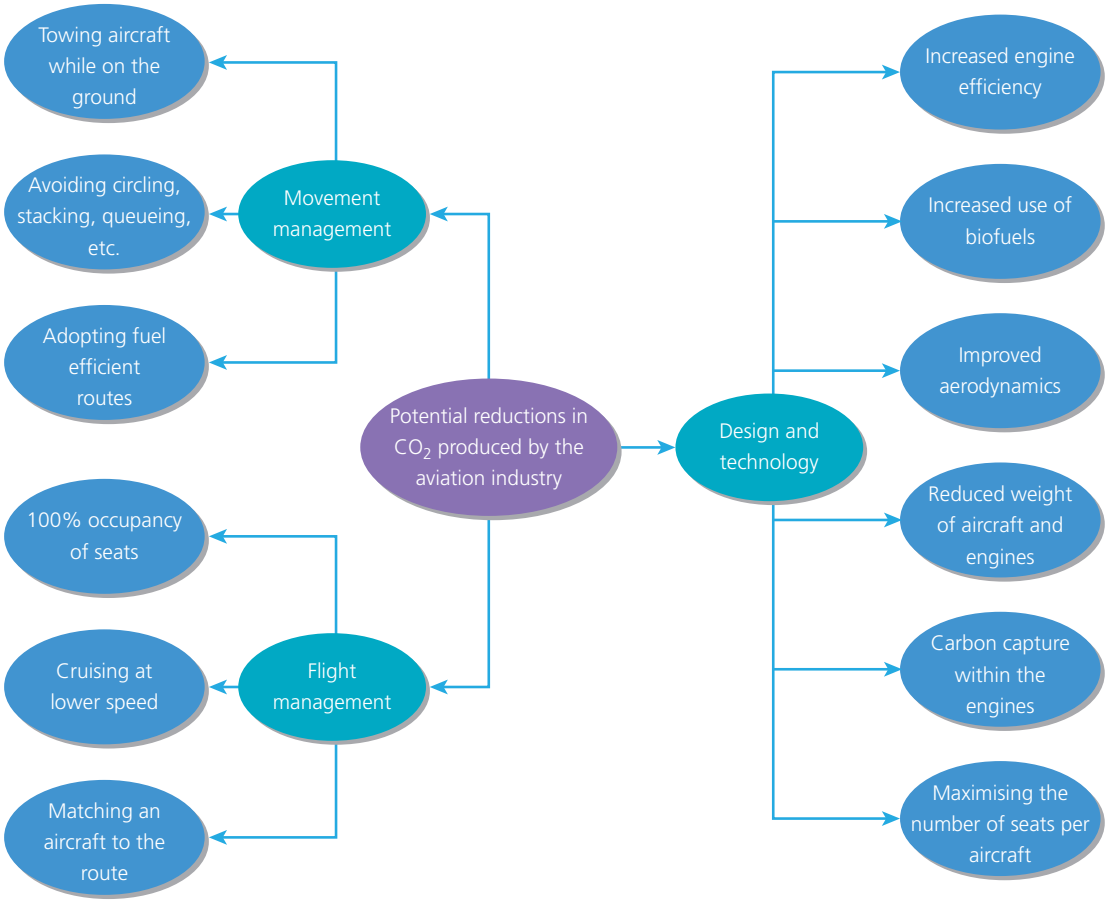


Figure 1.62 CO_2 mitigation within the aviation industry

Review questions

10 Summarise the attempts to mitigate the impacts of climate change.

11 How successful have these attempts been?

Case study of a tropical rainforest setting: Water and carbon in the Amazon

The Amazon Basin (Figure 1.63), at 670 million hectares, is the world's largest rainforest and one of the most biodiverse. Its 300 billion trees and 15,000 species store one-fifth of all the carbon in the planet's biomass. Tropical forests have been present in South America for millions of years and were at one point spread over most of the continent. In the past the forest has shrunk back and then advanced again as ice ages came and went. Today's Amazon rainforest is spread across nine countries and is home to 34 million people who depend on the resources it provides.

Figure 1.63 The location of the Amazon rainforest

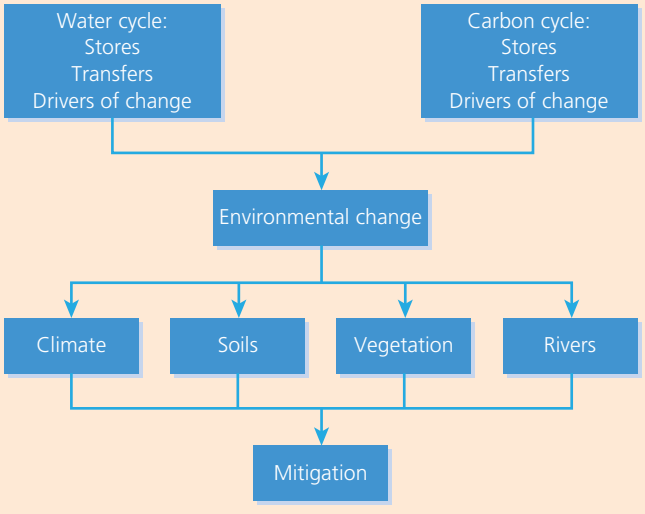


Figure 1.64 Key themes shown in this study

Carbon

There is little consensus amongst the scientific community regarding the actual amounts of carbon stored, emitted and captured by tropical forests. The Helmholtz Centre for Environmental Research (UFZ) estimated that the Amazon stored 76 billion tons of carbon in 2019. Along with other tropical rainforests, it forms a carbon sink of 1–3 GtC/year. Studies have shown that these forests have been increasing in above-ground biomass by 0.3–0.5 per cent per year and that the rising productivity of tropical forests is due to increasing CO_2 concentrations in the atmosphere (see Figure 1.64). The resulting negative feedback seems to offset any rising atmospheric levels of CO_2.

UFZ have also estimated that in 2019 the Amazon forest was a carbon sink, absorbing around 600 million tonnes per year; this is much less than its estimated peak of two billion tonnes of CO_2 each year in the 1990s.

It appears that an increase in atmospheric CO_2 led to a growth spurt for the Amazon basin's trees; in time, however, the growth stimulation feeds through the system, causing trees to live faster and die younger. This has led to a surge in the rate of trees that are dying across the Amazon.

Water

The average discharge of water into the Atlantic Ocean by the Amazon is approximately 175,000 m^3/s, or around 15 per cent of the fresh water entering the oceans each day.

The Rio Negro, a tributary of the Amazon, is the second largest river in the world in terms of water flow, and is 100 m deep and 14 km wide near its mouth at Manaus, Brazil.

Average rainfall across the whole Amazon basin is approximately 2,300 mm annually. In some areas of the north-western portion of the Amazon basin, it can exceed 6,000 mm. Up to half of this rainfall may never reach the ground. It is intercepted by the forest canopy and re-evaporated into the atmosphere. Additional water evaporates from the ground and rivers or is released into the atmosphere by transpiration from

plant leaves. Of the rainfall that is evapotranspired back into the atmosphere, about 48 per cent falls again as rain. Only about 30 per cent of the rainfall actually reaches the sea. The rest is caught up in this constant closed system loop.

Drivers of change

- Between 2000 and 2007, the Brazilian Amazon was deforested at a rate of 19,368 km² per year. During this time, an area of forest larger than Greece was destroyed.
- Brazil is the world's fourth largest climate polluter, with 75 per cent of its greenhouse gas emissions attributed to deforestation and land use change; 59 per cent of this is from loss of forest and burning in the Amazon.
- The removal of forest was done using slash and burn techniques which:
  - reduces the retention of humidity in the soil's top layer down to a depth of one metre
  - facilitates sudden evaporation of water previously retained in the forest canopy
  - increases albedo (reflectiveness) and temperature
  - reduces porosity of soil, causing faster rainfall drainage, erosion and silting of rivers and lakes.
- Any moisture that evaporates from deforested areas forms shallow cumulous clouds which usually do not produce rain.
- Forests emit salts and organic fibres along with water when they transpire. These act as condensation nuclei and assist in cloud and rain formation. Their loss inhibits the formation of cloud and reduces rainfall.
- If destroyed, the vast carbon store will be released into the atmosphere.
- There are a number of differences between tropical rainforest and the pasture land it is generally replaced with.
  - Forests absorb approximately 11 per cent more solar radiation.
  - The average temperature in the rainforest is approximately 24.1°C; in pastures it is 33°C.
  - The daily temperature variation of Amazon forest soils at 20 cm did not exceed 2.8°C, though under pastures it was 8°C.
  - The moisture content in the upper one metre of pasture soil is about 15 per cent less than under nearby forest.
  - Deeper forest roots can pump more soil moisture to the surface, producing 20–30 per cent more air humidity and consequently 5–20 per cent more precipitation than pastures.

Climate changes

Studies investigating all tropical rainforest regions found a mean temperature increase of 0.26°C ± 0.05° every ten years since the mid-1970s and predict that by the year 2050, temperatures in the Amazon will increase by 2–3°C. There has also been evidence of more frequent and increased extremes in temperature.

Amazonia experienced falling amounts of rainfall between the 1920s and the 1970s but since then it appears there has been no significant change.

Vegetation changes

There has been a massive net loss of forests in the Amazon Basin. WWF estimates that 20 per cent of the Amazon forest has already been lost and that will rise to 27 per cent by 2030 if the current rate of deforestation continues. Although most of this has been caused by deliberate deforestation, a significant amount has been as a result of climate change. Some species are limited by their tolerance to temperature change, drought and seasonality. Climate change can affect species sustainability by directly altering the conditions needed to grow and survive. Droughts and unusually high temperatures in the Amazon in recent years may also be playing a role in killing millions of trees although the tree mortality increases began well before an intense drought in 2005.

A 2009 study concluded that a 2°C temperature rise above pre-industrial levels would see 20–40 per cent of the Amazon die off within 100 years. A 3°C rise would see 75 per cent of the forest destroyed by drought over the following century, while a 4°C rise would kill 85 per cent.

Soil changes

Amazonian soils contain from 4 to 9 kg of carbon in the upper 50 cm of the soil layer, while pasturelands contain only about 1 kg/m^2. When forests are cleared and burned, 30–60 per cent of the carbon is lost to the atmosphere; unburned vegetation decays and is lost within ten years. The soil fungi and bacteria that used to recycle the dead vegetation die off.

When forest clearance first occurs, the soils are exposed to the heavy tropical rainfall. This rapidly washes away the topsoil and attacks the deep weathered layer below. Most of the soil is washed into rivers before the forest clearance has caused a reduction in the rainfall.

River changes

Changes in total precipitation, extreme rainfall events and seasonality may:

- lead to an overall reduction in river discharge
- cause an increase in silt washed into the rivers, which could disrupt river transport routes
- lead to flash flooding
- destroy freshwater ecosystems; this could remove a source of protein and income to local inhabitants
- destroy water supply which fulfils the needs of Amazonian peoples.

Warming water temperatures may:

- kill off temperature-dependent species
- change the biodiversity of the river system by introducing new species and killing others
- reduce water-dissolved oxygen concentrations, which could destroy eggs and larvae, which rely on dissolved oxygen for survival.

Mitigation

Figure 1.65 describes some of the strategies to reduce the effects of environmental change in Amazonia.

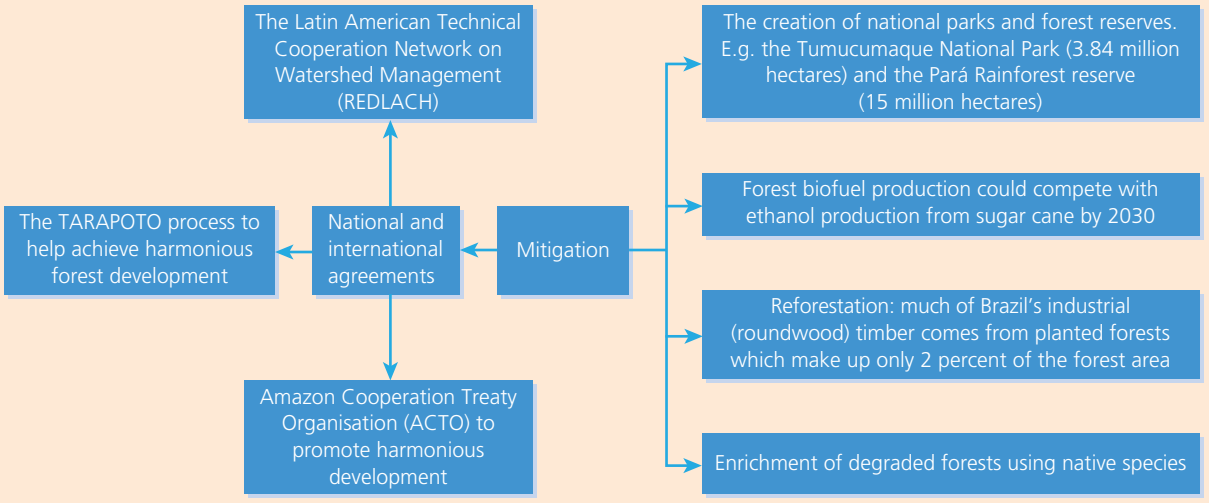


Figure 1.65 Strategies to reduce the effects of environmental change in Amazonia

Links

There are links between what is happening in the Amazon basin and external forces operating at different scales (government policies or the decisions of transnational corporations or the impacts of international or global institutions).

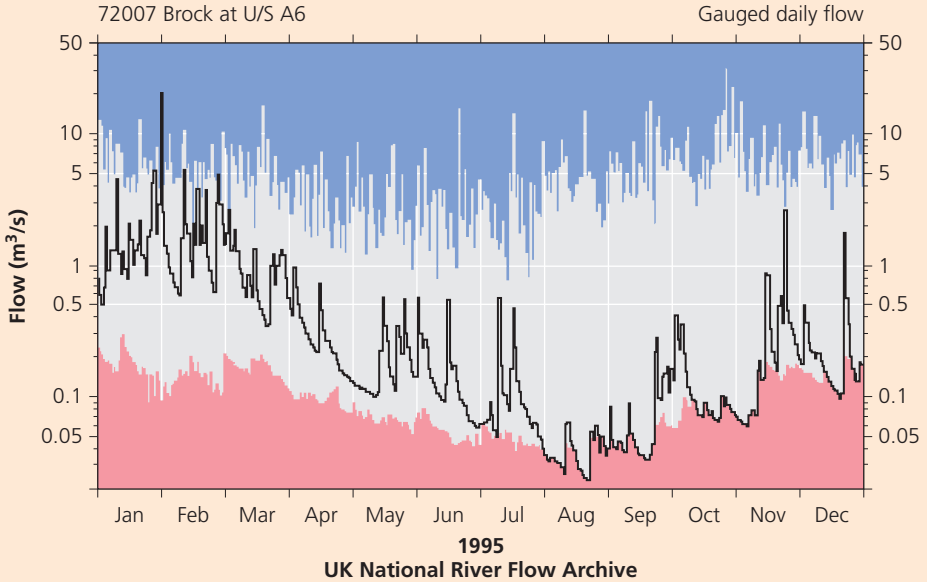
Case study of a river catchment at a local scale: River Brock, Lancashire

The River Brock (Figure 1.66) drains an entirely rural catchment. It is 17.8 km long, rising in the hills of the Forest of Bowland. The headwaters drain steep millstone grit hills in the north and carboniferous limestone in the south. There is peat on the high moors while the lower catchment is overlain by boulder clay. The flow is almost completely natural. Myerscough Agricultural College have a licence to abstract a maximum of 45.46 m³ daily with an annual maximum of 16,592.9 m³. At present this is being withheld.

The regime of the River Brock

This depends upon the pattern of rainfall, which can be highly variable. Table 1.2 indicates that in 1995 rainfall for the Brock catchment was relatively low. Figure 1.67 shows a lower than average flow all year with a late winter maximum and low flow in the summer. On the other hand, 2012 was a particularly wet year as shown by Table 1.3. See figures on page 50.

Figure 1.66 OS map extract showing the location of the River Brock (Map is reproduced at half of OS Landranger map scale of 1:50,000. Each grid square is 1 km^2). © Crown copyright and database rights, 2021, Hodder Education, under licence to Ordnance Survey. Licence number 100036470

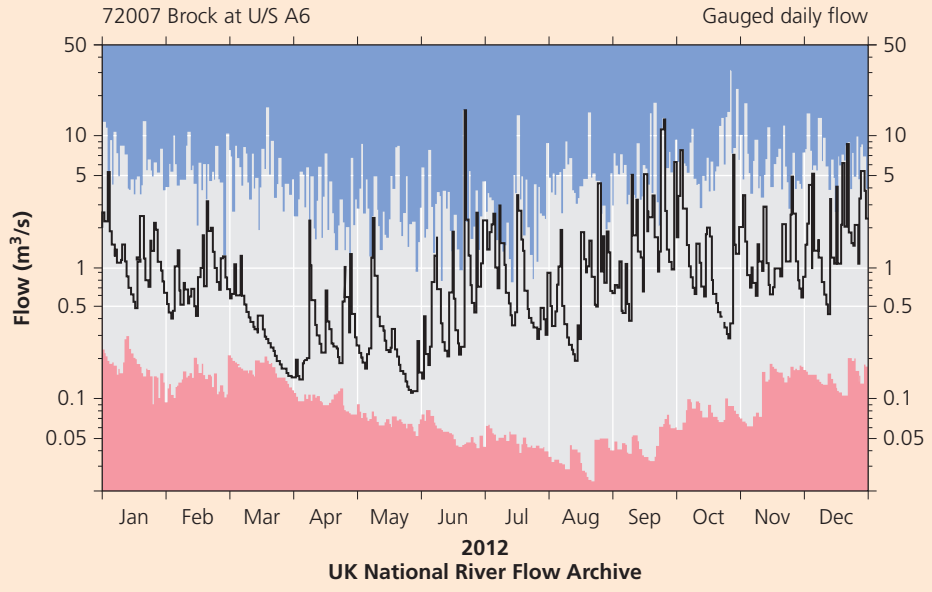


Key:
Red and blue envelopes represent lowest and highest flows on each day over the period of record.

Figure 1.67 Flow data of the River Brock at Brock, 1995
Source: Centre for Ecology & Hydrology

Table 1.2 Monthly rainfall totals for Chipping, Lancashire, 1995

Month	Rainfall
Jan	45
Feb	107
Mar	50
Apr	97
May	77
Jun	63
July	92
Aug	110
Sep	67
Oct	147
Nov	154
Dec	68
Total	1077



Key:
Red and blue envelopes represent lowest and highest flows on each day over the period of record

Figure 1.68 Flow data of the River Brock at Brock, 2012
Source: Centre for Ecology & Hydrology

Table 1.3 Monthly rainfall totals for Chipping, Lancashire, 2012

Month	Rainfall
Jan	188
Feb	124
Mar	31
Apr	98
May	79
Jun	213
July	166
Aug	159
Sep	313
Oct	195
Nov	190
Dec	225
Total	1981

Skills focus

You should be able to interpret the logarithmic scale used for discharge.

What statistical technique could you use to investigate the relationship between monthly rainfall and maximum (or minimum) discharge. Why use that technique?

Downstream changes in the discharge

The River Brock is short enough to be studied almost from source to mouth. Measurements of discharge variations downstream could be compared to other drainage basin characteristics.

Flooding

The most serious flooding to affect the village of St Michael's on Wyre was in October 1980 when there was a combination of intense and prolonged rainfall. As a result comprehensive flood protection was put in place:

- In the area of St Michael's to the north of the River Wyre the Environment Agency constructed a flood storage basin at Catterall for a one-in-fifty-years flood. It has a capacity to store 1.7 million m³ of floodwater.
- Embankments were raised and strengthened in the lowland area close to the River Wyre.
- Warnings are issued for one-in-five-years floods and sandbags given out by the local council, which maintains a store of 2,000.
- In the area of St Michael's to the south of the River Wyre there is a 1 in 100 years probability of flooding with warnings and sandbag distribution.

The River Brock contributes to the flooding of the River Wyre.

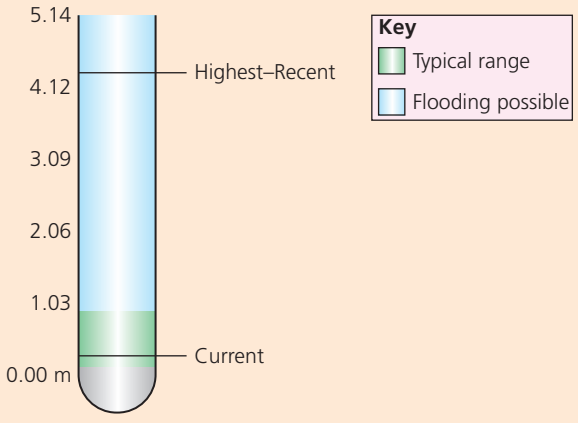


Figure 1.69 The discharge profile of the River Brock at 4.30 on 26/05/2015

Looking at Figure 1.69 we can see that the river level at or near the confluence with the River Wyre at 04.30 on the 26/05/2015 was 0.29 m. The typical river level range for this location is between 0.12 m and 0.93 m. It has reached a highest recorded level of 4.29 m on the 21/01/2008.

The steep sided nature of the Brock Valley means that flooding is not a problem until the river approaches the confluence with the River Wyre. Here the land is much flatter, the result of it being a former glacial outwash plain. Although the river is embanked before reaching St Michael's on Wyre, the agricultural nature of the land use makes it low priority for the Environment Agency.

Review questions

12 The source area of the River Brock is on west facing slopes. How will this influence the amount of rainfall (input) that enters the drainage basin?

13 The source area is peat moorland and rough grazing. How might this affect the transfer of water following precipitation?

14 How might the reintroduction of the Myerscough Agricultural College licence to abstract water affect:
- the discharge of the river
- the ability of the river to support a wide biodiversity?

15 How does the annual flow of the River Brock for 2012 vary from that of 1995?

16 Account for the differences in the two patterns of flow.

17 How might the discharge of the river be affected by the following potential climatic changes?
- An increase in summer temperatures with a reduction in precipitation.
- An increase in the intensity of rainfall events in winter.

18 The greatest discharge does not coincide with the worst floods. Why might this be the case?

Fieldwork opportunities

There are a variety of research questions that could be asked on this topic which would provide a sound basis for fieldwork. These include:

- How do infiltration rate and infiltration capacity vary with slope, steepness, soil texture, geology and/or land use?
- How do rainfall and evaporation affect the level of water, velocity and discharge of a river?
- How do different sections of a river respond to a period of rain?
- How does the amount of interception vary with vegetation type and rainfall intensity?
- How do different land uses affect the amount of carbon stored in the vegetation and soils?

Further reading

The monthly 'Hydrological Summary of the UK' published by the National Hydrological Monitoring Programme (NHMP) https://nrfa.ceh.ac.uk/monthly-hydrological-summary-uk

Typical snapshots of UK hydrology. A series of reports that could be used for research into changes in the water balance.

An up-to-date example of carbon capture and storage can be found at https://ccsnorway.com

This is a general overview of the state of the carbon capture industry in the UK. It looks at the economic argument for carbon capture as well as its possible future in the UK.

'Soil organic carbon; the hidden potential' published by UN FAO www.fao.org/3/a-i6937e.pdf

Provides an overview of current knowledge and knowledge gaps on soil organic carbon.

Air Transport Action Group (ATAG) www.atag.org/our-publications/latest-publications.html

Information on air transport, environment and infrastructure studies.

Global Environment Outlook GEO6: Environment for the Future We Want, United Nations Environment Programme, first published in 2019, www.unep.org/resources/global-environment-outlook-6

GEO-6 outlines the current state of the environment, illustrates possible future environmental trends and analyses the effectiveness of policies.

Question practice

A-level questions

1 Outline the ways in which feedback processes affect the carbon cycle. [4 marks]

2 Figure 1.70 shows the changing amounts of carbon dioxide production by source. Using the figure, analyse the changes shown. [6 marks]

3 Study Figure 1.55 (page 41), which shows the predicted increases in surface air temperature between 1960 and 2060. Using the figure and your own knowledge, assess the impact of these temperature changes on life on Earth. [6 marks]

4 With reference to an area of rainforest you have studied, assess the extent to which the water and carbon cycles in that area may be related to attempts at global governance. [20 marks]

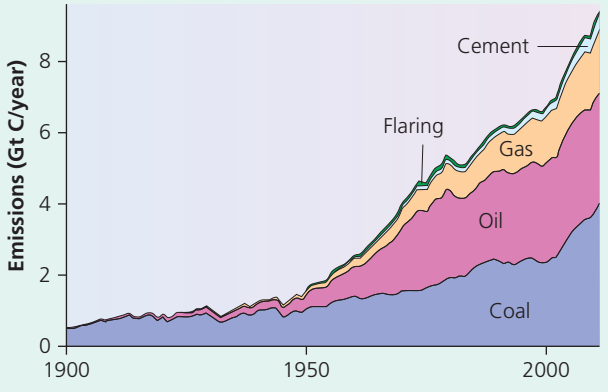


Figure 1.70 World carbon dioxide production by source

Source: US Department of Energy

Chapter

2 Hot desert systems and landscapes

Drylands are found at all latitudes around the globe, but most hot desert environments are arid regions of the world located at mid and low latitudes on or around the tropics. These areas are arid, receiving little or no rainfall, with sparse or no vegetation. Due to their extreme climate hot desert environments are sparsely populated.

In this chapter you will study:

- the global distribution of mid- and low-latitude deserts and their margins
- the climate, soil and vegetation characteristics of hot deserts and their margins
- the processes leading to aridity in hot deserts and their margins
- the systems and processes leading to the development of hot desert landscapes
- desertification – its causes, consequences and alternative futures for local populations.

2.1 Introduction to hot deserts and their margins

Deserts as natural systems

The introduction to Chapter 1: Water and carbon cycles outlined how many aspects of physical geography can be studied using a **systems** approach and you should read that section carefully before reading further here. This chapter explores the landscapes of hot desert environments and their margins and we will see that they are **open systems** with a range of **inputs**, **stores/components**, **flows/transfers** and **outputs** that combine to create one of the most distinctive types of landscapes on Earth. When there is a balance between the inputs and outputs then the system is said to be in a state of **dynamic equilibrium**. If one of the elements of the system changes, for example, there is increased sediment blown into the desert but there is no corresponding change in the amount of sediment being removed from the desert, then some desert landforms may change and the equilibrium is upset. This is called **feedback**. Figures 2.1a and 2.1b illustrate the open system of a desert landscape and how desertification can be viewed as an example of a positive feedback system. They highlight many of the ideas that are explored in detail in the rest of the chapter.

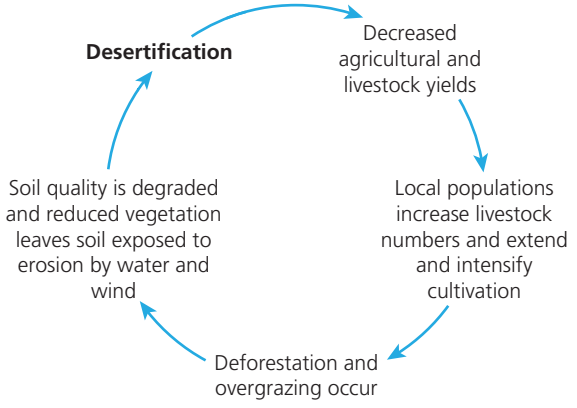


Figure 2.1b Positive feedback in a hot desert environment – desertification

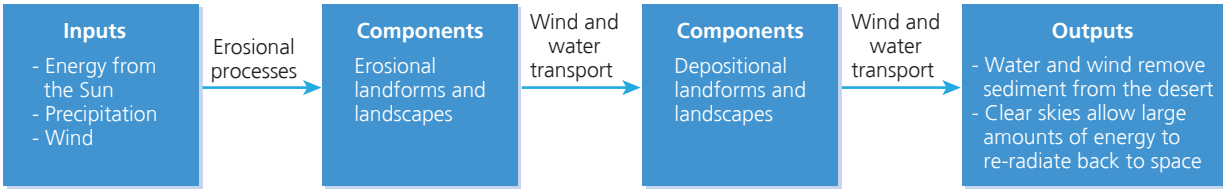


Figure 2.1a A hot desert landscape as an open system

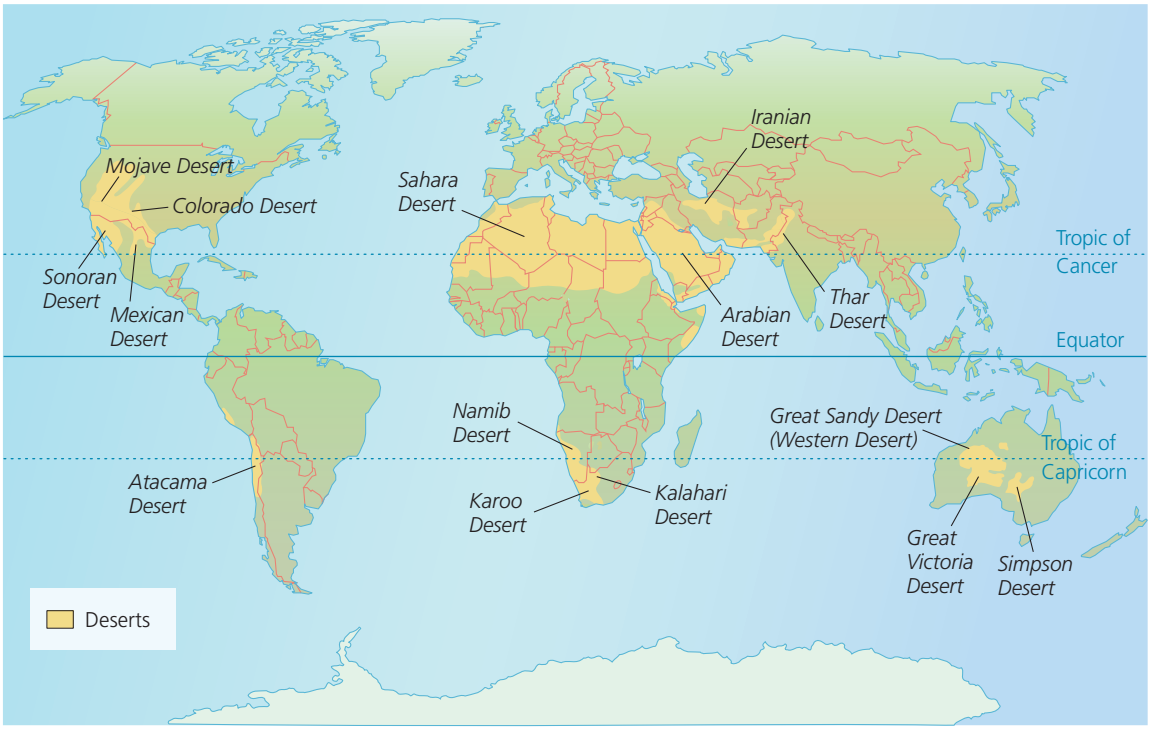


Figure 2.2 The location of major hot desert environments

Like all physical environments, the combination of erosional and depositional landforms that combine to produce characteristic hot desert landscapes are constantly changing elements (or components) of the geomorphological systems of which they are a part. As the system tends towards dynamic equilibrium, driven by the input of energy, the geomorphological processes present in hot deserts continually shape the desert landforms, creating their distinctive landscape features.

Global distribution

Deserts cover over one-fifth of the surface of the Earth, however hot desert environments and their margins generally run in parallel belts north and south of the Equator in hot **arid** and **semi-arid** mid and low-latitude locations.

Almost all hot desert landscapes in the northern hemisphere are located towards the west of continents. The major hot desert areas of the northern hemisphere (Figure 2.2) include:

- the Mojave, Sonoran, Great Basin and Chihuahuan deserts in North America
- the Sahara in North Africa and Somali-Ethiopian deserts in the Horn of Africa
- the Arabian, Iranian and Thar deserts stretching from the Red Sea through the Middle East to Pakistan and India in Asia.

In the southern hemisphere hot desert areas include:

- the Atacama desert in South America
- the Namib, Kalahari and Karoo deserts in southern Africa
- the Great Sandy (Western), Great Victoria and Simpson deserts, and others, in Australia.

Key terms

Arid – The climate of an area that receives less than 250 mm of precipitation per year.

Aridity index – The ratio between precipitation (P) and potential evapotranspiration (PET).

Continentality – The impact of increasing distance from the coast on the climate of an area.

Desert – An arid environment receiving very low levels of rainfall.

Semi-arid – The climate of an area that receives between 250 and 500 mm of precipitation per year.

Water balance – The relationship between the annual precipitation received and the amount of water lost to potential evapotranspiration.

Water and aridity index and the climate, soil and vegetation characteristics of hot deserts

Aridity index

A desert is most simply defined as an arid environment receiving very low levels of rainfall. Typically hot deserts are categorised as arid receiving less than 250 mm precipitation per year and semi-arid receiving 250–500 mm precipitation per year. However, this is very simplistic and the categorisation of deserts is varied and complex. Deserts are classified as hot and cold, sandy and rocky, foggy and sunny, and barren, and are found at both high and low extremes of altitude and in both coastal and continental interiors. Almost all definitions are linked to the lack of water, and the **water balance**, which compares the mean annual **precipitation (P)** received with the mean annual **potential evapotranspiration (PET)**. PET is the amount of water that could be lost from the soil by plant transpiration and direct evaporation from the ground. The ratio of P and PET is known as the **aridity index**.

Hot deserts and their margins are considered to be areas described as **hyper-arid** to **arid**, and **semi-arid** on the aridity index as illustrated in Figure 2.3.

Classification	Aridity Index	Global land area (million km²)	Global land area (%)
Hyper-arid	AI <0.05	10.0	7.5
Arid	0.05<AI<0.20	16.2	12.1
Semi-arid	0.20<AI<0.50	23.7	17.7

Figure 2.3 The aridity index

Places with greater P than PET will have an aridity index of greater than 1.0, while those with PET greater than P will have an aridity index of less than 1.0. Of the hot deserts in Figure 2.2 (page 54), the Sahara has the highest aridity, followed by those in Arabia, East Africa, Australia and those of South Africa, with lower levels in the Thar and the deserts of North America.

Climate

Although hot deserts are warm throughout the year with very hot summer temperatures, they have climates characterised by extremes:

- Temperatures vary wildly both annually (between the hottest and coldest months) and diurnally (between daytime and night-time temperatures).
- With very low humidity levels in hot deserts, especially those in subtropical latitudes, cloud levels are extremely low. Clear skies allow significant amounts of insolation to reach the surface, almost twice that at more humid latitudes closer to the Equator or at higher latitudes.
- Daytime temperatures can commonly rise to over 30°C, with highs of 45–55°C in many locations. One claim for the highest temperature ever recorded, 57.7°C, was at Al Aziziyah in Libya in 1922.
- The cloudless skies explain the high diurnal range, as temperatures drop very rapidly at night, with hot deserts losing twice as much heat at night as the more humid latitudes.
- Night-time temperatures can often drop below 0°C, with records of around –18°C not being rare. Therefore, the diurnal range can be over 30°C, with extreme ranges of over 40°C in places (Figure 2.4).

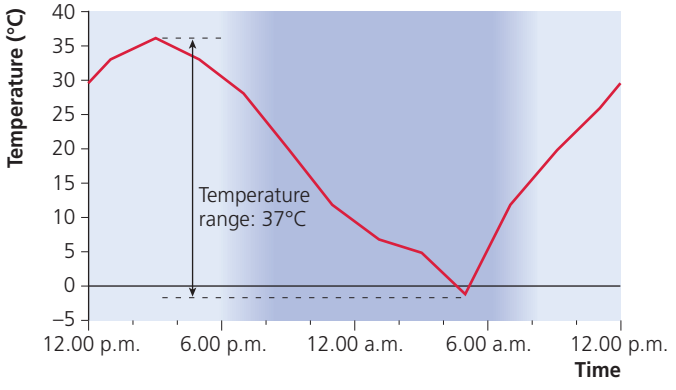


Figure 2.4 The diurnal temperature range in a hot desert

Semi-arid regions are also warm all year, with hot dry summers, average temperatures of around 21–27°C, temperatures rarely going above 38°C, and night-time temperatures being cool, around 10°C. As distance from the tropics increases so does the annual temperature range in hot deserts due to the increased seasonality of the climate. This means that desert margins have huge variations in temperatures, with those on the poleward side of arid areas being significantly colder than those on the equatorial side.

The reality is that there is huge variety in the temperature patterns experienced in hot desert areas around the globe due to differences in their geographical situation. For example, those closer to the sea, like the Atacama, have a cooler, less extreme temperature range than those in continental interiors.

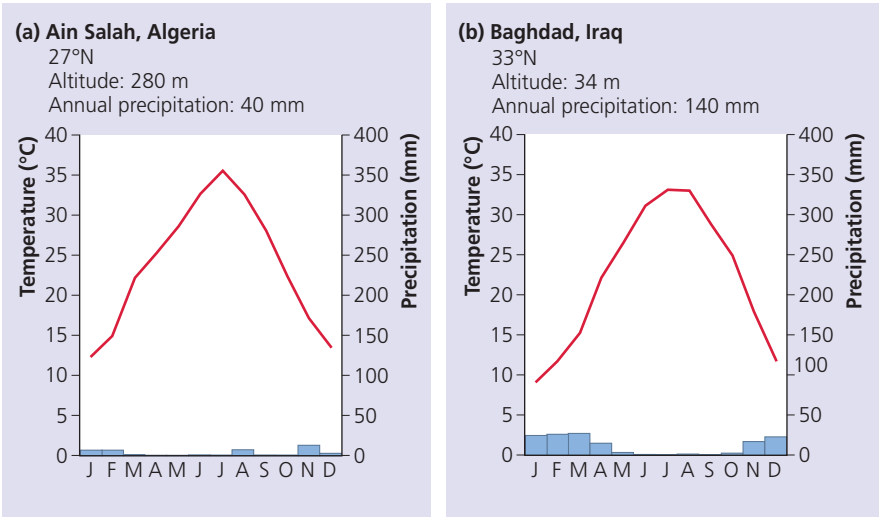


Figure 2.5 Climate graphs for two desert areas

The amount of rainfall is usually extremely low, or occurs in rarer, short downpours following long dry periods (Figure 2.5). These cloudbursts can often deliver the total annual rainfall in a number of hours. The rate of evaporation often exceeds the rate of rainfall. Hot deserts are classified as receiving less than 250 mm of rainfall, but rates vary greatly around the world. Places like Iquique in the Atacama Desert in Chile receive almost no rain and are the driest, averaging less than 15 mm a year, with the central Sahara receiving a similar amount. American deserts receive slightly higher amounts at around 280 mm. Desert margins or semi-arid areas are often classified as receiving 250–500 mm a year, but like hot deserts they are generally dry and rainfall events are unpredictable and can occur as sudden downpours. Due to the slightly lower temperatures less moisture is lost due to **evapotranspiration**, and in areas with low night-time temperatures the condensation of dew can equal or exceed the amount of rainfall received in some areas.

Skills focus

You can easily find annual rainfall and temperature data for a hot desert area and could use this to construct and interpret a climate graph.

Soil

Around 17 per cent of the Earth's surface has desert soils. The extent of desert soils varies in different continents: Australia has about 44 per cent, Africa 37 per cent and Eurasia only 15 per cent.

When soils do develop in deserts they are often infertile with a thin **soil profile** and tend to be alkaline and quite saline. Rates of soil development are extremely slow in hot deserts due to:

- lack of moisture
- extremely high temperatures and high rates of evaporation
- sparse vegetation and limited organic material.

Desert soils are often characterised by a thick accumulation of basic mineral salts (often calcium or sodium compounds) at or near the surface. This occurs due to the process of **capillary movement** where any moisture in the soil or subsoil moves upwards through the tiny spaces between soil particles (capillaries). This capillary action is most effective when evaporation exceeds precipitation, hence its prevalence in arid environments.

The main zonal soil type for hot deserts and their margins is the soil order of **aridisols**. This is an order of soils including infertile alkaline and saline soils of desert areas characterised by accumulations of mineral salts at or near the surface (Figure 2.6). These aridisols fall into two main categories **sierozems** (in semi-arid areas) and **raw mineral soils** (in arid environments), although there are many other groups and sub-groups of soils also known as aridisols.

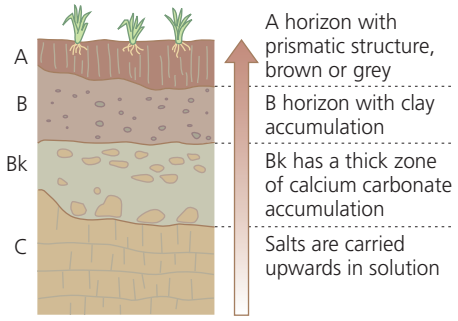


Figure 2.6 Soil profile of a typical aridisol

Aridisols range from reddish-yellow to grey-brown in colour depending on the nature of the parent material and are generally very thin (usually less than 100 cm deep).

The **raw mineral soils** of arid areas have a coarse texture and are often rocky or gravelly due to **physical weathering**. **Chemical weathering** also occurs and where water exists in the subsoil, capillary action leads to an accumulation of calcium and sodium salts near the surface. With a lack of moisture there is little downward movement of minerals (**leaching**), and arid soils tend to have limited variation between horizons. These soils tend to be slightly alkaline and are often unproductive rather than infertile, as minerals and nutrients are available but extremes of temperature and a lack of water are the main limiting factors for vegetation. If concentrations of salts are high enough hard salt crusts can form at the surface.

In arid to semi-arid areas receiving around 250 mm of rainfall, **grey desert soils** or **sierozems** can form. The darker colour suggests the presence of some organic material as they often develop below 'desert-shrub' vegetation. These soils are often used for cultivation and with continued irrigation calcium-rich B horizons can develop beneath the thin A horizon, with rich accumulations of calcium carbonate or gypsum giving a lighter colour.

Vegetation

Hot desert environments often look barren, desolate and lifeless, however other than the **sand sea** areas that are constantly mobile, arid and semi-arid areas have a remarkable diversity of plants. Plants are usually ground-hugging shrubs or short woody trees (Figure 2.7). However, even in the least hospitable, arid environments, vegetation cover remains sparse, and depending on the temperature and rainfall, **net primary productivity** values can range from near 0 to 120 g/m²/year.

Figure 2.7 Low-growing desert-shrub vegetation

The limited availability of water, extreme temperatures (in terms of heat and range) and the intense constant sunlight during the day, mean that plants have developed a range of physical and behavioural adaptations in order to:

- maximise use of, and limit the loss of moisture
- store moisture in their stem or leaves
- procure water with extensive and/or deep root systems
- respond rapidly to sporadic rainfall followed by rapid life cycles.

Figure 2.8 Saguaro cactus with adaptations for surviving in arid environments

Most plants in desert areas are **xerophytes**. Plants in desert areas have a number of features to help them survive the dry conditions (Table 2.1, page 58).

Table 2.1 Desert plant characteristics and adaptations

Feature	Description of adaptation
Succulence	• Succulent plants, or xerophytes, contain compounds or cells in fleshy leaves, stems or roots where they can store water, for example, all cacti and non-cacti like aloes

Procuring water after brief rainstorms	• Water does not penetrate deep into the soil or remain wet for very long so succulents have the ability to rapidly absorb huge amounts of water as plant roots can only take up water when the soil is more moist than the interior of the roots
	• Succulents have shallow (less than 10 cm) and very wide extensive root systems. For example:
	  • the roots of the 12–18 m-tall saguaro cactus (Figure 2.8, page 57) can extend horizontally to the same distance as the height of the plant
	  • the cholla cactus is 60 cm tall and has roots that extend up to nine metres
	  • agaves are the exception to this adaptation as they have a limited root system, but have a leaf rosette that channels rain water directly to the base of the plant

Conserving water	• Thick waxy cuticles and closed stomates, making the leaf surface waterproof • Small, spiky or waxy (or no) leaves to reduce surface area to limit transpiration
Protection from thirsty animals	• Being spiny, bitter or toxic • Living in inaccessible locations

Drought tolerance	• Some plants are **phreatophytes** with very deep root systems reaching water deep underground
	• Many plants are **ephemerals** with adaptations including:
	  • becoming dormant or losing their leaves during droughts
	  • appearing 'dead' or dormant for months or even years

Drought avoidance	• A number of plants, including some ephemerals, have extremely short life cycles • They come into bloom rapidly after rainfall • They channel all their life energy into producing seeds • Their seeds lay dormant until a future rainfall event
Salt tolerance	• Due to high rates of evaporation, desert plants, such as the saltbush, need to be **halophytes** with cells that have adapted to deal with high levels of elements, such as sodium and chlorine • Plant tissue adapts to survive in a saline environment

Fieldwork opportunities

The northern Sahara Desert is close to the town of Erfoud to the east of Marrakesh in Morocco. During a field visit it would be quite simple to collect basic information about the characteristics of this hot desert environment without requiring large amounts of equipment. Field observations could be made on the following:

- climate
- soils
- vegetation.

The causes of aridity in hot deserts

There are a number of factors that contribute to the arid nature of these regions:

- the general pattern of atmospheric circulation
- distance from oceans or continentality
- relief
- cold ocean currents
- wind (page 62).

Global atmospheric circulation

Figure 2.9 illustrates how the latitudes where hot deserts tend to be located, between about 20–30° north and south of the Equator, are generally dominated by high pressure throughout the year.

At the Equator there is a net gain in energy as large amounts of incoming solar radiation, **insolation**, is received because the Sun is directly overhead with a high **angle of incidence**. Air in contact with the Earth is heated and begins to rise, cooling, causing the water vapour within it to condense, form clouds and lead to precipitation. The rising air is replaced by air rushing in from the north and south creating an area of low pressure known as the **inter-tropical convergence zone (ITCZ)**. The rising air begins to cool and track polewards. At around 20–30° north and south this now cool denser air descends. As it does, it warms and expands, meaning little cloud formation occurs, giving the clear skies responsible for the heat and aridity of these latitudes. These two

circulation cells are known as the **Hadley cells**. Figure 2.9 also shows that the mid-latitude cells (**Ferrel cells**) have air descending at these mid-latitudes; the combination of the sinking air from both cells is what creates the high pressure at the surface.

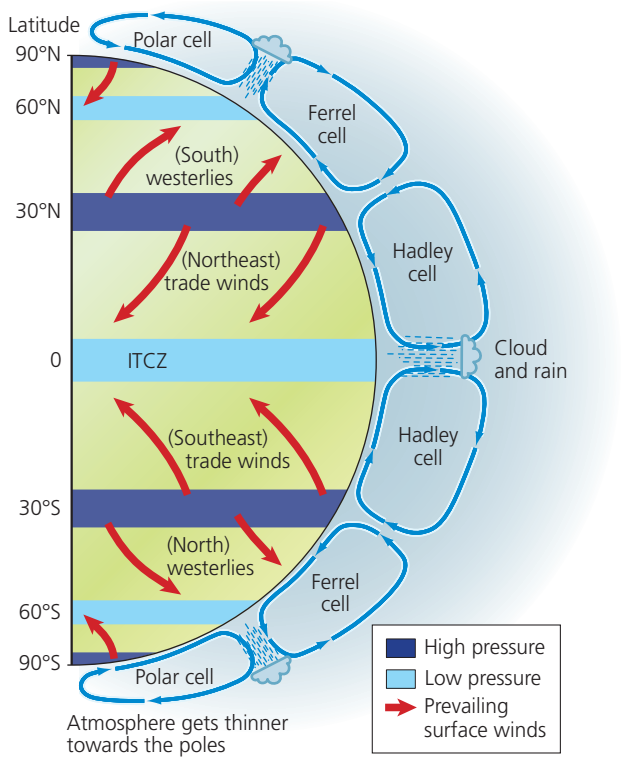


Figure 2.9 The tri-cellular model of atmospheric circulation

Distance from the sea

Continentality has an effect on both the rainfall and temperature received by maritime areas and continental interiors.

As Figure 2.10 illustrates, coastal locations have a more moderate climate with little annual variation in temperature and receive higher levels of rainfall. With distance from the sea, temperatures are far more extreme, can exhibit higher extremes and are generally much drier as there is little moisture available for cloud formation. For example, parts of the Sahara in the centre of North Africa are over 2,000 km from the sea.

Relief

Around mid-latitude areas there are a number of extremely dry regions on the leeward side of mountain ranges that experience what is known as the **rain shadow** effect, for example, Arabia to the west of the Himalayas and central Australia to the west of the Eastern Ranges. Figure 2.11 (page 60) illustrates the effect of the southeast **trade winds** from the southern Atlantic meeting the Andes and the effect on the Atacama desert in the west of South America.

If moist air brought inland by prevailing winds meets a range of mountains, it is forced to rise, leading to cooling, condensation, cloud formation and **relief rainfall** on the windward side. Once over the summit the air descends on the leeward side of the range.

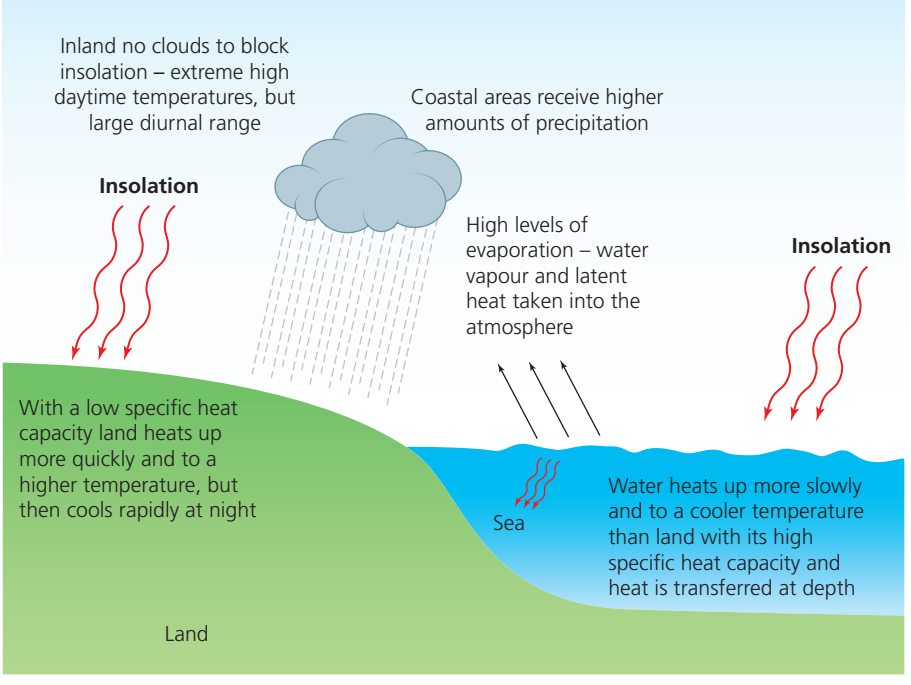


Figure 2.10 The effect of continentality

Figure 2.11 Rain shadow effect in South America

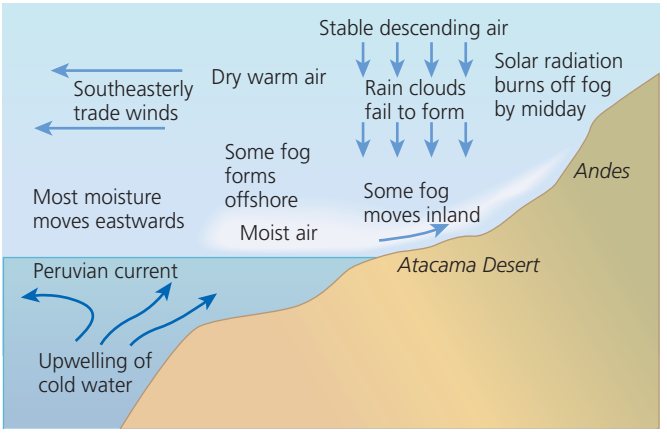


Figure 2.13 The influence of the Peruvian current

This sinking air warms leading to a drop in **relative humidity** and clear skies with little, or no, rainfall.

Cold ocean currents

The Atacama Desert in the sub-tropical latitudes of western South America illustrates the effect of cold ocean currents. There are a number of such currents that form part of the **global oceanic circulation** (Figure 2.12).

Any wind moving over these cold waters is cooled, which causes the relative humidity to increase and eventually moisture is condensed to create fog and mist offshore. As the land heats up quicker than the sea, this may generate gentle onshore breezes that will take the fog and mist inland. However, the intense heating from the Sun directly overhead soon burns this away. As the air is cool, it cannot hold much moisture so cloud formation is unlikely, and precipitation is very rare.

Some vegetation has even developed adaptations to take advantage of the moisture condensing on it as dew, such as in northern Chile where this may be the only moisture it receives for years at a time (Figure 2.13).

Review questions

1 Outline examples of feedback in hot desert landscapes.

2 What reasons can you suggest for the distribution of hot deserts?

3 What are the causes of aridity in hot deserts?

4 Why do hot deserts have such a large diurnal temperature range?

5 Why are soil profiles in hot deserts different to those in wetter environments such as the UK?

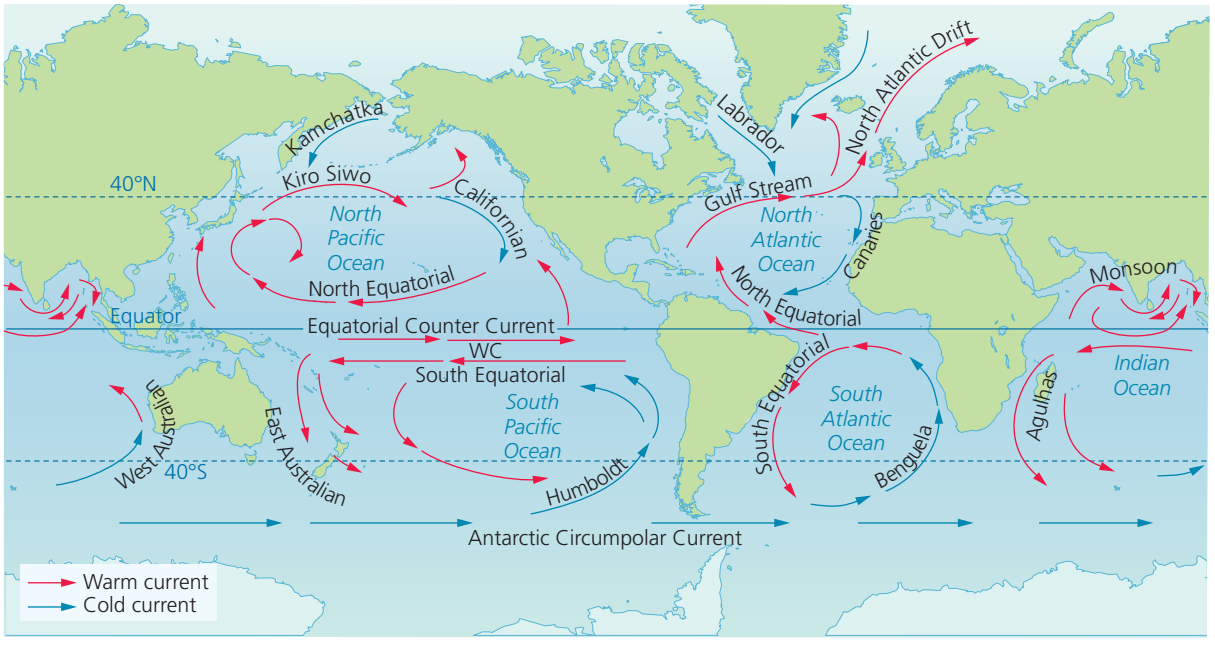


Figure 2.12 Oceanic circulation

2.2 Systems and processes in hot deserts

This book highlights how most aspects of physical geography can be thought of as natural systems with inputs, stores/components, flows/transfers and outputs; with geomorphological processes, driven by the input of energy, constantly changing the elements of these systems as they tend towards dynamic equilibrium. Hot deserts are no different and before exploring the detail of the processes that shape the landforms that combine to create their characteristic landscapes, a simple summary of the open system in hot deserts is included below.

- Inputs: energy from insolation; water and wind; sediment (mostly sand-sized particles).
- Components: characteristic erosional and depositional desert landforms.
- Outputs: water and wind remove sediment and clear skies allow large amounts of energy to be re-radiated back to space.

You should also read the relevant material at the beginning of Chapter 1 under the heading 'Systems frameworks and their application' (page 2).

Key terms

Abrasion – Where material carried by moving wind or water hits exposed rock surfaces, thus wearing them away. Often referred to as a *sandblasting* or *sandpapering* effect.

Aeolian – Relating to the action of the wind.

Chemical weathering – The processes leading to the decomposition or breaking down of rocks due to chemical reactions. This most often requires the presence of water, and/or exposure to the air.

Deflation – Where wind removes dry, unconsolidated (loose) sand, silt and clay particles from the surface and transports them away.

Deposition – Occurs when the velocity of the wind decreases until it can no longer transport the grains it is carrying.

Endoreic streams – Where rivers occupy drainage basins that are closed and do not flow out to the sea or other rivers, but instead end inland in lakes or swamps.

Ephemeral streams – Streams that flow intermittently in hot desert areas following heavy thunderstorms.

Erosion – The wearing away of the Earth's surface by the mechanical action of processes of glaciers, wind, rivers, marine waves and wind.

Exfoliation – A process of mechanical weathering that results in the breaking, splitting or *peeling-off* of the outer rock layers. Also commonly known as 'onion skin weathering'.

Exogenous streams – Rivers that originate external to the desert in adjacent highlands and more humid environments, flow from outside of the desert and pass through it.

Insolation – The incoming solar radiation that reaches the Earth's surface.

Mass movement – The movement of material downhill under the influence of gravity, but may also be assisted by rainfall.

Saltation – A process where sand-sized particles are transported by bouncing and hopping along the surface.

Sediment – Any naturally occurring material that has been broken down by the processes of erosion and weathering and has then been transported and subsequently deposited by the action of ice, wind or water.

Sediment budget – The balance between the input and output of sediment in hot deserts.

Surface creep – Where saltating particles return to the surface and hit larger particles that are too heavy to hop; they slowly creep (slide or roll) along the surface from a combination of the push of the saltating grain and the movement of the wind.

Suspension – Transportation by wind where the smallest particles, generally less than 0.2 millimetres, are held in the air.

Thermal fracture – The weathering of rock resulting from its rapid and repeated heating and cooling.

Transportation – The processes that move material from the site where erosion took place to the site of deposition.

Weathering – The breakdown and/or decay of rock at or near the Earth's surface creating regolith that remains *in situ* until it is moved by later erosional processes. Weathering can be mechanical, biological/organic or chemical.

Sources of energy in hot deserts

Many of the landform features of hot deserts are created by the action of water and wind. To fully explain how these processes work it is important to understand where the energy comes from to drive the various **weathering** and **erosion** processes. **Insolation**, wind and run-off provide the energy to drive hot desert geomorphological systems. Together with sediment (mostly sand-sized particles), these are the main **inputs** in desert systems.

Insolation

Many processes that occur in hot deserts result from changes or differences in temperature. Therefore energy from the Sun is a vital **input** in desert systems. We have already explored why hot deserts are arid, but why are they hot? In the low- to mid-latitude hot desert environments the Sun is almost directly overhead during the day throughout the year – meaning there are 12 hours of daylight. Also, as the Sun is high in the sky in these latitudes, the **angle of incidence** of the insolation is very high (Figure 2.14) meaning that any given amount of radiation is concentrated on, and so warms, a smaller area of Earth than at higher latitudes.

In more moist environments the Sun's energy is used to evaporate water on and under the surface, taking heat energy up into the atmosphere as **latent heat**, so keeping the ground cooler than it otherwise might be (Figure 1.12, page 9). In more arid areas the lack of moisture means the ground becomes excessively warm, making more energy available to heat the air in contact with the ground.

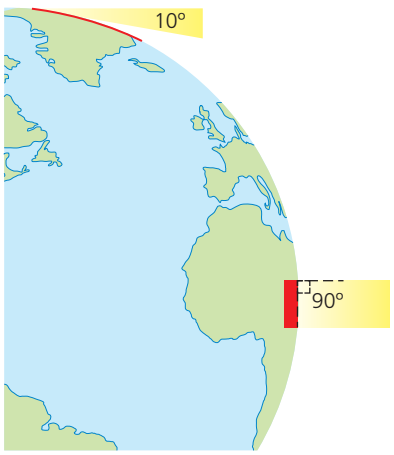


Figure 2.14 The height of the Sun in the sky – angle of incidence

Winds

Wind is also an important input in hot deserts as it is a source of energy and driver of processes in desert systems as they tend towards dynamic equilibrium. With their mid-latitude location along the sub-tropical high-pressure belt, many large hot desert areas are subject to localised winds blowing outwards towards the edge of the desert. These winds then act as an agent of erosion and transport. Examples of these local winds include:

- the *kharif* in the Somali-Chalbi desert in the Horn of Africa, blowing June to September
- the *irifi* easterly winds that blow in the Western Sahara around March
- famously, the *harmattan* in the Sahara blowing in winter months from November to March
- the *shamal* over the Arabian Gulf from late May to early July
- there are also more localised winds generated in the Sudanese Sahara known as *haboobs*.

Other hot desert areas have their own local winds. Although winds in desert areas are on average generally no stronger than elsewhere, it is their effects that are most notable, as we will see later in the chapter.

Run-off

Rainfall in deserts is often described as 'spotty', in that it is often very localised and spatially and temporally unpredictable. Despite arid hot desert areas being defined as having less than 250 mm of rainfall annually, and some hyper-arid areas recording almost none, rare intense storms can produce huge amounts of rainfall in very localised areas (Table 2.2).

Together with the baked ground and limited vegetation cover, rapid **overland flow (run-off)** can be a significant agent of erosion and transportation of sediment.

Sediment sources, cells and budgets in hot deserts

Hot desert landscapes are often dominated by loose **sediment**. Deserts are a store of this sediment and much of it is transported and deposited within the desert boundary itself, forming a sediment cell. There are a variety of origins, or **inputs**, of sediment into hot deserts:

- Sediment may simply be derived from the weathering of the underlying parent material.
- Sediment may be **fluvial** in origin; rivers may bring the sediment into deserts.

Table 2.2 Extremes of precipitation in deserts

Location	Mean annual precipitation (mm)	Highest precipitation in a single storm (mm)	Storm date
Chicama (Peru)	4.0	394.0	March 1925
Luderitz (Namibia)	16.7	102.0	16–22 April 2006
Masirah Island (Oman)	70.0	430.6	13 June 1977
Diego Aracena (Chile)	1.3	9.1	6 July 1992
El Djem (Tunisia)	275.0	319.0	25–27 Sept 1969
Ziyaratgah (Iran)	28.0	30.0	6 March 1997
Oodnadatta (Australia)	173.4	200.0	9 February 1976
El Cajoncito (Mexico)	220.7	237.5	29 Sept 1982

Source: Middleton, N. 2009, OUP: Deserts: A Very Short Introduction, *p.17*

  - If they are **ephemeral**, or dry up, then sediments may be left behind in dry riverbeds.
  - Or, as rainfall is episodic and often heavy, desert rivers may flood and deposit material on the desert surface.
- Other sediments originating beyond desert margins may be **aeolian** and be transported in and deposited by wind as **loess**.

Hot deserts have a range of landscapes and landforms depending on the combination of processes operating there.

- Areas dominated by erosion are a **source** of sediment and the system has a **net sediment loss**.
- In those dominated by the transportation and **deposition** of sediment, the system has a **net sediment gain**.
- The balance between the **input** and **output** of sediment in hot deserts is the **sediment budget**.

Large deserts can be significant sources of sediment themselves on a global scale. In fact satellite imagery often shows large clouds of *dust* being blown five thousand miles across the Atlantic to the Americas (Figure 2.15). Also the large loess deposits of central Asia and northern China, although not formed in hot deserts, were certainly transported in from adjacent deserts.

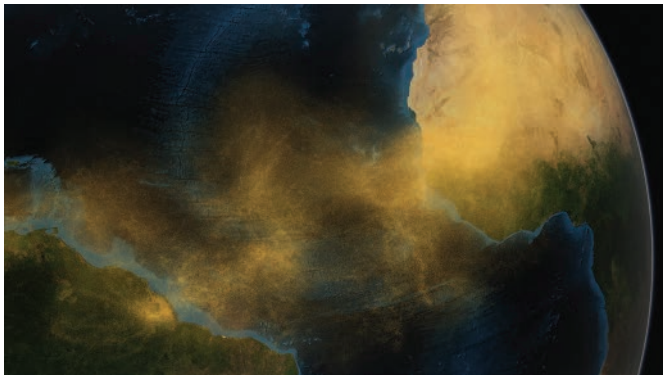


Figure 2.15 Saharan sediment being blown across the Atlantic

Geomorphological processes

As hot desert geomorphological systems develop over time and tend towards dynamic equilibrium, a variety of processes take place. Driven by the input of energy, these are what change the characteristics of the components of the hot desert landscape.

Weathering

Weathering is the breaking down or decaying of material where it is (*in situ*). Unlike erosion it does not involve the transporting away of the broken down material, but instead produces material known as **regolith**. Weathering can be subdivided into a number of types:

- mechanical weathering
- **chemical weathering**
- organic (or biological) weathering.

Mass movement

Mass movement is the movement of material downhill under the influence of gravity, but may also be assisted by rainfall. However, because of the low levels of moisture and little vegetation there is limited soil in hot desert environments. Therefore where mass movement occurs it is dominated by:

- **rock falls**, where small blocks of rock become detached from an exposed cliff face and fall freely to the base of the cliff
- **rock slides**, where there is a failure throughout the rock as a whole and the material collapses *en masse*, rather than as individual blocks.

These create generally steeper slopes with the accumulation of coarse, angular material at their base.

Steep solid rock cliffs are quite common in hot deserts. If vertical jointing is a common feature of the rock

forming these cliffs and this is underlain by rocks that are more easily eroded, then rock fall and cliff retreat will occur (Figure 2.16).

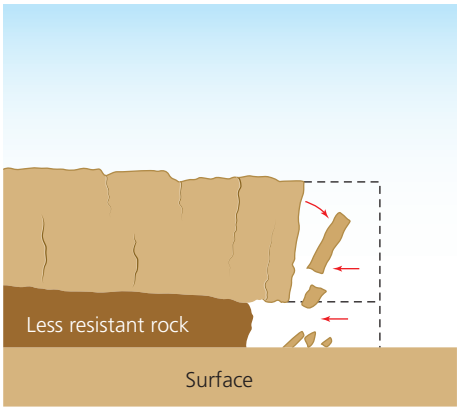


Figure 2.16 Rock fall and cliff retreat in vertically jointed rocks

Where the rock forming cliffs has more horizontal bedding planes, then the cliffs retreat in a more uniform manner due to the usual weathering and erosion processes. However, if the bedding planes of the geology dip then a ridge can form with a gentle slope forming parallel to the dip (a dip slope) and a steep cliff on the other side (Figure 2.17).

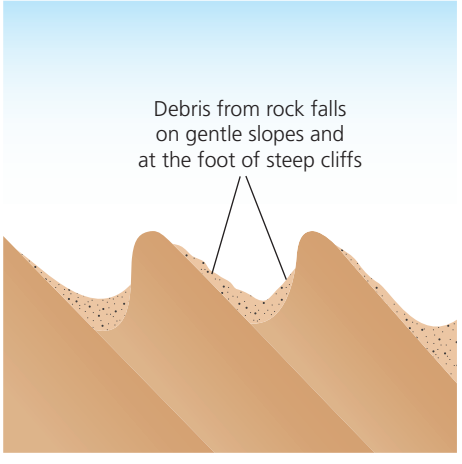


Figure 2.17 Mass movement in rocks with dipping bedding planes

Erosion

Erosion refers to processes that wear away the land surface by mechanical action. In general the agents of erosion include glaciers, rivers, wind and ocean waves and currents. Obviously wind and water are the only ones present in hot deserts. What distinguishes erosion from the processes of weathering is the removal of the *denuded* surface material by processes of wind and water transport (see below). The processes of erosion result in the **denudation**, or degradation of the surface of the Earth.

Transportation

Transportation refers to the processes that move material from the site where erosion took place to the site of deposition. In hot deserts the two agents that can perform transportation are moving water and wind. The material that is being transported is called the **load**. Generally water transports its load in the following ways:

- **Traction**: larger stones and boulders are rolled along the channel bed by the movement of water downstream. Traction occurs with higher levels of energy and **discharge**.
- **Saltation**: smaller stones and pebbles are bounced, or leap frog, along the channel bed. Discharge and energy levels need to be high enough to *pick up*, or thrust, small particles up into the water column, however the flow is not fast enough to keep the particle off the bed so it falls back to the bottom, having been moved a short distance downstream. Particles landing back on the bed may now dislodge other particles up into the moving water, thus repeating the bouncing action.
- **Suspension**: the smallest particles, generally silt and clay, are carried along in the moving water column by the flow of the water. Generally the turbulence of the moving water allows the river to not only pick up this load but also carry it downstream, only depositing the finest particles when the water is almost stationary.
- **Solution**: the transport of dissolved minerals within the mass of the moving water itself.

Deposition

As moving water and wind are not continuously present in hot desert environments depositional features are common. Water and wind both deposit – put down – their load when they experience a reduction in the amount of material they can carry, or a reduction in **capacity**. With a loss of energy both water and the wind will deposit their load. A common cause of deposition in hot deserts is the evaporation of water and streams simply dry up leaving their load behind.

Distinctively arid geomorphological processes

Weathering in hot deserts

Weathering processes are some of the main ways that typical desert landscapes are shaped. Most weathering

in hot deserts is believed to be mechanical or physical. However, it is due to a combination of factors and the reasons for the prevalence of weathering include:

- the regular heating and cooling of rocks
- the presence of moisture, even in the smallest quantities
- the presence of living organisms.

The role of different weathering processes is explored below, but it is clear that variations in the input of energy and water are important factors in controlling what changes take place.

Thermal fracture

A type of mechanical weathering, **thermal fracture**, results from the rapid heating and cooling of rocks in deserts, which is accentuated by the extremes of temperature experienced. With air temperatures during the day rapidly rising to in excess of 40°C the surface layers of rocks can get much hotter, up to 80°C in very exposed locations. At night, temperatures fall very rapidly, generally to below 10°C, but below 0°C in places is not uncommon. This huge **diurnal range**, and the extremes, creates a regular rhythm of heating and cooling that causes the exposed rock to also regularly expand during the day and contract at night. However, this is a rather simplistic understanding of how mechanical weathering operates on rocks in hot deserts and it is now believed that it is a combination of a range of weathering processes that cause rock to disintegrate (degrade) in a variety of ways.

Exfoliation

The mechanical weathering process of **exfoliation** results in the breaking, splitting or peeling off of the outer rock layers (Figure 2.18). This 'onion-skin weathering' was often believed to have simply resulted from pressure changes in the rock as only the outer layers were exposed to the cycles of heating and cooling. However, the reality is more complex. Exfoliation most often affects reasonably uniform coarse-grained crystalline igneous rocks. At depth the rock is under considerable pressure, and as erosion and weathering occur at the surface, removing surface material, the pressure on rocks lower down is decreased. This creates tension and to relieve this, cracks form, running parallel to the surface. As the rock is exposed to the heating and cooling outlined above, salt-rich water is drawn to the surface by **capillary action**. These dissolved salts are deposited in the cracks and together with chemical weathering, the cracks are enhanced and slabs of rock detach from the surface.

Figure 2.18 Examples of rocks affected by 'onion-skin' (or exfoliation) weathering

Chemical weathering

The rate of chemical weathering is slow in hot deserts because they have little soil and moisture levels are low. In more temperate and humid environments it is chemical weathering that helps in the development of soils, so where it is bedrock that is present at the surface in these arid environments, different features are formed. Due to the lack of organic material any soil that does form in a desert tends to be a similar colour to the parent rock. Most of the chemical weathering that does occur in hot deserts results from the deposition of salts precipitated from rain water or brought to the surface by capillary action. Processes of chemical weathering in deserts that contribute to the mechanical breakdown of rocks include:

- **Crystal growth**: this weathering process is now seen as one of the main agents contributing to some of the mechanical processes outlined above. Most simply, as water that is present in spaces (pores, joints, bedding planes) evaporates, salts that were dissolved in the water are deposited. Over time larger and larger crystals develop. These salts have different thermal capacities to the surrounding rock so it is often believed that the differential heating and cooling, and so expanding and contracting, of these salts compared to the rock itself can assist in the physical breaking down of the rock.
- **Hydration**: this is another chemical process that is also linked to the mechanical breakdown of rocks in hot deserts. Some rocks are able to absorb any water that is available, even small amounts of morning dew. As they absorb the water the rock may physically swell, causing the initiation of stresses in the rock, making it vulnerable to other mechanical breakdown. Where hydration causes salt minerals

Figure 2.19a Hardpan

Figure 2.19b Desert varnish

in certain rocks to alter chemically they can become weaker and thus more vulnerable to other kinds of chemical weathering. When water is added to the mineral anhydrite, for example, gypsum is formed.

Other chemical weathering processes:

- **Hydrolysis**: this occurs where mildly acidic water reacts or combines with minerals in the rock to create clays and dissolvable salts; this itself degrades the rock, but both are likely to be *weaker* than the parent rock making it more susceptible to further degradation.
- **Oxidation**: this is the breakdown of rocks by oxygen and water and leads to a common feature of many hot deserts, the red-brown colour of many surface rocks, where minerals of iron have been oxidised.

As any available moisture is continually drawn to the surface and evaporated, many of these chemical weathering processes lead to accumulations of salts on or near to the surface. Many of these become *cemented* and create layers of **duricrust**, or **hardpans** (Figure 2.19a). The chemical nature of the underlying rock can lead to the formation of different kinds of **crusts**: in calcium-rich limestone areas gypsum is created which itself leads to the formation of **gypcrete**, while where lime is present **calcrete** is deposited.

Where iron and manganese oxides have been weathered from rocks at depth and are drawn to the surface in solution, the desert heat evaporates the water and a deep red stain, glaze or **desert varnish** (Figure 2.19b) is left on all exposed surfaces.

Block and granular disintegration

In rocks that are either heavily jointed, like granites, or have prominent bedding planes like limestone, the various mechanical processes outlined above can make masses of rock break down into large blocks

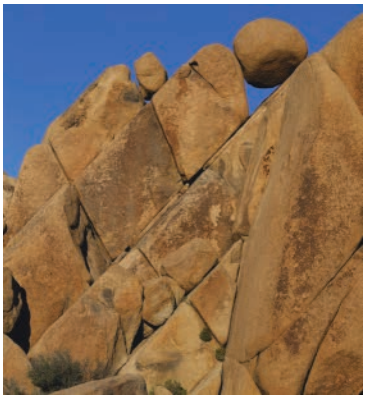


Figure 2.20 Weathering of well-jointed rock

(Figure 2.20). Many igneous rocks like granite have a uniform structure and form in large masses deep below the Earth's surface. When they form, regularly spaced fractures develop and as the weight of the rock above is reduced by erosion of the overlying rock and it is slowly brought towards the surface by uplift, these fractures open slightly becoming joints. It is then that even the smallest amounts of available water in the hot deserts can enter these cracks. Together with the constant heating and cooling, the water can now chemically weather the rock along these faults leading to blocks of rock literally breaking off (**block disintegration**). If temperatures do drop below freezing there is the possibility that **freeze thaw** weathering can occur, where during the day liquid water enters the joints, then freezes at night as temperatures drop below 0°C. As water freezes it expands by about 9 per cent, thus exerting pressure on the surrounding rock. Repeated cycles of freezing and thawing can lead to blocks of rock being dislodged. (Traditionally this process was known as **congelifraction**.)

Where rocks have a more granular structure they can be broken down into separate grains both by freeze thaw or differential thermal expansion and contraction

caused by the huge temperature ranges in hot deserts. Even when only small amounts of moisture are available, possibly even from dew on cold desert mornings, it can find its way into pore spaces within the granular rock. If the temperature drops below freezing it expands, thus putting pressure on the surrounding grains of rock. With repeated freezing and thawing individual rock particles are dislodged. Where rocks have a granular structure different minerals will heat up and cool down at different rates during the diurnal heating and cooling cycle. This leads to individual particles of rock expanding and contracting at different rates leading to individual grains being broken off (**granular disintegration**).

The role of wind in hot deserts

As one of the main inputs in hot desert systems, wind is important as a driver of change and contributes to a range of geomorphological processes occurring there.

Wind and **aeolian** processes are a very common feature of hot desert environments and their margins because:

- the cloudless skies and high angle of incidence means air at the surface is heated and rises, so cooler air moves in to replace it. It is this movement of air that results in winds
- many desert regions are relatively barren with few surface features to create friction to reduce wind speed, or provide shelter. Therefore, winds can blow unimpeded for considerable distances.

As wind has energy and is moving it has the ability to erode, transport and deposit sediment, and so has a significant impact on the landscape of hot desert regions.

Erosion

There are two processes of erosion by wind:

- **Abrasion:** this is often referred to as a sandblasting or sandpapering effect, where the material carried by the wind hits exposed rock surfaces and creates a range of erosional features. A number of factors affect the rate of abrasion: the direction, frequency and velocity of the wind, the **lithology** of the rocks and the size and nature of the particles carried by the wind.
- **Deflation:** wind removes dry, unconsolidated sand, silt and clay particles from the surface and transports it away. Wind only removes the finer material creating a surface covered in a concentration of coarse and fine pebbles known as **reg** or **desert pavement** (Figure 2.21).

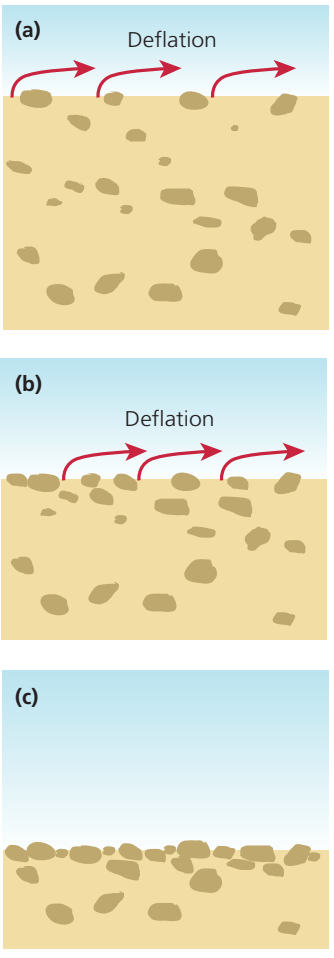


Figure 2.21 Deflation and the evolution of reg

Wind can remove very significant amounts of material from the surface creating **deflation hollows** (Figure 2.22, page 68). Some of the largest deflation hollows are found in North Africa, where they can extend over hundreds of square kilometres. The Qattara Depression in Egypt is the deepest, reaching 134 m below sea level at its deepest point. There is some debate about the exact combination of processes that forms these huge desert depressions, but wind has undoubtedly contributed to removing the millions of tonnes of sand and other material that was once there. The depth of deflation hollows is controlled by the level of underlying groundwater.

Transportation

As wind is an almost ever-present feature of many hot desert environments and there are often large amounts of loose surface material, wind transport is extremely important in shaping the landscape. The transportation of sediment by wind is not only an important agent of change, in terms of processes of erosion and

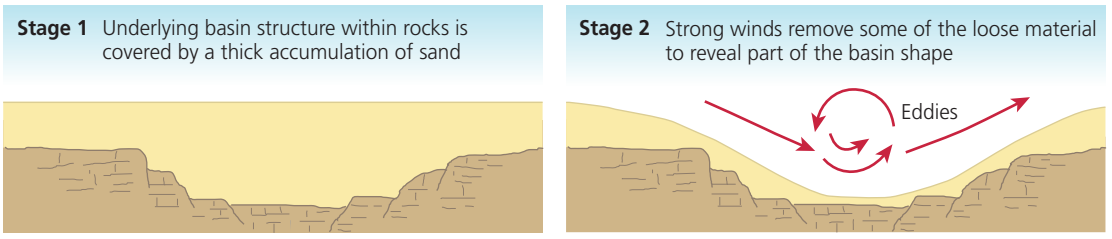


Figure 2.22 The formation of a deflation hollow

deposition, but represents a significant flow/transfer of material. The movement of particles by wind is dependent on a range of factors:

- the wind strength and direction
- the amount of turbulence
- the duration of the wind
- the relief and surface features
- the amount and nature of any vegetation.

Aeolian processes are able to transport material in three main ways. The nature, speed and effectiveness of this depends on the wind strength and the size of the particles (Figure 2.23).

- **Suspension:** this is where the smallest particles, generally less than 0.2 mm, are held in the air. The moving air is able to support the weight of these small particles and carry them indefinitely and appears as a layer of dust or haze close to the surface. If the winds are strong and persistent enough, clouds of dust can be carried to significant altitudes and over hundreds or thousands of kilometres, with 'red rain' of Saharan sand occasionally falling in the British Isles. With very high velocity winds vast amounts of silt-sized particles are lifted creating sand/dust storms.
- **Saltation:** this is a process where sand-sized particles are transported by bouncing and hopping along the surface. When there is turbulence, particles are lifted by a gust of wind, usually only a few centimetres into the air (but sometimes as high as 2 m) then carried a short distance downwind before dropping back to the surface. Grains hitting the surface may hit other particles and cause them to jump into the air, be caught by the wind themselves and so continue the saltation. Particles are usually transported at one-half to one-third of the speed of the wind.
- **Surface creep:** this is where saltating particles return to the surface and hit larger particles that are too heavy to hop, they slowly creep (slide or roll) along the surface from a combination of the push of the saltating grain and the movement of the wind. It is believed that surface creep accounts for as much as 25 per cent of all grain movement in hot deserts.

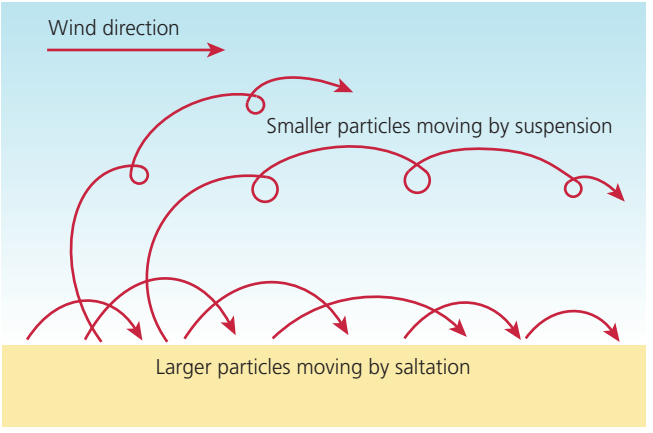


Figure 2.23 Wind transport

Deposition

Deposition occurs when the velocity of the wind decreases until it can no longer transport the grains it is carrying. Wind-deposited materials can often give clues to the scale and direction of winds in the past. For example, analysis of the huge wind-blown deposits of silt (loess) on the Loess Plateau in China, and similar but smaller accumulations in Europe and the Americas can help map the pattern of previous surface winds.

On a more local level, deposition may occur as the wind meets an obstacle and is slowed on the downwind side. This can lead to the formation of dunes (Figure 2.24).

Water in hot deserts

As one of the main inputs in hot desert systems, water is an important driver of change and contributes to a range of erosional and depositional processes that shape the landscape. Water can also transport a large

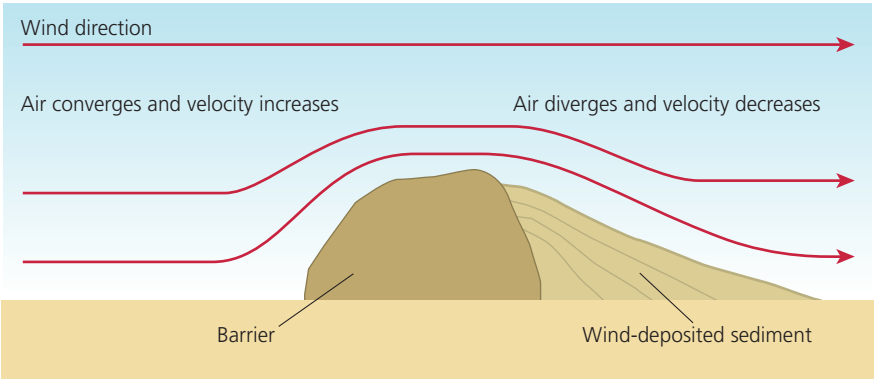


Figure 2.24 Wind deposition around obstacles

amount of sediment so is responsible for a significant flow/transfer of material.

Although rainfall and surface waters are rare in hot desert environments, streams and rivers originating outside of the desert or at higher elevations, along with occasional sudden isolated heavy downpours, can produce enough water to have an effect on the landscape. At higher altitudes in desert mountains and on high plateaux, water acts mostly as an agent of erosion, while in flatter lowland areas depositional landforms are more common.

Sources of water

The origins of water flowing in hot deserts fall into the following categories:

- **Exogenous**: rivers originate external to the desert in adjacent highlands and more humid environments and flow from outside of the desert and pass through them. The volume of these rivers is large enough that they are perennial and maintain a flow all year, although they often do respond to seasonal rainfall (or snow melt) patterns in the source area. These exogenous (or exotic) sources, including ground water and rivers, account for the greatest volume of water flowing in hot deserts. The Nile, flowing from its sources in the Ethiopian highlands and central Africa, flows north across the arid regions of North Africa to the Mediterranean at the Nile Delta. The Colorado River originates from the Rocky Mountains then flows across the arid regions of southwest USA to reach the Pacific Ocean (Figure 2.25). These rivers have been, and continue to be, of significant importance to the populations of the regions they flow across due to their historically reliable and constant flow.

Figure 2.25 The Colorado River

- **Endoreic**: this is where rivers occupy drainage basins that are closed and do not flow out to the sea or other rivers, but instead end inland in lakes or swamps. Examples include the River Jordan flowing into the Dead Sea and the Okavango that flows into Botswana and splits into a number of **distributaries** to form the Okavango Delta.
- **Ephemeral**: these are rivers that flow on the surface only periodically. They are a common source of water in hot desert areas and appear intermittently after heavy rainstorms. The suddenness, rapidity and intensity of the downpours can generate huge amounts of overland flow and very significant increases in channel discharge with extremely short **lag times**. This is the result of a number of factors:
  - the scale of the precipitation event itself
  - the lack of vegetation means that little interception occurs
  - following extended periods of drought, and the removal of much loose surface material by aeolian processes, the desert surface is often baked hard, so the torrential nature of the rainfall exceeds its infiltration capacity.

The episodic role of water

The episodic nature of rainfall in hot desert regions is exemplified well in North Africa. Subtropical deserts like the Sahara remain arid most of the time due to the dominance of high-pressure conditions, giving average rainfall of 25 mm/year generally and only rising to 100 mm in the high plateaux of Ahaggar and Tibesti. In fact parts of western Algeria and southwest Egypt can go anywhere between two and five years without receiving more than 0.1 mm of rain in any 24-hour period. However, single rainfall events can deliver two to three times the annual average in short intense desert storms. Some west-facing slopes in parts of Algeria have recorded intensities of between 8.7 mm in 3 minutes and 46 mm in just over an hour.

Most of this variability occurs in the Sahara during the summer months as the West African Monsoon moves north towards the area and can on occasion allow areas of moist south-westerly air to push further north into the desert itself, giving short-lived areas of low-pressure.

Disruption of the high-pressure cells along the sub-tropical high pressure belt can also lead to troughs, or *gaps*, in the high pressure at the surface, allowing more moist air to move north and create low pressure areas and rainfall on the southern edges of the Sahara.

In desert areas where there are local water sources to provide moisture for the atmosphere, if conditions are favourable, the intense heating of the desert surface can lead to localised warming and rising of air and localised low pressure systems developing. With such intense heating and rapid rising and cooling of the air, large releases of latent heat during condensation can fuel uplift, creating large but short-lived thunder storms.

Sheet flooding

During intense torrential downpours, both the impermeability of some desert surfaces and the sheer speed and amount of rainfall can lead to **sheet floods**. Initially the impact of the rain begins to dislodge and to move loose material across the surface (known as sheet wash or sheet erosion).

As the volume of water increases, this **overland flow** intensifies and can rapidly develop into extensive but relatively shallow floods. This huge amount of overland flow can remove large volumes of material as sheet erosion, thus creating erosional features, but equally contributing to depositional features downstream.

Channel flash flooding

If this overland flow is funnelled into steep-sided narrow valleys, common in some hot desert areas (known as **wadis**, or **arroyos** in the Americas), **channel flash flooding** occurs. These can also move significant loads of sediment and can be very powerful due to the concentration of the water in such a narrow valley, creating a deep, fast-flowing torrent of water with huge potential for erosion and transport of sediment, and thus significant depositional features when the flow recedes.

Review questions

6 Explain why low- to mid-latitude hot deserts have almost equal lengths of daylight and night-time.

7 Why does weathering play such an important role in hot deserts?

8 What are the reasons for sudden intense thunderstorms in hot deserts?

2.3 Arid landscape development in contrasting parts of the world

Origins and development of landforms of mid- and low-latitude deserts

Aeolian landforms

The processes of wind erosion, transport and deposition outlined above can involve very significant amounts of material, so wind action (**aeolian processes**) plays a significant role in shaping desert landscapes. Although similar processes happen in most desert environments, the resulting landscapes will be unique in each location, reflecting the way erosional, transport and depositional processes interact with the local geology and climate.

Ventifacts

Ventifacts are exposed rocks lying on the desert surface that have been abraded or shaped by wind-blown sediment (Figure 2.26). This is usually sand, or finer silt and clay-sized particles and is often likened to a **sand-blasting** effect. Ventifacts can

range in size from pebbles as small as a centimetre to huge boulders many metres in diameter. They are characterised by smooth and/or flattened sides (or facets) and sharp edges or ridges. There are many recognised 'faceted-types', with different numbers of flattened sides and sharp edges.

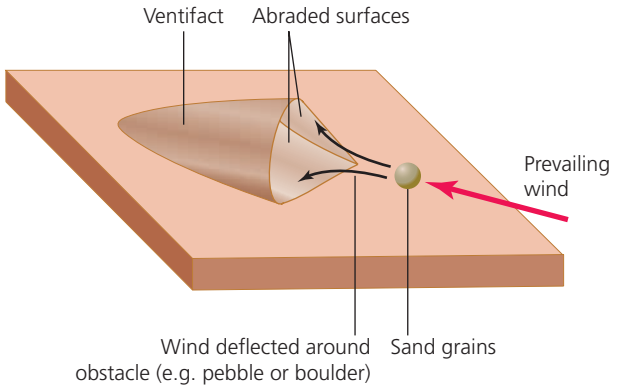


Figure 2.26 Ventifact

It is not completely understood how some ventifacts have multiple smoothed facets, but it is probable that either the rock changed position, thus presenting a different side to the prevailing wind, the prevailing wind changed direction over time, or there are areas with seasonal reversals of wind where the prevailing wind direction is different in summer and winter.

Yardangs

A **yardang** is a streamlined parallel ridge of rock often described as resembling the upturned keel of a boat. They are aligned in the direction of prevailing winds and neighbouring yardangs are separated by a wind-scoured groove or trough (Figure 2.27).

The streamlined shape, with a rounded upwind face and long tapering ridge downstream, suggests that aeolian erosional processes are the dominant processes of formation. It is agreed that abrasion plays a key role in shaping and smoothing the ridge, and in eroding material from the trough. However, as many larger yardangs will have taken millions of years to take their current form it is thought agents such as deflation, sporadic fluvial incision, weathering and mass movement may also have played a role. Some yardangs are found in areas where the geology has horizontal rather than vertical bedding.

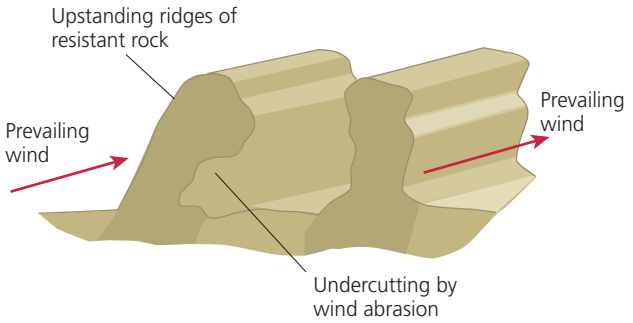


Figure 2.27 Yardangs

Yardangs found in contrasting arid environments vary in size quite considerably, with ridges ranging from a few centimetres high and several metres long, through all sizes to so-called 'mega-yardangs' measuring tens of metres high to many kilometres long. Several regions of the central Saharan desert have large numbers of these large features, including to the west of the Tibesti Mountains in northern Chad. These mega features, in some of the remotest deserts in the world, are so large and occur in such large concentrations that they are easily visible in remotely sensed Landsat satellite imagery.

Zeugen

Zeugen is the term that has come to collectively refer to features such as rock pillars, rock pedestals and rock mushrooms, or yardangs that exhibit considerable undercutting. This range of names suggests a range of different characteristics, but all have been formed where less-resistant rock underlies a layer of more resistant geology at the surface (Figure 2.28a and Figure 2.28b). As wind transports huge amounts of sediment, mainly within two metres of the surface, abrasion is the main agent of erosion. Rock pedestals, mushrooms or pillars tend to be smaller features with dimensions of several metres, however where the features are elongated in the direction of the prevailing wind they can have similar dimensions to yardangs.

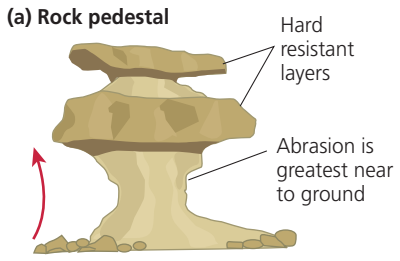


Figure 2.28a Rock pedestal

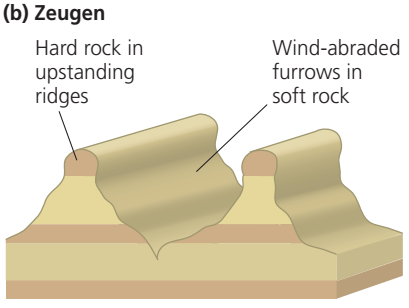


Figure 2.28b Zeugen

Barchans and seif dunes

Sand dunes are the main depositional landforms in hot desert environments and require certain conditions for formation to occur:

- a ready supply of sand
- prevailing winds strong enough to transport sediment, but not strong enough to remove more than is deposited
- steady winds from one dominant direction
- an obstacle, such as vegetation or rocks, to trap the sediment and encourage deposition to occur.

Where the supply of sand is great enough and a large extent of dunes develop, the landscape is referred to by the Arabic term **erg**. The specific size and shape of dunes varies widely depending on:

- the direction and strength of the prevailing winds
- the supply of sediment
- the morphology of the landscape where the dunes form
- the presence of any other winds.

Barchans are crescent-shaped sand-dunes with a gently sloping convex windward side with two *horns* or *wings* extending laterally which curve in a downwind direction (Figure 2.29). The leeward, downwind face (or slip face) of the dune is a steeper concave slope. Barchans only form where winds almost only come from one direction. As material is moved up the windward slope by saltation and creep and then slides, falls and flows down the leeward side, the whole dune moves downwind across the desert landscape. The leeward slope is kept steep by eddying of the wind as it descends down the dune. The height of barchans ranges from a few millimetres to giant dunes of between 500 m and 1,000 m. The

Figure 2.30 Dunes in the Namib Desert

wings advance more rapidly than the centre of the dune, as the rate of advance is inversely proportional to the height of the dune. Individual barchans do exist, but they often occur in groups or linear formations to form a constantly shifting 'sand-sea'.

Seif-dunes derive their name from an Arabic term *sword-dune*. These knife-edged ridges of sand or longitudinal dunes form long parallel ridges of sand separated by wind-scoured depressions. Groups, or chains, of seif dunes can extend over 100 km and be over 200 m high. The tops and sides of such dunes are often serrated due to local wind action and eddying.

Other contemporary classifications of dunes:

- Transverse dunes: resembling large-scale sand ripples, these are large ridges of sand with a steep downwind face and form in large groups or 'fields' of dunes. They occur in areas with an abundant, constant supply of sand, and can be the result of barchan dunes merging if their supply of sand increases significantly.

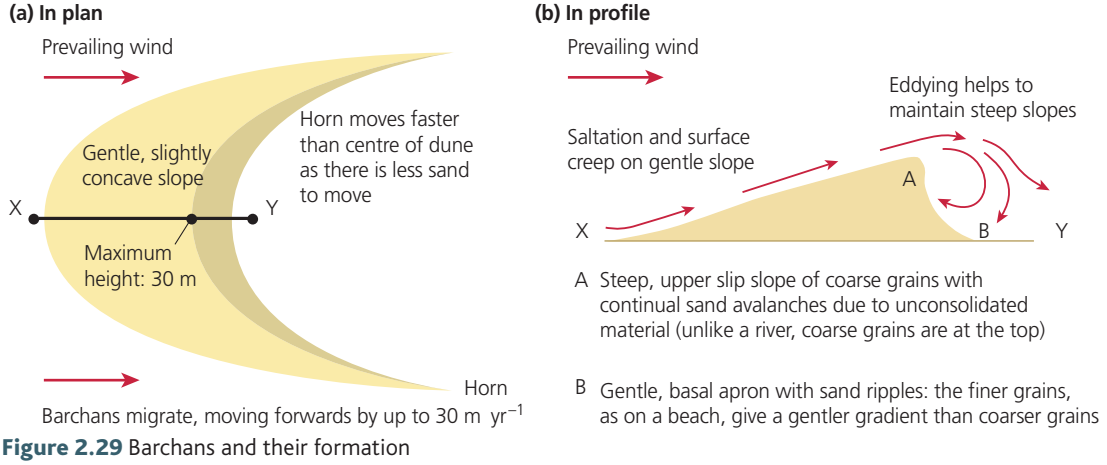


Figure 2.29 Barchans and their formation

- Parabolic (or blowout) dunes: a horseshoe-shaped dune where the open end faces upwind. These are generally relatively stable features that have been colonised by vegetation, and occur where there is a constant supply of sand, often in coastal locations.
- Star dunes: these have variable slip face directions as they occur in areas with an abundant supply of sand but variable wind directions. These are often quite massive and permanent dome-like features extending over hundreds of kilometres and being hundreds of metres high, creating a **draa** landscape.

Skills focus

Construct an annotated sketch or diagram to show the characteristics and formation of an aeolian desert landform.

Landforms created by water

Even though most rainfall in hot desert environments is episodic and low intensity, there are occasional, sudden isolated heavy downpours which can have a significant effect on the development of desert landscapes. In mountainous upland areas and plateaux water most often acts as an agent of erosion, while in lowland regions depositional features dominate.

Wadis

Wadis are steep-sided, wide-bottomed, gorge-like valleys formed by fluvial erosion in arid and semi-arid regions. As the wadis are rarely filled with water the valley walls are steep and often covered with thick layers of weathered material, with a build-up of sediment on the valley floor.

These valleys are either permanently dry, due to climate change and the drying up of former rivers, or only occasionally occupied by ephemeral streams, which can flow as a torrent with very high levels of discharge following sudden storms. They can range from small channels a matter of metres in length to complex channel systems over 100 km long. When ephemeral channels do appear on the valley floor they are often **braided** due to the large amounts of previously deposited material that they have to find a pathway through. The stream pattern will change following each flash flood (Figure 2.31).

Bajadas and pediments

Where there is a distinct break in gradient as highland regions meet a more gently sloping lowland area or depression in hot desert environments, a **pediment** is found (Figure 2.32, page 74). These are gently sloping areas (less than 7°) of bare rock and debris. They are believed to be formed by a range of processes including fluvial erosion, alongside the deposition of material washed down from the uplands and deposited as the **capacity** and **competence** of the water is suddenly reduced by the abrupt change in slope angle.

If wadis, or valleys, containing ephemeral or perennial streams run down from uplands, depositional material may build up on these pediments due to the rapid loss of energy upon meeting the much gentler slopes. The deposited load forms **alluvial fans** as distributaries run out from wadis onto the pediment. The deposited material is **graded**, with the coarsest material left at the upstream end of the alluvial fan and the smallest material on the downstream of the fan. Alluvial fans can extend for a few metres to several kilometres from the mouth of the wadi.

In areas where a linear mountain range has several parallel wadis in close proximity, the alluvial fans may coalesce forming a **bajada** covering, or just beyond, the pediment.

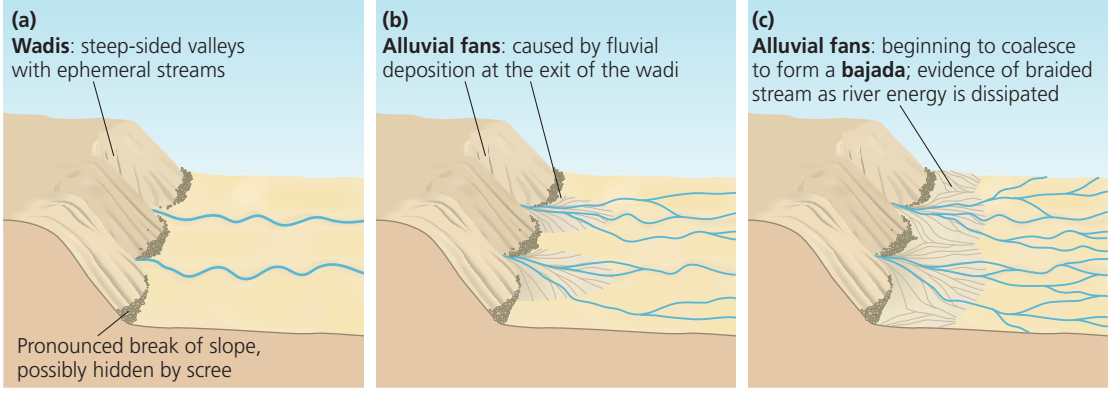


Figure 2.31 The features and formation of a bajada

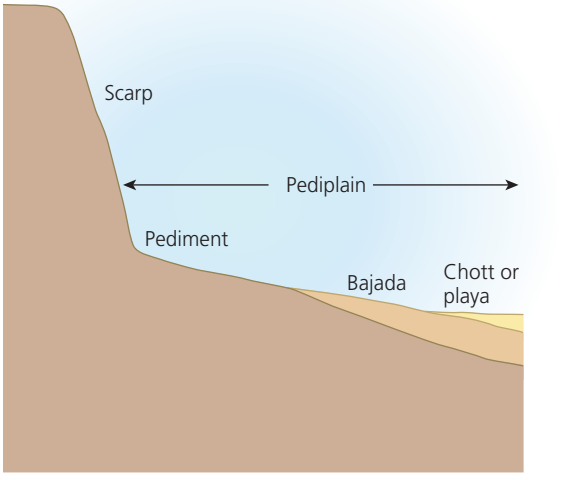


Figure 2.32 Location of the pediment

Playas

Permanent standing bodies of water or lakes are rare in hot desert regions because of the intermittent nature of the rainfall. Where ephemeral streams flow into inland basins or depressions **salt lakes**, **chotts** or **playas** form (Figure 2.33). Following the rare periods of rainfall the water quickly evaporates due to the intense heat, leaving a dry lakebed behind.

These lakebeds consist of salts (evaporites) that were carried in by the streams and precipitated during the evaporation. Sodium chloride (NaCl) is the most common precipitate, giving the dry lakebed a colour resembling a beach (the possible origin of their name, as *playa* means beach in Spanish). Other salts common in these crusts include calcium sulphate (gypsum), sodium sulphate, magnesium sulphate, and potassium and magnesium chlorides. People have commercially exploited these rich salt deposits throughout history.

Figure 2.33 A salt lake in the Namib desert

Chott el Djerid in southern Tunisia occupies an intermontane basin, is over 100 km wide and is almost entirely bare salt flats.

Inselbergs and mesas and buttes

In regions where the bedrock is sedimentary with horizontal bedding planes with a resistant cap rock, fluvial erosion creates **mesas** and **buttes**.

Derived from the Spanish for table, mesas are isolated flat-topped plateaux or hill-like features, with steep slopes or cliffs on at least one side, often falling away to a wadi or canyon. Buttes closely resemble mesas in all ways except that they are much smaller. They are thought to be the remains of heavily dissected plateaux where the water has eroded all but a thin isolated rock pillar.

The lower slopes of both are covered by **scree** resulting from mechanical weathering and rock fall from the upper sections. Monument Valley National Park in Arizona famously is home to some of the most spectacular examples (Figure 2.34).

Figure 2.34 Mesas and buttes in Monument Valley National Park, USA

Inselbergs are more rounded steep-sided hills that rise abruptly from a lowland plain. They are characteristic of tropical savanna zones, but are found in a range of climatic zones and are generally composed of solid crystalline rocks such as granite. There has been considerable debate about the processes that formed them, but most agree that they are relic features of previous geomorphological processes.

One theory suggests that the surrounding slopes retreated in parallel as pediments encroached into the remaining highlands in a process called **pediplanation**. Others argue they are the result of deep chemical weathering when the resistant rock was

buried beneath the surface and has since been exposed by the subsequent removal of the surrounding regolith. Both mechanisms suggest that the desert environments which inselbergs are found in today were once more humid, supporting the theory that deserts experienced wetter periods (**pluvials**) in the past.

Fieldwork opportunities

The Grand Canyon National Park in southwestern USA is a fantastic location to study the landforms of a hot desert landscape. A fieldtrip would allow you to investigate a whole range of features including:

- open deserts
- plateaux
- mesas and buttes.

Studying the mighty Colorado River provides the opportunity to explore the role of water in a hot desert landscape, including the Grand Canyon itself.

Characteristic desert landscapes through time

Perhaps the one feature that initially comes to mind when thinking about **characteristic desert landscapes** is sand. However, hot desert landscapes include some of the most spectacular landscapes in the world, alongside others that are exceptionally monotonous. From seemingly endless flat plains to fields of massive dunes the size of mountain ranges, some hot deserts are also home to remarkable features that attract millions of tourists each year, including Uluru in the Red Centre of Australia and the Grand Canyon in southwestern USA. What separates a desert landscape from others is the lack of vegetation, leaving their geological fabric exposed and allowing the structure of their geomorphological features to be clearly visible.

No two hot desert landscapes are the same because of the unique interaction of processes, time and landforms in different locations. Each hot desert landscape is formed by the processes explored earlier in this chapter and is composed of unique assemblages of the features and landforms also discussed above. Some of the factors relating to time, process and landforms that lead to this fascinating diversity of hot desert landscapes include:

- the speed and nature of weathering. Rocks will be broken down quicker in areas with greater diurnal temperature ranges, where there are more rapid cycles of heating and cooling of the different minerals in different rocks
- the presence of moisture. The amount, and availability, of water in different places will lead to the formation of different landscapes, as it can speed up weathering processes, including the growth of salt crystals or frost shattering, especially if temperatures fluctuate rapidly. Where moisture is present, hot deserts are more likely to have landscapes with biological and organic features, from lichens and bacteria, to grasses and cacti.

The speed at which both of the above shape the desert landscape is heavily dependent on time. Some deserts will have extreme temperature fluctuations on a daily basis, while others will have more moderate diurnal ranges; some will regularly have moisture present, while others will be completely arid for years.

We saw above how the landscape of a hot desert may be dominated by erosional or depositional landforms, and that those landforms could be either aeolian or fluvial. Again, the rate of landscape formation will vary as the presence of wind and flowing water may be constant, seasonal or sporadic, depending on the climate.

It is also important to remember that water and wind not only have the power to shape the desert landscape of the present, but they may also have played a fundamental role in the geological history of the landscape.

- Some landscapes will have features that record a much wetter past, where great rivers may once have flowed through areas that are now almost completely arid, for example, the presence of inselbergs, indicating that an area was once far wetter than it is today.
- Some desert landscapes may be dominated by features that are relatively recent, like alluvial fans at the mouth of wadis that may have developed over hundreds to a few thousands of years.
- Other desert landscapes are truly ancient, like the great 'shield deserts', including much of central Australia that is dominated by vast almost featureless plains, with rare outcrops of resistant rock, like Uluru.

Wind and water have been slowly wearing away these landscapes since the ancient super-continents, like Gondwana, broke up between 100 and 200 million years ago. In truth no two deserts are exactly alike.

Badlands

However, water action can lead to the formation of one of the most distinctive features of semi-arid regions, **badlands** (Figure 2.35). In areas with relatively impermeable, less-resistant geology, run-off from heavy but sporadic rainfall moulds a distinctive landscape. Vegetation cover is never able to establish itself due to the aridity so there is little holding the regolith and bedrock together, so flowing water is able to create dramatic erosional and depositional landforms.

Badlands have a range of general features:

- wadis of all shapes and sizes with steep sides and debris-covered bottoms
- unstable slopes that regularly collapse as gullies erode headward into hillsides
- regular mass movement as slope failure and slumping are common
- as the many wadis and small gullies emerge onto lowland areas, alluvial fans and bajadas are common
- as water carves out surface cracks, pipes are formed which are eroded into larger caves as surface water flows into them
- as caves erode further over time, aided by weathering processes, natural arches form.

Due to their striking landscapes and often inaccessible locations, many badlands, especially in the USA, have proved popular with adventure tourists, as they remain undeveloped due to the lack of options for other viable economic activity. The unique landscape of some of the badlands of southern Tunisia near Matmata has been popular with film-makers for some time, with sequences for the original *Star Wars* film being filmed here, keeping the location popular with tourists to the present day.

Figure 2.35 Badlands in South Dakota, USA

This section shows just how unique all hot desert landscapes are, with different factors giving rise to different combinations of landforms and features to give each desert its own characteristic landscape. Unlike badlands, not all assemblages of landforms form a named landscape type. Table 2.3 gives a summary of the dominant features of some of the world's most well-known hot deserts.

Table 2.3 Physical features of different desert landscapes

Desert	Characteristic landscape features
Arabian (Arabian Peninsula)	Almost entirely sandy, with some of the largest sand dune systems in the world
Australian Great Sandy, Simpson and Great Victoria Gibson and Sturt	Mostly sandy plains Mostly stony surfaces
Chihuahuan (Arizona, New Mexico, Texas and north central Mexico)	High flat plateau with some stony surfaces and sandy soil, broken by mountain ranges and distinctive mesas
Kalahari (southwestern Africa)	Extensive sand dunes interspersed with gravel plains
Mojave (Arizona, California, Nevada)	A varied landscape including sandy soils, gravel pavements and salt flats
Sahara Northern Africa	Vast ranges of dunes amongst mountains and rocky areas, and gravel plains and salt flats
Thar India and Pakistan	Mostly sand dunes with areas of gravel plains

Skills focus

It is important that you are able to describe and explain the features of characteristic hot desert landscapes. Use the library or the internet to research two contrasting hot deserts, then complete the following tasks.

a Identify the different landforms in each desert.
b Explain the processes that created these landforms.
c Try to identify the age and origin of the features, for example, do they result from current processes or are they relic features that are being modified in the present day?
d Draw and annotate a sketch of the landscape to describe the characteristic features.

Review questions

9 Explain how aeolian processes form different landforms in hot deserts.

10 Explain how water can shape landforms in hot deserts.

11 Why do different hot deserts have different landscape features?

12 When trying to understand the current landscape features of hot deserts, why is it important to understand what processes have happened in the past?

2.4 Desertification

There have and continue to be a range of definitions of **desertification**, and it is generally accepted that it is not simply the 'advance of deserts', but it could include sand dunes encroaching on land. However, a UNESCO source in 2015 defines it as 'the persistent degradation of dryland ecosystems by human activities and by climate change'.

Key term

Desertification – The persistent degradation of dryland ecosystems by human activities and by climate change.

The introduction to this chapter indicated that hot deserts are open systems and that the process of desertification is an example of positive feedback. It is also possible to think of desertification as an open system with inputs, components and outputs, although this is quite simplistic and the issue is developed in greater detail later in the chapter. This section goes on to explore the changing extent of hot deserts during the Holocene and the causes and impacts of desertification, alongside the possible future impacts of climate change on the landscapes and people in areas at risk of desertification in the twenty-first century.

The changing extent and distribution of hot deserts over the last 10,000 years

As the Earth has experienced colder and warmer periods during the **glacials** and **interglacials** of the Pleistocene, the extent and distribution of desert environments have altered considerably. Records show the distribution of deserts has changed markedly since the last glacial maximum about 18,000 years ago (Figure 2.36 – NB not all areas shown were hot deserts, but all were arid).

We clearly see that during the height of the last glacial maximum about 18,000 years ago, the extent of arid areas was huge, however almost all of these areas would not have been hot, as global temperatures were on average about 12°C lower than today. By about 8,000 years ago, during the 'Holocene Climate Optimum', temperatures were much warmer, with similar temperatures to today, but the extent of deserts was confined to very small areas, with present day deserts extending over considerably larger areas.

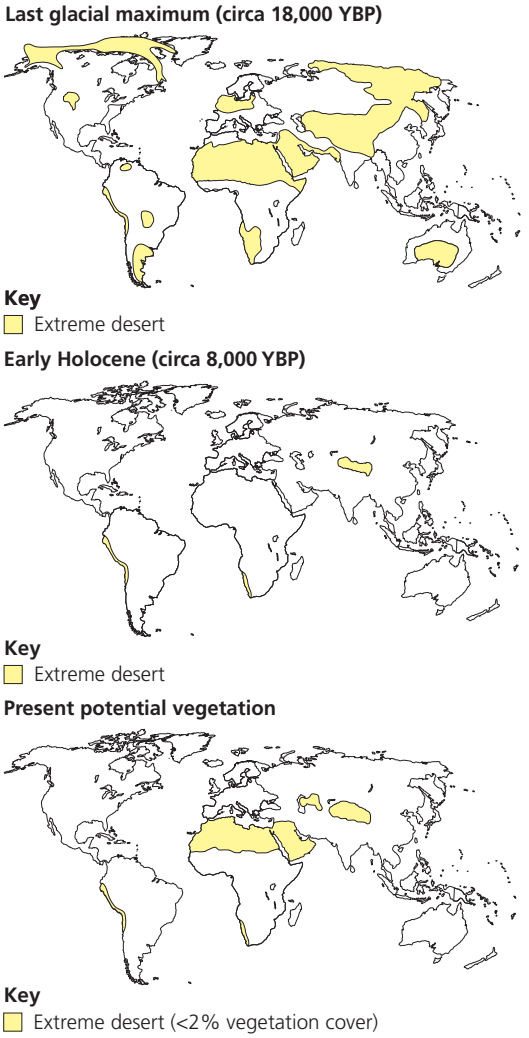


Figure 2.36 Changing extent of deserts over the last 18,000 years

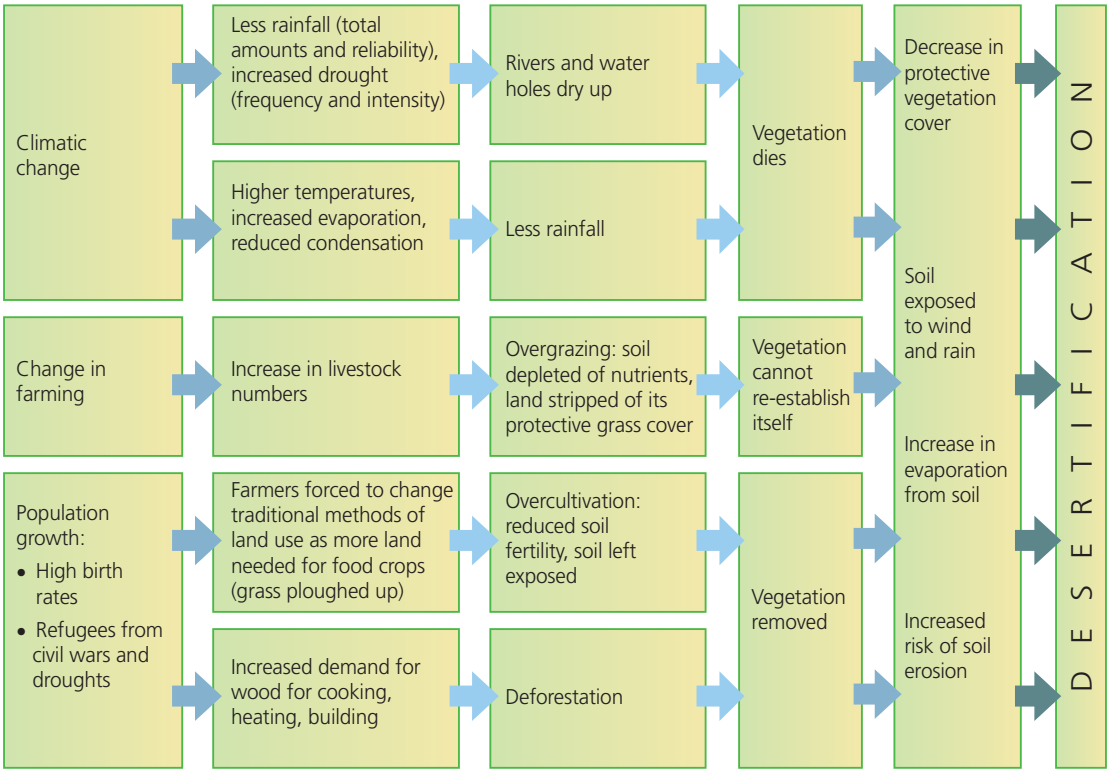


Figure 2.37 The causes of desertification in areas such as the Sahel

The causes of desertification

In the context of the definition and history given above, the causes of desertification are a combination of human factors intensified by physical processes and natural and human enhanced climate change. Concern about desertification is not a new phenomenon, but some would argue that it was the first global environmental issue to be acknowledged as occurring on a global scale.

Due to debate among some scientists about what is the most important cause of desertification: climatic factors or direct human action, there appears to be slow movement on agreeing the best way to tackle it. However, it is clear that whatever the causes of it, recent enhanced climate changes, together with direct human activities, are exacerbating desertification.

Links

See Chapter 1, Water and carbon cycles – The carbon cycle, pages 24–41, to understand the relationship between desertification and the changes in the carbon cycle.

Climate change

The IPCC states that areas like the Sahel are experiencing:

- less and more unpredictable rainfall – in the last 25 years of the twentieth century the Sahel had the greatest loss of rainfall anywhere on the planet, and rates of decline were the fastest since instrumental measuring began
- higher temperatures – without available moisture local evapotranspiration is limited, and in desert areas this is often the only local water source for precipitation
- a reduced water supply as rivers dry up and the water table falls.

Human causes

Historically population levels in many hot deserts and their margins remained relatively low and stable and were well adapted to the environmental conditions. However, in recent decades people are having a direct effect on rates of desertification due to:

- population growth from natural increase (possibly as small local populations begin to move through stages of demographic transition) and migration as populations flee environmental problems or human conflicts in neighbouring areas

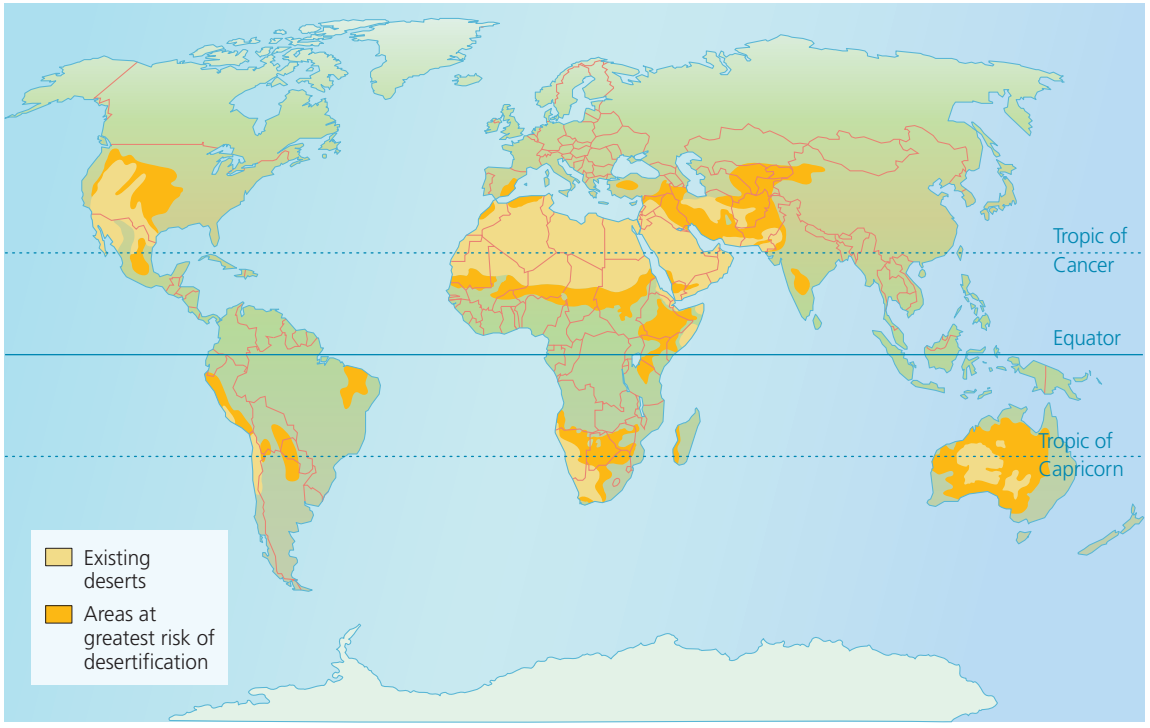


Figure 2.38 Areas at risk from desertification

- population pressure on the land because of:
  - intensification of agriculture to feed larger local populations
  - over-cultivation, overgrazing and deforestation, for a huge range of reasons in different locations. All remove vegetation which cannot then get re-established due to soil erosion following the exposure of the surface to wind and rain
  - the increasing demand for fuelwood in some locations which cannot be underestimated.

All the above lead to **land degradation** and ultimately desertification.

Areas at risk of desertification

The areas of the Earth at greatest risk of desertification are shown in Figure 2.38.

According to the UN, approximately 25 per cent, or 3.6 billion hectares, of the Earth's surface is currently desertified. The UN also reports that around one billion people in some 100 countries are at risk of desertification, with around 12 million hectares of land being lost to desertification annually.

Impact of desertification on ecosystems, landscapes and populations

The impacts of desertification are extensive and can vary considerably from place to place, but general impacts on ecosystems, landscapes and populations are outlined below with some place-specific issues covered in the case study at the end of this chapter.

Climate change and possible futures for local populations in hot deserts

Predicted climate change and its impacts

It is clear that the area of degraded land and deserts has increased over recent decades and is continuing to do so at present. It is also clear that the majority of climate scientists agree that during the last quarter of the twentieth century and start of the twenty-first, the Earth has experienced significant climate change with unprecedented enhanced rates of global warming. Some sources suggest that the average temperature of the planet rose between 0.3 and 0.6°C during the last century, and that with every doubling of the concentrations of atmospheric CO_2 drylands could experience temperature increases of between 2 and 5°C.

Future desertification is not a simple problem to predict or solve. Desertification is not a problem with a single simple cause or equally with a single simple solution (Figure 2.39, page 80). Figure 2.40 (page 81) attempts to show how future climate change and desertification are both components of major feedback loops. The inner loops show how desertification and biodiversity loss are linked to climate change via soil erosion while the outer

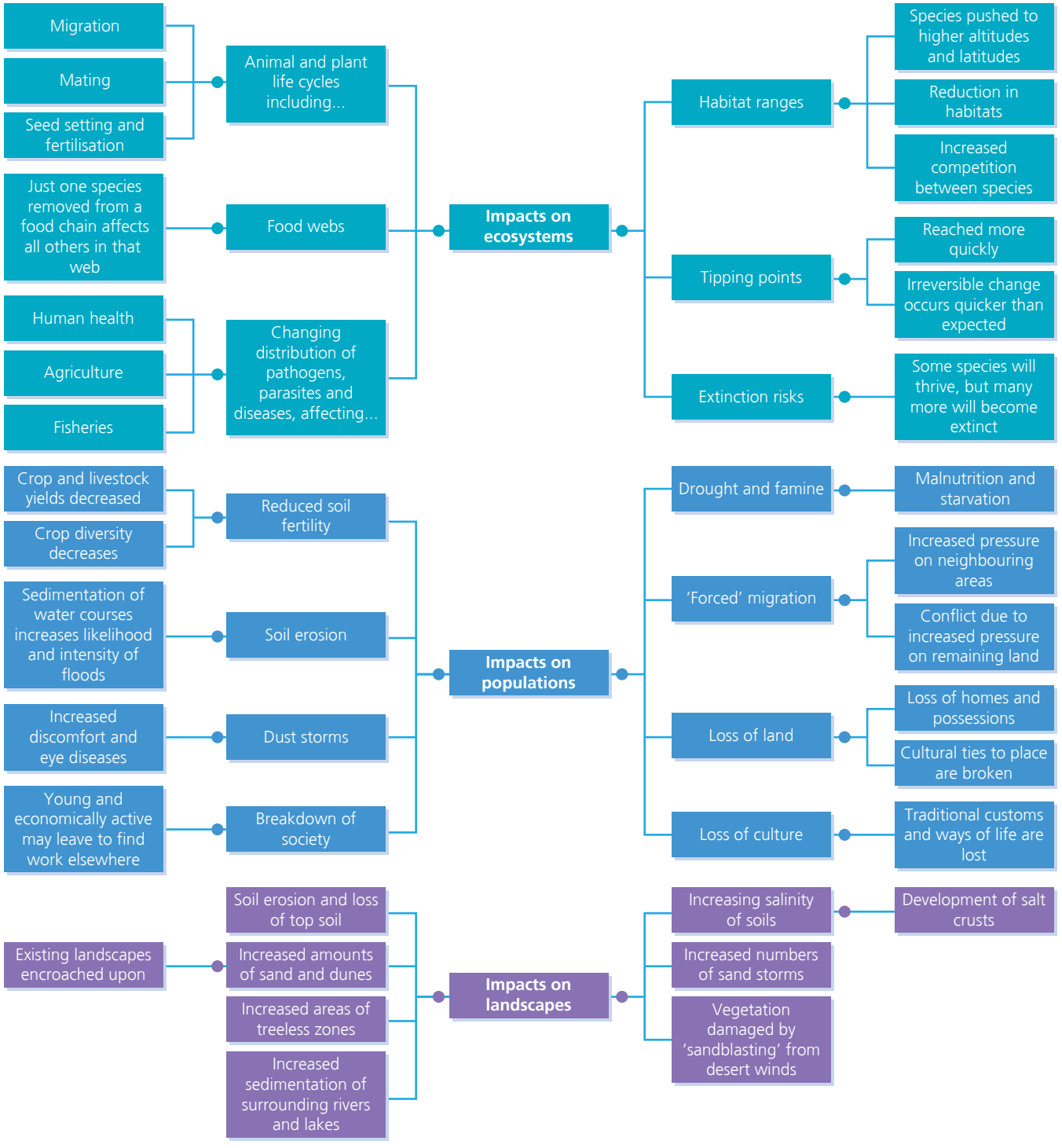


Figure 2.39 Impacts of desertification on ecosystems, landscapes and populations

loop illustrates the link between climate change and biodiversity loss.

The problems relating to future desertification are likely to affect many of those identified as the **least developed countries** by the UN, as many of these are located in or have areas of hot deserts at their margins. The problems of very low levels of HDI (**Human Development Index**) will exacerbate and be exacerbated by the impacts of desertification, especially as many of these countries also have the most pressing population pressures. However, the pattern and rates of desertification are as varied as the rates of poverty and poverty-related problems. This suggests that other than global-level solutions to climate change, solutions to the problems posed by desertification are going to be found at **grassroots** levels by the local populations at risk.

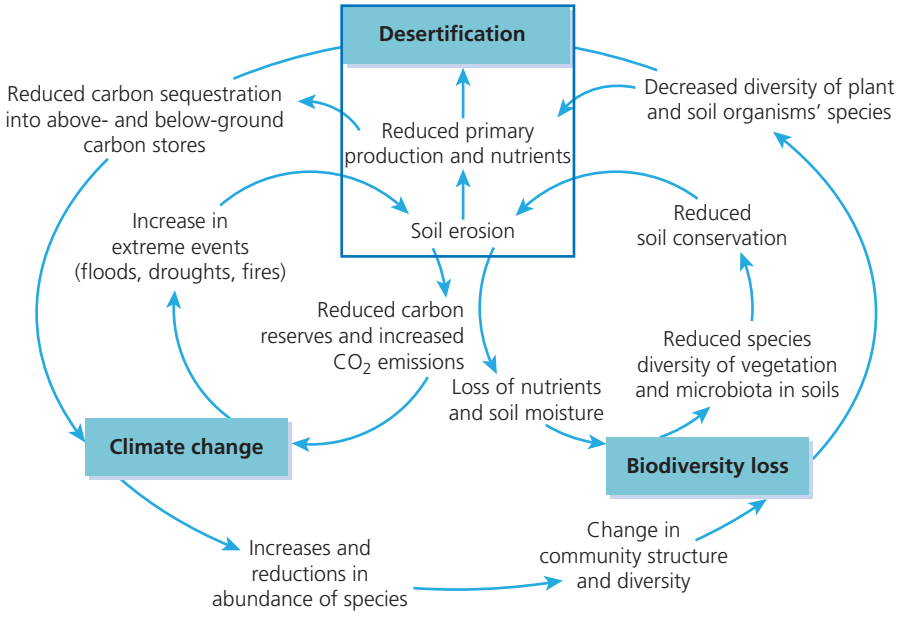


Figure 2.40 Links and feedback between desertification, global climate change and biodiversity loss

Alternative possible futures for local populations

The Millennium Ecosystem Assessment commissioned by the UN in 2005 said that looking to the future is difficult and combatting desertification will depend on the development path the world follows. They suggested four scenarios, two of which see a world of increasing **globalisation** and two of which see a world of increasing **regionalisation**. In both scenarios they see opportunities for both **top-down** and **bottom-up** solutions including actions that are *reactive* to problems that already exist and actions that are *proactive* to try to limit future problems (Table 2.4).

Global officials may have mapped out four possible routes to dealing with desertification, but the consensus is that the area of desertified land will increase in the future. However, there are some key points on desertification and the future:

- Poverty and unsustainable land use will continue to be the main causes of desertification.
- Relieving the human pressure on at-risk areas is strongly linked to poverty reduction.
- Population growth and increased demand for food will lead to an expansion of cultivated land and inevitable deforestation and land degradation.
- Climate change and desertification are linked, but the future impacts and mitigation of both are uncertain.
- Hot deserts and their margins are predicted to experience more extreme weather events such as droughts and floods.
- Proactive approaches are likely to be more effective in addressing economic problems and desertification.

Table 2.4 Scenarios for addressing the problems of desertification

		Path of world development	
		Globalisation	**Regionalisation**
Approach to ecosystem management	**Reactive**	**Global orchestration** Suggests an increasingly globalised world where ecosystem management is reactive to current issues. 'Development' initiatives are top-down and focus on equity, economic growth and public goods such as infrastructure and education.	**Order from strength** In a more regionalised world, countries, regions and communities are empowered to react to environmental problems as the 'development' focus is more bottom-up but concentrates on security and economic growth. However, it does not look to issues that may arise in the future.
	Proactive	**TechnoGarden** In a globalised world, technology is shared to allow populations affected by future desertification to be proactive in prevention by using green technologies. The problems of those affected can be alleviated, thus alleviating future global-scale problems from developing.	**Adapting mosaic** A bottom-up proactive scenario where the emphasis is on empowering the development of local adaptations and learning.

- If local knowledge and local priorities are incorporated into future plans to help mitigate the vulnerability of local populations and ecosystems, it is less likely that 'solutions' will be proposed that simply make the problem worse or move it elsewhere.
- Local populations need to be helped to adapt to current and future pressures to reduce current and future vulnerability and improve their resilience.
- The extension of intensive irrigation systems will have a negative impact, increasing desertification.
- Current pressures on ecosystems due to climate change, overgrazing, over-cultivation and large-scale irrigation will get worse if they are not 'reacted' to.
- It is not inevitable that globalisation will lead to increased desertification.
- Globalisation could lead to increased co-operation and resource transfers to support those on the front line in areas affected.
- Organisations in a globalised world can be proactive in promoting the protection of local property rights and encourage all levels of governance to integrate environmental issues into their planning.
- In a fragmented regionalised world, the role of global organisations will be limited and there could be reduced interest in the exchange of technology and resources, and in issues beyond national borders.

Skills focus

In what ways do you think geographical information systems (GIS) and remotely sensed data could be useful for monitoring desertification?

Fieldwork opportunities

The Toubkal National Park and surrounding area, two hours drive from Marrakesh in Morocco, is accessible and provides the opportunity to study water management in the margins of a hot desert environment. Also, due to its location on the edge of the Saharan desert, it is possible to study the impacts and potential future impacts of desertification on the local populations and their way of life.

Review questions

13 What are the main reasons for the increasing rates of desertification in hot deserts?

14 How is desertification linked to feedbacks associated with current global climate change?

2.5 Quantitative and qualitative skills in hot deserts

Quantitative and **qualitative** techniques in geography are defined and explored in Chapter 12, and you should read this section carefully.

It is very unlikely that you will be fortunate enough to visit and undertake fieldwork in a hot desert environment as part of this course, however it is important to reflect on how a range of quantitative and qualitative skills could be utilised in a study of hot desert landscapes. As most hot deserts are inaccessible for most of us, the case study on page 84 illustrates how a number of quantitative techniques could be employed in a sand dune landscape which, like hot deserts, is also dry and dominated by aeolian processes.

Quantitative data and hot deserts in the twenty-first century

Satellite technology has proved to be one of the most powerful tools for collecting quantitative data about hot deserts, especially the ongoing process of desertification over the last 30 years. The **Landsat** collaboration between NASA and the United States Geological Survey (USGS) began in 1972, with Landsat 8 (2013) now providing some powerful evidence for the spread of desertification (Figure 2.41). Landsat 7 (launched in 1999) is being replaced with Landsat 9 in 2021.

Landsat can provide images of the same geographical area over time. If the images are taken at the same point in the growing season they can indicate where significant changes to the landscape have taken place. Powerful geographical information systems (GIS) and geospatial mapping techniques can use high-tech software to apply different filters to the satellite imagery to identify changes to the desert surface, including vegetation cover, soil characteristics, water content/cover and surface temperatures, among others.

Increasingly scientists are working towards creating global databases of field observations, where researchers on the ground record changes in land use, which can be used alongside remotely sensed data to record the extent and impacts of desertification.

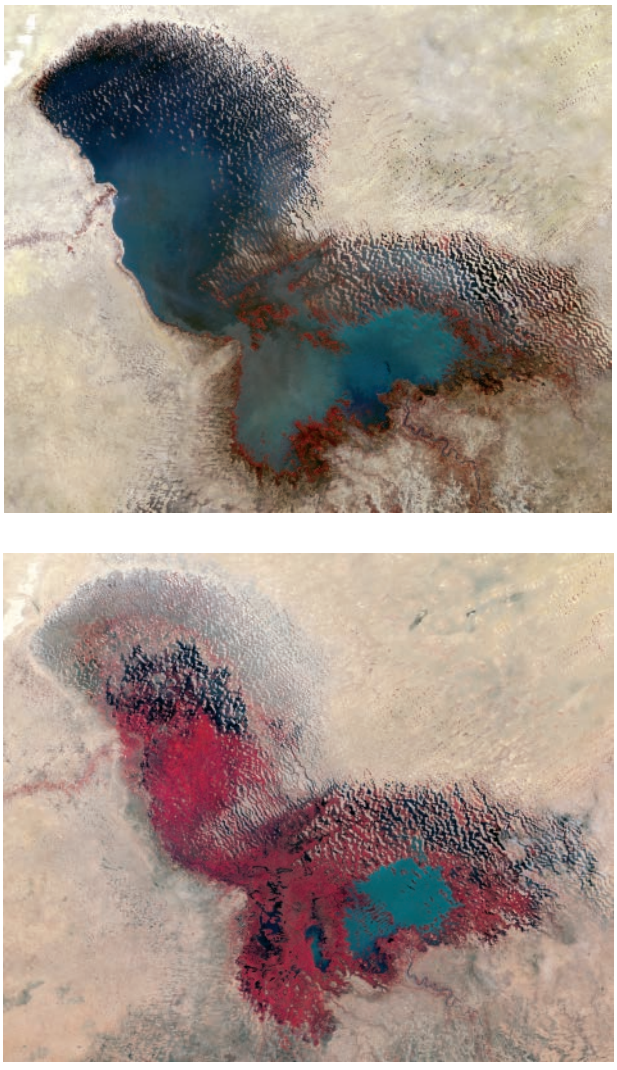


Figure 2.41 Landsat images show the shrinking of Lake Chad in the Sahel between 1973 and 2017

Source: NASA

Qualitative data and hot deserts in the twenty-first century

This chapter illustrates how understanding the causes of desertification is complex and a number of authors suggest that this provides opportunities to employ more qualitative research techniques. Some suggest that it is crucial to understand and explain the links between social and economic factors and local development pathways that produce and exacerbate land degradation and desertification. Quantitative empirical data can provide excellent data on the existence and extent of desertification but cannot explain the social and economic factors that drive local populations to pursue unsustainable activities that degrade the land.

One study on Lesvos in Greece promoted the use of in-depth interviews which, although prepared in advance, did not have a specific research hypothesis (Figure 2.42). Without using leading questions, researchers were able to collect a large range of information about the experiences and thoughts of the respondents (Iosifides and Politidis, 2005).

Figure 2.42 Soil degradation in Lesvos, Greece

The interview schedule used with local farmers in Greece focused on five themes:

- characteristics of the respondent and their farming practices
- views on rural development and development processes
- views on local socio-economic conditions
- community resilience (alternative employment and income opportunities and social relations)
- views of the local community on the problems of desertification and land degradation.

In-depth interviews generate huge amounts of information which then needs to be **coded** into different categories so that a detailed analysis can take place. The researchers suggest that it is impossible to formulate solutions to the problems of desertification without fully understanding the motivations and circumstances of the local populations of the areas affected, and that this can only be achieved using qualitative techniques.

Review question

15 Why would you use quantitative methods rather than qualitative methods to assess the extent of desertification?

Case study of a hot desert environment setting: Shifting dunes threaten settlements – measuring aeolian processes that shape sand dunes

Geographical context

In areas like the Sahel, the semi- to hyper-arid zone running in a belt east to west across Central Africa along the southern fringes of the Sahara desert, the notion of huge areas of sand dunes rapidly spilling out of the desert and swallowing up huge expanses of land is generally agreed to be a misconception. However, it is well documented that dunes have, at times, advanced towards the city of Nouakchott, the capital city of Mauritania on the western edge of the Sahara (Figure 2.43).

For areas concerned about encroaching sand dunes scientists would obviously need to take measurements to establish the size of the dunes and the rate at which the sand is advancing to establish the nature of the threat.

Figure 2.43 Dunes advance towards Nouakchott

Response

As geography students we could use a range of techniques to gather field data to measure and quantify some of the aeolian processes at work in shifting sand dunes. As most of us will not be able to visit sand dunes in mid- and low-latitudes, the same techniques could be employed in coastal sand dunes closer to home where similar aeolian processes are at work. This case study includes examples of the types of fieldwork data and techniques that could be conducted in a local sand dune setting.

Investigating aeolian processes in a local sand dune setting

Fieldwork aim

To investigate how wind speed and direction affect the rate of sediment movement on a stretch of sand dunes.

Background

A fieldwork investigation was completed along a stretch of dunes along the Drigg Coast. Drigg is located on the Cumbrian coast in northwest England (Figure 2.44a and 2.44b).

Collecting field data

At regular 5 m intervals along the dune front, students could measure the rate at which the wind is removing/transporting sand. At each collection point a small plastic sheet is secured on the sand surface. Using an anemometer and weather vane, the wind speed and wind direction are recorded. A known amount of sand is placed on the plastic sheet and a stop watch is used to record the length of time taken to remove the sand (Figure 2.45).

Using systematic sampling the wind direction and wind speed could be recorded at 5 m intervals along the dune front. A measured amount of sand is placed on the plastic sheet and the time taken for all the sand to be transported away recorded in the table below.

This primary data could then be analysed to assess areas of the dunes where sand is being transported most rapidly. Or, knowing the quantity of sand and the full length of the dunes, a calculation/estimation could be made of the amount of sand being moved along the

Distance along dunes (m)	0	5	10	15	20	25	30	40	45	50
Wind direction										
Wind speed (m/s)										
Time (s)										

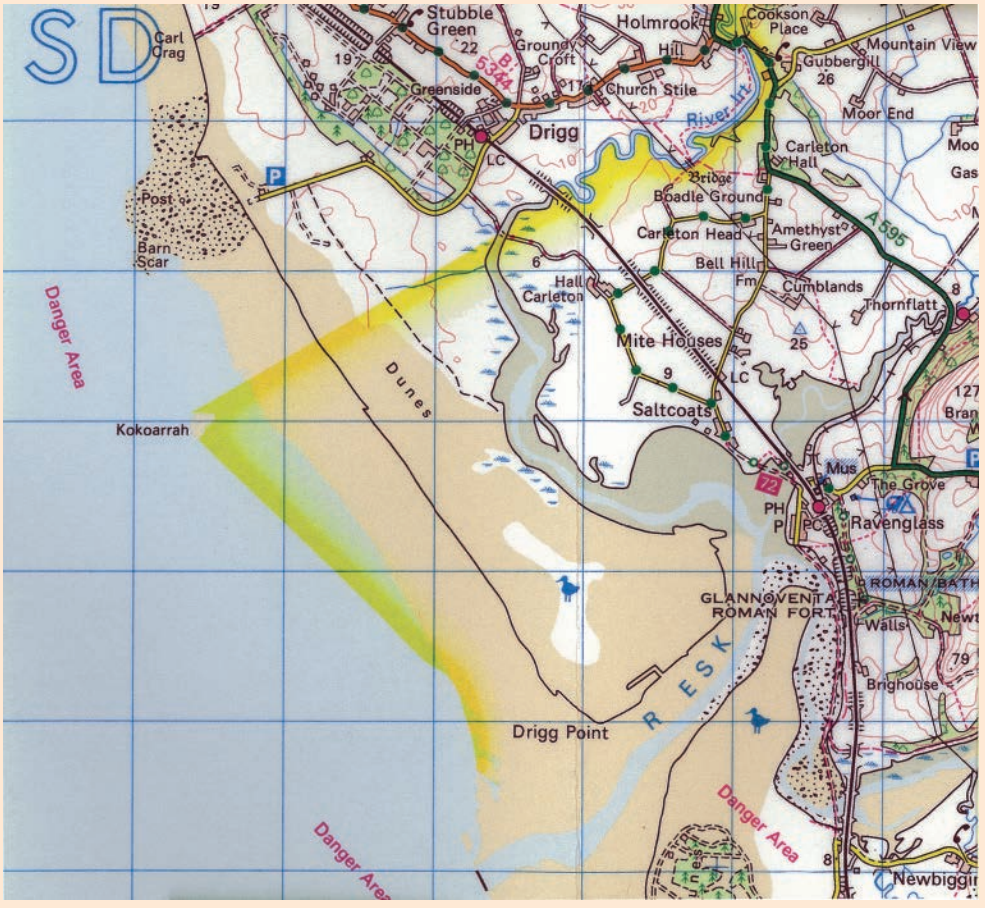


Figure 2.44a The Drigg Coast, Cumbria. © Crown copyright and database rights, 2021, Hodder Education, under licence to Ordnance Survey. Licence number 100036470

Figure 2.44b Sand dunes on The Drigg Coast, Cumbria

Figure 2.45 Measuring the effect of wind on sand dunes

dunes in a given period, for example, in grams of sand per second per km.

Students could also investigate if the slope angle or profile of the dunes made any difference to the rate of sand transportation. So at each 5 m interval along the dunes a measuring tape, ranging poles and clinometer could be used to measure the profile of the dunes. The tape measure is placed at the high tide mark on the beach and taken perpendicular to the beach inland across the dunes in a straight line. Ranging poles are placed at 2 m intervals along the tape measure and the clinometer used to measure the angle between each pole (Figure 2.46, page 86).

Figure 2.46 Measuring the dune profile

Having recorded this data in a table like the one on page 84, students could then use it to construct cross sections through the dunes. A correlation could then be conducted to see if there is a relationship between the steepness of the dunes and the rate of sand transportation from the earlier investigation.

Skills focus

What techniques could you use to present and analyse this field data back in the classroom?

Case study at a local scale: Causes, impacts and implications of desertification for sustainable development

Aims of case study:

- to illustrate how population pressure can lead to desertification
- to analyse the relationships between people and landscape in areas affected by desertification
- to show how sustainable solutions can help people adapt to and mitigate the effects of desertification.

Figure 2.47 Location maps - Adrar, Touat region, Southern Algeria

Close to Adrar in the Touat region of southern Algeria (Figure 2.47) people have turned to a traditional approach to improve their **resilience** to the arid climate and **adapt** to and **mitigate** the impacts of the pressures of desertification on their precious water supply.

Geographical context

For centuries efficient and sustainable irrigation systems utilising the limited water supplies of the **oases** of this area of Northern Sahara have enabled local populations to survive in the hostile arid conditions. Small settlements traditionally developed around a small oasis, utilising ground water to sustain themselves and small-scale agriculture (Figure 2.48). However, in recent decades water supplies, soil quality and the sustainability of these oases have come under threat due to:

- population growth – Algeria's population is growing at almost 2 per cent per annum and life expectancy has increased to 78 for females and 76 for males in 2020
- intensification of agricultural production and over-cultivation – for example, moving towards monoculture of high yielding modern seed varieties (some of which are cash crops for export) and the use of chemical fertilisers and modern machinery (Figure 2.49)
- most importantly the use of modern powerful water pumps removing huge quantities of water direct from groundwater stores. This is then used for jet irrigation where most is lost to high rates of evaporation while there is only an average of 50 mm of rainfall a year.

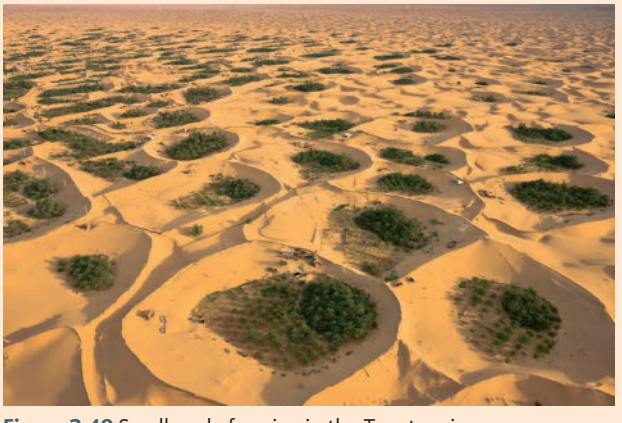


Figure 2.48 Small-scale farming in the Touat region

Figure 2.49 More intensive farming in the Adrar region

This has created a number of issues:

- further intensification of agriculture to try and counter declining yields
- deeper and deeper wells have been dug to tap even deeper groundwater supplies
- more intensive use of fertilisers.

Response

Over the last three decades (prior to the events of the Arab Spring in North Africa), and with research continuing in 2017 and 2018, there has been a collaboration between a network of **non-governmental organisations (NGOs)** from the Mediterranean region and Algeria itself to develop a strategy to protect the oasis ecosystem of this region (including the Algerian association Touiza and The Mediterranean Co-operation of the MED Forum). Four oases were selected where activities took place, such as restoring palm trees to provide shade and help prevent soil erosion and campaigns to raise public awareness about preserving the oases. Ultimately this led to schemes to rehabilitate and preserve the traditional irrigation schemes known as **foggaras**.

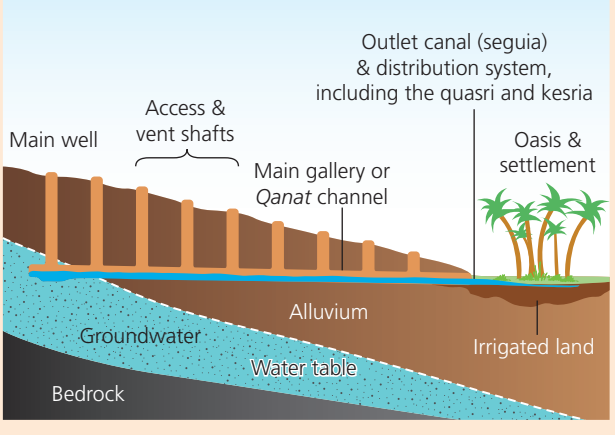


Figure 2.50 Diagram of a foggaras irrigation system

Foggaras are a series of man-made galleries and tunnels that harvest groundwater. They work where the oasis and village are located on the floor of a gentle-sided valley, or at the foot of a gentle slope. Wells are dug about 2.5 km *upstream* from the oasis and underground pipes then run almost horizontally under the surface, at a very slight incline of about 1–2 mm/m. Water then flows gently downstream to the oasis and settlement. The galleries and access wells (located about every 20–30 m) are constructed out of locally quarried stone and palm trunks held together with a cementing mix of clay and straw (Figure 2.50).

The advantages of foggaras include:

- low gradient and gravity flow prevent erosion by fast-flowing water and do not require expensive diesel-driven pumps
- water flowing underground prevents evaporation and water loss
- locally sourcing building materials provides work and reduces the need for imported materials.

Once the water reaches the oasis it flows into an open-air canal called a *seguia*. Here it is collected in a small basin, or reservoir, called a quasri (Figure 2.51).

Figure 2.51 Water collected from a foggaras system in the quasri

With water being such a valuable commodity for the community, they set up a 'water assembly' to decide how much water those who have water rights receive. A comb-like stone device (kesria) is used to divide the water from the quasri into various irrigation channels to be used freely by each of those with water rights (Figure 2.52, page 88). The 'water deciders' hold responsibility for distributing the water.

Figure 2.52 Water distributed via a kesria

Maintaining the foggaras system can be a dangerous and difficult job, so if galleries or wells do collapse, more modern surveying techniques and equipment are used to help repair or dig new wells to direct water to the system. In oases where electricity is available, any spare/excess water downstream of the oasis can be gathered and pumped back to the main well at the head of the foggaras to maintain the supply and prevent the groundwater supply running low.

Studies in 2017 and 2018 suggested that foggaras projects have been increasing in the region. Although, their simple construction left them susceptible to damage during some recent rare flood events.

Local populations also benefit from:

- workshops on:
  - saving water
  - combating pollution and desertification more generally
  - preserving palm trees
  - rehabilitating existing but currently redundant foggaras
- being able to take their skills learnt on their own foggaras projects to other communities in the region and thus create a trickle-down effect.

Further reading

The texts below develop many of the ideas set out in this chapter. Some are quite academic and specialist texts. Your local library may be able to find copies should you wish to take your studies further.

Middleton, N. (2009) *Deserts: A Very Short Introduction*. Oxford University Press

Hill, M. (2002) *Access to Geography: Arid and Semi-arid Environments*. Hodder Education

Harvey, A. (2012) *Introducing Geomorphology: A Guide to Landforms and Processes*. Dunedin

Grotzinger, J. & Jordan, T. (2014) *Understanding Earth*. Freeman

Laity, J. (2008) *Deserts and Desert Environments*. Wiley-Blackwell

Thomas, D.S.G. (2011) *Arid Zone Geomorphology: Process, Form and Change in Drylands*. Wiley-Blackwell

Goudie, A.S. (2013) *Arid and Semi-arid Geomorphology*. Cambridge University Press

Granger, A. (2015) *The Threatening Desert: Controlling Desertification*. Routledge

Williams, M. (2014) *Climate Change in Deserts: Past, Present and Future*. Cambridge University Press

More information about foggaras may be found by completing an internet search for 'Algeria foggaras' or visiting www.unesco.org/mab/doc/ekocd/algeria.html

Iosifides, T. & Politidis, T. (2005), 'Conducting Qualitative Research on Desertification in Western Lesvos, Greece' In: *The Qualitative Report*, Volume 10, Number 1, March 2005, pp. 143-62. www.nova.edu/ssss/QR/QR10-1/iosifides.pdf

Question practice

A-level questions

1 Summarise characteristics of sources of water in hot deserts. [4 marks]

2 Ethiopia is a country in East Africa that is already experiencing desertification. This is expected to continue during the twenty-first century due to climate change. Analyse the data shown in Figure 2.53 and Figure 2.54. [6 marks]

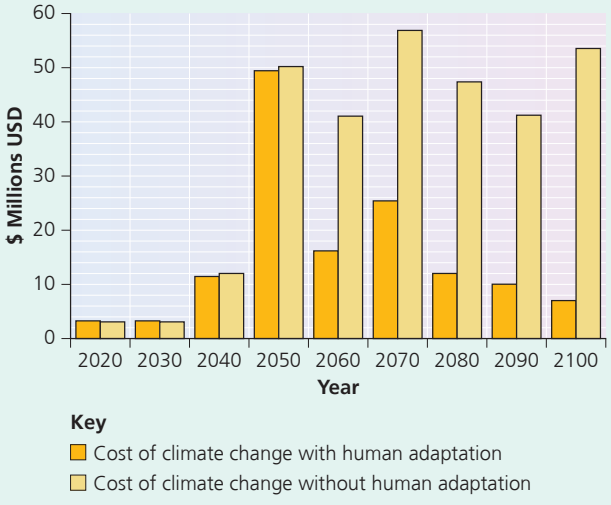


Figure 2.53 Possible predicted costs of climate change in Ethiopia – with and without human adaptation to the changes

Source: Engineering Project Organization Journal

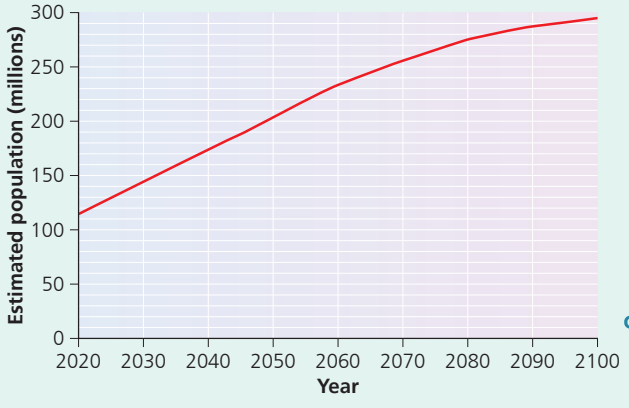


Figure 2.54 Estimated predicted population change in Ethiopia

Source: United Nations

3 Using Figure 2.55 and your own knowledge, assess the extent to which vegetation living in hot deserts has adapted to allow it to survive. [6 marks]

Figure 2.55 Vegetation growing in the hot desert of Saguaro National Park in Arizona

4 'The causes of aridity vary for hot deserts in different parts of the world.' To what extent do you agree with this view? [20 marks]

5 'Human-induced changes to the carbon cycle during the twenty-first century will only have negative impacts on the lives of people.'

To what extent do you agree with this view, with reference to a local scale desert landscape you have studied? [20 marks]

Chapter 3

Coastal systems and landscapes

Coastal zones are dynamic environments with distinctive landscapes formed by the interaction of a range of wind, marine and terrestrial processes. Coastal environments are important to the human race. About half of the world's population live on coastal plains with over 50 per cent of the population living within 150 km of the sea.

In this chapter you will study:

- coasts as natural systems
- the systems and processes leading to the development of coastal landscapes
- the management of coastal landscapes, including coastal flooding and erosion
- sustainable approaches to managing coastal systems in the future.

3.1 Introduction to coastal systems and landscapes

Coasts as natural systems

Links

The introduction to Chapter 1, Water and carbon cycles, outlined how many aspects of physical geography can be studied using a systems approach and you should read that section carefully before reading further here.

This chapter explores coastal environments, which are **open systems** with a range of **inputs**, **components/stores**, **flows/transfers** and **outputs** that combine to create distinctive landscapes. When there is a balance between the inputs and outputs then the system is said to be in a state of **dynamic equilibrium**. If one of the elements of the system changes, for example, there is increased deposition on a beach but there is no corresponding change in the amount of sediment removed from the beach, then the beach features may change and the equilibrium is upset. This is called **feedback**. Figure 3.1a illustrates the open system of coastal landscapes and Figure 3.1b is an example of how a negative feedback mechanism on the coastal zone acts to stabilise coastal morphology and maintain a dynamic equilibrium. The figures also introduce many ideas that are explored in detail in the rest of the chapter.

Coastal landscapes, like all geomorphological environments, consist of a constantly changing assemblage of erosional and depositional landforms; they are the result of continuous change in the elements (or components) of the coastal system. As this system tends towards **dynamic equilibrium**, driven by the input of energy, the processes operating in coastal systems continually shape the coastal landforms, creating their distinctive landscape features. However, the coastline itself consists of a series of different zones where specific

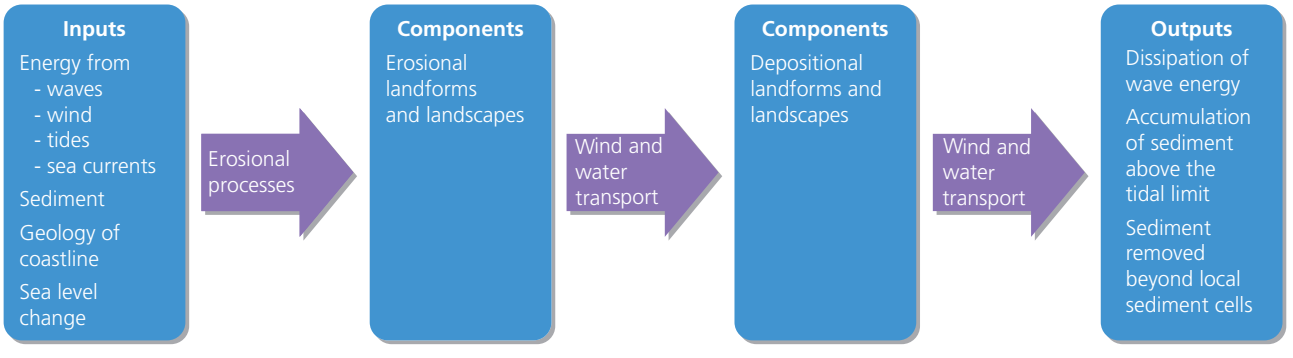


Figure 3.1a Coastal landscapes as open systems

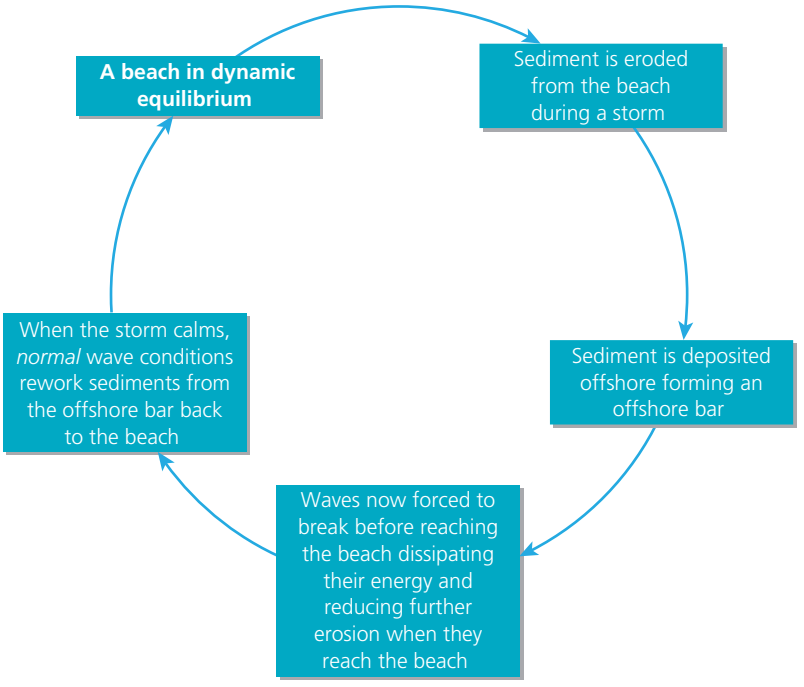


Figure 3.1b A negative feedback mechanism in a coastal environment

conditions prevail that depend on factors such as tides, wave action and the depth of the sea. Figure 3.2 illustrates how these zones relate to each other.

- **Backshore** is the area between the high water mark (HWM) and the landward limit of marine activity. Changes normally take place here only during storm activity.
- **Foreshore** is the area lying between the HWM and the low water mark (LWM). It is the most important zone for marine processes in times that are not influenced by storm activity.
- **Inshore** is the area between the LWM and the point where waves cease to have any influence on the land beneath them.
- **Offshore** is the area beyond the point where waves cease to impact upon the seabed and in which activity is limited to deposition of sediments.

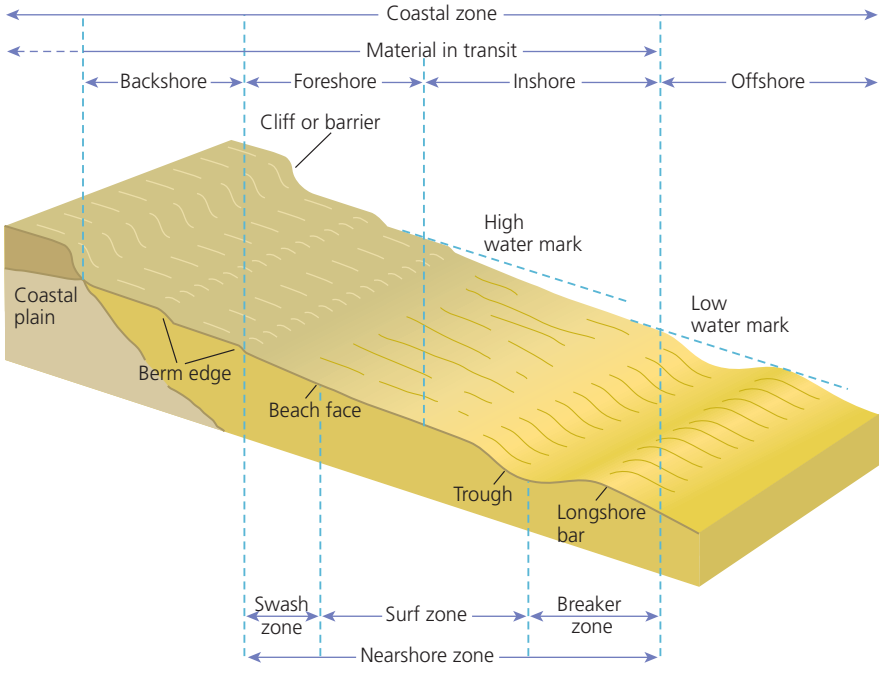


Figure 3.2 Coastal zones

- **Nearshore** is the area extending seaward from the HWM to the area where waves begin to break. It includes the following:
  - **Swash zone:** the area where a turbulent layer of water washes up the beach following the breaking of a wave.
  - **Surf zone:** the area between the point where waves break, forming a foamy, bubbly surface, and where the waves then move up the beach as **swash** in the swash zone.
  - **Breaker zone:** the area where waves approaching the coastline begin to break, usually where the water depth is 5 to 10 m.

Fieldwork opportunities

There is an opportunity here to visit a beach and identify these features for yourself.

Review questions

1 Outline an example of positive feedback in a coastal system.

2 Can you think of other examples of feedback in coastal systems?

3.2 Systems and processes in coastal environments

This book highlights how most aspects of physical geography can be thought of as natural systems with inputs, stores/components, flows/transfers and outputs; with geomorphological processes, driven by the input of energy, constantly changing the elements of these systems as they tend towards dynamic equilibrium. Coastal environments are no different and before exploring the detail of the processes that shape the landforms that combine to create characteristic coastal landscapes, a summary of a simple coastal system is included below.

- Inputs:
  - energy from waves, winds, tides and sea currents
  - sediment
  - geology of the coastline
  - sea level change
- Components:
  - characteristic erosional and depositional coastal landforms
- Outputs:
  - dissipation of wave energy
  - accumulation of sediment above the tidal limit
  - sediment removed beyond local sediment cells.

You should also read the relevant material at the beginning of Chapter 1 (page 2) under the heading 'Systems frameworks and their application'.

Many of the characteristic features of coastal environments are created by the action of wind, waves, tides and sea currents. To fully explain how these processes work it is important to understand where the energy comes from to drive the various **weathering**, **erosion** and **mass movement** processes.

Key terms

Erosion – The wearing away of the Earth's surface by the mechanical action of processes of glaciers, wind, rivers, marine waves and wind.

Fetch – Refers to the distance of open water over which a wind blows uninterrupted by major land obstacles. The length of fetch helps to determine the magnitude (size) and energy of the waves reaching the coast.

Mass movement – The movement of material downhill under the influence of gravity, but may also be assisted by rainfall.

Weathering – The breakdown and/or decay of rock at or near the Earth's surface creating regolith that remains *in situ* until it is moved by later erosional processes. Weathering can be mechanical, biological/organic or chemical.

Sources of energy in coastal environments

The energy to drive the coastal system is provided by waves, wind, tides and currents.

Wind

Wind is a vital input into the coastal system as it is a primary source of energy for other processes, but is also an important agent of erosion and transport itself. Features of wind as an input into the coastal system:

- Spatial variations in energy result from variations in the strength and duration of the wind. Where wind speeds are persistently high and uninterrupted, wave energy is likely to be higher. Although local weather patterns may influence short-term changes in wind speed and direction, most coastlines will have a **prevailing wind direction**. That is, the wind will generally reach the coast from one direction. This is important as it is one factor that controls the direction that waves approach

the coastline, and also the direction of the transport of material in the coastal zone.

- **Fetch** refers to the distance of open water over which a wind blows uninterrupted by major land obstacles. The length of fetch helps to determine the magnitude (size) and energy of the waves reaching the coast.
- Wind plays a vital role in wave formation. Waves are created by the transfer of energy from the wind blowing over the sea surface (referred to as the 'frictional drag' of the wind). The energy acquired by waves depends upon the strength of the wind, the length of time it is blowing and the fetch.
- Wind acts as an agent of erosion as it can firstly pick up and remove sediment from the coast (like sand from a beach) and use it to then erode other features. The most common type of wind erosion is **abrasion**, where the wind uses the material it carries to wear away landscape features. (These processes occur in hot deserts and are covered in detail in Chapter 2.) As wind is able to pick up and move material it is an important agent of moving sediment along the coast, inland from the shoreline and beyond the tidal zone (see below).

Waves

Once created and driven by the wind, waves are the primary agent of shaping the coast. Figure 3.3 illustrates some of the key features of waves. Wave characteristics include:

- **wave height** or **amplitude:** this is the height difference between a wave crest and the neighbouring trough
- **wavelength:** this is the distance between successive crests
- **wave period:** this is the time for one wave to travel the distance of one wavelength, or the time between one crest and the following crest passing a fixed point.

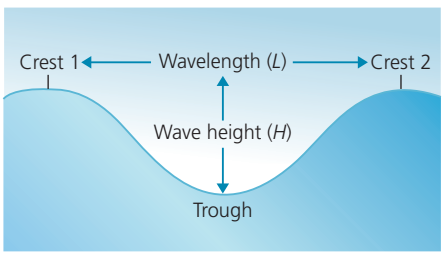


Figure 3.3 Wave terminology

As waves approach shallow water, friction with the seabed increases and the base of the wave begins to slow down. This has the effect of increasing the height and steepness of the wave until the upper part plunges forward and the wave 'breaks' onto the shore

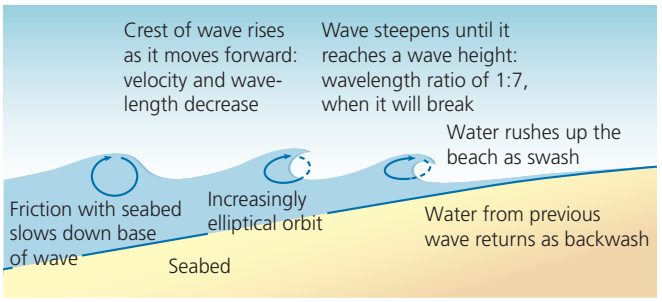


Figure 3.4 Wave movement
Source: Waugh, D. An Integrated Approach

(Figure 3.4). The rush of water up the beach is known as **swash** and any water running back down the beach towards the sea is the **backwash**. Waves can be described as **constructive** or **destructive waves**.

Key terms

Backwash – The action of water receding back down the beach towards the sea.

Constructive waves – Waves with a low wave height, but with a long wavelength and low frequency of around 6–8/min. Their swash tends to be more powerful than their backwash and as a consequence beach material is built up.

Destructive waves – Waves with a high wave height with a steep form and high frequency (10–14/min). Their backwash is generally stronger than their swash, so more sediment is removed than is added.

Swash – The rush of water up the beach after a wave breaks.

Constructive waves

Constructive waves (Figure 3.5) tend to have low wave height, but with a long wavelength, often up to 100 m. They have a low frequency of around 6–8/min. As they approach the beach, the wave front steepens only slowly, giving a gentle spill onto the beach surface. Swash rapidly loses volume and energy as water percolates through the beach material. This tends to give a very weak backwash which has insufficient force to pull sediment off the beach or to impede swash from the next wave. As a consequence, material is slowly, but constantly, moved up the beach, leading to the formation of ridges (**berms**).

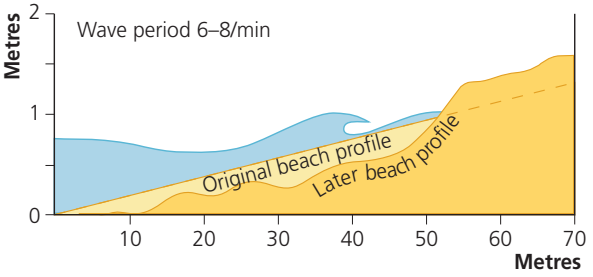


Figure 3.5 A constructive wave

Destructive waves

Destructive waves (Figure 3.6) have high wave height with a steep form and high frequency (10–14/min). As they approach the beach, they rapidly steepen and, when breaking, they plunge down. This creates a powerful backwash as there is little forward movement of water. It also inhibits the swash from the next wave. Very little material is moved up the beach, leaving the backwash to pull material back down the beach. Destructive waves are commonly associated with steeper beach profiles. The force of each wave may project some shingle well towards the rear of the beach where it forms a large ridge known as the **storm beach**.

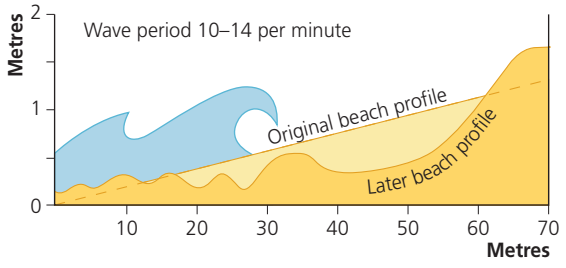


Figure 3.6 A destructive wave

Most beaches are subject to an alternating cycle of constructive and destructive waves (Figures 3.7a and 3.7b). Constructive waves build up the beach resulting in a steeper beach profile. This encourages waves to become more destructive (as destructive waves are associated with steeper profiles). With time, though, destructive waves move material back towards the sea, reducing the beach angle and encouraging more constructive waves. So the pattern repeats itself. This is another example of **negative feedback** that should maintain the state of **dynamic equilibrium**, but this is often impossible as other factors, such as wind strength and direction, are not constant.

Fieldwork opportunities

There is an opportunity here to visit a local beach and identify the kinds of waves present there.

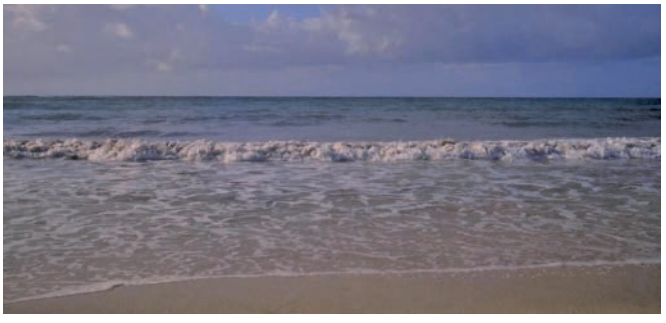


Figure 3.7a A constructive wave

Figure 3.7b A destructive wave

Wave refraction

Another crucial factor in determining the effects of wave action is the topography of the coastline. When waves approach a coastline that is not a regular shape, they are **refracted** and become increasingly parallel to the coastline. Figure 3.8 shows a headland separating two bays and illustrates the process of **wave refraction**. As each wave approaches the coast, it tends to drag in the shallow water which meets the headland. This increases the wave height and wave steepness and shortens the wavelength. That part of the wave in deeper water moves forward faster, causing the wave to bend. The overall effect is that the wave energy becomes concentrated on the headland, causing greater erosion. The low-energy waves spill into the bay, resulting in beach deposition. As the waves pile against the headland, there may be a slight local rise in sea level that results in a **longshore current** from the headland, moving some of the eroded material towards the bays and contributing to the build-up of the beaches.

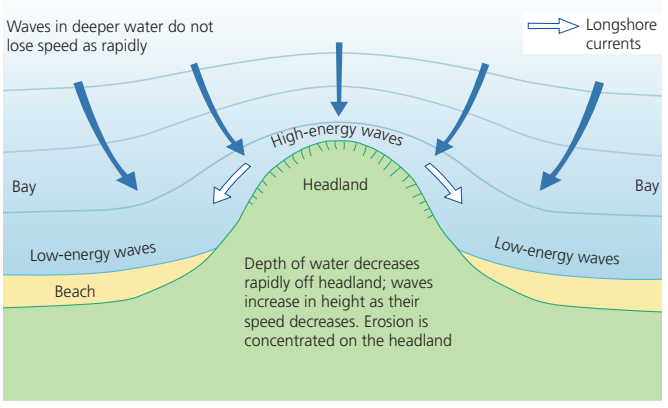


Figure 3.8 Wave refraction

Currents

The term **current** refers to the permanent or seasonal movement of surface water in the seas and oceans. There are three currents covered here:

- **Longshore currents** (sometimes known as **littoral drift**) occur as most waves do not hit the coastline 'head on' but approach at an angle to the shoreline. This generates a flow of water (current) running parallel to the shoreline. This not only moves water along the surf zone but also transports sediment parallel to the shoreline.
- **Rip currents** are strong currents moving away from the shoreline. They develop when seawater is *piled up* along the coastline by incoming waves. Initially the current may run parallel to the coast before flowing out through the breaker zone, possibly at a headland or where the coast changes direction. These can be extremely hazardous to swimmers and small boats.
- **Upwelling** is the movement of cold water from deep in the ocean towards the surface. The more dense cold water replaces the warmer surface water and creates nutrient-rich cold ocean currents. These currents form part of the pattern of global ocean circulation currents.

Key terms

Longshore or littoral drift – Where waves approach the shore at an angle and swash and backwash then transport material along the coast in the direction of the prevailing wind and waves.

Wave refraction – When waves approach a coastline that is not a regular shape, they are **refracted** and become increasingly parallel to the coastline. The overall effect is that the wave energy becomes concentrated on the headland, causing greater erosion. The low-energy waves spill into the bay, resulting in beach deposition.

Tides – The periodic rise and fall of the level of the sea in response to the gravitational pull of the Sun and Moon.

Tides

Tides are the periodic rise and fall in the level of the sea. They are caused by the gravitational pull of the Sun and Moon, although the Moon has much the greatest influence because it is nearer. The Moon pulls water towards it, creating a high tide, and there is a compensatory bulge on the opposite side of the Earth (Figure 3.9). In areas of the world between the two bulges, the tide is at its lowest.

As the Moon orbits the Earth, the high tides follow it. Twice in a lunar month, when the Moon, Sun and Earth are in a straight line, the tide-raising force is strongest. This produces the highest monthly tidal range or **spring tide**. Also twice a month, the Moon and Sun are positioned at 90° (perpendicular) to each other in relation to the Earth. This alignment gives the lowest monthly tidal range, or **neap tides**. At this time the high and low tides are between 10 to 30 per cent lower than the average.

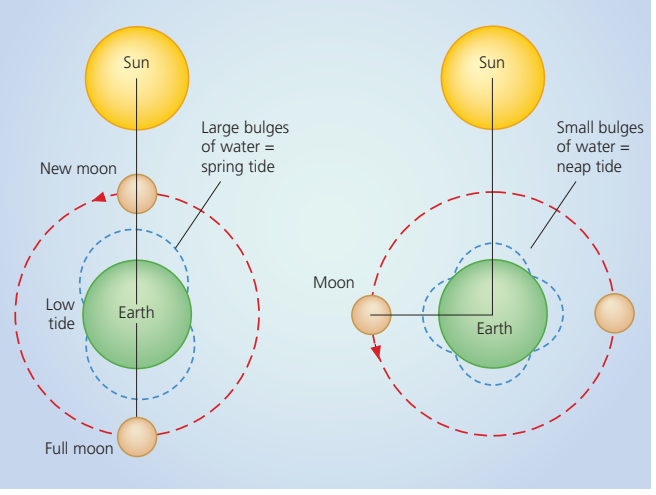


Figure 3.9 The causes of tides

The regular pattern of tides illustrated above is in fact significantly modified in individual locations by the morphology of the seabed, the proximity of land masses and the impact of the spinning forces of the Earth (**coriolis** force). Also, **tidal range** can be a significant factor in the development of a coastline. Tidal range is simply the difference in height of the sea water at high and low tide. This is not fixed due to the tide cycles outlined above. Tidal ranges determine the upper and lower limits of erosion and deposition and the amount of time each day that the littoral zone is exposed and open to **sub-aerial weathering**. Along the coasts of the Mediterranean Sea, tidal ranges are low. This restricts wave action to a narrow width in the coastal zone. In parts of the British Isles, however, tidal ranges are high. This gives a wide zone of wave attack, resulting in the formation of wide wave-cut platforms

in many places. Table 3.1 shows a simple classification of tidal range.

Table 3.1 Classification of tidal range

Macrotidal	More than 4 m
Mesotidal	2 to 4 m
Microtidal	Less than 2 m

Tidal surges

Tidal, or **storm surges** are occasions when meteorological conditions give rise to strong winds which can produce much higher water levels than those at high tide. One area affected by this phenomenon is the North Sea and east coast of Britain. Depressions (intense low pressure weather systems) over the North Sea produce low pressure conditions that have the effect of raising sea levels. The sea level can rise by about one centimetre for every one millibar drop in pressure. Strong winds drive waves ahead of the storm, pushing the sea water towards the coastline. This has the effect of *piling-up* water against the coast. The shape of the North Sea means that often water is increasingly concentrated into a space that is decreasing in size (funnelling). High tides, especially those of a spring tide, intensify the effect.

Famously the North Sea was affected by a tidal surge in 1953. However, the storms and associated surges of December 2013 and January 2014 brought some places along the east coast of England higher water levels than 60 years earlier. Figure 3.10 shows the size and power of the waves on Cromer pier and promenade during 5–6 December 2013.

Figure 3.10 Tidal surge at Cromer, Norfolk, December 2013

Skills focus

Visit the Met Office website (link below) to find out more about the winter storms of 2013 and study the synoptic charts in the document.
www.metoffice.gov.uk/weather/learn-about/past-uk-weather-events#y2013

Low energy and high energy coasts

Key terms

Coastal sediment budget – The balance between sediment being added to and removed from the coastal system, that system being defined within each individual sediment cell.

High energy coast – A coastline where strong, steady prevailing winds create high energy waves and the rate of erosion is greater than the rate of deposition.

Low energy coast – A coastline where wave energy is low and the rate of deposition often exceeds the rate of erosion of sediment.

Sediment cell – A distinct area of coastline separated from other areas by well-defined boundaries, such as headlands and stretches of deep water.

The processes outlined above combine to create **low energy coasts** and **high energy coasts** depending on local conditions. The following simple classification in Table 3.2 suggests the features of each.

Table 3.2 Features of low and high energy coasts

Low energy coasts	High energy coasts
• Coastlines where wave energy is low • The rate of deposition often exceeds the rate of erosion of sediment • Typical landforms include beaches and spits • Examples include many estuaries, inlets and sheltered bays. The Baltic Sea is one of the best examples due to its sheltered waters and low tidal range	• Coastlines where strong, steady prevailing winds create high energy waves • The rate of erosion is greater than the rate of deposition • Typical landforms include headlands, cliffs and wave-cut platforms • Examples of high energy coasts are the exposed Atlantic coasts of northern Europe and North America, including the north Cornish coast in south-west England

Sediment sources, cells and budgets

Alongside energy, sediment is also an important input in coastal systems. Coastal sediment comes from a variety of sources, including:

- streams or rivers flowing into the sea
- estuaries
- cliff erosion
- offshore sand banks
- material from a biological origin – including shells, coral fragments and skeletons of marine organisms.

The source of sediment which leads to the build-up of certain depositional features around the British coast is in dispute. Research has suggested that sediment movements occur in distinct areas or **sediment cells**, within which inputs and outputs are balanced. Along the coastline of England and Wales, 11 of these sediment (or littoral) cells have been identified (Figure 3.11).

Figure 3.11 The sediment cells around England and Wales

These are distinct areas of coastline separated from other areas by well-defined boundaries, such as headlands and stretches of deep water. In theory, sediment cells can be regarded as closed systems from which nothing is gained or lost. However, in reality it is easy for fine sediments to find their way around headlands and into neighbouring cells.

Sediment cells vary in size. The larger ones are divided into smaller sections (**sub cells**) to allow closer study and management. An example of a sub-cell is the one that operates between Flamborough Head and the Humber Estuary on the east coast of England (Figure 3.12).

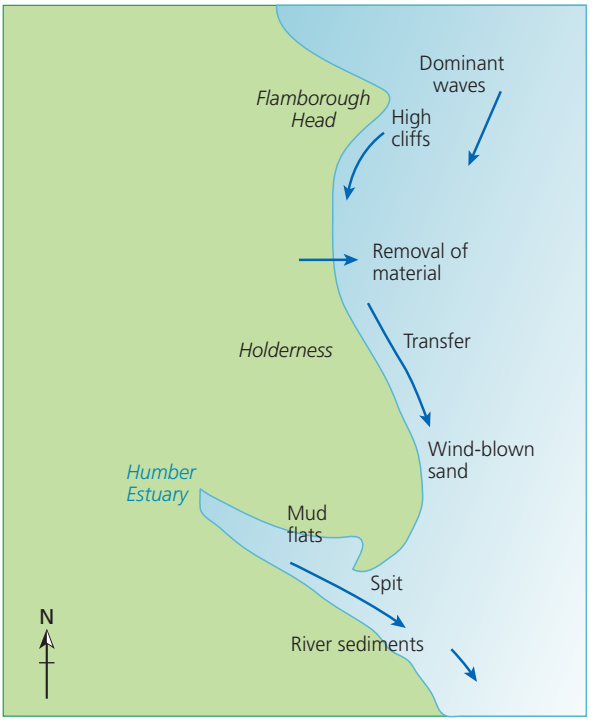


Figure 3.12 Detail of the Flamborough Head – Humber Estuary sub cell

Coastal sediment budgets

For a sandy beach or muddy shoreline to exist, sediment has had to come from somewhere and the combination of wave, current and tide action to be such that the material remains there. This introduces the concept of the **coastal sediment budget.** The coastal sediment budget is best thought of as being similar to a bank account, and is defined as the balance between sediment being added to and removed from the coastal system, that system being defined within each individual sediment cell. Figure 3.13 (page 98) illustrates the concept of the coastal sediment budget.

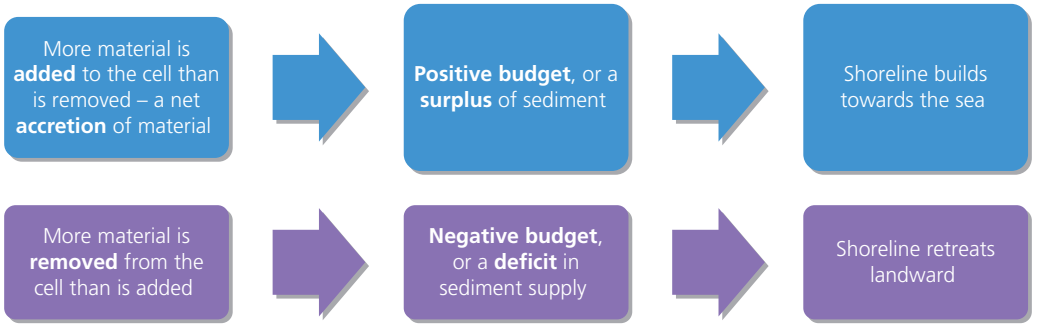


Figure 3.13 Coastal sediment budget

The processes of coastal erosion control the level of the deficit as they remove material from the shoreline, and if more sediment is removed than is added then the coastline will recede. The sediment budget can also be used to identify the sources that deliver sediment to the cell and to the sites where sediment is stored, or sediment sinks. Therefore, calculating the sediment budget for a cell requires the identification of all the sediment sources and sinks, and an estimation of the amount of sediment added and removed each year. Calculating budgets is extremely difficult and most efforts rely on complex models and estimations.

Coastal geomorphological processes

Key terms

Marine processes – Operate upon a coastline and are connected with the sea, such as waves, tides and longshore drift.

Sub-aerial processes – Includes processes that slowly (usually) break down the coastline, weaken the underlying rocks and allow sudden movements or erosion to happen more easily. Material is broken down *in situ*, remaining in or near its original position. These may affect the shape of the coastline, and include weathering, mass movement and run-off.

As coastal geomorphological systems develop over time and tend towards dynamic equilibrium a variety of processes take place. Driven by the input of energy these are what change the characteristics of the components of coastal landscapes.

Coastlines are affected by two sets of geomorphological processes:

- **Marine processes** are those that operate upon a coastline that are connected with the sea, such as waves, tides and longshore drift.
- **Sub-aerial processes** operate on the land but affect the shape of the coastline, such as weathering, mass movement and run-off.

Figure 3.14 introduces the main components of these geomorphological processes that shape the coastline.

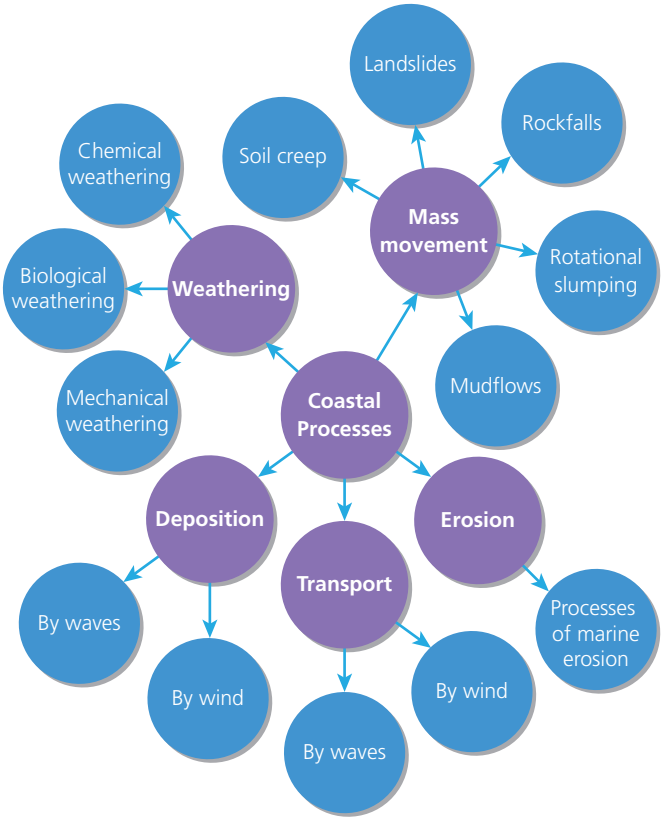


Figure 3.14 Coastal geomorphological processes

The interaction of these processes along the coastline is distinctive, creating the range of uniquely coastal landscapes.

Distinctively coastal processes

Processes of marine erosion

Waves often break on a coastline with considerable energy. It has been estimated that waves breaking against the foot of a cliff can generate energy of 25–30 tonnes m^{-2}. Coastlines are eroded in a number of ways:

- **Hydraulic action** refers to the impact on rocks of the sheer force of the water itself (without debris). This can exert enormous pressure upon a rock surface, thus weakening it. Such activity is sometimes referred to as **wave pounding**.
- **Wave quarrying:** A breaking wave traps air as it hits a cliff face. The force of water compresses this air into any gap in the rock face, creating enormous pressure within the fissure or joint. As the water pulls back, there is an explosive effect of the air under pressure being released. The overall effect of this over time is to weaken the cliff face. Storms may then remove large chunks of it. Some authors also use the term **cavitation** from the study of the effects of pressure changes in areas of rapid flow in rivers for this process.
- **Abrasion/corrasion:** The material the sea has picked up also wears away rock faces. Sand, shingle and boulders hurled against a cliff line will do enormous damage. This is also apparent on inter-tidal rock platforms where sediment is drawn back and forth, grinding away at the platform.
- **Attrition:** The rocks in the sea which carry out abrasion are slowly worn down into smaller and more rounded pieces.
- **Solution (corrosion):** Although this is a form of weathering rather than erosion, it is included here as it contributes to coastal erosion. The process of solution includes the dissolving of calcium-based rocks (for example, limestone). It is unlikely that seawater itself is the agent of this, as its pH is generally stable between 7.5 and 8.5, making it slightly alkaline. If the pH of seawater was generally below 7 then it would potentially kill much marine life. However, in localised areas where fresh water interacts with seawater, conditions for solution may occur, or carbon-based rocks at the coast may be broken down by water flowing from the land or rainwater which may be slightly acidic. Some geographers also include the effects of the evaporation of salts from water in the rocks to produce crystals. These expand when they form and put stress upon the rocks. Salt from seawater spray is capable of corroding several types of rock. This is most definitely a form of weathering and is expanded upon below.

The rate of coastal erosion is affected by a number of factors (Table 3.3):

Table 3.3 Factors affecting coastal erosion

Factor	Description
Wave steepness and breaking point	• Steeper waves are high energy waves and have greater erosive power than low energy waves • The point at which waves break is also important; waves that break at the foot of a cliff release more energy than those that break some distance from the shore
Fetch	• How far the wave has travelled determines how much energy has been generated within it
Sea depth	• A steeply-shelving seabed at the coast will create higher and steeper waves
Coastal configuration	• Headlands attract wave energy through refraction
Beach presence	• Beaches absorb wave energy and can therefore provide protection against marine erosion • Steep, narrow beaches easily dissipate the energy from flatter waves, while flatter, wide beaches spread out the incoming wave energy and are best at dissipating high and rapid energy inputs • Shingle beaches deal with steep waves as energy is rapidly dissipated through friction and percolation
Human activity	• People may remove protective materials from beaches (sand and shingle), which may lead to more erosion, or they may reduce erosion by the construction of sea defences • Sea defences in one place, however, may lead to increased rates of erosion elsewhere on the same coastline

There is one more especially important factor that determines the nature of the erosional processes that take place on a particular coastline: geology.

Geology

Lithology refers to the characteristics of rocks, especially resistance to erosion and permeability. Very resistant rocks, such as granite, and to a lesser extent chalk, tend to be eroded less than weaker materials such as clay. Some rocks, like limestone, are well-jointed, which means that the sea can penetrate along lines of weakness, making them more vulnerable to erosion. Variation in the rates at which rocks wear away is known as **differential erosion**.

The **structure** and variation of the rocks also affect erosion. When rocks lie parallel to the coast, they produce a very different type of coastline than when they lie at right angles (perpendicular) to the coast. Figure 3.15 (page 100) shows two contrasting types of coastline that are found close to one another in Purbeck (southern England).

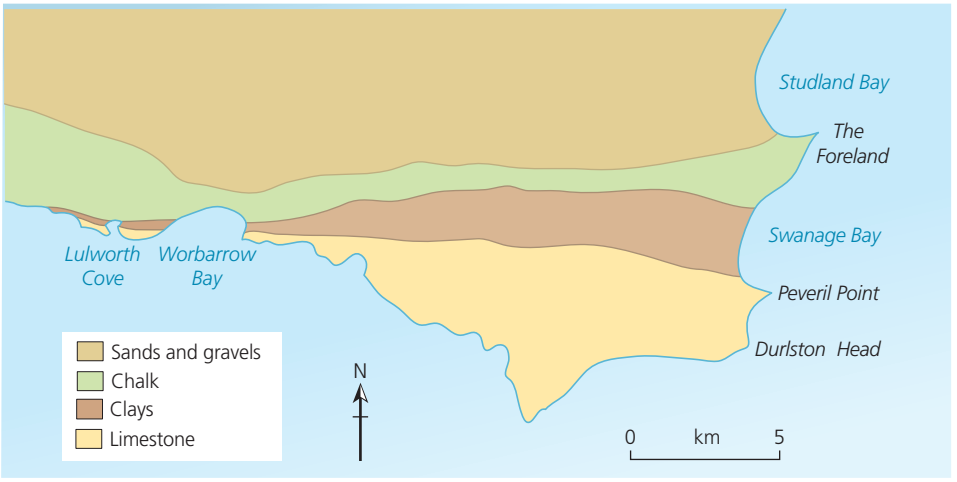


Figure 3.15 The Purbeck Coast

The southern part of the coast has the rocks running parallel to it – known as a **concordant** coastline. Here the resistant Portland limestone forms cliffs, and these have protected the coast from erosion, only allowing the sea to break through in a few places (the large area of Worbarrow Bay and the small area of Lulworth Cove) to the clay behind.

To the east, the rocks run at right angles to the coast (known as a **discordant** coastline), allowing the sea to penetrate along the weaker clays and gravels and produce large bays (for example, Swanage Bay) flanked by outstanding headlands (The Foreland and Peveril Point).

The **dip** of the rocks is also a major factor. The steepest cliffs tend to form in rocks that have horizontal strata or which dip gently inland, whereas rocks that dip towards the coast tend to produce much more gently sloping features (Figure 3.16).

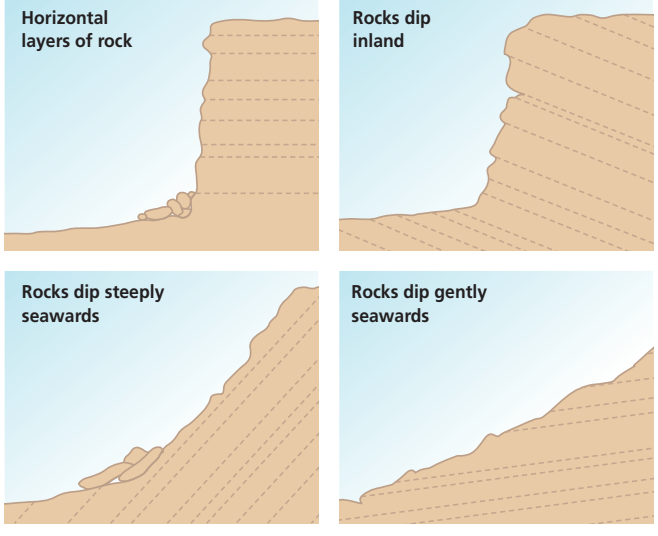


Figure 3.16 The influence of rock strata on coastlines

Processes of marine transportation

Wave and tidal energy that is not used for erosion or not lost through friction with the seabed can be used to transport material at the coast. The transportation of sediment by the sea is not only an important agent of change, in terms of processes of erosion and deposition, but also represents a significant flow/transfer of material.

Figure 3.17 illustrates that material is transported by seawater in a number of ways. These include:

- **Traction:** Large stones and boulders are rolled and slid along the seabed and beach by moving seawater. This happens in high energy environments.
- **Saltation:** Small stones bounce or leapfrog along the seabed and beach. This process is associated with relatively high energy conditions. Small particles may be thrust up from the seabed only to fall back to the bottom again. As these particles land they in turn dislodge other particles upwards, causing more such bouncing movements to take place.
- **Suspension:** Very small particles of sand and silt are carried along by the moving water. Such material is not only carried but is also picked up, mainly through the turbulence that exists in the water. Large amounts of suspended load, especially near estuaries, can cause a milky or murky appearance of the sea.
- **Solution:** Dissolved materials are transported within the mass of moving water.

As waves rarely approach the coastline head-on the processes of transport described above combine to move material along the coast.

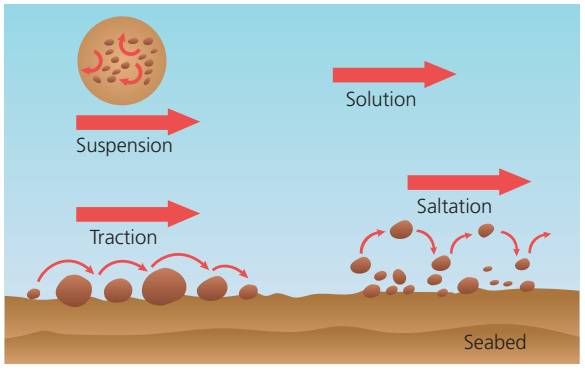


Figure 3.17 How material is transported by the sea

Longshore (or littoral) drift occurs when waves approach the shore at an angle and material is pushed up the beach by the **swash** in the same direction as the waves approach. As the water runs back down the beach, the **backwash** drags material down the steepest gradient, which is generally perpendicularly back down the beach, where it is picked up by the following incoming wave. Over a period of time, sediment moves in this zigzag fashion along the coast (Figure 3.18). If the material is carried some distance it will become smaller, more rounded and better sorted.

Obstacles such as groynes (wooden breakwaters) and piers interfere with this drift and accumulation of sediment occurs on the windward side of the groynes, leading to entrapment of beach material. Deposition of this material also takes place in sheltered locations, such as the head of a bay and where the coastline changes direction abruptly – here spits tend to develop.

Fieldwork opportunities

There is an opportunity here to visit a local beach and investigate how material is being transported along it.

Processes of marine and aeolian deposition

Marine deposition often takes place where the waves are low energy or where rapid coastal erosion provides an abundant supply of material. The sea lays down material when there is a reduction in energy resulting from a decrease in velocity or volume of water.

Situations when deposition occurs include:

- when sand and shingle accumulate faster than they are removed
- as waves slow following breaking
- as water pauses at the top of the swash before backwash begins
- when water percolates into the beach material as backwash takes it back down the beach.

Aeolian deposition

Aeolian processes refer to the entrainment, transport and deposition of sediment by wind. Wind plays an important role in shaping many coastlines and is an almost constant feature of most, not just because of the general pattern of prevailing winds that drives the waves. During the day, the wind on the coastal fringe is generally from the sea. Air moves in response to the small pressure differentials set up by the warmer land and colder sea. When there is a large tidal range, large amounts of sand may be exposed at low tide thus providing a supply of sediment to be entrained (picked up) by the wind.

Sand-sized sediment is the most significant in terms of depositional features at the coast. Once entrained, sand is generally transported close to the ground and over relatively short distances. Sand is transported by wind

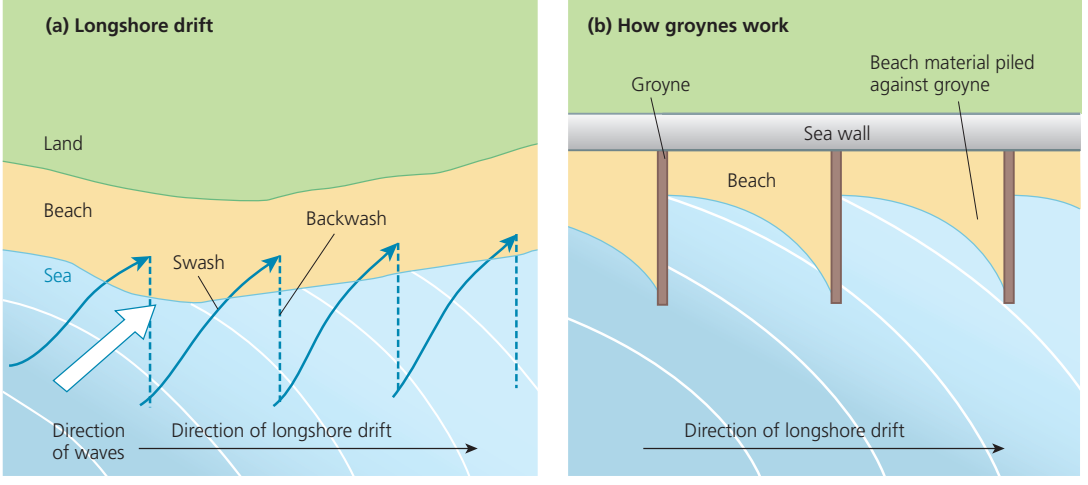


Figure 3.18 Longshore drift

in two ways, which are dependent on wind speed and how dry or moist the source of sand is.

- **Surface creep:** A process similar to traction, where wind rolls or slides sand grains along the surface.
- **Saltation:** Where the wind is strong enough to temporarily lift the grains into the airflow to heights of up to one metre for distances up to 20 to 30 m.

Wind action can shape and form a range of landforms at the coast, but the most obvious is a beach itself and perhaps the most distinctive are dunes, which are explored later in the chapter.

Sub-aerial processes

As mentioned above, as well as marine processes there are also **sub-aerial** (land-based) **processes** which shape the coastline. These come under the general headings of **weathering** and **mass movement**.

Sub-aerial weathering

Sub-aerial weathering includes processes that slowly (usually) **break down** the coastline, weaken the underlying rocks and allow sudden movements or erosion to happen more easily. Material is broken down *in situ*, remaining in or near its original position. Weathering processes are common at the coast due to the presence of air and water and cycles of wetting and drying, and can be categorised as:

- mechanical/physical weathering
- biological weathering
- chemical weathering.

Table 3.4 Categories of sub-aerial weathering

Factor	Description
Mechanical/ physical weathering	These processes depend on the nature of the climate. • **Freeze-thaw:** in latitudes where temperatures fluctuate above and below freezing, freeze-thaw action is common, especially as there is a ready supply of water. Water that enters cracks in the rocks freezes as temperatures remain below 0°C. As it freezes, the water expands by almost 10 per cent, meaning the ice occupies more space and so exerts pressure on the surrounding rock. As the process repeats and continues, the crack widens, and eventually pieces of rock break off. • **Pressure release:** where processes of erosion, weathering and mass movement remove overlying material, the rock beneath is said to experience pressure release. As the overlying mass is unloaded, mechanisms within the rock cause it to develop weaknesses, or cracks and joints, as it is allowed to expand. This makes the rock susceptible to other processes of erosion and weathering.
Biological weathering	These are processes that lead to the breakdown of rocks by the action of vegetation and coastal organisms. Biological weathering is quite active on coastlines. • Some marine organisms, such as the piddock (a shellfish), have specially adapted shells that enable them to drill into solid rock. They are particularly active in areas with chalk geology where they can produce a sponge-like rock pitted with holes. • Seaweed attaches itself to rocks and the action of the sea can be enough to cause swaying seaweed to prise away loose rocks from the seafloor. • Some organisms, e.g., algae, secrete chemicals capable of promoting solution. • Animals can weaken cliffs as they burrow or dig into them, such as rabbits or some cliff-nesting birds.
Chemical weathering	This occurs where rocks are exposed to air and moisture so chemical processes can break down the rocks. • **Solution** is the main chemical process and was included above (page 99) as it combines with erosion to produce many distinctive features. • **Oxidation** causes rocks to disintegrate when the oxygen dissolved in water reacts with some rock minerals, forming oxides and hydroxides. It especially affects ferrous, iron-rich rocks, and is evident by a brownish or yellowish staining of the rock surface. • **Hydration** makes rocks more susceptible to further chemical weathering as it involves the physical addition of water to minerals in the rock. This causes the rock to expand, creating stress, which can itself cause the rock to disintegrate. The process weakens the rock and can create cracks or widen joints, allowing further chemical weathering to occur. • **Hydrolysis** is where mildly acidic water reacts or combines with minerals in the rock to create clays and dissolvable salts; this itself degrades the rock, but both are likely to be weaker than the parent rock, thus making it more susceptible to further degradation. • **Carbonation** occurs where carbon dioxide (CO_2) dissolved in rainwater makes a weak carbonic acid (H_2CO_3). This reacts with the calcium carbonate ($CaCO_3$) in rocks like limestone and chalk to create calcium bicarbonate ($Ca(HCO_3)_2$) which then dissolves easily in water. Carbonation is more effective in locations with cooler temperatures as this increases the amount of carbon dioxide that is dissolved in the water. • **Acid rain:** as well as naturally occurring CO_2 there are increasing levels of other gases associated with industry and the burning of fossil fuels in the atmosphere that also react with rainwater making it mildly acidic. The presence of sulphur dioxide and nitric oxides can create rainwater with weak sulphuric and nitric acids. This acid rain then reacts with various minerals in different rocks, weakening or even dissolving them.

Fieldwork opportunities

Visit a local beach and investigate which are the dominant weathering processes and why.

Mass movement

Mass movement is common on coastlines, especially those that are steep. The nature of the mass movement experienced on a particular coastline is dependent on a number of factors:

- the level of cohesion within the sediment
- the height of the slope and slope angle
- grain size within the sediment
- temperature and level of saturation.

Mass movements are generally either rapid sudden failures of the slope or the effects of processes that develop over some time. Types of mass movement:

- **Landslides:** Occur on cliffs made from softer rocks or deposited material, which slip as a result of *failure* within it when lubricated, usually following heavy rainfall.
- **Rock falls:** These occur from cliffs undercut by the sea, or on slopes affected by mechanical weathering like frost action.
- **Mudflows:** Heavy rain can cause large quantities of fine material to flow downhill. Here the soil becomes saturated and if excess water cannot percolate deeper into the ground, surface layers become very fluid and flow downhill. The nature of the flow is dependent on the level of saturation, type of sediment and slope angle. On relatively gentle slopes the flows are often referred to as 'solifluction', creating lobe-like features towards the base of the slope.
- **Rotational slip, or slumping:** Where softer material overlies much more resistant materials, cliffs are subject to slumping. With excessive lubrication, whole sections of the cliff face may move downwards with a slide plane that is concave, producing a rotation movement. Slumps are a common feature of the British coast, particularly where glacial deposits form the coastal areas, for example, east Yorkshire and north Norfolk.

Figure 3.19 shows a typical rotational slump in an area where glacial deposits form cliffs on top of an impermeable layer.

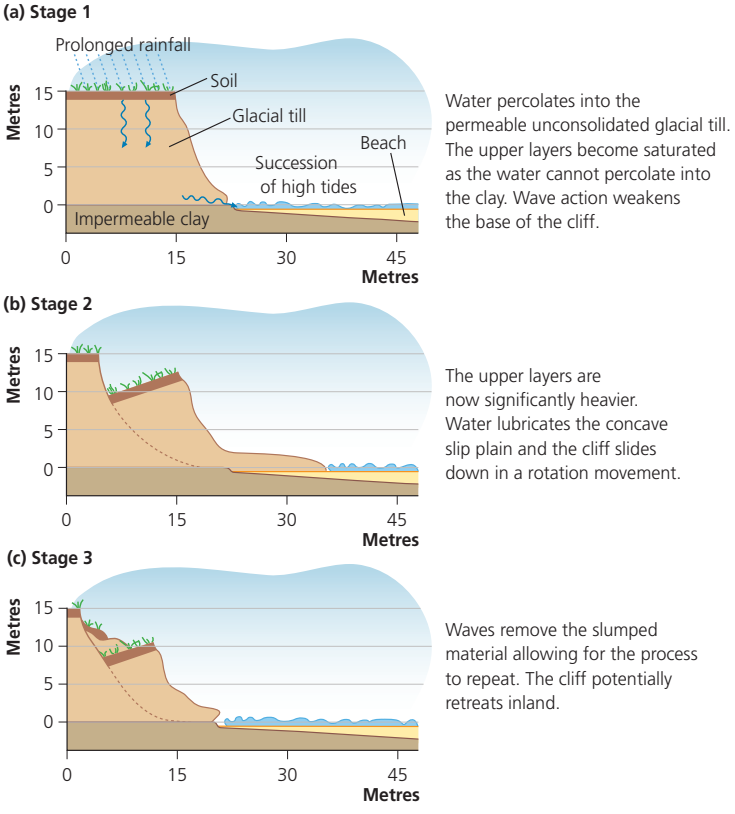


Figure 3.19 Rotational slumping

- **Soil creep:** This occurs where there is a very slow, almost imperceptible, but continuous movement of individual soil particles downslope. There is some uncertainty about the exact causes of creep, but most geographers agree that the presence of soil moisture is important, together with a range of weathering and other processes.

Run-off

Run-off is another important process that operates on coastlines. It may take the form of a stream emerging in a bay, taking with it large quantities of load during times of flood, or it may be a stream cascading over a cliff, excavating a V-shaped groove as it does so. It can also simply be water that flows over the surface to reach the coastline, the presence of which will also assist many of the mass movement processes above.

Review questions

3 Explain the important role played by tides in shaping coastal landforms.

4 Explain how the process of longshore drift contributes to the formation of different coastal landforms.

5 To what extent do you think the geology of the coastline is the most important factor in determining what coastal landforms will be found there?

3.3 Coastal landscape development

Characteristic coastal landscapes

As the coastline is simply the zone where the land meets the sea, there is not one single *typical* landscape type that typifies all coasts. Around the world individual coastal locations will have a coastal landscape with features that reflect the interaction of a range of factors, including, among other things:

- coastal geology and lithology
- climate
- nature of tides and waves.

However, it is possible to place coastlines in the following simple classification. They can be:

- concordant or discordant, or
- a cliffed coast, flat coast or graded shoreline, or
- an emergent or submergent coastline.

The characteristics of a particular coastal landscape will depend on whether it is:

- a high or low energy coast
- dominated by processes of erosion or deposition
- more or less intensely managed by people.

One thing is certain, and that is that coastal landscapes are not static and have many characteristic features that change over a range of timescales. Waves and tides continually shape and rework coastlines on a daily basis, yet as noted above, the nature of these waves and tides is not static. Over longer periods of time processes of erosion, transportation and deposition create landscapes that we become familiar with, which over human timescales we classify as characteristic of a particular location. For example, the cliffs of parts of the Yorkshire coast in Eastern England are seen as *characteristic* of that place, and the long sandy beaches and barrier reefs of Queensland, Australia are *characteristic* of that place. However, over even longer geological timescales coastal landscapes respond to more significant changes. For example, the Yorkshire cliffs mentioned above are composed of material deposited during recent glacial periods, before which that material would not have been there. Equally, the timescales followed by glacials and interglacials of ice ages create coastlines that have emerged from formerly higher sea levels, like the raised beaches of the Atlantic coast of Ireland, or others that have become submerged as sea levels have risen, like the flooded glacial valleys forming the fjords of Norway and New Zealand.

This section explores a range of landforms that are part of characteristic coastal landscapes from around the world, including some from the British Isles. It is important that you understand the diversity of coastal landscapes that exist globally and, accordingly, study examples from a wide range of places.

Skills focus

a For each landform explored here use the internet to search for another example in a contrasting part of the world.

b For each landform in this chapter, practise drawing an annotated sketch to describe its characteristics and explain its sequence of formation.

Origin and development of landforms of coastal erosion

The section above illustrated how the nature of the geology of the coastline can influence the landforms found at a coastline, and where this leads to differential rates of erosion, one of the most recognisable combinations of landforms is created: headlands and bays.

Headlands and bays

Figure 3.21 shows the impact of geology on a coastline. Areas with alternating more and less resistant rocks are a common feature of many coastlines. Initially erosional processes predominate in areas with less-resistant rock, forming bays, leaving the more resistant rock protruding out to sea as headlands. Because of refraction, the headlands then receive the highest-energy waves and are more vulnerable to the forces of erosion than the bays. The bays experience

low-energy waves that allow sediment to accumulate and form beaches. These then act to protect that part of the coastline.

Figure 3.20 Concordant coast of San Francisco Bay

When geology runs parallel to the coast, as in Purbeck, Dorset, or the San Francisco Bay area of the Californian coast, it is possible for marine processes to create headlands and bays. In San Francisco Bay (Figure 3.20), sea level rise over the last 10,000 years has slowly inundated a series of valleys running parallel to the coastline that correspond to the alignment of the complex geology of the Californian coast. Figure 3.21 illustrates how the differential erosion of alternating geology has also led to the formation of headlands and bays on the Purbeck coast.

Remember that it is important to think of landforms at places like Lulworth as being the result of systems operating at the coast, where the main inputs include:

- the geology and lithology of the coast
- the angle of the dip of the coastline in front of the headland
- the nature of the waves approaching the coast
- the direction and strength of the prevailing wind.

The components (or processes) include:

- the differential rates of erosion of the different rocks
- wave refraction
- erosion of the headland
- deposition in the bay.

The outputs include:

- the characteristic features of the resulting landscape including:
  - the headland and bay
  - the erosional features of the headland
  - the depositional features in the bay.

Cliffs and wave-cut platforms

When high and steep waves break at the foot of a cliff their energy and erosive action is concentrated into a small area of the rock face. With erosion concentrated at its base, the cliff begins to be undercut, forming a feature known as a **wave-cut notch**. Further erosion increases the stress on the cliff above and over time it will collapse; how long this takes will depend on the characteristics of the cliffs and waves.

The cliff line will begin to retreat and after successive collapses a gently sloping (less than 5°), relatively smooth, **wave-cut platform** is formed at the base of the cliff (Figure 3.22, page 106). If the platform is still in the tidal zone and continually exposed to cycles of marine erosion and sub-aerial processes, weaknesses

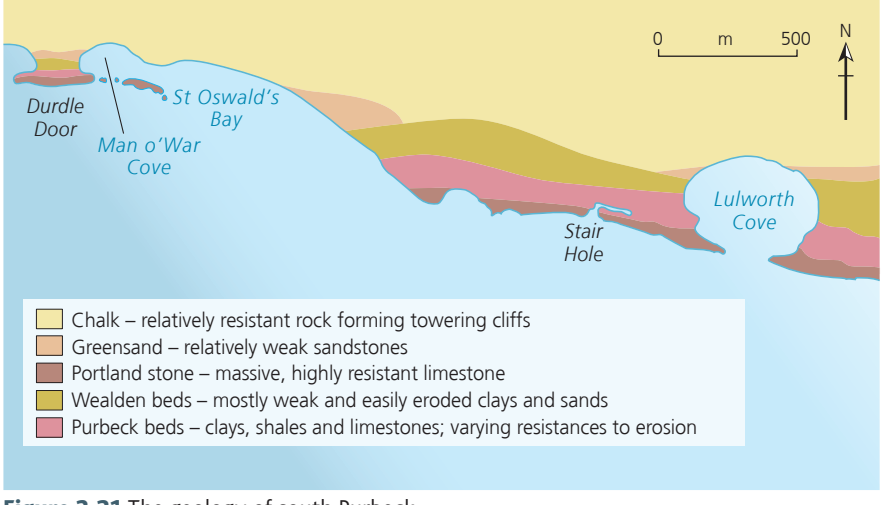


Figure 3.21 The geology of south Purbeck

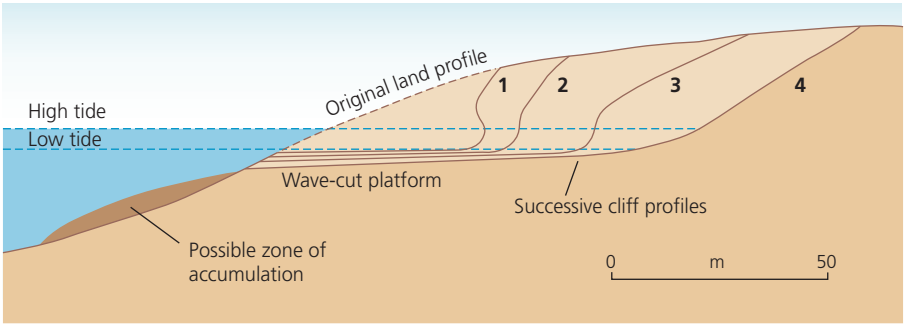


Figure 3.22 Formation of a wave-cut platform

in the rock surface may be exploited, and on closer inspection what appears to be a smooth rock surface may in fact be quite rough and jagged, with features like rock pools and fissures.

The platform continues to grow and, as it does, the waves break further out to sea and have to travel across more platform before reaching the cliff line. This leads to a greater dissipation of wave energy, reducing the rate of erosion on the cliff and slowing down the rate of growth of the platform. There tends, therefore, to be a limit to how big the feature can grow and some experts have suggested that growth beyond 0.5 km is unusual.

Key term

Wave-cut platform – A gently sloping (less than 5°), relatively smooth, marine platform caused by abrasion at the base of the cliff.

Geos, caves, blowholes, arches, stacks and stumps

These features are all independently observable on coastlines around the world, but they also represent a sequence of events in the erosion of a cliff or headland. On any cliff line it is the weakest parts, such as cracks, fissures, joints and bedding planes that are attacked by the sea. Along a joint the sea will cut inland, widening the crack to form a narrow, steep-sided inlet known as a **geo**.

Figure 3.23 Erosional features of Cape Dyrhólaey, southern Iceland

In other circumstances the cliff is undercut and a **cave** is formed, usually from a combination of marine processes. If erosion continues vertically upwards, it is possible for the cave to be extended to the top of the cliff where a **blowhole** will form. Much more likely, if the cave is on the side of a headland, is that it will extend backwards to meet another on the other side, due to refraction around the headland. Eventually the conjoining of the caves will create a hole all the way through the headland, known as an **arch**.

As the cliff recedes and the wave-cut platform develops, the arch will eventually collapse due to gravity, following a combination of marine erosion from below aided by sub-aerial processes weakening the arch. This leaves an isolated portion of rock as a **stack** standing above the platform. In time, the sea will exploit the wave-cut notch at the base of the stack, leading eventually to its collapse. A small portion of the wave-cut platform may be left marking the former position of the stack. This is known as a **stump**.

There are many famous examples of coastlines around the world where these features stand out. Parus Rock, or Sail Rock, in the Black Sea and the Twelve Apostles in Victoria, Australia are well-known stacks. The south coast of Iceland has some excellent examples of many of these features including the caves, arches and stumps associated with the impressive headland at Dyrhólaey, the most southerly point of the island (Figure 3.23).

On Purbeck, in Dorset, in the Portland stone (a highly resistant limestone) the sea has cut the famous arch of Durdle Door. Also on Purbeck the chalk escarpment culminates in The Foreland and its detached pieces that are known as Old Harry Rocks. Figure 3.24 is a sketch of this area, where the sequence from headland to stack can be seen. Marine erosion and

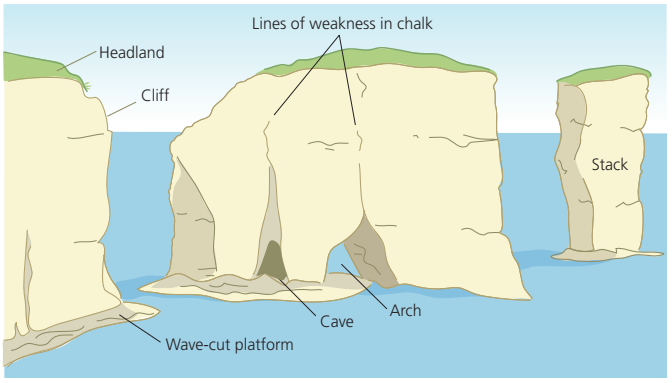


Figure 3.24 Coastal erosion features seen at Old Harry Rocks, Purbeck

sub-aerial processes will eventually reduce the upstanding parts of this area to a wave-cut platform.

Skills focus

This chapter has already highlighted the importance of thinking about coastal features using a systems approach. Look back at the section on headlands and bays. Now use a systems approach to help you understand the formation of the other landforms of coastal erosion. Make sure you are as specific as possible and include details from located examples you have studied. Landforms include:

- cliffs and wave-cut platforms
- geos, caves, blowholes, arches, stacks and stumps.

Origin and development of landforms of coastal deposition

Landforms of coastal deposition occur on coastlines where sand and shingle accumulate faster than they are removed. It often takes place where the waves are low energy or where rapid coastal erosion provides an abundant supply of material.

Beach characteristics

Beaches are found at the point where the land meets the sea and represent the accumulation of sediment deposited between low spring tides and the highest point reached by storm waves. They are mainly composed of sand and shingle. Sand tends to produce beaches with a more gentle gradient (usually below 5°) because its small particle size means the sand becomes compact when wet, and allows very little percolation. Most of the swash therefore returns as backwash, little energy is lost to friction, and material is carried down the beach. This leads to the development of **ridges** and **runnels** in the sand at the low-water mark. These run parallel to the shoreline and are broken by channels that drain the water off the beach (Figure 3.25).

Shingle may make up the whole or just the upper parts of the beach. The larger the size of the material, generally the steeper is the gradient of the beach (around 10–20°). This is because water rapidly percolates through shingle, so the backwash is somewhat limited in its ability to transport material back down the beach. This, together with the uneven surface, means that very little material is eroded from the beach.

At the back of the beach, strong swash at spring high tide level will create a **storm beach**. This is a ridge composed of the biggest boulders thrown by the largest waves, above the usual high tide mark. Below this will be a series of ridges marking the successively lower high tides as the cycle goes from spring to neap. These beach ridges are known as **berms** and are built up by constructive waves. **Cusps** are semi-circular-shaped depressions which form when waves break directly on to the beach and swash and backwash are strong. They usually occur at the junction of the shingle and sandy

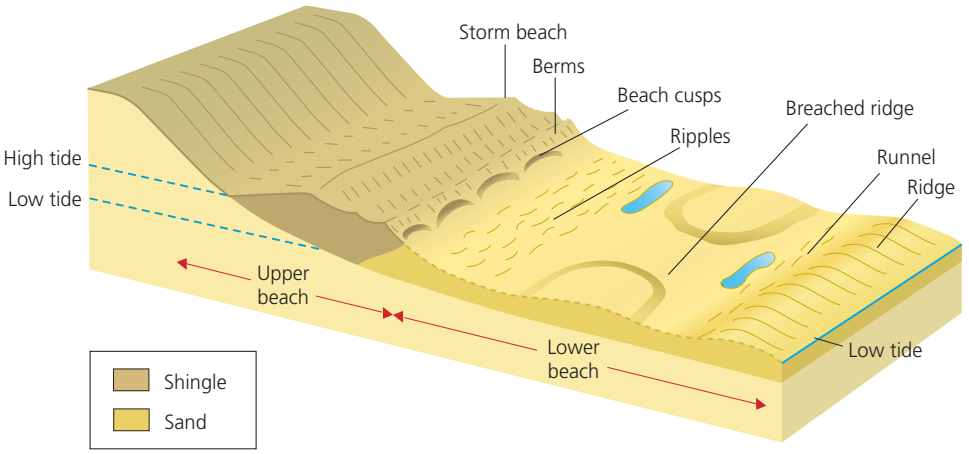


Figure 3.25 Beach features

beaches. The sides of the cusps channel incoming swash into the centre of the embayment and this produces a stronger backwash in the central area which drags material down the beach, deepening the cusp. Below this, **ripples** are developed on the sand by wave action or tidal currents.

The angle at which waves generally approach a coastline will determine the nature of the kinds of beach features that develop. Some authors use the terms **swash-** and **drift-aligned beaches** as a broad classification of beach types to reflect the nature of equilibrium between erosion and deposition a beach tends towards.

Swash-aligned beaches:

- are generally oriented parallel to the incoming wave crests
- experience minimal longshore drift
- can be found on irregular coastlines where longshore drift is impeded, and waves hit sections of the coast head-on.

Drift-aligned beaches:

- are generally oriented parallel to the direction of dominant longshore drift
- can have considerable amounts of sediment transported long distances along them
- initially develop where a section of coastline is fairly regular, or where the predominant wave direction is at an angle to the beach
- can extend out from the coastline if there is a sudden change in the direction of the coastline, for example, upon reaching an estuary.

The idea of equilibrium is especially significant for drift-aligned beaches as without a continual supply of sediment, the longshore drift would remove the sediment faster than it was deposited.

Fieldwork opportunities

There is an opportunity here to visit a local beach and investigate the characteristics of sediment in different locations on the beach.

Spits, tombolos, bars and barrier-beaches

A **spit** is an elongated, narrow ridge of land that has one end joined to the mainland and projects out into the sea or across an estuary, usually on a drift-aligned coast. Like other depositional features, it is composed of sand and/or shingle and the mixture is very much dependent upon the availability of material and the wave energy required to move it.

Figure 3.26 shows the formation of a spit.

1. The prevailing winds and maximum fetch are from the southwest, so material will be carried from west to east along the coast by the process of longshore drift.
2. Where the coastline changes to a more north–south orientation, there is a build-up of sand and shingle in the more sheltered water in the lee of the headland. As this material begins to project eastwards, storms build up more material above the high-water mark, giving a greater degree of permanence to the feature.
3. Finer material is carried further eastward into the deeper water of the estuary, and as the water loses its capacity to transport it further, is deposited, extending the ridge (spit) into the estuary.
4. The end of the spit begins to curve round as wave refraction carries material round into the more sheltered water. The second most dominant wind direction and fetch may contribute to this, pushing the spit material back towards the mainland. The spit cannot grow all the way across the estuary as the material is carried seaward by the river and the deeper water at the centre inhibits growth.

Figure 3.26 The formation of a spit

Figure 3.27a Spurn Head in East Yorkshire, illustrating some of the features of a simple spit

Figure 3.27b Cape Cod, Massachusetts, USA – a compound spit

A simple classification of spits includes **simple** or **compound spits**. Figure 3.27a illustrates the characteristics of a simple spit. Simple spits:

- are either straight or recurved
- do not have minor spits, or recurved ridges, along their landward edge.

Figure 3.27b is an example of a compound spit. Compound spits:

- may have similar features to simple spits
- have a number of recurved ridges, or minor spits, along their landward side, possibly marking the position where they terminated in the past.

As spits mature, **sand dunes** can also develop as deposited sand dries out and is blown to the landward side of the spit, where it can accumulate and become stabilised by vegetation as species like marram grass get established. Also, as the spits increase in size, an increasingly large, more sheltered area develops between the land and the spit. Low-energy, gentle waves enter this area and deposit finer material such as silt and clay. These deposits build-up and are colonised by vegetation to become **salt marshes.**

Some examples of spits have been mentioned above but other well-known spits around the British coast are found at Borth (west Wales), Dawlish Warren (Devon), Orford Ness (East Anglia) and Blakeney Point (Norfolk). Other famous spits from around the world include Farewell Spit, New Zealand and Homer Spit in Alaska.

A spit that joins an island to the mainland is known as a **tombolo** (Figure 3.28a). The best example in Britain is Chesil Beach on the south coast of England. This links the Isle of Portland to the mainland and is about 30 km long. One of the most beautiful and often photographed tombolos from around the world has to be the Angel Road of Shodo Island, Japan.

If a spit develops across a bay where there is no strong flow of water from the landward side, it is possible for the sediment to reach across to the other side. In this case, the feature is known as a **bar** (Figure 3.28b).

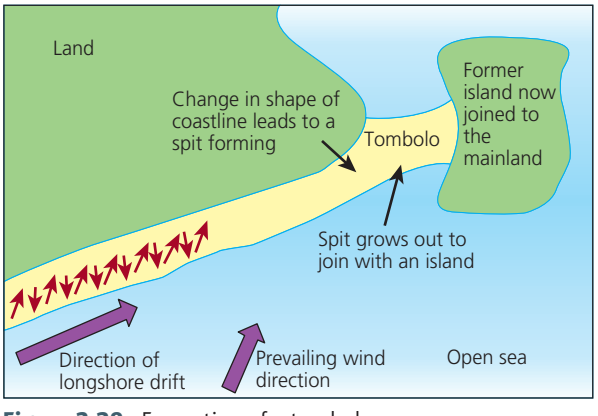


Figure 3.28a Formation of a tombolo

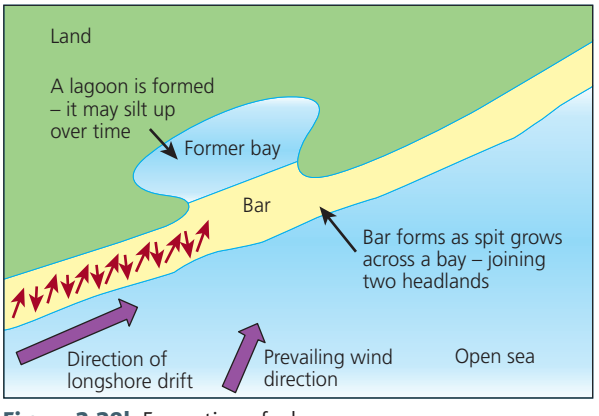


Figure 3.28b Formation of a bar

Some bars, however, may simply be the result of the onshore migration of material from offshore as sea levels rose following the last ice age. Like beaches, these bars may be completely or partially submerged by incoming tides. Slapton Ley, a bar found in Devon, is believed to have formed in this way. Some research suggests that it is a combination of this onshore migration and longshore drift that also formed Chesil Beach. The longest spit in the world, Arabat Spit in the Sea of Azov, is a bar, and joins towns in Ukraine to the north and Crimea to the south.

Where a ridge of beach material that remains semi-submerged accumulates seaward of the breaker zone then it is known as an **offshore bar**.

A **barrier beach** or **barrier island** is an elongated bank of deposited sand or shingle lying parallel to the coastline and not submerged by incoming tides. Where the bank is high enough to allow sand dunes to develop it is known as a barrier island. Often the sheltered area between the barrier beach and the land becomes a lagoon or coastal marsh, or in more tropical locations mangrove swamps may develop. One of the most famous barrier beaches is the Lido of Venice. They may be formed by the breaching of a spit or by constructive waves pushing a bar towards the land.

Skills focus

In the previous section you used a systems approach to help understand the formation of the various landforms of coastal erosion. Now do the same for the depositional landforms. These include beaches, spits, tombolos, bars and barrier beaches.

Process, time, landforms and landscapes in coastal settings

It is important to view each unique coastal landscape as an assemblage of features and landforms that combine in that place to give it its own characteristic landscape. In order to understand why that landscape looks the way it does today, it is crucial to look beyond the processes that are happening in the present and think about how different processes may have shaped its features through time.

There is certainly not one combination of landforms that creates one distinct coastal landscape. Where the sea and land interact there is a huge diversity of characteristic coastal landscapes including:

- bays
- estuaries
- beaches
- deltas
- dunes
- mud flats and salt marshes.

The list could go on and more landforms are explored below, but it is also important to recognise that with the large and ever-changing amounts of energy available from wind-driven waves, tides and currents, coastal landscapes are dynamic and constantly changing.

The characteristic coastal landscapes that we see today are a result of the factors that shaped them during the Holocene (roughly the time since the end of the last glacial period some 12,000 years ago). Therefore a range of factors has been important in producing the present coastal landscape features, and continues to do so today. These include:

- local tectonic processes
- sea level change – global and local
- climatic change – natural and that enhanced by human activity
- changing ocean currents and wave regimes
- natural disasters or events – including, for example, storms or tsunamis
- changing sources, types and amounts of sediment
- the changing nature of human activity.

All of the above, and more, have continually changed over the millennia, but all will have left their mark as features in the coastal landscapes of the present day, where contemporary coastal processes will inevitably continue to alter and modify them further.

Coastal sand dunes

Coastal sand dunes are accumulations of sand shaped into mounds by the wind. They represent a dynamic landform. Like other coastal landforms sand dunes can be studied using a systems approach. Important inputs include:

- a plentiful supply of sand
- strong onshore winds
- a large tidal range

- an obstacle to trap the sand
- vegetation growth to encourage further growth of the dune.

Sand is mostly moved inland by the process of saltation. Due to the differential heating of the land and the sea, localised differences in atmospheric pressure mean that during the day, the wind on the coastal fringe is generally from the sea. Air moves towards the area of slightly lower atmospheric pressure generated over the land as air is warmed by and rises over the warmer land, from slightly higher pressure over the colder sea. Where there is a large tidal range, large amounts of sand are exposed at low tide, and this further contributes to dune formation. The vegetation succession and development of the sand dune **psammosere** ecosystem is explored in detail in Chapter 5, but here the sequence of the development of sand dunes themselves (Figure 3.29) is as follows:

- Sand may become trapped by obstacles (seaweed, rock, driftwood, litter) at the back of the beach, possibly on the highest berm or storm beach.
- The first dunes to develop are known as **embryo dunes**. They are suitable for colonisation by grasses. These are able to grow upwards through the accumulating wind-blown sand, stabilising the surface. As a result low, hummocky dunes are formed. The long roots and underground shoots (rhizomes) of marram grass also help to bind the sand together. The presence of plants adds organic matter to these dunes, which aids water retention.
- Upward growth of embryo dunes raises the height to create dunes that are beyond the reach of all but the highest storm tides. These **foredunes** (**mobile dunes**) are initially yellow, because they contain little organic matter, but as vegetation cover increases, humus is added to the sand. As a result, the dunes look more grey in colour and in places can reach heights of 20 m.
- The dunes gradually become **fixed**. An organic layer develops as other types of vegetation colonise alongside the marram grass.
- In places **dune slacks** develop. These are depressions within the dunes where the water table is on or near the surface and conditions are often damp.
- Behind the yellow and grey dunes, the supply of sand is gradually cut off, giving smaller dune features. These areas may be referred to as **wasting dunes**, or are areas of **dune heath**.
- Within this system, it is possible to find **blowouts** where wind has been funnelled through areas and has removed the sand. Wildlife or human activity can often be a catalyst for the formation of blowouts.

Fieldwork opportunities

There is an opportunity here to visit a sand dune and investigate how their characteristics change along a transect from the beach inland.

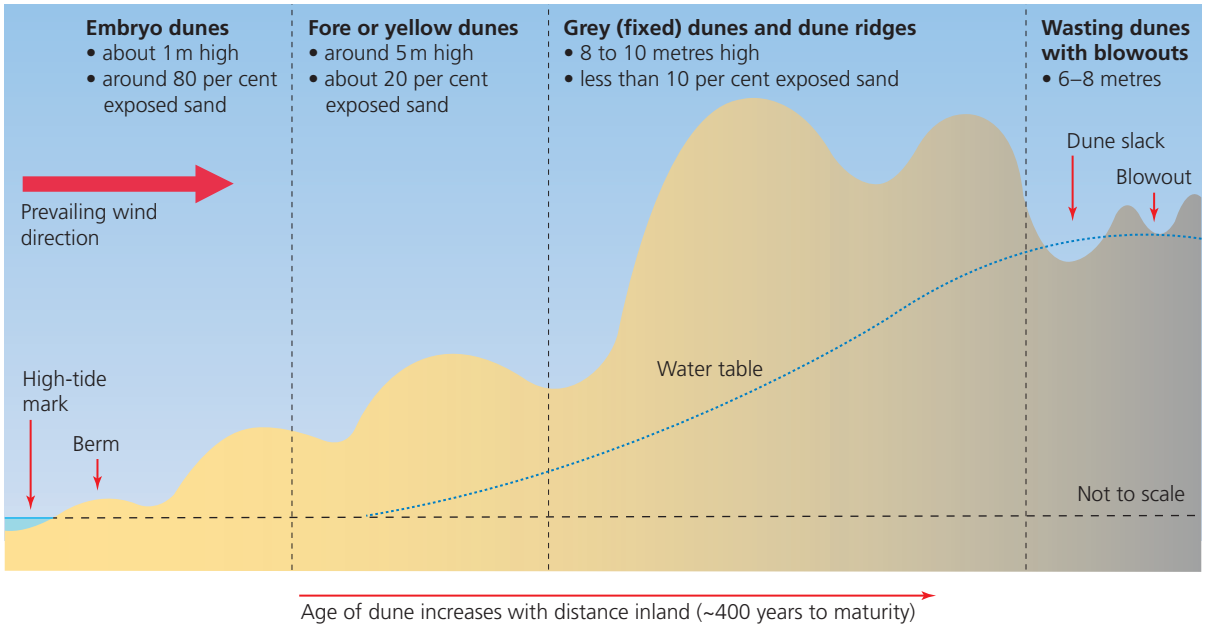


Figure 3.29 A typical sand dune transect

Landscapes associated with estuarine mudflats and salt marsh environments

Mudflats only develop on sheltered shorelines that are not exposed to powerful waves. Mudflats have the following features.

- They are often located in estuaries or on the landward side of a spit.
- They are low-lying areas of the shore that are submerged at high tide and are composed of silt and clay.
- They develop in estuaries where the flow of fresh water out of the river is slow and the sea water flows into the river mouth with each high tide and out with each low tide.
  - The saltwater (from the slow-moving tidal/sea currents) brings large amounts of fine sediments and meets the river which is also carrying its own load of fine silts and clays.
  - As the two flows meet the fine particles settle out of suspension by the process of **flocculation**, where the individual clay particles aggregate together to form larger, heavier particles that can sink to the bed.
- At low tide the inter-tidal area of mud is left exposed. Water is only left in permanent channels and the generally smooth surface shows evidence of tidal action where the flowing water has carved into, or shaped the surface (Figure 3.30).
- Mudflats can be very extensive covering many tens of square kilometres.

Figure 3.30 Mudflats of the Keer Estuary, in Morecambe Bay, northwest England

- They are not necessarily a permanent feature as they are very susceptible to:
  - changes in sea level
  - wave action
  - changes in discharge levels in the river
  - changes to tidal flows.

In the shelter of Morecambe Bay in northwest England, four river estuaries are home to the largest single area of continuous mud- and sandflats in the country. Mudflats are a common feature of coasts globally. Sizeable examples in the USA are the extensive mudflats in Cape Cod and Plymouth Bays off Massachusetts.

Over time mudflats can develop into **salt marshes**. Like the development of sand dunes, these ecosystems change slowly over time (Figure 3.31

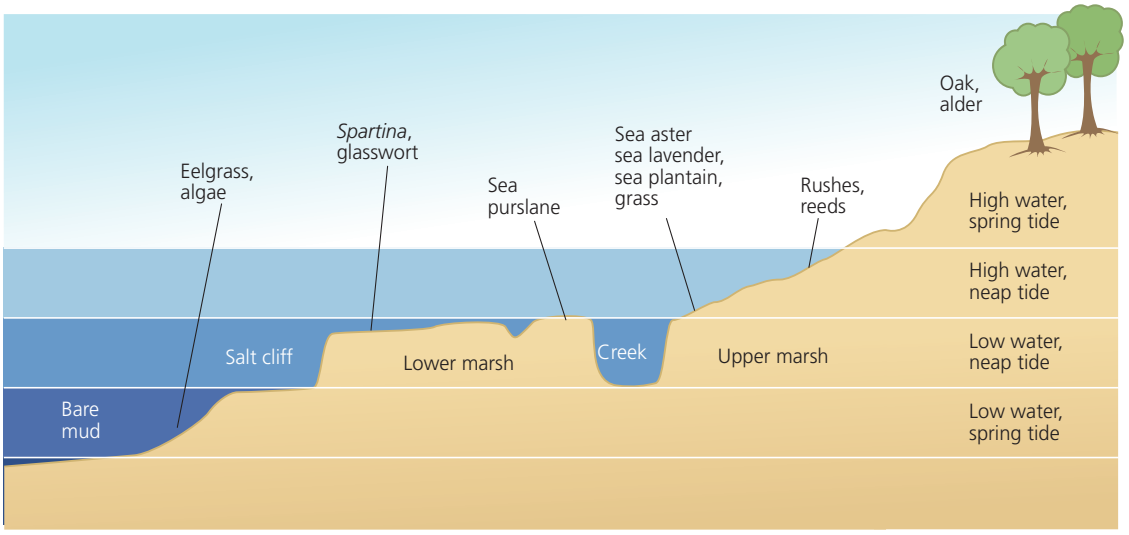


Figure 3.31 The structure of a salt marsh

Figure 3.32 Salt marsh in Morecambe Bay, northwest England

and Figure 3.32). The vegetation succession that develops is known as a **halosere** (tolerant of salty conditions) and follows these stages (Figure 3.33):

Low-lying vegetation, such as eelgrass, begins to grow on the mudflats. This slows the currents further and leads to increased, more uneven, deposition.

↓

Pioneer plants begin to colonise the area. These tolerate salt and periodic submergence by the sea. They are known as **halophytes** and examples include glasswort, sea-blite and *Spartina*.

Spartina has two root systems – a fine mat of surface roots to bind the mud and long, thick, deep roots that can secure up to two metres of deposited material. This enables the plant to trap more mud than other pioneers and so it has become the dominant vegetation on tidal flats in the British Isles.

↓

The pioneers gradually develop close vegetation over the mud allowing colonisation by other plants such as sea aster, marsh grass and sea lavender.

These form a dense mat of vegetation up to 15 cm high. This vegetation slows the tidal currents further and this, together with the vegetation's ability to trap particles, leads to more mud and silt accumulation.

Dead organic matter also builds up on the surface, which grows in height between 1 and 25 mm per year.

↓

As mud levels rise, complex creek systems develop that channel the tides. These deepen as the marsh becomes higher.

Hollows may form where seawater becomes trapped and evaporates, leaving salt pans where the salinity is too great for plants to survive.

As the land rises above sea level, rushes and reeds become established. Eventually trees such as alder, ash and then oak grow to complete the succession.

This land is now rarely covered by the sea.

Figure 3.33 Stages in the formation of a salt marsh

Skills focus

It is important that you are able to describe and explain the features of characteristic coastal landscapes. After visiting coastal landscapes (or by using the library or internet to research contrasting coastal landscapes) complete the following tasks.

a Identify the different landforms in each landscape.

b Explain the processes that created the landforms.

c Try to identify the age and origin of the features; for example, do they result from present-day processes or are they relic features that are being modified today?

d Draw and annotate a sketch of the landscape to describe the characteristic features.

Sea level change

There are two ideas relating to sea level change that are explored in this chapter:

- **eustatic** change
- **isostatic** change.
- **tectonic** change.

Key terms

Eustatic change – A global change in sea level resulting from an actual fall or rise in the level of the sea itself.

Fjord – Former glacial valley drowned by rising sea levels.

Isostatic change – Local changes in sea level resulting from the land rising or falling relative to the sea.

Raised beaches – Areas of former wave-cut platforms and their beaches which are at a level higher than the present sea level.

Ria – Former river valley drowned by rising sea levels.

Tectonic change – The rise or fall in sea level resulting from tectonic processes.

Tides are responsible for daily changes in the levels at which waves break on to the land, but the average position of sea level relative to the land has changed significantly through time. Many changes took place during the glacial and interglacial times of the Quaternary ice age. However, the current sea level reflects changes that have occurred since the last glacial maximum 18,000 years BP, and in the last 10,000 years in particular. Figure 3.34 (page 114) illustrates what the world looked like during the last glacial maximum when sea levels were on average 110 metres below current levels.

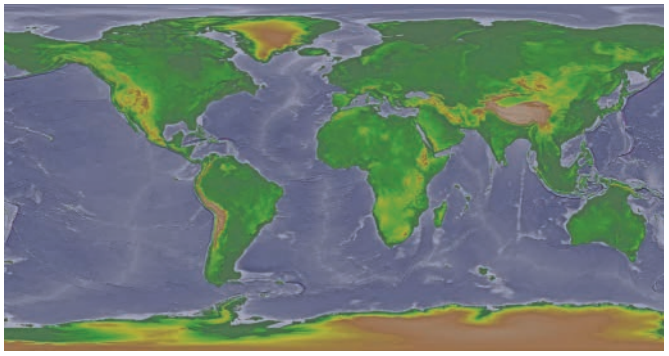


Figure 3.34 Global sea levels 18,000 years ago

A typical sequence of sea level rise to reflect the advance and retreat of the ice would have run as follows:

- **Stage 1:** As the climate begins to get colder, marking the onset of a new glacial period, an increasing amount of precipitation falls as snow. Eventually, this snow turns into glacier ice. Snow and ice act as a store for water, so the hydrological cycle slows down – water cycled from the sea to the land (evaporation, condensation, then precipitation) does not return to the sea. As a consequence, sea level falls and this affects the whole planet. Such a worldwide phenomenon is known as a **eustatic** fall.
- **Stage 2:** The weight of ice causes the land surface to sink. This affects only some coastlines and then to a varying degree. Such a movement is said to be **isostatic** and it moderates the eustatic sea level fall in some areas.
- **Stage 3:** The climate begins to get warmer. Eventually the ice masses on the land begin to melt. This starts to replenish the main store and sea level rises worldwide (eustatic). In many areas this floods the lower parts of the land to produce **submergent** features such as flooded river valleys (**rias**) and flooded glacial valleys (**fjords**).
- **Stage 4:** As the ice is removed from some land areas they begin to move back up to their previous levels (**isostatic** readjustment). If the isostatic movement is faster than the eustatic, then **emergent** features are produced such as **raised beaches**. Isostatic recovery is complicated as it affects different places in different ways. In some parts of the world it is still taking place as the land continues to adjust to having masses of ice removed. Today, the southeast of the British Isles is sinking while the northwest is rising. This reflects the fact that the ice sheets were thickest in northern Scotland and that this was the last area in which the ice melted.

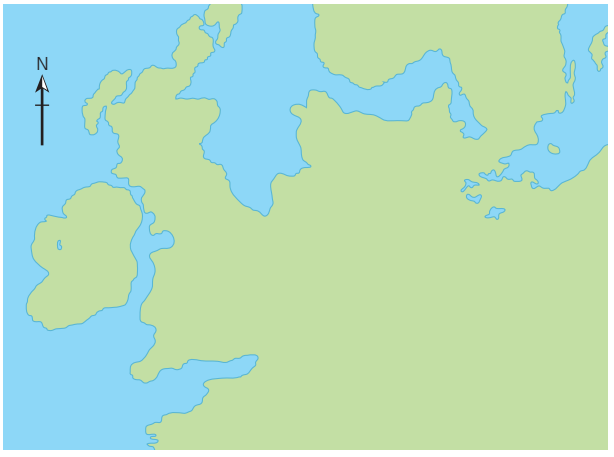


Figure 3.35 The coastline of Western Europe 10,000 years ago

Figure 3.35 shows how the sea level rise from 18,000 to 10,000 years BP was beginning to make the coastline of Western Europe recognisable. However, there were still significant land bridges, especially the extensive grassy plain that stretched between the southern Baltic and eastern Britain (known as Doggerland).

Skills focus

There is an opportunity here to research the British coastline to identify examples of emergent and submergent sections of coast.

Coastlines of submergence and emergence

As mean global temperatures continue to rise, there is an inevitable consequence for sea levels. As more standing ice melts, particularly in Antarctica and Greenland, fresh water will be released into the oceanic store. This could have serious implications for many islands in the Indian and Pacific Oceans and for low-lying coastal areas.

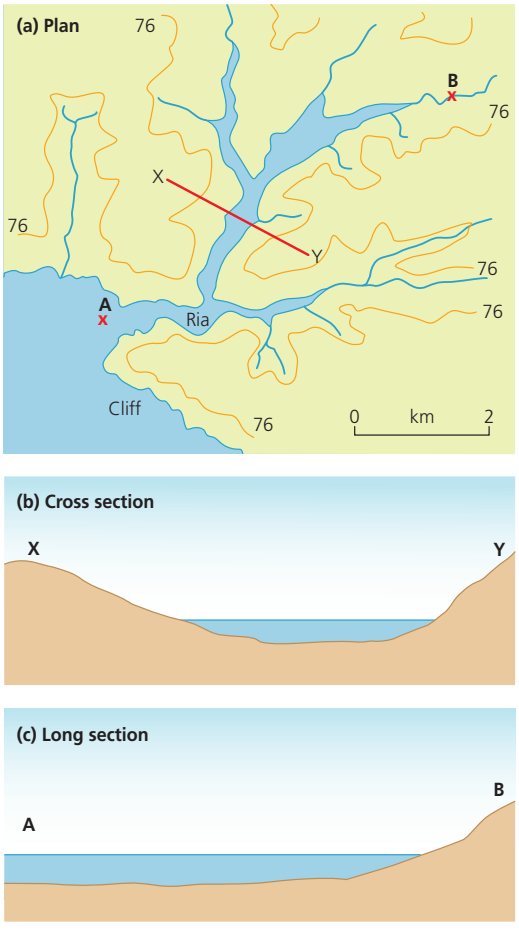


Figure 3.36 A ria

Submergent features

Rias are created by rising sea levels drowning river valleys. The floodplain of a river will vanish beneath the rising waters, but on the edges of uplands only the middle- and upper-course valleys will be filled with sea water, leaving the higher land dry and producing this feature. In Devon and Cornwall, for example, sea level rose and drowned the valleys of the rivers flowing off Dartmoor and the uplands of Cornwall. Good examples are the Fowey estuary in Cornwall and the Kingsbridge estuary in south Devon. Rias have a long section and cross profile typical of a river valley, and usually a dendritic system of drainage (Figure 3.36).

Fjords are drowned glacial valleys typically found on the coasts of Norway, southwestern New Zealand, British Columbia in Canada, southern Chile and Greenland. The coast of western Scotland contains fjords which are not as well developed as those in the areas above because the ice was not as thick and did not last for the same length of time.

Fjords have steep valley sides (cliff-like in places) and are fairly straight and narrow (Figure 3.37). Like glacial valleys, they have a typical U-shaped cross section with hanging valleys on either side. Unlike

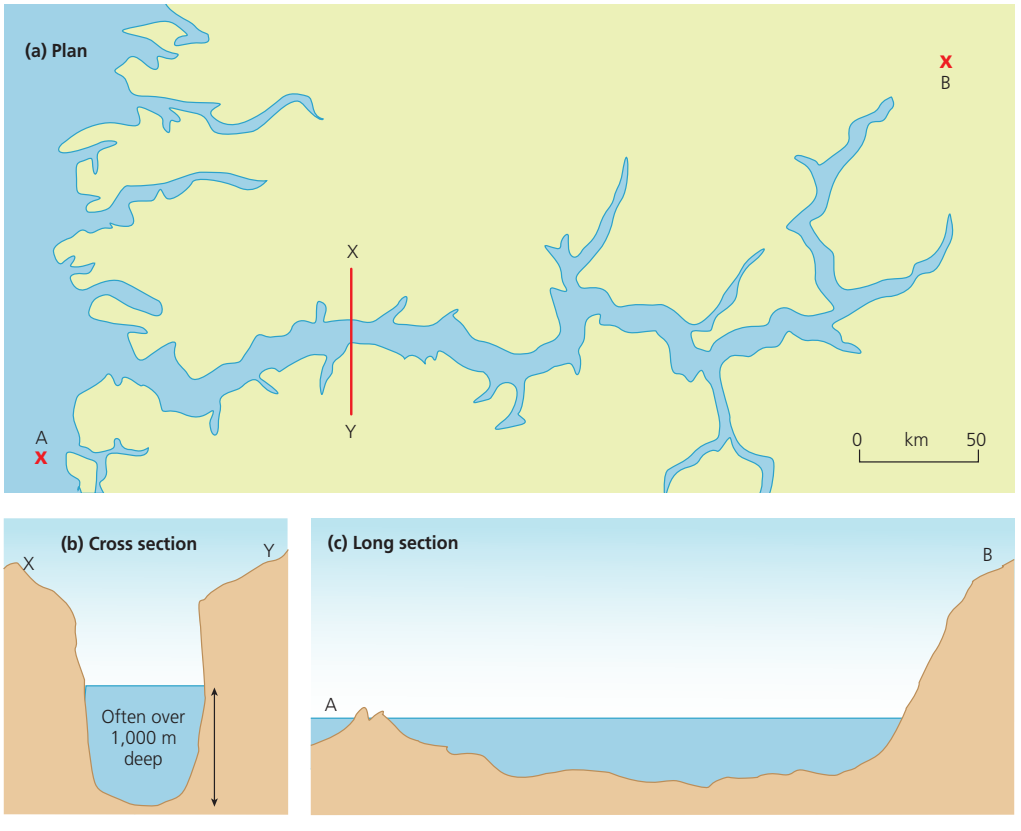


Figure 3.37 A fjord

rias, they are not deepest at the mouth, but generally consist of a glacial rock basin with a shallower section at the end, known as the threshold. They were formed when the sea drowned the lower part of glacial valleys that were cut to a much lower sea level. The threshold is thought to be due to reduced glacial erosion as the glacier came in contact with the sea and the ice became thinner. Good examples include Sogne Fjord in Norway, which is nearly 200 km long, and Milford Sound in New Zealand (Figure 3.38).

Figure 3.38 The entrance to Milford Sound fjord, New Zealand

Where the topography of the land runs parallel to the coastline and becomes flooded by sea level rise a **Dalmatian coast** is formed. Deriving their name from the Croatian coast in the Adriatic, the flooded valleys run parallel to the coast rather than at *right-angles* to it as in the case with fjords and rias. Here islands and peninsulas are aligned parallel to, but just offshore from, the mainland.

Skills focus

Use an atlas to find a map of the Adriatic Sea and draw and annotate a sketch map of the Dalmatian coastline of Croatia.

Emergent features

Raised beaches are areas of former wave-cut platforms and their beaches, which are at a level higher than the present sea level. Behind the beach it is not unusual to find old cliff lines with wave-cut notches, sea caves, arches and stacks. Raised beaches are common around the coasts of western Scotland where three levels have been recognised, at 8 m, 15 m and 30 m. Because of differential uplift these are only approximate heights.

On the west of the Isle of Arran there is a well-developed raised beach north of Drumadoon. This has a **relict cliff,** arches, stacks and caves, including the well-known King's Cave. This beach is around four or five metres above present sea level and is probably the equivalent of the lower raised beach in Figure 3.39. It was clearly produced when the sea was at that level, which initially suggests that the sea has fallen to its present level. However, we know that sea levels have *risen* considerably (eustatic) since the end of the last ice age, so the beach must have reached its raised position by isostatic rising of the land. The land locally must have risen faster than the eustatic rise in sea level to create this phenomenon.

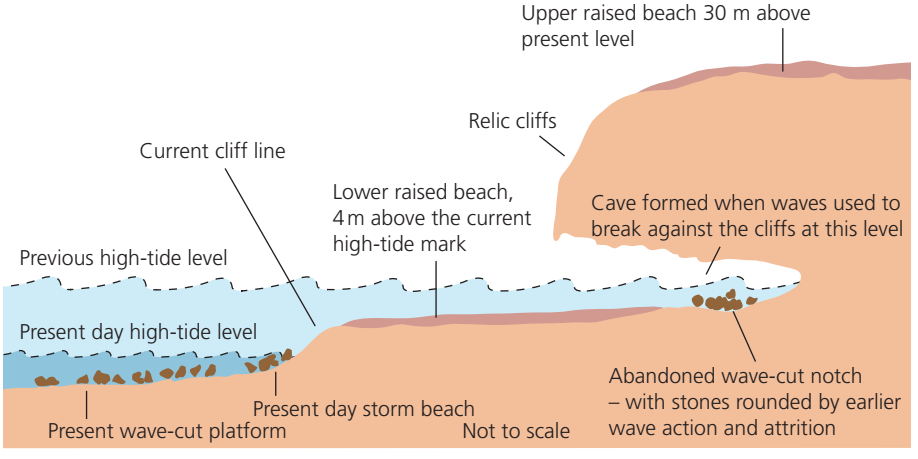


Figure 3.39 A simplified diagram of a raised beach

Where a greater expanse of gently sloping formerly submerged land has been exposed by uplift or the lowering of sea levels, some geographers refer to this feature as a **marine platform**, or marine terrace. (Some see these terms, along with coastal terrace and perched coastline as interchangeable with 'raised beach'.) One distinction between a raised beach and a marine platform may be that in a marine platform what is now exposed is part of a gently sloping continental shelf, whose gentle gradient is now continued for some distance both offshore and inland.

Impacts of recent and predicted climate change on coasts

Sea level has risen consistently over the last 10,000 years. However, rates have increased significantly in recent decades. The rate of future sea level rise is uncertain, but most predictions range between 30 cm and 110 cm by 2100 and continued significant increases into the twenty-second century (Table 3.5).

Table 3.5 Rates of sea level rise

Time period	Average sea level rise per year
1901–1990	1.4 mm
1970–2015	2.1 mm
1993–2015	3.6 mm
2100	15 mm (possible)
22nd century	Several centimetres (possible)

Source: IPCC

Changes in sea level are the result of two processes:

- increases in the volume of the ocean
- subsidence of the coast.

Links

See Chapter 1, Water and carbon cycles, for links between changes in sea level and Water, carbon, climate and life on Earth (pages 41–6).

There are two ways in which the volume of the ocean is increasing. First, and most obviously, as the Earth's climate warms, both naturally and due to current warming related to greenhouse gas emissions and a human-induced enhanced greenhouse effect, water currently stored on the surface as ice is released into the oceans as it melts. Second, as temperatures warm there will be thermal expansion of the oceans. Scientists have been improving their understanding of the impacts of both thermal expansion and the melting of ice caps and smaller valley glaciers, however, what is proving more difficult to predict is how the massive ice sheets will behave during future warming. There is also significant uncertainty about how well people will be able to mitigate their contributions to climate change. This, alongside the complexity of the modelling involved, leads to quite a range in the various predictions of sea level rise above 1990s' levels. Figure 3.40 shows how the National Oceanic Atmospheric Administration (NOAA) summarises this range.

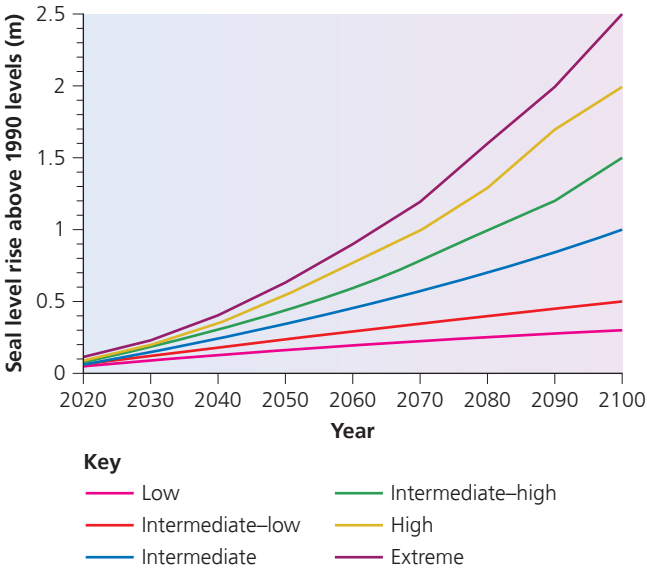


Figure 3.40 Range of possible predicted sea level rise scenarios
Source: NOAA

Rising sea levels could have many adverse effects, including increased coastal flooding and erosion, unless action is taken to mitigate it. In certain parts of the UK, notably the southeast, the rise in sea relative to the land may be greater than average, owing to localised subsidence. Since the last ice age, glacier load removal in the north and west of the UK has resulted in a gradual uplift. At the same time, gradual subsidence has occurred on the margins of the North Sea basin in the east and southeast. This

has resulted in the loss of numerous villages from low-lying east coast areas since the compilation of the Domesday Book.

A number of areas around the British coastline are potentially at risk from rising sea levels. These include major conurbations, such as London, Hull and Middlesbrough, and high-grade agricultural land. Major road and rail links near the coast are also at risk, and several power stations are situated on low-lying ground.

As well as direct effects such as coastal erosion or the flooding of coastal areas, higher mean sea levels could have an impact on underground water resources. The zone where seawater mixes with fresh water in rivers is dynamic and a rise in sea levels can cause it to move upstream. A similar effect can occur between fresh water contained in rocks under the land and salt water in sea sediments, leading to intrusion of salt water beneath the land. This would adversely affect some points along the lower reaches of rivers where water is abstracted for domestic and irrigation purposes. These abstraction points would have to be moved upstream or become intermittent to avoid abstracting saline water.

Review questions

6 Explain the formation of an emergent coastal landform.

7 Outline the role of deposition in the formation of landforms in estuarine environments.

8 Can you explain why there is uncertainty about the predicted sea level change by 2100?

3.4 Coastal management

Human intervention in coastal landscapes

Some coastal landscapes, including human and natural environments, are coming under increasing pressure from both natural processes and human activities. In response to this, a range of protection and management strategies have been put in place in many coastal areas. These solutions are often successful but, in some cases the solutions themselves cause other problems. Coastal management has two main aims:

- to provide defence against, and mitigate the impacts of flooding
- to provide protection against, and mitigate the impacts of coastal erosion.

Other aims of management include:

- stabilising beaches affected by longshore drift
- stabilising sand dune areas
- protecting fragile estuarine landscapes.

Management strategies can work either with or against natural processes. Working *with* nature means allowing the natural processes of erosion to occur (**managed retreat**) and not spending money on the defence of the coastal area. This is now applied to large stretches of coastline in the UK where there are few settlements. **Soft engineering** techniques such as beach nourishment are said to work with nature.

Working *against* nature usually occurs where there is significant capital investment – buildings and communications – in the coastal region that has to be protected. Protection involves constructing sea walls, revetments, groynes and other examples of **hard engineering**. The costs of such defences are justified by the potential expense of replacing sea-damaged buildings and infrastructure if they were not in place.

Key terms

Hard engineering – Making a physical change to the coastal landscape using resistant materials, like concrete, boulders, wood and metal.

Soft engineering – Using natural systems for coastal defence, such as beaches, dunes and salt marshes, which can absorb and adjust to wave and tide energy.

There are an ever-increasing range of approaches to coastal defence, some of which are outlined and illustrated in the next section.

Traditional approaches to coastal erosion and flood risk

Hard engineering

Hard engineering generally involves making a physical change to the coastal landscape using resistant materials, like concrete, boulders, wood and metal. Often each strategy is large scale and costs a

Table 3.6 Examples of hard engineering

Hard engineering	Description
Sea walls (sometimes recurved)	• Aim to dissipate wave energy. The recurved structure throws waves back out to sea into the path of the next incoming wave, reducing its wave impact • They also provide a physical barrier to flooding by raising the height of the coastline and must have a continuous facing because any slight gap will be exploited by waves • They also need drain outlets so that any water that does get over them does not accumulate inland
Rock armour (rip-rap)	• Consists of large boulders placed in front of a cliff or sea wall to take the full force of the waves • The boulders are deliberately left angular in appearance to present a large surface area to the waves and create gaps for water to filter through, to mitigate their impact on the coast • Usually the boulders are not secured in place, so energy is taken out of the most powerful waves by rocking or slightly moving the massive rocks
Gabions	• Operate on the same principle as rip-raps • Smaller boulders are contained within steel wire mesh cages, each of which can be joined together to form larger structures or walls
Revetments	• Concrete or wooden structures placed across a beach or coastline to take the full force of the wave energy, preventing further erosion of the coast
Groynes	• Wooden, stone or steel breakwaters built nearly at right angles to waves (usually 5 to 10° from the perpendicular to prevent scouring on the down-drift side) • They are built to control longshore drift by trapping sediment to create higher and wider beaches which will then also dissipate wave energy • Groynes also break up the waves as they hit the coast • Halting the bulk of longshore drift in an area may have serious effects down the coast where it will cut off the supply of beach material and could leave the coast there exposed to erosion
Cliff fixing	• Often done by driving metal bars into the cliff face, both to stabilise it and to absorb some wave power
Offshore reefs	• Force the waves to break offshore, which reduces their impact on the base of cliffs • Sometimes redundant ships have been deliberately sunk parallel to the shore to both slow down approaching waves and to act as a substructure for reef material to begin to colonise
Barrages	• Large structures built to prevent flooding on major estuaries and other large sea inlets • They act as a dam across an estuary and prevent incursion of seawater • Good examples of barrages are those that are part of the Delta Plan in the Netherlands and the Cardiff Bay barrage in Wales, which was completed in 1999

significant amount of money. Each type of hard sea defence is built with a specific purpose (Table 3.6 and Figure 3.41).

Hard engineering strategies are generally long lasting and effective over their planned lifespan; however they do have several disadvantages, including:

- structures can be expensive to build and to maintain (to build a sea wall can cost over £7,000/m)
- defence in one place can have serious consequences for another area of the coast
- structures are sometimes an eyesore, spoiling the landscape and physically disrupting natural habitats.

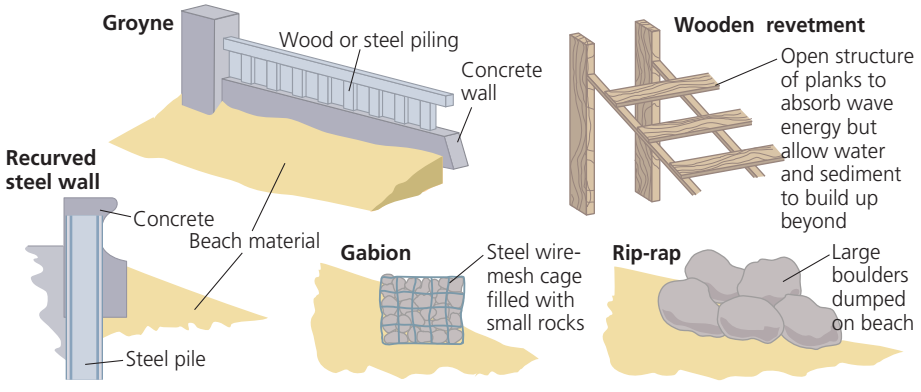


Figure 3.41 Examples of hard engineering solutions

Hard engineering: Heysham and Morecambe

The coastal town of Morecambe in Lancashire is used as an illustrative example of the use of hard engineering strategies. Lancashire County Council, like all others on the UK coast, has four defence options:

- **Hold the line** – retain the existing coastline by maintaining current defences or building new ones where existing structures no longer provide sufficient protection.
- **Do nothing but monitor** – on some stretches of coastline it is not technically, economically or environmentally viable to undertake defence works. The value of the built environment here does not exceed the cost of installing coastal defences.
- **Retreat the line** – actively manage the rate and process by which the coast retreats.
- **Advance the line** – build new defences seaward of the existing line.

Figure 3.42 illustrates the choice of hard engineering techniques used along the coast at Heysham and Morecambe in Lancashire.

Background

The current defences along the 8.5 km stretch of coastline are the result of a comprehensive improvement scheme of the existing structures during a multi-phase programme between 1989 and 2018 costing £30 million. It includes a mix of traditional hard engineering strategies with more contemporary methods to improve the potential for sustainable management.

Strategy 1 – Rock armour/rip rap to enhance and protect the existing sea wall

Boulders of locally sourced limestone were placed along the majority of the existing promenade and sea walls from the western end of the promenade to about one kilometre east of the town centre.

Stones range from 0.77 to 7 tonnes, averaging 3.5 tonnes. In total 436,000 tonnes of rock armour was installed (Figure 3.43).

Figure 3.43 Traditional sea wall protected by rock armour with new wave reflection wall on the promenade

Strategy 2 – Breakwaters or rock groynes

Around ten breakwaters and rock groynes were built at intervals in front of the town. These included a number of fish-tail breakwaters (Figure 3.44). Just under a million tonnes of locally sourced limestone boulders were used.

Figure 3.44 Fish-tail breakwater (rock groyne)

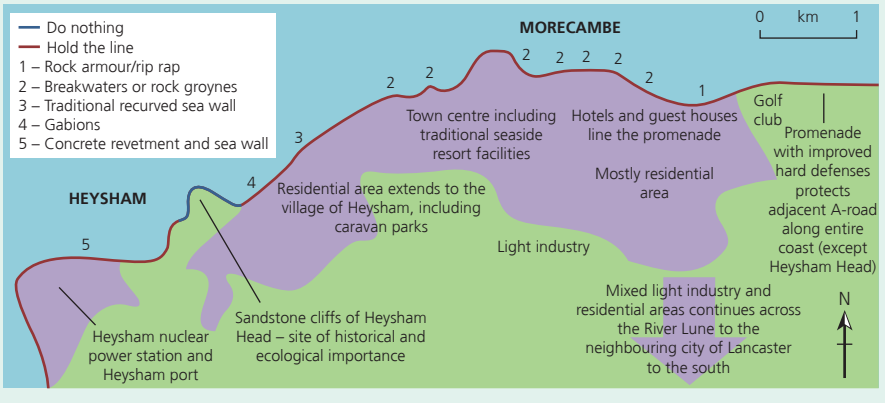


Figure 3.42 Heysham and Morecambe coastal defences

Strategy 3 – Repaired and replaced sea walls

The traditional recurved sea wall fronting the promenade was repaired and reinforced (Figure 3.45). Between 2015 and 2018, four km of 30-year-old sea wall running along the promenade was replaced with a state-of-the-art and decorative wave reflection and flood wall. This cost £11 million and protects 11,400 homes and 2,246 commercial properties. It is designed to last 100 years.

Figure 3.45 Traditional recurved sea wall improved protection on the promenade

Strategy 4 – Gabions

Over 500 cages, measuring 2 × 1 × 1 m, filled with small limestone boulders were used in various locations to reinforce the coast (Figure 3.46).

Figure 3.46 Gabions

Strategy 5 – Concrete revetment and sea wall

To the west of Heysham Head, extending to the port and power station, the existing sea wall and large concrete revetments were repaired and left in place (Figure 3.47).

Figure 3.47 Concrete revetment

Summary

Although a range of traditional hard engineering strategies were implemented, they were designed in ways that were sympathetic to Morecambe Bay's classification as a **Site of Special Scientific Interest (SSSI)** and **Special Area of Conservation (SAC)**, and aimed to support the sustainable development of the shoreline. For example, the local sourcing of limestone for the rock armour and rock groynes meant that it was of the same geological origin as the rocks of the coastline. As most of the quarries were less than 10 km away this also reduced the carbon footprint of its transportation. The location of the breakwaters and rock groynes is such that it has created a series of small bays to encourage beach formation, to act as a soft engineering strategy to reinforce the hard defences.

As well as encouraging beach formation between the breakwaters, 89,000 tonnes of sand and 19,000 tonnes of shingle were used for **beach nourishment** between Heysham and the eastern end of the promenade (another soft strategy).

The whole project has been designed to be aesthetically attractive and improve social amenities. It includes combined cycle and walking paths, playgrounds and various artistic elements.

Soft engineering

Soft engineering uses natural systems for coastal defence, such as beaches, dunes and salt marshes, which can absorb and adjust to wave and tide energy. It involves manipulating and maintaining these systems, without changing their fundamental structures. Table 3.7 outlines the different soft engineering approaches.

The Sefton Coast is used to illustrate how some soft engineering strategies have been used in an attempt to protect another stretch of the Merseyside and Lancashire coast.

Table 3.7 Examples of soft engineering

Soft engineering	Description
Beach nourishment	• Aims to replace material that has been lost through longshore drift • Local councils may move material from one end of a beach to the other before the start of the tourist season

Dune regeneration	• Fragile sand dune environments are easily disrupted by human activities
	• Most damage is caused by the removal of vegetation by agriculture (overgrazing) or tourism (trampling the dunes). This can lead to blowouts where large amounts of sand may be carried inland and deposited on valuable agricultural land
	• Management strategies to regenerate dunes include:
	  • replanting vulnerable areas with plants such as marram grass and stabilising the surface with sacking or wire mesh
	  • afforestation with quick-growing conifers
	  • selective grazing
	  • restricting access by fencing off areas
	  • providing boardwalks for tourists
	  • giving tourists information about potential damage

Managed retreat	• Abandoning the current line of sea defences and developing the exposed land in some way, perhaps with salt marshes, to reduce wave power • If old sea defences, such as walls, are abandoned, low-lying land will be flooded. This will be reclaimed naturally by marsh plants. The new area of marsh will act as a defence against rising sea levels • The scale of hard engineering sea defences can be reduced • There have been proposals in some areas to ban new developments on the coast. In California, for example, there are already requirements on some stretches of coastline that buildings must be a certain distance from the shore
Land-use management	• Even if erosion or flooding of a coastal area in the future is inevitable, a local authority may be able to mitigate the impact through land-use management • It involves addressing people's behaviour and educating the local community • Officials help people plan for the future by encouraging landowners to think about how they can continue to use land that is at risk. For example, caravan parks on cliff tops can provide an income from the land but be moved and re-sited quickly when the time comes • Giving land at risk of flooding over to grazing rather than growing crops means the sheep or cattle can be moved if storms are forecast • It only works if the local population agree to having their use of land limited, and it cannot remedy damage that has already been done
Do nothing	• In the 21st century, a school of thought has grown up that asks whether the coast should be protected • Hundreds of millions of pounds are spent annually in the UK on coastal protection. It might be cheaper to let nature take its course and pay compensation to those affected • The storms of December 2013 and January 2014 illustrated how trying to control the power of nature is often futile, with the traditional hard sea defences of places like Dawlish in Devon and Aberystwyth in west Wales proving ineffective against the powerful wind and waves • Some argue that the limited funds available for coastal protection should be targeted to places that have significant infrastructural or economic value for large numbers of people • This debate is not new; the House of Commons Select Committee on Agriculture suggested in 1998 that large tracts of land should be 'surrendered to the sea' as trying to protect them is a waste of money • Those living in places deemed not worthy of protection may view the debate quite differently

Coastal erosion and soft engineering at Formby Point on the Sefton Coast

The Sefton Coast, north of Liverpool, has the largest dune area in England, extending over 17 km, and ranging from 200 m to 4 km in width (Figure 3.48). The sand dune system around Formby Point experienced continual erosion during the twentieth century, losing 700 m from 1920 to 1970. Rates of erosion average around 4 m/year.

Table 3.8 Causes of coastal erosion at Formby Point

Physical (natural)	Human
Periodic storms combined with high tides	Dredging of beach material for the local foundry trade and glass industry
	Building of hard/fixed sea defences to the north (Birkdale to Southport) and south (Hightown to Crosby)
	Activities relating to the development of ports in Liverpool and Preston
	Spoil dumping to the north
	Human access (walking on the frontal dunes)
	Use of off-road vehicles (breaking up the dunes and destroying vegetation)
	Afforestation of a conifer plantation in the mid-twentieth century

The area around Formby Point and Ainsdale still attracts large numbers of visitors to the beaches, sand dunes and pine forests. The local populations of rare red squirrels and natterjack toads are also an attraction.

Some recent climate models suggest an average sea level rise of 0.3 m over the next 60 years and an increase in maximum wave height could lead to a significantly increased risk of erosion. In this scenario the dunes could play a vital role in mitigating flood risk for the local area. The immediate impacts of the erosion of Formby Point are not constrained to the immediate area. As the coastline at the Point becomes straighter, material is being transported to the north and south where sand is deposited. In places like Crosby this can be problematic where sections of coastal footpaths and access roads are occasionally buried under considerable accumulations of sand.

The **Sefton Coast Management Scheme** was developed in the late 1980s and in the Formby and Ainsdale areas includes:

- 'planting' used Christmas trees on the seaward edge of the dunes to trap more sand and encourage dune regeneration
- fencing off areas of sand dunes to restrict pedestrian access
- wooden posts placed in front of the dunes to encourage dune regeneration

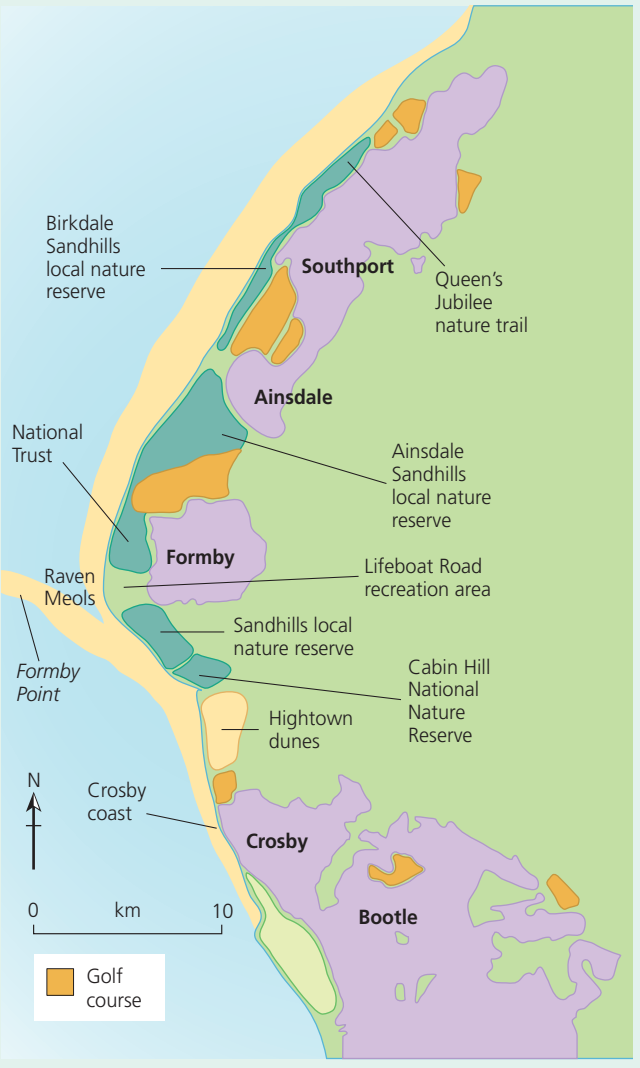


Figure 3.48 The Sefton Coast

- building boardwalks to stop people trampling the dunes
- signage to direct people to the beach via routes avoiding the dunes
- ranger services to educate local children about the protection and conservation of the dunes
- banning off-road vehicles from the dune system
- controlling the extraction of sand for commercial purposes
- a debate about the impact the pine plantation has had on the natural dune landscape and whether removing the plantation and scrub cover would help to re-establish and maintain a broader spectrum of habitats.

All of these measures seek to protect the area for future use, in other words to be sustainable.

Sustainable approaches to coastal flood and erosion management

Since the 1990s approaches to the management and mitigation of coastal flood risk and coastal erosion have focused on sustainability. This section explores the policies used by the UK Government to make coastal management more sustainable around the British Isles, and the policies developed at a global level which seek to create a common and integrated approach around the world.

Shoreline management plans

Figure 3.11 (page 97) shows that the movement of coastal sediment around the British Isles mainly occurs within distinct cells, the boundaries of which rarely coincide with the administrative boundaries of the Local Authorities. To avoid a piecemeal approach to coastal management an integrated system of **Shoreline Management Plans (SMPs)** was introduced in 1995. There are 22 SMPs around the coast of England and Wales corresponding to the sediment cells and sub-cells discussed earlier. SMPs are designed to identify the most sustainable approach to managing the flood and coastal erosion risks to the coastline. They aim to plan for the:

- short term (0–20 years)
- medium term (20–50 years)
- long term (50–100 years).

The management objectives of SMPs are continually reviewed. In the 2010s, a second generation, SMP 2s, were generated.

The key features of SMPs are that they:

- provide an assessment of the risks associated with the evolution of the coast
- provide a framework to address the risks to people and to the developed, historic and natural environment
- address risks in a sustainable way
- provide the policy agenda for coastal defence management planning
- promote long-term management policies for the twenty-second century
- aim to be technically sustainable, environmentally acceptable and economically viable
- ensure management plans comply with international and national nature conservation and biodiversity legislation
- incorporate a 'route map' to allow decision makers to make changes to the short- and medium-term plans to ensure long-term sustainability is maintained

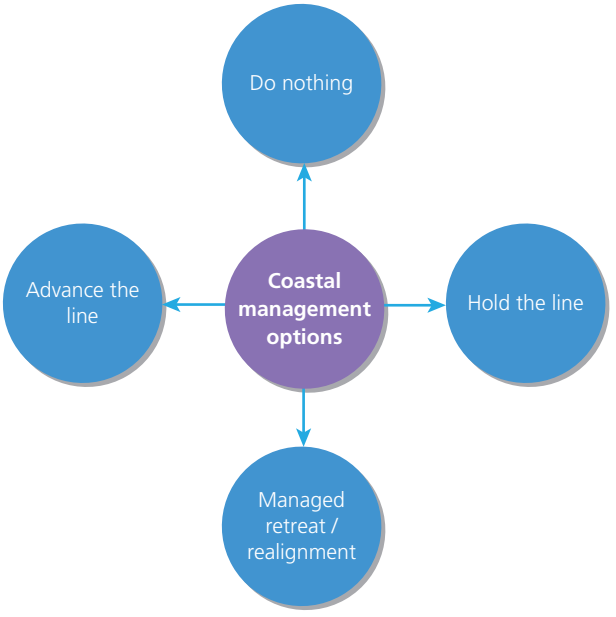


Figure 3.49 Options for coastal management

- provide a foundation for future research and the development of new coastal management strategies in the future
- are 'live' working documents to be continually reviewed and updated, setting new targets for future management objectives.

Each SMP describes how each **management unit**, or stretch of coastline covered by the plan, is to be managed. Figure 3.49 is a reminder of the four management options available for each unit.

Figure 3.50 shows the location of the 22 SMPs in England and Wales.

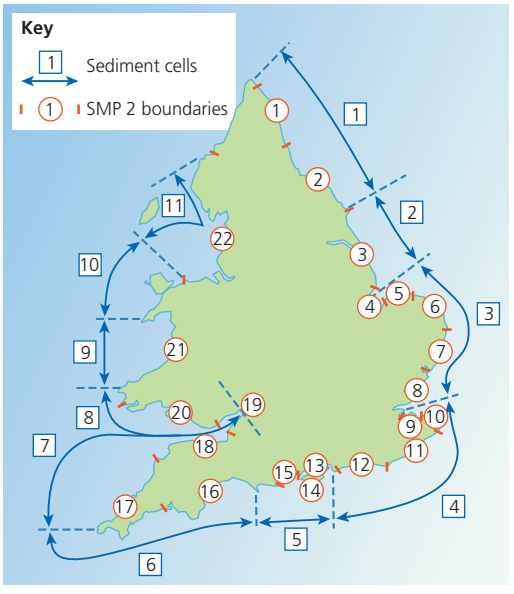


Figure 3.50 The SMP areas of England and Wales

Skills focus

Visit the web link below and find the SMP for the coastline nearest to you and produce a summary of the key objectives for the management of that area.
www.gov.uk/government/publications/shoreline-management-plans-smps

Integrated coastal zone management

The term '**integrated coastal zone management**' (ICZM) originated from the UN Earth Summit of Rio de Janeiro in 1992. The guidelines about how ICZM should operate were set out in the **Agenda 21** documents released following the summit. In subsequent decades different organisations with responsibility for the governance of coastal zones have developed their own policies and approaches for integrated coastal management. On the environment pages of the website of the European Commission, ICZM is described as follows:

> Integrated coastal management aims for the coordinated application of the different policies affecting the coastal zone and related to activities such as nature protection, aquaculture, fisheries, agriculture, industry, offshore wind energy, shipping, tourism, development of infrastructure and mitigation and adaptation to climate change. It will contribute to sustainable development of coastal zones by the application of an approach that respects the limits of natural resources and ecosystems, the so-called 'ecosystem-based approach'.
>
> Integrated coastal management covers the full cycle of information collection, planning, decision-making, management and monitoring of implementation. It is important to involve all stakeholders across the different sectors to ensure broad support for the implementation of management strategies.

Following Rio, the European Commission developed its framework for ICZM to promote **integrated coastal management**. The background to this initiative includes:

- coastal zones are some of the most ecologically productive areas in the world
- the natural assets of coasts have for millennia made them popular for:
  - settlements
  - tourist destinations
  - business centres
  - ports
- around 200 million people live near Europe's coastline.

The Commission feels that this concentration of people and economic activity is putting great pressure on our coastal environment and creates excessive exploitation of natural resources, all of which is leading to:

- biodiversity loss
- habitat destruction
- pollution
- conflicts between stakeholders
- overcrowding in some locations.

Some argue that these impacts of human activity are not unique to the coastal zone. The commission suggests that the revised ICZM initiative is required because:

- they are some of the areas most vulnerable to climate change and natural hazards
- coasts are especially at risk of:
  - flooding
  - erosion
  - sea level rise
  - extreme weather events
- the lives of people in many coastal communities are already changing due to the impacts of these issues.

For the well-being and economic prosperity of many coastal communities the condition of the natural environment is often crucial, therefore officials believe the long-term plans proposed by ICZM are vital to ensure sustainable development for generations to come.

Links

For links between coasts and lived experience, see Chapter 8, Changing places, page 354.

In the past coastal management was often viewed as having a *sectoral approach*, where the local authorities made decisions for their stretch of coastline and other agencies and interest groups tried to manage their particular environment, cause or social group. In the view of the European Commission this led to

decisions that undermined each other, an inefficient use of resources and failure to meet sustainability objectives. This is why ICZM is designed to integrate the interests of all stakeholders to avoid such problems. ICZM aims to co-ordinate policies that affect the coastal zone and the activities that take place there, including:

- nature conservation/protection
- aquaculture (farming marine organisms, for example, a lobster farm)
- fishing
- agriculture
- industry
- offshore wind energy
- shipping
- tourism
- infrastructure development
- developments to adapt to and mitigate climate change.

ICZM aims to contribute to sustainable development by employing an '**ecosystem-based approach**' that operates within the limits of natural resources and ecosystems.

If conducted properly ICZM should operate in a cycle with each stage generating feedback to be addressed by the following stage (Figure 3.51).

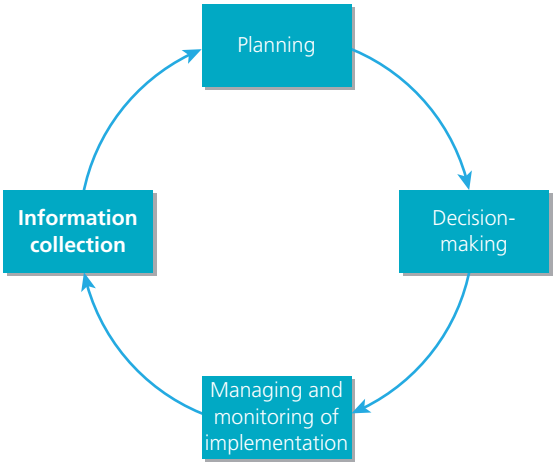


Figure 3.51 The cycle of integrated coastal zone management

Since 2013, building on and working alongside ICZM, the European Commission has developed a framework for maritime spatial planning (MSP). This adopts land-use planning, powerful GIS and geo-spatial mapping techniques. The Commission website highlights the importance of MSP:

> Competition for maritime space – for renewable energy equipment, aquaculture and other uses – has highlighted the need to manage our waters more coherently. Maritime spatial planning (MSP) works across borders and sectors to ensure human activities at sea take place in an efficient, safe and sustainable way.

The advantages of MSP are seen to include:

- reduced conflicts between stakeholders
- encouragement of investment
- an increase in cross-border co-operation
- protection of the environment.

Review questions

9 Explain why a local authority might still adopt traditional hard engineering strategies to protect a stretch of coastline.

10 What are the main features of Shoreline Management Plans?

11 Why is it important to co-ordinate the needs of different stakeholders when developing coastal management strategies?

3.5 Quantitative and qualitative skills in coastal landscapes

Quantitative and **qualitative** techniques in geography are defined and explored in Chapter 12 and you should read this section carefully.

There are many opportunities to visit, and undertake fieldwork, in a coastal landscape as part of this course, and it is important to reflect on how a range of quantitative and qualitative skills could be utilised in a study of coastal systems and landscapes. Pages 130–1 illustrate how a number of quantitative techniques can be employed in a beach landscape.

Quantitative data and coastal systems and landscapes

Coastal landscapes provide a wealth of opportunities to employ a range of observational and measurement techniques to collect quantitative data. A good starting point is completing annotated field-sketches. Others include measuring:

- characteristics of waves, for example, identifying constructive and destructive waves

- the movement of material by longshore drift
- the size and shape of beach sediment
- the gradient of beaches and sand dunes – beach and dune profiles
- wind direction and wind speed
- the dimensions and characteristics of coastal landforms.

Other quantitative techniques could be used to identify, measure and map the types of sea defences used along a stretch of coast.

Quantitative investigations of coastal landscapes and systems can produce a large amount of useful data that readily lends itself to manipulation and analysis, with a range of data presentation and statistical methods, as explored in Chapter 12. Coastal fieldwork also produces quantitative data that is easily mapped using a range of geo-spatial mapping techniques to help give a clear visual representation of the findings of the research.

Qualitative data in coastal landscapes

This chapter illustrates how predicting the impacts of climate change on future sea levels is complex. Even more challenging will be assessing the increased challenges that this may pose for many vulnerable coastal populations around the world. Equally those tasked with the challenge of mitigating the worst impacts of any sea level rise, while improving the resilience of communities and helping them adapt to their changing circumstances will require a more holistic assessment of many aspects of their situations. Qualitative techniques such as **participant observation** and **in-depth interview** schedules would be one way of engaging with the views of the populations involved to inform the development of bottom-up solutions to future issues. The case study on page 131 exemplifies a region where such techniques could be employed – with marginal coastal communities of southern Bangladesh.

Perhaps on a more local level, questionnaires and interviews could be used with coastal communities in the British Isles to assess the views of various stakeholders about any proposed sea defences for a particular stretch of coast. The qualitative results of such surveys could then be analysed alongside more empirical data about factors such as financial cost and impacts on natural coastal processes to produce thorough cost–benefit or SWOT analyses of any proposed coastal management schemes, that include the views of all stakeholders.

Fieldwork opportunities

Here is an opportunity to compile a questionnaire that could be used with stakeholders to gather information about the views of the stakeholders involved in a coastal management scheme.

Case study of a coastal environment at a local scale: The coastal processes, landscape and sustainable management of Pevensey Bay, East Sussex

Geographical context

Pevensey Bay, on the East Sussex coast, occupies a low-lying area of softer sediments between sandstone geology to the east at Hastings and chalk of the South Downs to the west (Figure 3.52, page 128). Over several thousands of years natural processes associated with sea level rise, and subsequent wave and longshore drift action from west to east, have created a natural shingle barrier. This shingle bank extends for nine kilometres between Eastbourne and Bexhill, and although the shape of the coastline has continually changed through geomorphological processes over hundreds of years, and continues to be modified today, it has remained a natural defence against coastal flooding. The management of this coastline is challenging, as the modern-day settlements require that the coastline remains in a relative fixed position. However, as sea levels continue to rise, wave action and longshore drift continue to move sediment in the littoral zone and storm frequency appears to increase, the shape of the coastline changes, and will continue to change, almost on a daily basis. So balancing the wishes of residents, economic costs and benefits and sustainability is an ongoing challenge.

The natural and man-made sea defences currently protect:

- the permanent flooding of a 50 km² area – including communities in Pevensey Bay, Norman's Bay, Langley, Westham and parts of Pevensey
- 10,000 properties
- recreational and commercial sites

Figure 3.52 Ordnance survey map showing the location of Pevensey Bay, East Sussex (Map is reproduced at half of OS Landranger map scale of 1:50,000. Each grid square is 1 km²). © Crown copyright and database rights, 2021, Hodder Education, under licence to Ordnance Survey. Licence number 100036470

- A259 coast road and railway line from Hastings to Portsmouth
- two nature reserves and a SSSI wetland site
- numerous livestock and arable farms.

Coastal processes

Like all natural coastlines the shore at Pevensey Bay constantly tries to adapt to minimise the effects of the power of the sea. The prevailing winds from the south-west move beach material towards the north-east creating beaches that are more parallel to the waves. During major storm events the beaches become flatter, creating beach profiles that reduce wave energy and prevent waves reaching the top of the shingle bank, thus reducing the flood risk. In the past the shingle bank and beach have been maintained naturally, however research showed that the movement of sediment was increasingly irregular and there was a net loss of 25,000 m³ of beach every year. As flooding, rather than erosion, is the biggest risk here, soft man-made defences constructed of sand and shingle are seen as most effective in absorbing and breaking up the power of the waves.

Current sea defences

The management of the sea defences in Pevensey Bay is unique as they are organised and maintained by the country's first public–private partnership (PPP) coastal management scheme. The project began in 2000 after a realisation in 1997 that a 1-in-20-year storm event could breach the shingle beach. The contract was awarded to a consortium called Pevensey Coastal Defence Ltd (PCDL), was to run for 25 years and had an initial budget of £30 million.

The defences centre on the need to manage and improve the existing shingle beach. PCDL has therefore based its management on working with the natural environment of the local sediment sub-cell. The management is environmentally and aesthetically acceptable and is an example of development which maximises opportunities for future recreational and environmental projects.

The shingle beach stretches the length of Pevensey Bay. The crest height of the defence is six metres and the shingle bank extends seawards an average of 45 m. Traditionally this had been maintained with the use of 150 ageing wooden timber groynes. Since 2000 the project has removed each groyne as it fails, with the aim of retaining about 10, and focuses on a more sustainable soft management approach to work with the coast processes.

Table 3.9 Beach maintenance strategies in Pevensey Bay

Strategy	Notes
Recharge	The natural movement of sediment west to east along the coast is interrupted at Sovereign Harbour. Longshore drift still removes 25,000 m³ of sediment annually from the bay, transporting it to Bexhill. Lorries transport 5,000 m³ around the harbour each year, but the remaining 20,000 m³ of gravel and sand is dredged from the sea floor just off shore. This is the natural source of the original beach material so is a close match (Figure 3.53). In recent years, rather than using bulldozers to push it into place (causing noise and disruption), all 20,000 m³ have been left in situ and redistribution was left to the natural process of longshore drift. This provided more natural sorting of the beach material and saved considerable amounts of time and money.
Recycling	The rate of drift of material along the bay varies from place to place. Some areas have net accumulation, some net loss and others do not change. Dump trucks and bulldozers are used to redistribute sand and shingle from areas of gain. Recycling sediment is required following especially stormy weather, such as Winter 2013–14, and when conditions do not allow the dredger to operate close to shore. If relatively small amounts of sediment need recycling, a 'box' or earth scraper is attached to the bulldozer to limit the use of larger plant machinery.
Bypassing	The southern arm of Sovereign Harbour interrupts the west to east movement of sediment causing it to accumulate. If this was allowed to happen it would eventually spill round into the harbour and silt-up the lock gates. During the winter months (so as not to disrupt the tourist season) trucks are used to transfer between 5,000 and 15,000 m³ of surplus shingle to the west (bypassing the harbour). This is either added directly to the beach or is stockpiled as an emergency source of beach material.
Reprofiling	Destructive waves in winter storms remove material from the crest of the shingle beach and deposit it lower down the beach. Over time wave action would naturally return the material back up the beach, but in the meantime areas of the coast may be left vulnerable. To prevent the shingle being transported away by longshore drift, a bulldozer simply pushes the material back up the beach, re-instating the beach crest. During winter this and recycling can be necessary on an almost daily basis (Figure 3.54).
Groynes	Restoring the 150 groynes along the bay would have cost 40 per cent of the 25-year budget, about £12 million. Instead, as each groyne fails, it is carefully removed so that it does not become dangerous for beach users or break apart and float out to sea. The wood is reused to either repair the remaining groynes or made available to local residents, businesses and farmers to be recycled and reused. Around 10 groynes will be retained to maintain important discontinuities in the coastline. This will create a more open natural beach.
Beach surveys	Twice a month a quad bike with a GPS receiver surveys the beach. The data is used to produce three-dimensional models of the shingle bank and beach. The system works like a satellite navigation system in a car but is accurate to between 15 and 30 mm. These surveys help to maintain the sustainability of the project as replenishment, reprofiling and recycling can be targeted to only where needed.

Figure 3.53 Dredger recharges the beach at Pevensey Bay

Further information and updates on current progress managing Pevensey Bay coast can be found at www.pevensey-bay.co.uk.

The Pevensey Bay sea defences project has won many awards over the years. By May 2019, it had won 15 awards, including gold and silver awards under the Considerate Constructors Scheme, which highlights how the scheme tries to manage the coastline sustainably.

Figure 3.54 Using a bulldozer to reprofile the beach after a storm

Review questions

12 Summarise which stakeholders would have an interest in any proposed coastal management scheme on the British coast.

13 Why do you think the authorities have chosen not to replace the old wooden groynes at Pevensey Bay?

14 What is the important role played by beach surveys and Geographical Information Systems in the management strategy at Pevensey Bay?

Investigating coastal processes

Background

As geography students it would be possible to collect field data to understand processes like those occurring at Pevensey Bay. An investigation could be conducted to investigate the characteristics of a beach and the impacts of processes like wave action and longshore drift. This could include an analysis of:

- the shape of the beach
- the amount of beach material in different locations
- the nature of beach material.

Collecting field data

At regular intervals along the length of the beach (5 or 10 m depending on the length of the beach, time available and number of students) mark out a straight line **transect** from the shoreline to the back of the beach. Record the length of the transect and use a compass to note its orientation. Its position should be recorded on a base map. Using a tape measure, data is collected at regular intervals up the beach. Again, the length of the interval will depend on the length of the transect and time available. The beach may have many obvious changes in slope angle which may be used instead, especially as this may give a more accurate profile when plotted back in class. Always start at the water's edge and work inland, as the tide may begin to come in.

Along the transect a beach profile could be measured and an analysis of the sediment completed. A summary of the method for each is outlined below.

Beach profiling and sediment characteristics

Two students are needed to complete each transect and could follow the method below:

- Student 1 stands at the seaward end of the transect with a ranging pole.
- Student 2 takes a second ranging pole further up the beach along the transect to the first break of slope.
- A tape measure is used to measure the distance between the ranging poles.
- Student 1 uses a clinometer to measure the angle between the tops of the two ranging poles to give the gradient of that section of the transect (Figure 3.55).

  (Also see Figure 2.46, page 86, for a similar technique on sand dunes.)
- This process is repeated for every break of slope to the top of the beach.

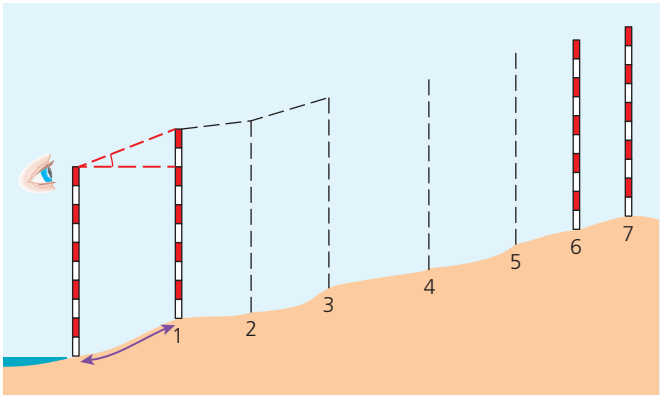


Figure 3.55 Measuring the gradient of the beach

In order to investigate the impacts of wave action and longshore drift on the sorting of sediment on the beach, data could be collected while measuring gradient along the transect. If a statistically significant number of transects are completed at 10 or more sites located systematically along the beach (for example, at 10 m intervals) an analysis of sediment change along the beach can be conducted. Collecting data along each transect will also allow an analysis of how waves and tides sort material up the beach.

- At each collection point a sample of at least 10 pebbles is selected.
- A quadrat could be placed adjacent to the ranging pole and the pebbles selected from within it.
- Using a ruler, or callipers, measure either the long-axis of each pebble or multiple axes to give an average size (Figure 3.56). (If the beach material is fine then a set of graduated sieves would be used to sort a sample of sediment from each quadrat to ascertain the average size of sediment at each site.)
- Students could also record pebble roundness, where each pebble is assessed against an exemplar Power's Roundness scale. (The scale ranges from very angular, sub-angular or rounded, to very rounded.)

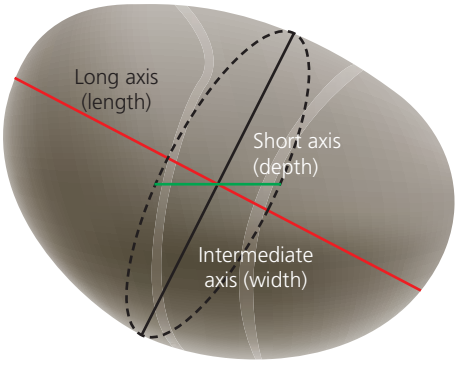


Figure 3.56 Measuring axes of pebbles

Table 3.10 Recording sheet for beach data

Date:		Weather conditions:										
Transect number:		Grid reference for start of transect:							Transect length:			
Beach profile	Break of slope on transect	Waters-edge	1	2	3	4	5	6	7	8	9	10
	Distance from previous break	NA										
	Gradient											
Sediment	Average pebble size (cm)											
	Average pebble roundness											

Findings

Back in class the results can be analysed to draw conclusions about the characteristics of the beach from waters-edge to the top of the beach. Also, the profile data can be used to calculate the cross-sectional area at each location, which could then give an assessment of the amount of beach material at that point. If time allows, measurements taken at the same location on different dates could give an assessment of how the amount of sediment at each location changes over time, to help draw conclusions about the effects of wave action and longshore drift. If data can only be collected on one visit, then an analysis of profile and sediment data for each transect along the beach could also allow conclusions to be drawn about the effects of longshore drift and wave action.

Case study of a contrasting coastline beyond the United Kingdom: The Sundarbans, Bangladesh

Links

For links between Sunderbans and lived experience, see Chapter 8, Changing places, page 354.

Background

The Sundarbans is a coastal zone occupying the world's largest delta, that extends over 10,000 km^2 of southern Bangladesh and India on the Bay of Bengal. The delta is formed from the sediment deposited by three of the world's great rivers, the Ganges, Brahmaputra and Meghna (Figure 3.57). The natural climax ecosystems of the Sundarbans are mangrove forests and swamps.

Figure 3.57 Map showing the Sundarbans Delta, Bay of Bengal

Coastal processes

Tidal action is the primary natural process that shapes this distinctive coastal landscape. A dense, well-developed, network of interconnecting river channels flows across the clay and silt deposits. Traditionally the location of the network of main channels remained relatively static due to the silts and clays being quite resistant to erosion. The larger channels are generally straight and up to two or more kilometres wide, flowing generally north to south due to the strong tidal currents. The extensive network of interconnecting smaller channels (or **khals**) drains the land with each powerful ebb tide.

The non-cohesive sediments like sand are washed out of the delta and deposited on banks, or chars at the river mouths, where the strong south-westerly monsoon winds then blow them into large ranges of sand dunes. With the

Figure 3.58 The Sundarbans of Bangladesh

protection of the dunes, finer silts washed into the bay are deposited, where wave action then adds and shapes further deposits of sand to form new islands. Vegetation establishes itself and eventually, if the natural succession can proceed, the dense mangrove forests, home to the endangered Royal Bengal Tiger for which the area is famous, can develop (Figure 3.58).

This is a unique coastal landscape where for millennia these great rivers have brought rich sediment that, once shaped and maintained by tide and wave action, has created the diverse mangrove forests that have sustained local populations for generations. However, the equilibrium of the natural processes that exist here is very delicate and the increasing pressures that are placed on it may be jeopardising its very existence.

Challenges and opportunities of the Sundarbans

The mangrove forests of the Sundarbans are an extremely important ecosystem for the local populations, providing many economic and environmental opportunities. If managed sustainably this coastal zone provides a wide range of goods and services (Figure 3.59).

Despite the wealth of goods and services outlined in Figure 3.59, many from outside view this area of coastal Bangladesh as uninhabitable. However, traditionally (and perhaps surprisingly) coastal erosion has not been a problem; the main challenges facing the local populations are outlined in Table 3.11.

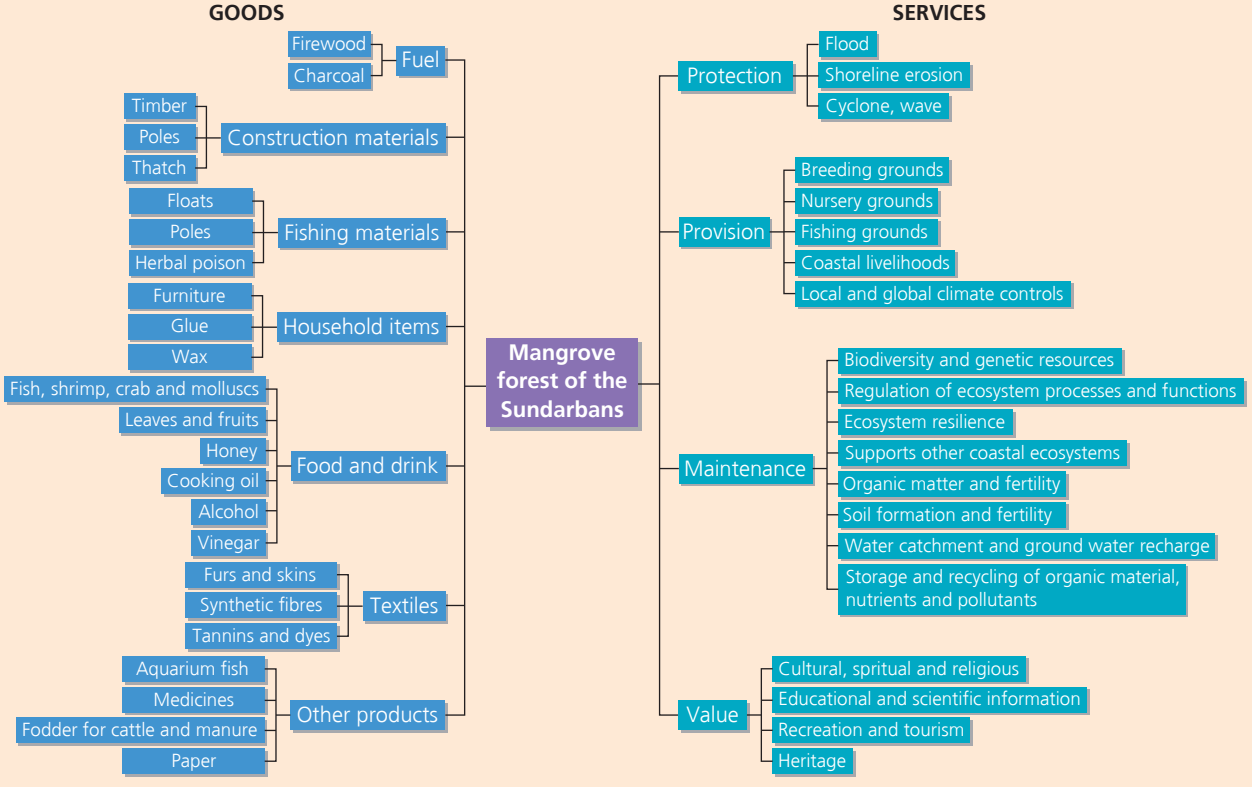


Figure 3.59 Goods and services of the mangrove forests of the Sundarbans

Table 3.11 Natural and human-induced challenges for the people of the Sundarbans

Natural	Human
• Coastal flooding • Cyclones • High levels of salinity in the soil • Instability of the islands • Accessibility and remoteness • Human-eating tigers	• Over-exploitation of coastal resources from vulnerable habitats • Conversion of wetlands to intensive agriculture and settlements • Destructive fishing techniques • Lack of awareness of the environmental and economic importance of the region • Resource-use conflicts • Lack of awareness of coastal issues by decision makers

The human response to challenges of the Sundarbans

Traditionally many local populations have lived successful productive lives in the Sundarbans of Bangladesh. Life is not always easy, but by living with their environment rather than against it they have coped with many of the natural challenges they face (Figure 3.60).

Figure 3.60 Women catching shrimp larvae in the Passur River, Chandpai Village, Sundarban National Park, Bangladesh

Resilience

The wealth of goods and services outlined in Figure 3.59 (page 132) has allowed local populations to remain resilient to the challenges of this low-lying coastal landscape. In fact the mangrove forests provide a high level of resilience as they provide significant protection and shelter against:

- storm winds
- floods
- tsunamis
- coastal erosion.

It is said that a density of 30 trees per 0.01 hectares can reduce the destructive force of a tsunami by up to 90 per cent. In fact the coastal communities of the Sundarbans are said to be more resilient to natural disasters than other coastal communities elsewhere in Bangladesh.

The fertility of the soil and ecological diversity also provide a plentiful supply of a large range of nutritious foods. Despite the constant threats of flooding and storms the forests have an economic value, even if used for traditional activities such as fishing, gathering crustaceans, or timber and tannin production. Some estimate that just one hectare of mangrove forest has an annual economic value of over $12,000. This suggests that unlike many other local communities in other parts of Bangladesh, one of the poorest countries in the world, the mangroves could provide resilience against poverty, and opportunities for sustainable economic development in the future for the people living there.

Mitigation

Although vulnerable to many challenges in the coastal Sundarbans, its people moderate their risk in a number of ways. Some suggest that it is the strength of each community's resources that helps to mitigate the worst impacts of this challenging environment. Primarily, prior to more recent human pressures like deforestation, overfishing, intensive agriculture and settlement growth, the communities of the Sundarbans could utilise a number of **open access** natural resources. These included:

- *khas* land (government-owned lands that are supposed to be protected for use by local populations)
- wetlands and fisheries
- forests.

The ever-present threat of natural disasters has meant that in some areas of the Sundarbans there has been significant investment in the physical infrastructure (roads, telecommunications, schools, hospitals, cyclone shelters, flood protection and tube wells) which also mitigate the risks they face. Many communities of the Sundarbans are said to have good levels of social capital from the legal frameworks and services provided by a number of formal government and NGO organisations, alongside the traditional laws, tenets and social sanctions that communities have used for generations to manage their use of the region. As noted above there is some economic value to the resources of the area and together with financial resources like access to micro-credit, some Sundarbans communities have a greater economic safety net than other vulnerable groups in Bangladesh. Figure 3.61 (page 134) summarises the **livelihood assets** that help mitigate the challenges of the Sundarbans.

Figure 3.61 Livelihood assets of communities in the Sundarbans

In recent decades the level of resilience provided by these livelihood assets may be decreasing as poverty and marginalisation of some coastal communities increases due to:

- shrinking of the open access resources
- degradation of ecosystems
- corruption of both local and national political institutions
- conflicts over land ownership
- increasing deaths by tigers (if wives are widowed they can often face very limited opportunities in this male-dominated society, especially as many marry very young before completing their formal education).

Adapting to the Sundarbans

In the past it appears that people of the Sundarbans have adapted well to their coastal environment and in some places many have prospered on the natural bounty. However, it is almost inevitable that the impacts of climate change may pose the greatest risk to current and future generations in southern Bangladesh. Table 3.12 summarises some of the future challenges facing the region alongside possible adaptations that could mitigate some of the risk and improve the community's resilience to climate change.

Generations of people have been sustained by the coastal mangrove forests of the Sundarbans. However, population pressure and the unpredictable impacts of future climate change may prove too hard to adapt to for many.

Table 3.12 Future challenges and adaptations in the Sundarbans

Challenges	Adaptations
• Increased frequency and intensity of floods • Flood waters remaining for longer • Permanent embankments built by the commercial shrimping industry are encouraging the deposition of silt, thus raising water levels in the rivers • Rising temperatures contribute to increased salinity levels in the soil • Increased pesticide and fertiliser use is affecting water quality • Changes to seasonal patterns of rainfall	• Grassroots NGOs run education programmes to encourage farmers to return to more traditional ecologically-friendly methods • NGOs provide education and training about community level preparation for natural disasters – planning and preparation can significantly mitigate the impacts of natural disasters • Organisations like USAID train communities to become resilient to future climate shocks, including providing training on improved agricultural techniques for tens of thousands of farmers • Build more multi-purpose cyclone shelters equipped with communication equipment, which also serve as primary schools and community facilities (Figure 3.62). (The government built 100 shelters in coastal Bangladesh between 2011 and 2015, 750 more were proposed in 2019) • New salt-tolerant rice varieties that can survive being submerged in sea water for over two weeks • NGOs build latrines on higher ground and educate communities about water and sanitation, and water-borne diseases • Installing storage tanks for rainwater in areas at most risk from inundation by salt water • Distributing water-tight containers to store important belongings and papers during floods

Figure 3.62 Raised cyclone shelter nears completion on the edge of the Sundarbans

Opportunities for sustainable development in the future

The threats to the mangrove forests and peoples of the Sundarbans may appear bleak, however with timely intervention and careful management there are opportunities for the sustainable development of this unique coastal zone in the future:

- The region has many international and national designations already in place.
- There are opportunities for eco-tourism in the mangrove forest and wetland areas.
- There has been recent investment to improve communications in the region.

Further research

Find out more about the Sundarbans at:

www.eoearth.org/view/article/156336/

http://uddin.digital.conncoll.edu/sundarbans/

www.usaid.gov/bangladesh/environment-energy-climate-resilience

Fieldwork opportunities in coastal landscapes

Studies of coastal systems and landscapes provide many opportunities for collecting field data. Fieldwork could be used to research a range of human and physical geographical enquiries. There is potential to collect a broad range of quantitative and qualitative data from primary and secondary sources (as outlined above). Coastal fieldwork activities could fall into one of the three categories described below (with some examples):

- Investigating coastal processes:
  - the characteristics of waves and their impacts
  - the nature and impact of longshore drift
  - factors affecting rates of coastal erosion
  - factors affecting sub-aerial processes
- Investigating coastal landforms:
  - beach profiles
  - sediment analysis
  - features of headlands and bays
  - features of wave-cut platforms
  - evidence of sea level change
- Investigating coastal management:
  - the methods of coastal management used
  - impacts of coastal management
  - the success, or cost–benefit analysis of coastal management
  - the views of different stakeholders about coastal management.

Further reading

The texts below develop many of the ideas set out in the chapter above. Some are quite academic and specialist texts but your local library may be able to find copies should you wish to take your studies further.

Harvey, A. (2012) *Introducing Geomorphology: A Guide to Landforms and Processes*. Dunedin Academic Press

Holden, J. (2012) *An Introduction to Physical Geography and the Environment, Fourth edition*. Pearson

Davidson-Arnott, R. *et al* (2019) ***Introduction to Coastal Processes and Geomorphology, Second Edition***. Cambridge University Press

Woodroffe, C. D. (2002) *Coasts: Form, Process and Evolution*. Cambridge University Press

Stiff, P. (2018) ***A-level Geography Topic Master: Coastal Landscapes***. Hodder

Question practice

A-level questions

1 Summarise a positive feedback mechanism in a coastal system. [4 marks]

2 Figure 3.63 shows data about the number of people expected to be flooded each year if global temperatures increase by 2°C above pre-industrial levels.

Figure 3.64 Shows data about the estimated annual costs of coastal flooding in 2100 if global temperatures increase by 2°C above pre-industrial levels.

Analyse the data shown in Figures 3.63 and 3.64. [6 marks]

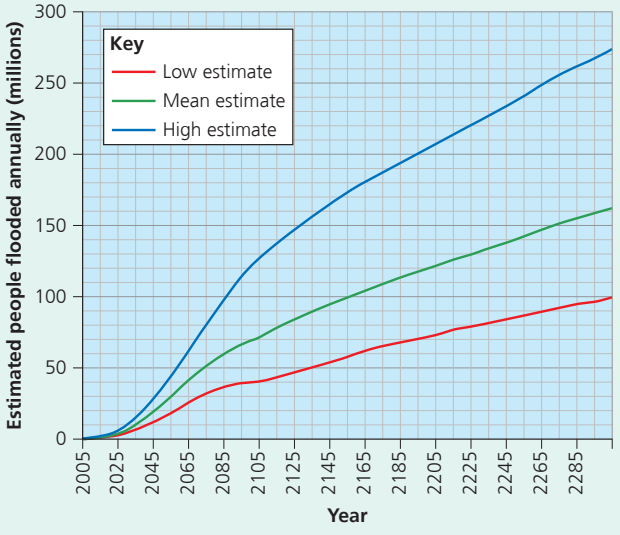


Figure 3.63 Estimated number of people affected by coastal flooding by 2295

Source: The Royal Society

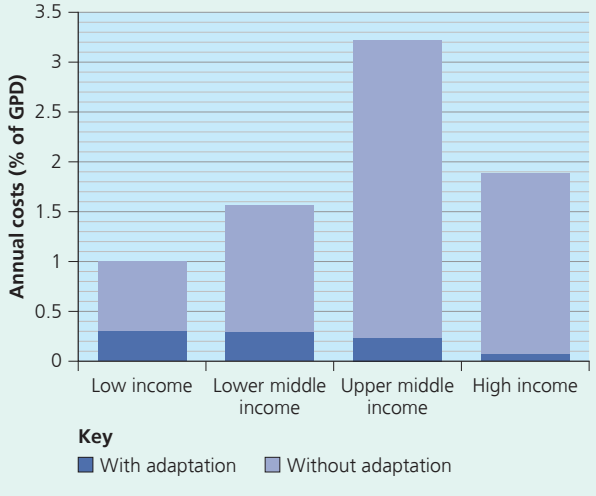


Figure 3.64 Estimated annual cost of coastal flooding in 2100 as a percentage of GDP for low income, lower middle income, upper middle income and high-income countries, with and without human adaptation

Source: Environmental Research Letters Vol. 13 – IOPscience

3 Using Figure 3.65a and Figure 3.65b and your own knowledge, assess the extent to which the coastal processes operating on these two stretches of coast are different. [6 marks]

Figure 3.65a A stretch of coast in south-west Iceland

Figure 3.65b A stretch of coast in west Wales

4 'Processes of coastal erosion are the most important factor in shaping coastal landscapes'.

With reference to a local scale coastal landscape you have studied, to what extent do you agree with this view? [20 marks]

5 'Future climate change will create more challenges than opportunities for people living in coastal landscapes.'

With reference to a coastal landscape beyond the UK that you have studied, to what extent do you agree with this view. [20 marks]

Chapter

4 Glacial systems and landscapes

Glacial landscapes are dynamic environments in which the landscape continues to develop through contemporary processes but which mainly reflect former climatic conditions associated with the Pleistocene era. The operation and outcomes of fundamental geomorphological processes and their association with distinctive landscapes are really observable. In common with the other physical sections of this book, glaciers can be seen as natural systems. Glacial landscapes present the limited number of human inhabitants with unique economic, social and political opportunities and challenges.

Many of these areas are increasingly referred to as being **fragile environments** with important issues facing their biodiversity and sustainability.

In this chapter you will study:

- the nature and distribution of cold environments
- glaciers as natural systems and glacial systems including glacial budgets
- geomorphological processes and glacial landscape development
- human impacts on glaciated landscapes.

4.1 Introduction to glacial systems and landscapes

Glaciers as natural systems

The introduction to Chapter 1: Water and carbon cycles outlined how many aspects of physical geography can be studied using a **systems** approach and you should read that section carefully before reading further here. This chapter explores glacial systems and landscapes and we will see that glaciers are **open systems** with a range of **inputs, stores/components, flows/transfers and outputs**. When there is a balance between the inputs and outputs then the system is said to be in a state of **dynamic equilibrium**. If one of the elements of the system changes, for example there is increased ice accumulation but there is no corresponding change in the amount of meltwater, then the mass of the glacier may change and the equilibrium is upset. This is called **feedback**. Figures 4.1a and Figure 4.1b (page 138) illustrate the open system of a glacier in a glacial landscape and are a simple example of a positive feedback system in a cold environment. They highlight many of the ideas that are explored in detail in the rest of the chapter.

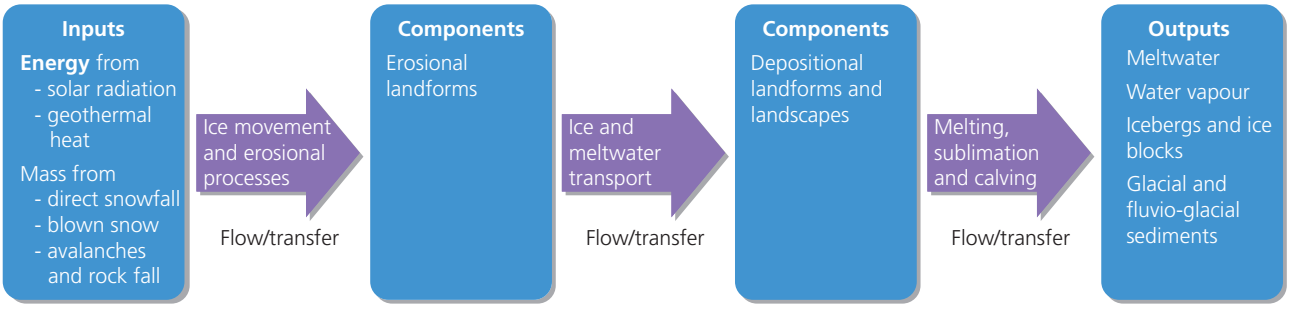


Figure 4.1a A glacier as an open system in a glacial landscape

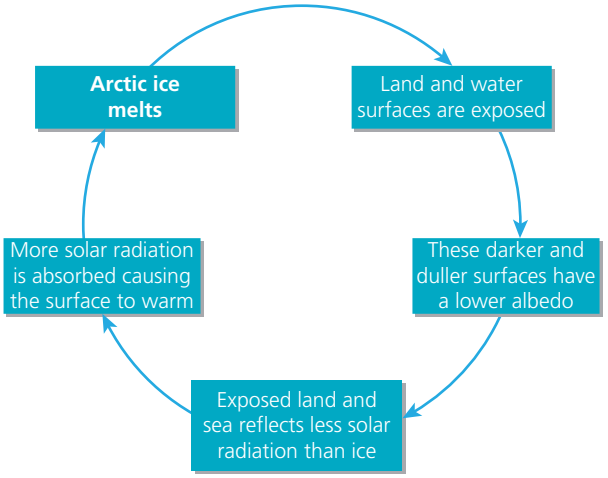


Figure 4.1b Positive feedback in a cold environment – the 'ice–albedo feedback'

Like all physical environments, characteristic glacial landscapes are a combination of a range of constantly changing erosional and depositional landforms within a geomorphological system. Landscapes are shaped and reworked by the action of physical processes driven by the input of energy as the system tends towards dynamic equilibrium. There are a variety of characteristic landscapes in cold environments, including combinations of landforms that form glacial, glaciated, fluvio-glacial and periglacial landscapes, each with their own unique assemblages of erosional and depositional features.

Review question

1 Can you summarise any other examples of feedback in glacial landscapes?

4.2 The nature and distribution of cold environments

Physical characteristics of cold environments

The term 'cold environments' covers a broad range of landscapes, climates and ecosystems. Some receive large amounts of precipitation as snow while others receive the same low levels of precipitation as some hot deserts. By definition all are 'cold', but some of the coldest parts of Antarctica can be below –90°C with almost no areas of the whole continent ever achieving above 0°C, while some have very seasonal temperature ranges and fluctuate around 0°C.

Latitude, altitude and climate characteristics

The distribution of cold environments is varied, and they are found in different locations with different climate characteristics dependent on their latitude and altitude. Cold environments can be grouped into the following classifications:

- **Polar:** Found at the Earth's northern and southern extremes, almost entirely within the Arctic and Antarctic circles, above approximately 66° north and south of the Equator.
  - The polar Arctic region of the north largely consists of the Arctic Ocean that encircles the North Pole, and some islands and most northerly extremes of Asia, Europe and North America. The central polar area is continuously covered by a drifting ice pack. This sea ice averages 3 m in thickness and doubles in extent in winter to around 13 to 14 million km². The Arctic area has a very wide range of climates, but it is continuously cold with mean annual temperatures ranging from 4°C to –28°C. There are similarly wide-ranging levels of precipitation, with the overall average being only around 100 mm, with the Atlantic area of the region receiving significantly more.
  - The southern polar region includes the massive frozen wilderness of the **ice sheet** and surrounding ice shelves of Antarctica which cover 13 million km², and the periodically frozen expanse of the Southern Ocean. With such a large expanse of uninterrupted ocean this southern zone experiences strong westerly winds. Together with the coldness of the ocean and large landmass, this gives Antarctica a much colder climate than the Arctic. The continuously cold waters of the Southern Ocean mean that sea ice develops rapidly in the winter from about 2 million km² in February to around 16 million km² in September. The fluctuation in the extent of the sea ice in the Arctic is much less, due to the influence of the surrounding continents. Another recently observed difference between the poles is that the extent of the winter sea ice around Antarctica is increasing while that in the north is shrinking. The interior of Antarctica includes mountains and high plateaux. The combination of the atmospheric circulation patterns and the high altitude means that the South Polar Plateau is the coldest location on Earth, with mean annual temperatures in places of

around –55°C. The coastal areas of Antarctica are much milder with average annual temperatures of around –10°C. Average precipitation for Antarctica as a whole is around 200 mm/year. Rates range from three to five times this in coastal regions to very arid conditions on the Polar Plateau.

- **Alpine:** These are the most instantly recognisable 'glaciated' landscapes. Commonly they are areas of high relief, with peaks often rising well over 3,000 m from valley floors. The most striking landscapes are also associated with tectonically active mountain ranges, including the Himalaya, Karakoram and Tibet plateaux in Asia, the Rockies, Cascades and Andes in the Americas, and the New Zealand Alps. It is the combination of tectonic uplift and rapid rates of erosion by water and ice that creates such spectacularly well-developed glaciated features. Older mountain ranges in more tectonically stable regions have lower relief and the features of their well-developed glaciated landscapes have often been *softened* and smoothed by subsequent weathering and erosion. Alpine features are found in areas that were once glacial, including the mountains of northern and western Britain, and those recently and currently occupied by glaciers including the European Alps and Norwegian mountains. Alpine landscapes can include small ice caps, mountain glaciers and tundra environments.
  - Alpine regions with active glaciers today occur at almost any latitude, but they are at high enough altitudes for snow and ice to remain throughout the year. Obviously temperatures drop below freezing for significant periods, often below –10°C, but in summer temperatures can rise well above freezing, with those in lower latitudes regularly above 20°C in summer months.
  - It is suggested that Alpine landscapes develop in a sequence following an *Ergodic* model (Flint, 1971), from a preglacial fluvial landscape which is then modified episodically during cycles of **glacials** and **interglacials** during **ice ages**. Recurring **glacial erosion** takes these landscapes through three stages to a mature alpine landscape with many of the features explored later in this chapter.
- **Glacial:** These are environments that are currently covered by ice sheets and **glaciers**. These are permanently cold enough for substantial masses of ice to remain on the landscape throughout the year.

The two most significant areas are the ice sheets and ice shelves of Antarctica and Greenland. At present about 10 per cent of the land area of Earth is covered by ice, including glaciers, ice caps and ice sheets covering over 15 million km². Table 4.1 summarises the distribution of this ice.

Links

Ice sheets, ice caps and glaciers are a global store of water. See Chapter 1, Water and carbon cycles – The major stores of water, pages 4–8.

Table 4.1 Land area covered by ice sheets and glaciers

Region	Estimated area (million km²)
Antarctica	13.5
Greenland	1.74
Rest of Arctic region	0.24
Asia	0.125
Alaska	0.041
Rest of North America	0.03
Andes	0.03
Scandinavia	0.003
Alps	0.004
Australasia	0.0012
Africa	0.0001
Total	15.7143

- **Periglacial** (including **tundra** regions): These are generally found in dry high-latitude areas that are not permanently covered by snow and ice, but are cold enough for periglacial processes to occur. Currently areas of northern Alaska and Canada, northern Scandinavia, Siberia and the islands of the Arctic Ocean such as Spitsbergen have active periglacial regions with tundra ecosystems. In these places the cold conditions create permanently frozen ground known as **permafrost**, which is often overlain by a more mobile unfrozen layer of soil called the **active layer**. Due to low levels of precipitation, cold temperatures and often thin poor soils, vegetation is sparse in periglacial environments. Some areas in lower latitudes but at higher altitudes do experience some periglacial processes.

The climate of cold environments

The main features of the climate in cold environments are summarised in Table 4.2 (page 140).

Table 4.2 Characteristics of climates in cold environments

Characteristics	Cold environment climate type	
	Polar	Tundra
Temperatures	• Mean monthly temperatures are below freezing all year • Extreme cold in winter with monthly average temperature below –50°C	• Long cold winters with temperatures averaging –20°C • Brief, mild summers with temperatures rarely above 5°C • At least eight months of the year the temperature remains below 0°C
Precipitation	• Precipitation below 150 mm/year; all of it falling as snow (at the South Pole, precipitation is around 50 mm/year)	• Small amounts of precipitation, less than 300 mm/year, most of which falls as snow
Winds	• Strong winds blow outwards from the centre of the continent • Winds of over 200 km/hr have been recorded • There is a substantial wind-chill factor • The winds whip up the powdery snow to create frequent blizzards and white-outs	• Frequent strong winds increase the wind chill

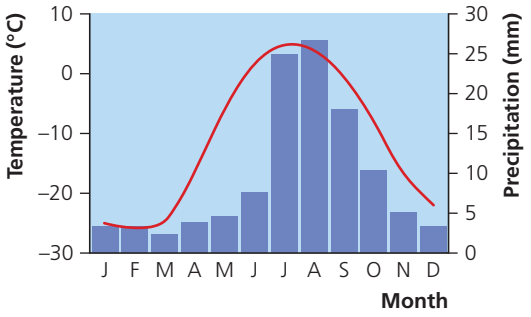


Figure 4.2a The climate of tundra regions: Barrow, Alaska

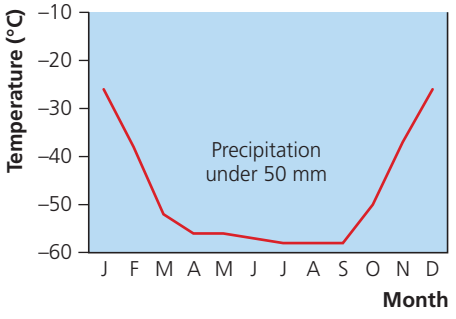


Figure 4.2b The climate of polar regions: Amundsen–Scott Station, South Pole

The causes of such climates are as follows.

- **The low level of insolation:** In summer the Sun remains at a low level in the sky even though there is a period of continuous sunlight. Because the rays strike the surface obliquely, there is a wider surface area to heat. In addition, the longer passage through the atmosphere allows for increased absorption, scattering and reflection of radiation. Less insolation therefore reaches the surface (see Figure 4.3). Continuous

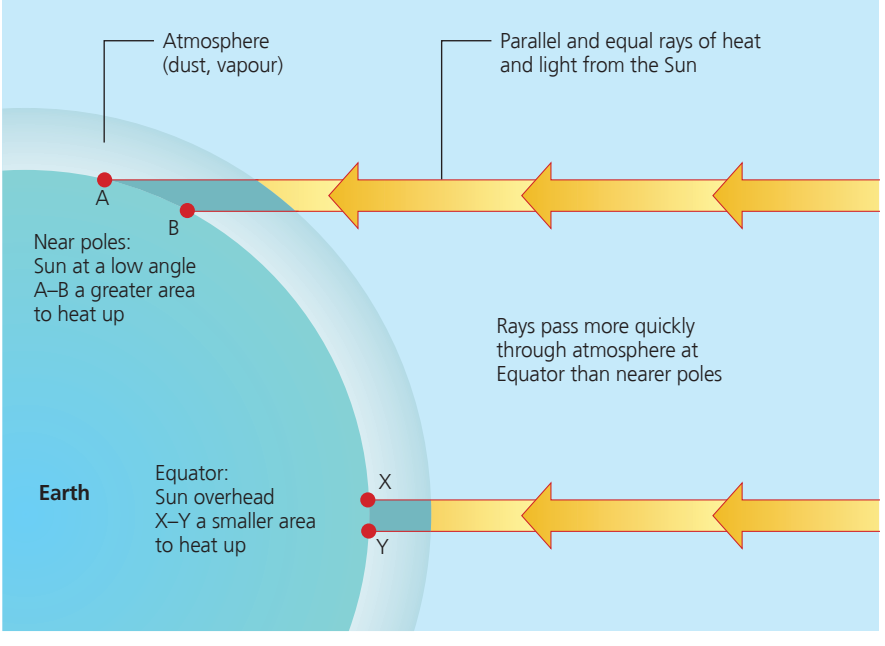


Figure 4.3 The effects of latitude on insolation

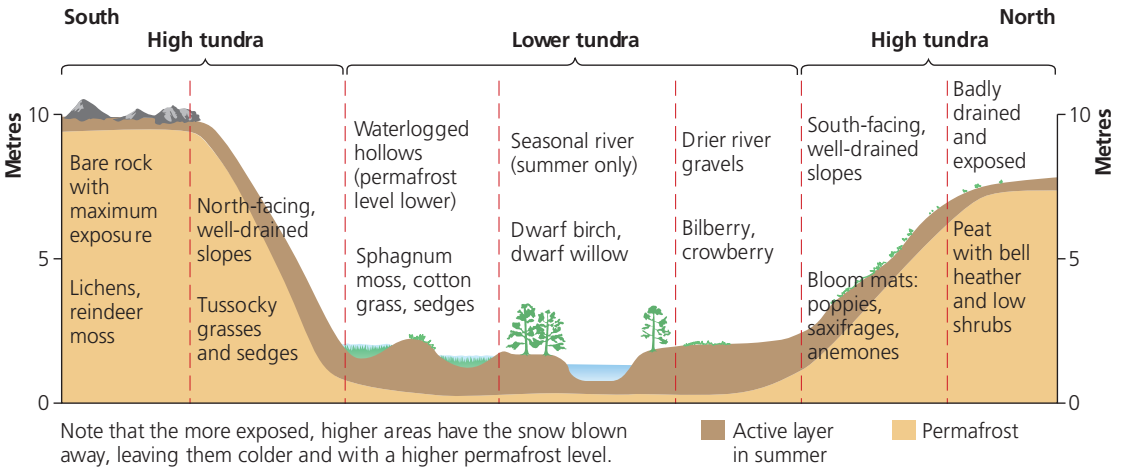


Figure 4.4 Small-scale variations in tundra vegetation

summer radiation ('the land of the midnight Sun') makes the tundra days seem pleasantly warm, but this only lasts for a short period. In winter there is a corresponding period when there is no incoming solar radiation as the Sun does not appear above the horizon.

- **The high albedo:** In areas of continuous snow cover, much of the incoming solar radiation is reflected off the ice/snow surface. This reduces the amount that can actually contribute to the warming of the atmosphere.
- **The high pressure systems** of polar regions means that frontal systems rarely penetrate these areas, giving low levels of precipitation.
- **The coolness of the air** is another factor in low precipitation. Very cold air can only hold low levels of water vapour. When precipitation occurs it is often little more than very light powdery snow.
- **Katabatic winds:** In Antarctica, masses of cold dense air flow down the valleys and off upland areas. Such movements are known as katabatic and are strong in Antarctica where there is a real difference between the interior and coastal areas. With few obstacles to hinder air movement, such winds can exceed 200 km/hr.

The vegetation of the Tundra

The main features of the vegetation are:

- a very low level of productivity
- very few plant species and therefore low biological diversity
- an absence of full grown trees (tundra means 'treeless plain' in Sami). There are, however, a number of dwarf varieties to be seen
- most flowering plants are perennials – they flower year after year. They have hardy seeds armoured by a thick seed case
- there are five types of plants, each occupying its own specialised niche: lichens, mosses, grasses, cushion plants and low shrubs (Figure 4.4). Lichens are the pioneer plants as they colonise bare areas. They have no roots and are able to absorb water and nutrients directly into their foliage. Together with mosses, they initiate soil formation.

The tundra has many features that make plant growth difficult, including:

- low levels of insolation
- a very short growing season of not more than 50 to 60 days
- low levels of precipitation and soil moisture frozen for much of the year
- waterlogged soils, as thawed water is unable to penetrate to depth because of the permafrost and with low temperatures, evaporation levels are low
- strong desiccating winds and permafrost not far below the surface.

Tundra plants have a number of features as a result of their adaptation to the environmental difficulties:

- ground hugging to avoid the cold, desiccating winds and to take advantage of any warmth close to the surface
- slow growth rates
- shallow roots because of the presence of permafrost
- the ability to photosynthesise at extremely low temperatures, even under a thin covering of snow
- reduced transpiration due to the small leaves and a thick cuticle

- low albedo of plant surfaces meaning that they can absorb more radiation than the surrounding surfaces
- perennial plants which can store food from year to year.

Tundra soils

A distinctive soil type has developed in tundra areas (Figure 4.5). Its major features are:

- a lack of clearly differentiated soil horizons (layers)
- a thin surface organic layer which is often very acidic
- a uniform blue grey colour
- waterlogged in summer
- gleyed (because of the waterlogging, iron compounds are reduced to their ferrous form which is grey – as opposed to ferric compounds which are red-brown).

Some reasons for the characteristics of tundra soils:

- Surface litter is restricted by the limited vegetation.
- The limited amounts of organic matter in the soil are due to the slow rate of decay of vegetation and the presence of only a few soil organisms.
- With few soil organisms to act as mixing agents, the soil horizons are poorly differentiated.
- The freezing of the soil in autumn leads to it being churned and the horizons become further distorted.
- Where the soil overlies bedrock, frost-heave raises fragments to the surface giving rise to typical periglacial features.

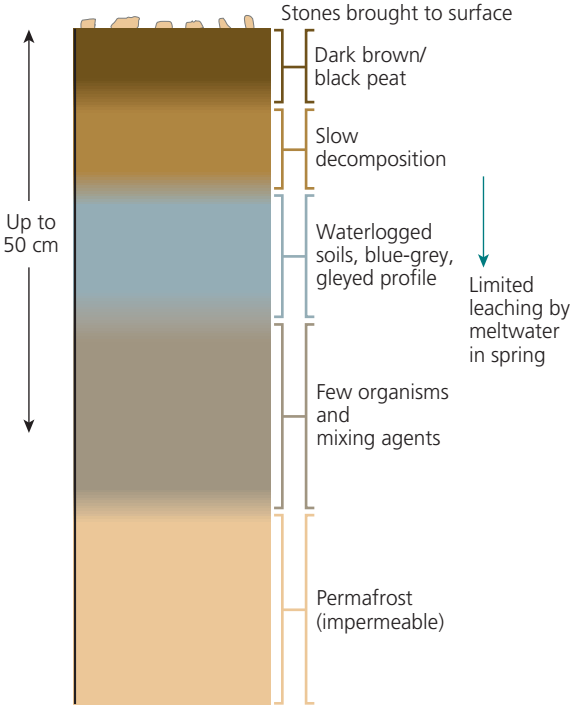


Figure 4.5 Tundra soil profile

Key terms

Fragile environments – Those that are easily damaged or disturbed and then difficult to restore once destroyed.

Glacier – A large mass of ice on land moving downhill due to the influence of gravity.

Ice ages – A common term for periods when there were major cold phases known as glacials and ice sheets covered large areas of the world, interspersed with warmer interglacials. The last 2 million years, the Quaternary Period, contains the Pleistocene epoch lasting from about 1.8 million years to about 11,500 BP (before present). This was characterised by a series of episodes of glaciation, the last major glacial beginning around 120,000 BP reaching the **Last Glacial Maximum** about 18,000 BP. At that time about a third of the Earth's surface was covered by ice, compared with just a tenth today.

Ice sheet – A continental-size mass of ice, covering at least 50,000 km² that is dome-shaped with flow of ice outward from the centre. Ice sheets can be up to 2,000 m thick today; those in the Quaternary glaciation could have been twice that thickness.

Insolation – The radiant energy emitted by the Sun; i.e. incoming solar radiation.

Periglacial – Processes and landforms associated with the fringe of, or area near to, an ice sheet or glacier.

Tundra – A Lapland term for the climate and vegetation type found across extensive areas of northern Siberia, Scandinavia, Canada and Alaska. Tundra lies between the region of continual snow and ice of the Polar regions and the northern limit of tree growth (or the *taiga* environments) and at high altitudes above the tree line in the Alps, Rockies, Andes and Himalayas. Similar ecosystems are found around the ice-free fringes of Antarctica. It is characterised by a stony or marshy surface with mosses, lichens and other low-lying vegetation including shrubs and dwarf trees. Tundra has short summers but mean monthly temperatures do not exceed 10°C, which is just warm enough for snow and the surface layer of the permafrost (active layer) to melt. Currently tundra covers about 8 per cent of the land area of Earth, but during the glacials of the Quaternary glaciation tundra covered huge areas of central North America and Europe, including the British Isles.

Global distribution of past and present cold environments

Throughout the Pleistocene glaciations (during the Quaternary Period) between 1.6 million and about 12,000 YBP (years before present) glacials and interglacials have seen cold environments expand and

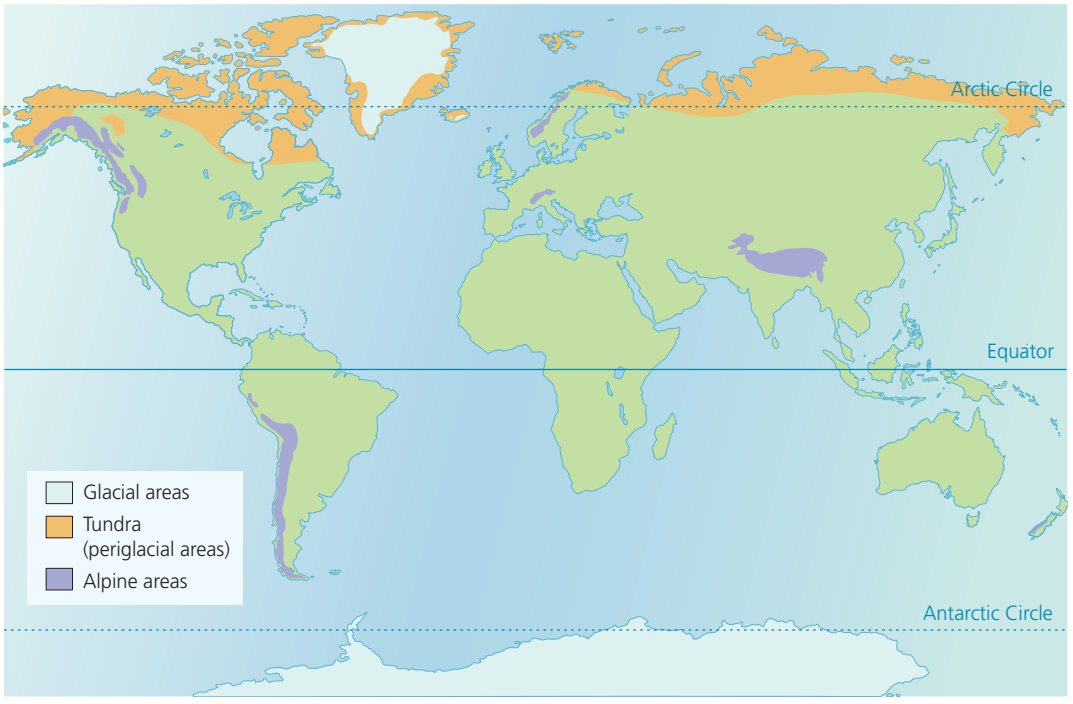


Figure 4.6 Cold environments

contract across the surface of the Earth. There is still some debate about the number of cycles and extent of ice coverage in some locations during this epoch. Figure 4.6 shows the current extent of cold environments.

Figures 4.7a and 4.7b show some recent research that describes how the extent of the distribution of ice on Earth has changed over the last 18,000 years or so, since the **Last Glacial Maximum**.

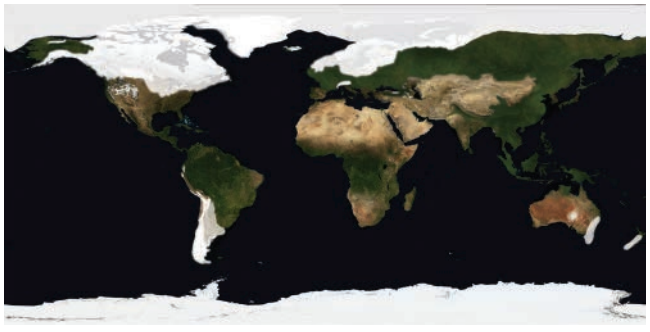


Figure 4.7a Ice coverage from c. 18,000 years ago

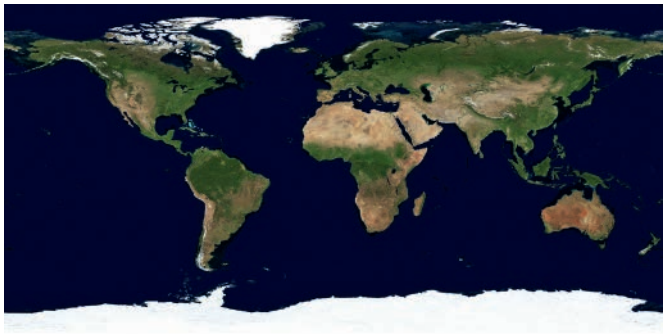


Figure 4.7b Ice coverage January 2019

Skills focus

Use the internet to investigate how scientists are using satellite technology and geo-spatial and digital imagery to map the ever-changing pattern of ice on Earth. Good places to start are: www.antarcticglaciers.org/glaciers-and-climate/glacier-recession/observing-glacier-change-space/

Review questions

2 Explain the distribution of different cold environments.

3 What are the relationships between climate, soils and vegetation in cold environments?

4.3 Systems and processes in glaciers and glacial landscapes

Glaciers as natural systems

In the introduction to Chapter 1: Water and carbon cycles, you will have read that many aspects of physical geography, water and coastlines, for example, can be studied using a systems approach. This is because certain aspects of physical geography are highly complex and geographers have attempted to simplify aspects of them so that relationships between components can be better understood. A system is an

assemblage of interrelated parts that work together by way of some driving process. Before you read on in this section about glacial systems and how they relate to the development of the landscape, you should also read the relevant material at the beginning of Chapter 1 under the heading 'Systems frameworks and their application'.

Glacial systems

In order to understand how, when and where glaciers and ice sheets form, and what effect they then have on the landscape it is useful to view glaciers as **systems**. Like other physical systems glaciers have **inputs**, **stores**, **transfers** and **outputs**. Figure 4.8 shows a cross-section through an ice sheet to illustrate an idealised diagram of the **open system** of a glacier or ice sheet.

Snow and ice are the most important inputs to the glacier system. This comes from:

- direct snowfall
- blown snow
- avalanches from slopes surrounding the glacier.

Together these inputs are known as **accumulation**. These inputs are then transferred down-valley, under the influence of gravity, by glacier movement. Mass is lost from the system by either melting and evaporation, or the *calving* of ice blocks or icebergs. Together these outputs are known as **ablation**.

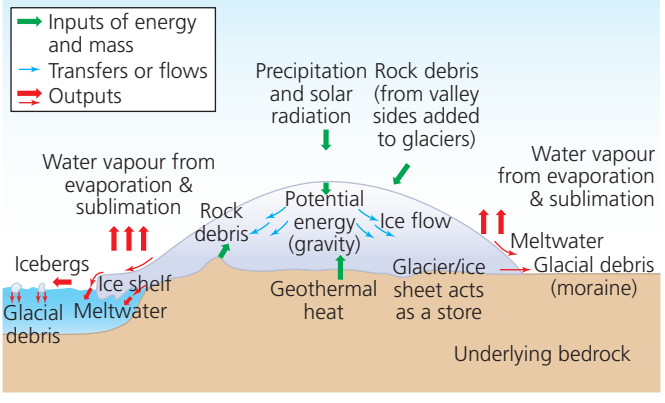


Figure 4.8 The glacial system

Links

There is a link between glacial budgets and changing global water stores. See Chapter 1, Water and carbon cycles, pages 4–10.

Glacial budgets

As glaciers are open systems their size and mass will increase or decrease depending on the balance of inputs and outputs. During glacials the climate starts to become colder and in winter more precipitation falls as snow. Summers also begin to shorten, so there is less time for the snow to melt allowing it to remain year on year. Initially, this leads to permanent snow cover in upland areas, the lower edge of which is known as the **snow line**. As annual mean temperatures continue to drop, the snow line moves down the slope.

Currently the snow line is at sea level in Greenland but at 6,000 m at the Equator. Britain's climate is currently too warm for there to be a snow line, but it is estimated that if the Scottish mountains were about 250 m higher there would be. The aspect of a slope determines the altitude of the snow line, for example, in the northern hemisphere it is lower on north-facing slopes and higher on south-facing slopes as they receive more insolation. In the European Alps the more habitable south-facing slope is known as the *adret* slope and the more inhospitable north-facing slope is called the *ubac* slope (Figure 4.9).

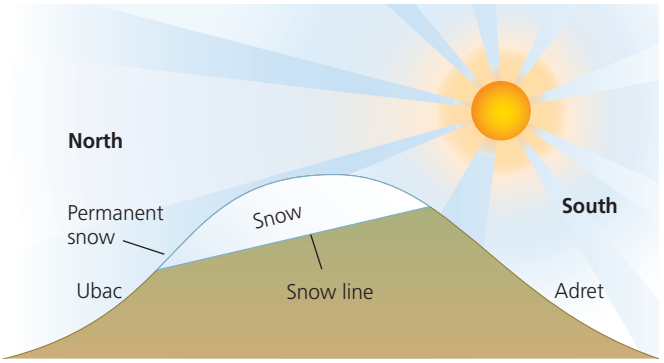


Figure 4.9 The snow line in the northern hemisphere

Snow initially falls as flakes with an open, feathery structure that traps air. As snow accumulates the lower layers are slowly compressed by the upper layers, turning this lower snow into a more compact form known as **firn** (usually after about two seasons of accumulation). The French term **névé** does refer to the layers of freshly accumulated snow but is often interchangeable with 'firn'. Meltwater also seeps into spaces between the snow particles and freezes, further compacting the mass. Further accumulation progressively squeezes air out of the lower layers by the additional weight of the upper layers and after some time (most experts suggest between 20 and 40 years) solid ice develops. As the ice forms the

mass changes colour from white, indicating the presence of air, to a bluish colour, indicating the air has largely been expelled. If the mass of ice becomes large enough it can develop into a glacier and begin to flow downhill.

Ablation and accumulation

The **zone of accumulation** refers to the upper part of the glacier where inputs exceed outputs and therefore more mass is gained than is lost over a year. Where outputs exceed inputs in the lower part of the glacier, and mass is lost rather than gained, this is known as the **zone of ablation**. The boundary between the ablation zone and accumulation zone, where net gain and net loss are balanced, is known as the **equilibrium line**. (This is not to be confused with the *firn line*, which is similar but different.) (Figure 4.10.)

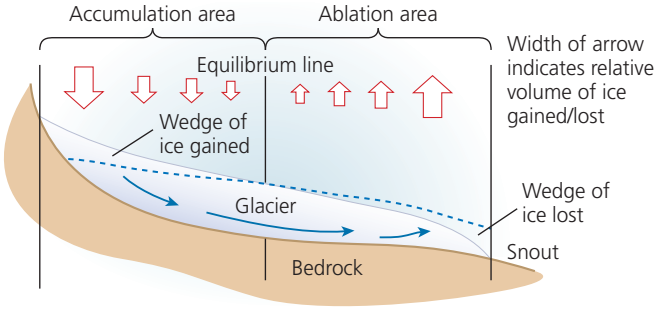


Figure 4.10 Glacier budget

Historical patterns of ice advance and retreat

The difference between the total accumulation and total ablation during one year is known as the **net balance** (Figure 4.11). **Temperate** glaciers in alpine areas have a negative balance in summer when ablation exceeds accumulation and a positive balance in winter when accumulation exceeds ablation. If the summer and winter budgets cancel each other out, the glacier appears to be stationary. In other words, the snout of the glacier remains in the same position while ice is still advancing down the valley from the zone of accumulation into the zone of ablation.

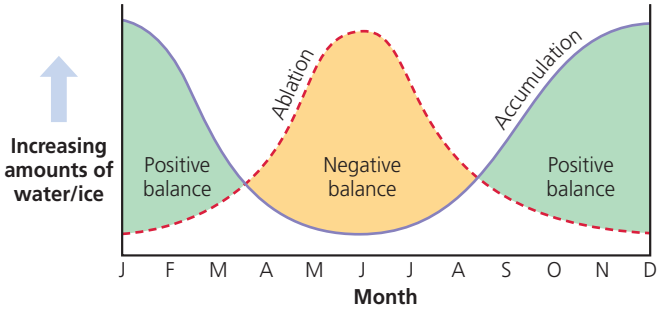


Figure 4.11 Net balance in a northern hemisphere temperate glacier

Figure 4.12 Ice coverage in Europe 22,000 years BP

If accumulation begins to exceed ablation rates, then the snout moves down the valley. This is known as **glacial advance**. When ablation rates exceed accumulation the glacier begins to shrink in size and the position of the snout moves up the valley. This is called **glacial retreat**. It is important to note that despite the *position* of the snout moving back up the valley, the (retreating) ice continues to move downslope from the upper parts of the glacier.

During the glacials and interglacials of the Quaternary glaciers have advanced and retreated many times. Evidence for this is found in the landscape. At present glaciers cover about 6,000 km² in Europe, but 22,000 years BP ice coverage was much more extensive (Figure 4.12). Individual glaciers not only advance and retreat in response to long-term climate changes, but also gain and lose mass in response to more localised, or short-term changes in climate. Figure 4.13 illustrates the pattern of advance and retreat observed in one European glacier over a much shorter time period.

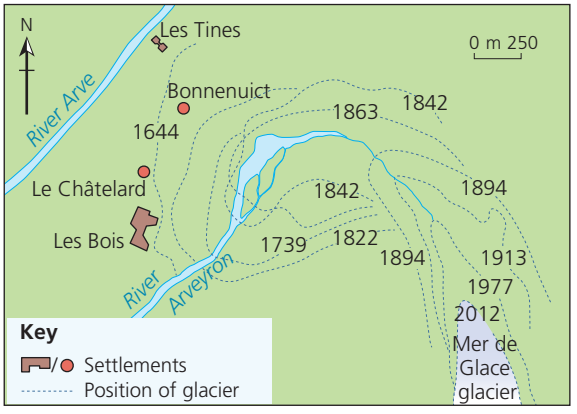


Figure 4.13 Changing snout position of the Mer de Glace, France

While most glaciers around the world are following the pattern of the Mer de Glace in Chamonix, where negative glacial budgets mean their snouts have retreated in the recent past, some glaciers in the Himalayas and Alaska in North America have positive annual net balances and their snouts are advancing, like Alaska's Hubbard Glacier (Figure 4.14).

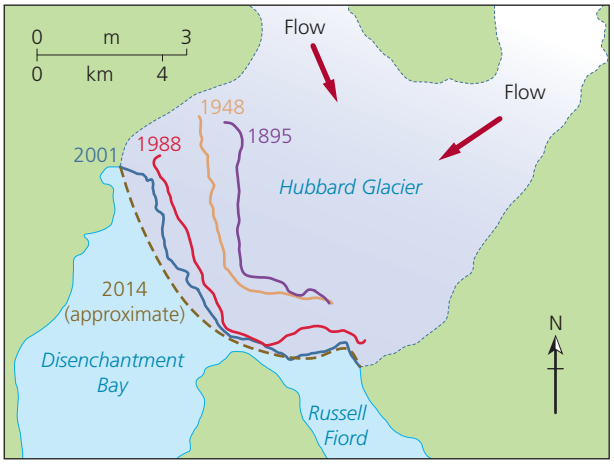


Figure 4.14 Alaska's advancing Hubbard Glacier

Warm- and cold-based glaciers

The characteristics of individual glaciers are unique. However, glaciologists have developed a simple three-fold classification of glaciers, including **temperate** or **warm-based**, **polar** or **cold** and sub-polar glaciers. Over the decades many authors have adapted this classification and adopted a range of other interchangeable terms for each type. For simplicity this chapter will concentrate on warm- and cold-based glaciers.

Warm-based, or temperate, glaciers (Figure 4.15) tend to be found in more temperate maritime locations. Good examples of this include the western coastal mountain ranges of North America, western Norway, southern Iceland, New Zealand and Patagonia (Chile). Here the alpine valley glaciers of the European Alps are included in the warm-based category. These tend to be smaller valley glaciers typically ranging from hundreds of metres to a few kilometres in width and hundreds of metres to tens of kilometres in length.

Figure 4.15 Franz-Joseph glacier in New Zealand

Cold-based or polar glaciers by definition are almost all found in areas of high latitude within the Arctic and Antarctic circles, in especially cold areas of Alaska, Canada and Antarctica (Figure 4.16). Here cold-based glaciers include some very cold valley glaciers but also refer to the much larger glaciers associated with the vast **ice caps** and **ice sheets** covering hundreds of km^2.

Figure 4.16 Commonwealth Glacier, Antarctica

The characteristics and development of temperate/warm-based glaciers

- Temperate glaciers are found in places with high winter snowfall rates and spring and summer temperatures high enough to create rapid summer melt rates.
- Meltwater acts as a lubricant and allows the glacier to be far more mobile than cold-based glaciers.
- Faster rates of glacial movement mean that warm-based glaciers are more likely to erode, transport and deposit material.

- The thin surface layer (a few metres) of more recent snow and firn is subject to seasonal temperature fluctuations so melts rapidly at around 0°C in the summer. This surface layer insulates the layers of ice beneath it.
- With increasing depth in a glacier the ice is under ever increasing pressure from the surrounding ice, which has the effect of lowering the melting point of the ice. This is known as the **pressure melting point**. For example, at the base of a 2,000 m deep glacier the melting point of the ice is –1.27°C.
- All ice in a temperate glacier is at or near the *melting point* because of the warmer atmospheric temperatures, the weight of the ice above, and as temperature glaciers may be relatively thin a greater proportion of the ice is influenced by geothermal heat at the bed (Figure 4.17).

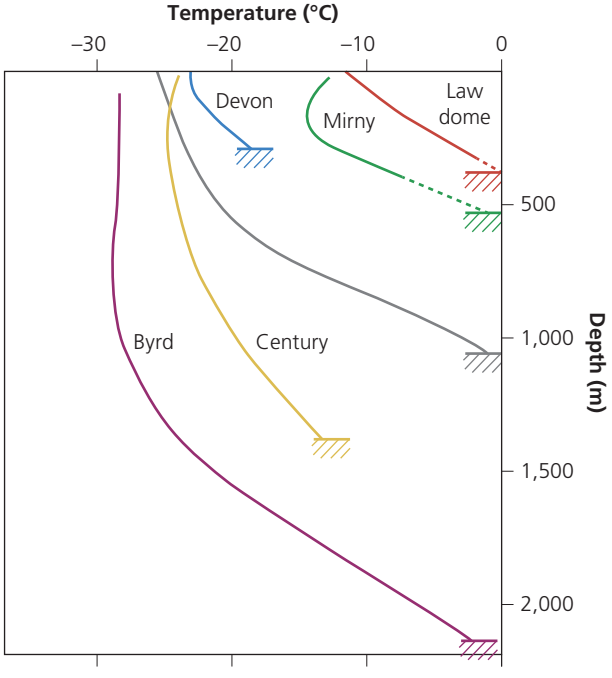


Figure 4.17 Temperature profiles of selected Greenland and Antarctic glaciers

The characteristics and development of polar/cold-based glaciers

- Cold glaciers are generally found in places with low precipitation rates, even arid conditions, so receive little new snow each year and accumulation rates are low.
- There is little or no melting of ice in cold glaciers and so the ice can be very old, with the ice at the base of some Arctic and Antarctic ice sheets dating back around 100,000 years.
- All ice, except the most upper surface layers that can be exposed to summer atmospheric temperatures of above 0°C, is below the *melting point*.
- There is little meltwater associated with these glaciers as atmospheric and sub-glacial (geothermal) heat sources are not great enough to reach melting point.
- Other than limited melting of the surface in summer the majority of ice loss is due to *sublimation* and the calving of icebergs/blocks.
- Movement in cold-based glaciers is much slower than in temperate glaciers as they are often frozen to their beds, thus most movement is due to **internal flow**.
- Much less erosion, transportation and deposition occurs.

While, as we will see later, many of the most striking features of the landscapes of many cold environments are the result of processes of erosion, transportation and deposition, it is important not to forget the role played by other geomorphological processes.

Geomorphological processes and glaciated landscape development

As glacial geomorphological systems develop over time and tend towards dynamic equilibrium a variety of processes take place. Driven by the input of energy these are what change the characteristics of the components of glacial and glaciated landscapes.

Weathering

Weathering is the breakdown and/or decay of rock at or near the Earth's surface creating regolith that remains *in situ* until it is moved by later erosional processes. Weathering can be mechanical, biological/organic or chemical. Although the low temperatures of most cold environments limit the impact of biological or organic weathering, mechanical and chemical weathering play a very important role in both shaping the environment and providing material for glacial erosion (Table 4.3, page 148).

Table 4.3 Weathering processes in cold environments

Weathering process	Characteristics
Frost action/ freeze–thaw cycles	• Leads to **frost shattering** • Occurs in areas where temperatures rise above 0°C during the day but drop below freezing at night for a substantial part of the winter • Water that enters cracks in the rocks freezes overnight. As it freezes, the water expands by just under 10 per cent, so the ice occupies more space and exerts pressure on the surrounding rock • As the process repeats and continues, the crack widens and eventually pieces of rock break off (Figure 4.18) • On steep slopes this leads to the collection of material at the base, known as scree • In a glacial valley much of this material falls from the valley side onto the edges of the glacier (an input) and some finds its way to the base of the ice via crevasses and moulins on the glacier's surface
Nivation	• A series of processes that operate underneath patches of snow in hollows, particularly on north- and east-facing slopes (in the northern hemisphere) • Freeze–thaw action, together with chemical weathering processes operating under the snow, causes the underlying rock to disintegrate • As snow melts in the spring, the weathered particles are 'flushed-out' of the hollow and moved downslope by the meltwater and solifluction (see below) • The accumulation of more debris, by repeated seasons of freezing and thawing, and the accumulation of more snow, leads to the formation of nivation hollows which, when enlarged, can be the beginnings of a **corrie** (cirque) as outlined below

Figure 4.18 Evidence of frost shattering, Ecrins National Park, France

Ice movement

As ice accumulates to greater depths it begins to move due to the influence of gravity. As the ice mass moves downhill it does not all behave in the same way. It has great rigidity and strength, but under steady pressure it behaves as a plastic (mouldable) body. In contrast, when put under sudden compression or tension, it will break or shear apart. This gives two zones within the glacier (Figure 4.19):

- the upper zone where ice is brittle, breaking apart to form crevasses
- the lower zone which has steady pressure. Here meltwater resulting from pressure melting and from *friction* with the bedrock allows a more rapid, plastic flow.

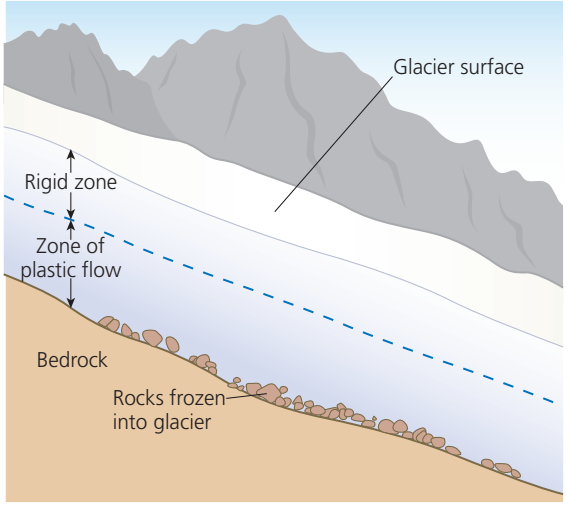


Figure 4.19 Zones within a glacier

The technical details of ice movement are extremely complex, but several main types of movement can be identified:

- **Internal deformation** including **creep** that occurs when stress builds up within a glacier, allowing the ice to behave with plasticity and flow. This is common when obstacles are met. *Internal flow* within the glacier involves the ice crystals orientating themselves in the direction of the glacier's movement and sliding past each other. Internal deformation and creep are the main feature of the flow of cold, or polar, glaciers as, without the presence of meltwater, they tend to be frozen to their beds.
- **Rotational flow** that occurs within the corrie (cirque), the birthplace of many glaciers. Here ice moving downhill can pivot about a point, producing a rotational movement. This, combined with increased

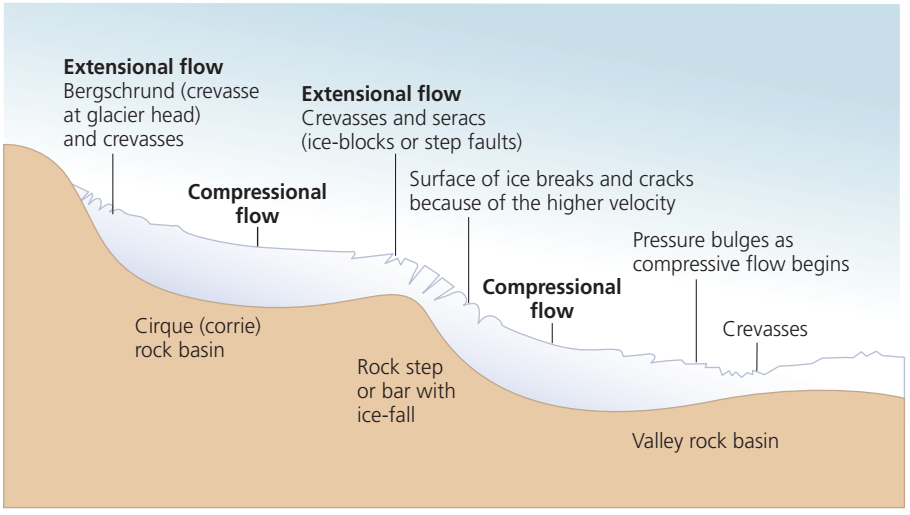


Figure 4.20 Extensional and compressional flow

pressure within the rock hollow, leads to greater erosion and an over-deepening of the corrie floor.

- **Compressional flow** that occurs where there is a reduction in the gradient of the valley floor leading to deceleration and a thickening of the ice mass. At such points ice erosion is at its maximum (Figure 4.20).
- **Extensional flow** that occurs when the valley gradient becomes steeper. The ice accelerates and becomes thinner, leading to reduced erosion (Figure 4.20).
- **Basal sliding (flow/slippage)** occurs because as the glacier moves over the bedrock there is *friction*. The lower levels of ice are also under a great deal of pressure and this, combined with the friction, results in some melting. The resulting meltwater acts as a lubricant, enabling the ice to flow more rapidly.

At certain times glaciers can move forward quite suddenly. These are known as **surges**.

- **Surges** occur when an excessive build-up of meltwater under the glacier leads to the ice moving rapidly forward, perhaps as much as 250–300 m in one day. Such surges represent a hazard to people living in the glacial valley below the snout. Following the increased sub-glacial volcanic activity in Iceland over recent years, glaciologists have been able to study these surges more closely.

Within a glacier itself there are different rates of movement. The sides and base of the glacier move at a slower rate than the centre surface ice (Figure 4.21). As a result, the ice cracks, producing crevasses on the surface. These also occur where extending flow speeds up the flow of the ice, where the valley widens or the glacier flows from a valley on to a plain like a *piedmont* glacier.

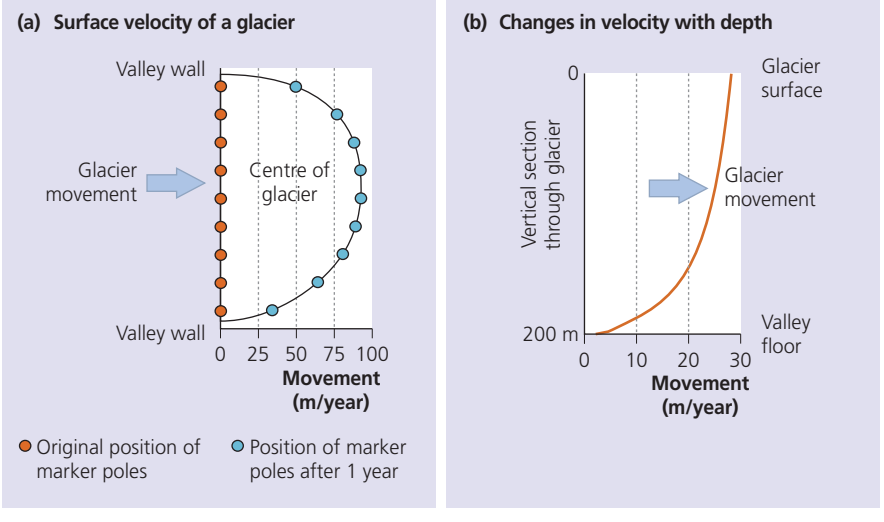


Figure 4.21 Differential rates of flow within a glacier

Glacial erosion

As warm-based glaciers are large bodies of ice and are moving downhill they have the ability to erode and shape the landscape. Glacial erosion tends to happen in upland regions and is carried out by two main processes (Table 4.4):

Table 4.4 Erosion processes in cold environments

Erosion process	Characteristics
Abrasion	• The material the glacier is carrying rubs away at the valley floor and sides • Often likened to the effect of sandpaper, as it is not the ice (the paper) that has the ability to actually wear away at the rock, but the material embedded in it (like the sand on the paper) • Coarser, harder material may leave scratches on the rock known as striations • Finer debris smooths and polishes rock surfaces • The debris involved is often worn down into a very fine material known as rock flour. This is taken away from the glacier by sub-glacial meltwater streams, turning them a milky colour
Plucking	• Involves the glacier freezing onto and into rock outcrops • As the ice moves forward it pulls away masses of rock • Mainly found at the base of the glacier where pressure and friction often result in the melting of the ice • Meltwater at the base of most glaciers is very close to freezing point, and often this will refreeze, attaching the glacier to the bedrock • Or simply, the weight and dragging force of the glacier breaks away rock fragments incorporating them into the base of the ice • Common in well-jointed rocks and those where the surface has been weakened by freeze–thaw action (frost shattering) • Leaves a very jagged landscape

Glacial transport and deposition

Sub-glacial debris, en-glacial and supra-glacial moraine

As well as eroding rock over which it is flowing, a valley glacier is also capable of transporting large amounts of debris. Some of this may also be derived from rockfalls from the valley side. Material is transported:

- on the surface of the glacier (**supra-glacial debris**)
- buried within the ice (**en-glacial**)
- at the base of the glacier (**sub-glacial**) and may include rock fragments that have fallen down crevasses or been washed down *moulins* as well as material eroded at the base.

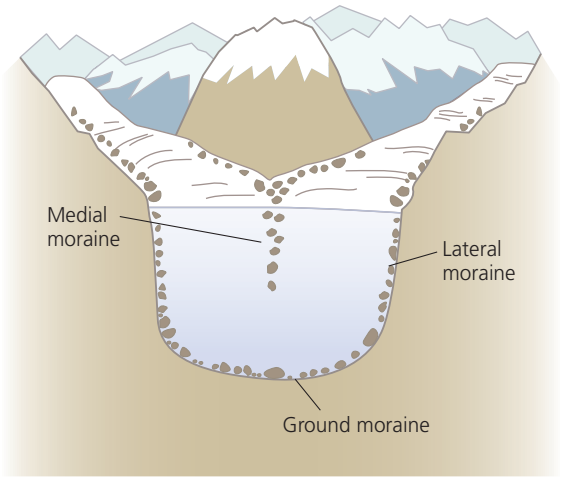


Figure 4.22 Transport of debris within a glacier

Figure 4.22 illustrates where material is transported by the glacier, including an alternative description of this material. Strictly speaking **moraine** applies to types of landform (explored below), but many textbooks now use the term in this way.

Glacial deposition

The huge amounts of material carried by a glacier will eventually be deposited. The majority of this will be the debris released by the melting of the ice at the snout. It is also possible for the ice to become overloaded with material, reducing its capacity to transport it. This can occur near to the snout, as the glacier melts, or in areas where the glacier changes between compressing and extending flow, thus reducing the mass of ice at that point. Material that is deposited directly by the ice is known as **till**; this has generally replaced the term **boulder clay** as this is quite specific in describing certain kinds of till.

The term till is used to describe an unsorted mixture of rocks, clay and sand that was mainly transported as supra-glacial or en-glacial debris and deposited when the ice melted. Individual stones tend to be angular to sub-angular, unlike river or beach material which is rounded and smooth. Till reflects the character of the rocks over which the ice has passed. The till deposited in south Lancashire, for example, includes rocks from the Lake District (for example, shap granite) and southern Scotland (for example, riebeckite from Ailsa Craig in the Firth of Clyde). In East Anglia the till contains pieces of granite from southern Norway. This indicates not only the passage of the ice but the fact that the sea level must have been considerably lower to allow ice to move over the area that later became the North Sea.

Review questions

4 What are the main differences between cold-based and warm-based glaciers?

5 Why might weathering processes have more impact in the mountains of Scotland than in the Antarctic?

Glaciated landscape development

Landforms of glacial erosion and deposition

Glaciers are large powerful forces and play a significant role in shaping the landscapes where they are found. There are a number of major landforms that are mainly produced by **glacial erosion**. These include:

- **roche moutonnée**
- **corries (cirques)**
- **arêtes**
- pyramidal peaks
- **glacial troughs**, including **hanging valleys** and **truncated spurs.**

Fieldwork opportunities

Erosional landscapes

There are several accessible areas of the United Kingdom where you can study all the above features of glacial erosion:

- the Lake District in north-west England (see Case study 1)
- Snowdonia in North Wales
- the Cairngorms in Scotland.

Further afield:

- the Chamonix area of France has alpine glaciers with deep glacial valleys
- the Lauterbrunnen Valley in Switzerland is readily accessible
- in Iceland, you can study glaciation as it is happening.

Depositional landscapes

Many lowland areas of the United Kingdom show significant features of glacial deposition. For example:

- the fringes of upland areas, such as the south Lake District, which has good examples of drumlins
- there are moraines and fluvio-glacial deposits in East Anglia; the highest point of Norfolk, Beacon Hill (105 m) is a glacial deposit and is part of the moraines of the Cromer Ridge
- deposits are also found in the Vale of York, drumlins in the Ribble Valley (Lancashire/Yorkshire) and within many of the soils of lowland Britain (north of the Thames Valley) you will find stones which are small erratics which were brought south by the ice sheets and glaciers which once covered the country.

Due to the enormous amount of material eroded and transported by glaciers they also create a range of **depositional landforms,** including:

- **drumlins**
- **erratics**
- **moraines** (lateral, recessional, push and terminal/end moraines
- **till plains.**

Erosional landforms

Corries and associated landforms

A **corrie** is an armchair-shaped rock hollow, with a steep back wall and an over-deepened basin with a rock lip. It often contains a small lake (**tarn**). In the British Isles, corries are mainly found on north-, northeast- and east-facing slopes where lower levels of insolation allowed increased accumulation of snow.

The dimensions of corries will depend on the topography of the area and the size of the glaciers that formed them, but in the English Lake District, Red Tarn has a depth of 25 m, a back wall rising 300 m to the summit of Helvellyn and is around 300–400 metres across (see case study on page 178).

If several corries develop in a highland region, they will jointly produce other erosional features.

- An **arête** is formed when two corries lie back to back or alongside each other. If enlarged over many glaciations a narrow, steep-sided ridge between the two hollows is formed. An example is Striding Edge above Red Tarn on Helvellyn in the Lake District (see case study on page 178).
- If more than two corries develop on a mountain, the remaining central mass will survive as a **pyramidal peak**, which often takes on a very sharp appearance due to frost shattering (Figure 4.23). A classic example is the Matterhorn in the Alps bordering Switzerland and Italy.

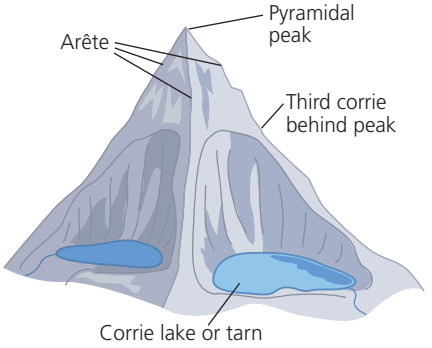


Figure 4.23 Arêtes and pyramidal peak

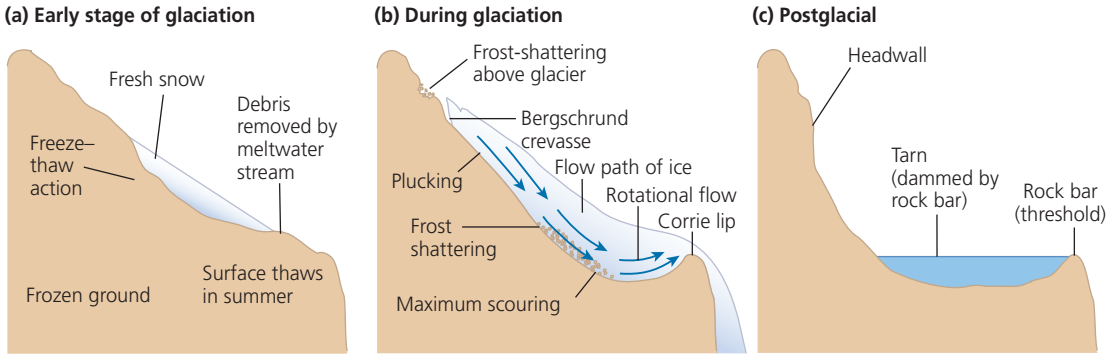


Figure 4.24 The formation of a corrie

Corries result from the interaction of several processes (Figure 4.24).

- Initially, **nivation** acts upon a shallow, pre-glacial hollow and enlarges it into an embryo corrie (this can be a slow process across several glacial periods within an ice age).
- As the hollow grows, the snow becomes deeper and is increasingly compressed to form firn and then ice.
- **Plucking** makes the back wall steeper.
- Due to the bowl shape of the hollow, the **rotational flow** of the ice, together with the debris supplied by plucking and frost shattering on the back wall, abrades the floor of the hollow, which over-deepens the corrie.
- As the hollow becomes deeper, the thinner ice at its edge does not produce the same amount of downcutting and a rock lip develops on the threshold.
- Some thresholds are increased in height by the deposition of moraines formed when the glacier's snout was last in that position.
- During interglacials when the ice has melted, the corrie fills with meltwater and rainwater to form a small lake (tarn).

Glacial troughs and associated landforms

Glaciers flow down pre-existing river valleys as they move from upland areas. They straighten, widen and deepen these valleys, changing the original V-shaped river feature into the U- or trough shape typical of glacial erosion. The action of ice, combined with huge amounts of meltwater and sub-glacial debris, has a far greater erosive power than that of the pre-glacial river that occupied the valley.

As both extending and compressing flow are present, the amount of erosion varies down the valley. Where compressing flow is present, the glacier will over-deepen parts of the valley floor, leading to the formation of **rock basins**. Scientists have suggested that this over-deepening is caused by increased erosion at the confluence of glaciers, areas of less-resistant geology or zones of well-jointed rocks.

Major features of glacial troughs:

- Generally straight with a wide base and steep sides – a U-shape.
- Stepped, long profile with alternating steps and rock basins (Figure 4.25).
- Some glacial valleys end abruptly at their heads in a steep wall, known as a **trough end**, above which lie a number of corries (Figure 4.26).
- Rock basins filled with **ribbon lakes**, for example, Wast Water in the Lake District.

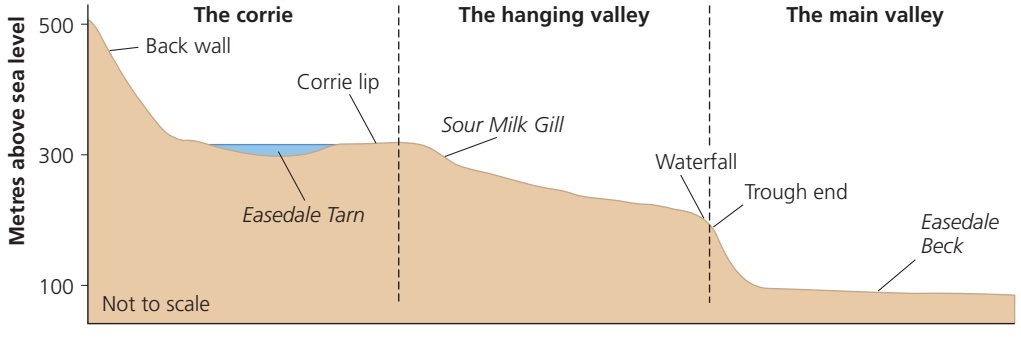


Figure 4.25 The long profile of a glacial valley: Easedale, Lake District

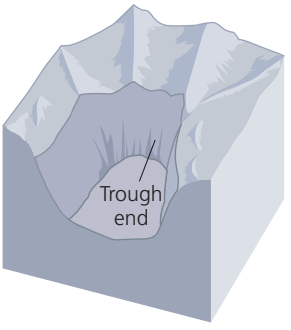


Figure 4.26 A trough end

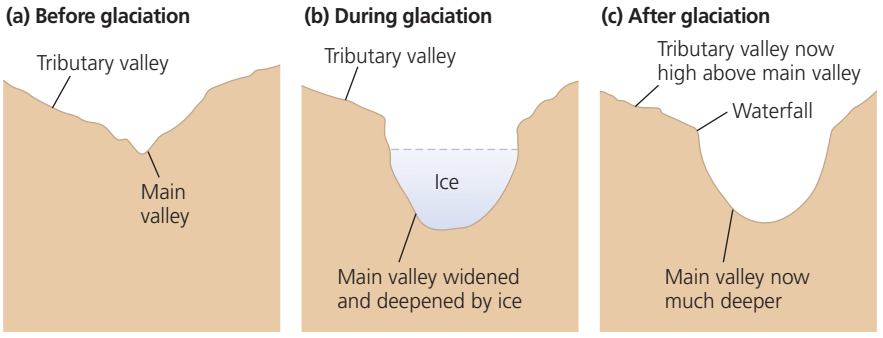


Figure 4.27 The formation of a hanging valley

- Over-deepening below the present sea level – this led to the formation of **fjords** when sea levels rose after the ice ages and submerged the lower parts of glacial valleys, for example, on the coasts of Norway and southwest New Zealand (Milford Sound).
- **Hanging valleys** on the side of the main valley (for example, the valley of Church Beck which flows down into Coniston Water in the Lake District). These are either pre-glacial tributary river valleys that were not glaciated, or smaller tributary glacial valleys (Figure 4.27 and Figure 4.28). The glaciers in the tributary glacial valleys would have been smaller, containing less ice and so eroding less than in the main valley. Consequently the floor of the tributary valley was left higher than that of the main valley when the ice retreated.

Figure 4.28 Lady Bowen waterfall in Milford Sound, New Zealand, is a hanging valley

- Areas of land projecting from the river-valley side (spurs) have been removed by the glacier, producing truncated spurs.
- Small mounds of resistant rock on the valley floor are not always completely removed and as the glacier moves over these, roches moutonnées are formed. These have an up-valley side (stoss slope) polished by abrasion and a down-valley side (lee slope) made jagged by plucking. These are relatively small features, up to a few tens of metres long and less than 5 m high (Figure 4.29).

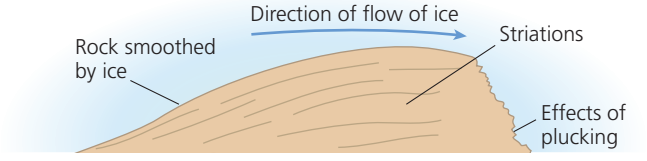


Figure 4.29 The formation of a roche moutonnée

- After ice retreat, many glacial troughs were filled with shallow lakes which were later infilled, and post-glacial frost-shattering modified their sides, developing screes altering the original glacial U-shape (for example, Great Landgdale, Lake District) (Figure 4.30).

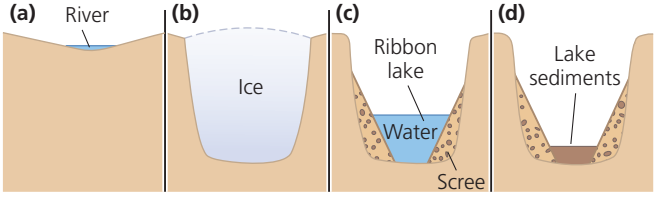


Figure 4.30 The formation of a glaciated valley

Depositional landforms

Glaciers can transport huge quantities of material, all of which is deposited following the melting of the ice. Two types of glacial deposit are recognised:

- **lodgement till:** sub-glacial material that was deposited by the actively moving glacier. A typical feature formed in this manner is a **drumlin**
- **ablation till:** produced at the snout when the ice melts. **Terminal (end)**, **push** and **recessional moraines** are typical features produced from ablation till.

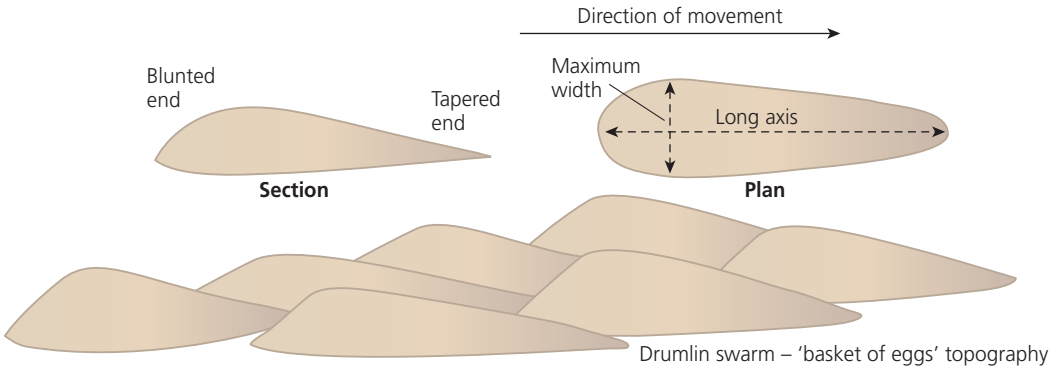


Figure 4.31 Drumlins

Drumlins

The term drumlin is derived from the Gaelic word *druim*, meaning a rounded hill, and has been adopted to describe these streamlined elongated 'whaleback' hillocks (Figure 4.31). Main features of drumlins:

- They are smooth, oval-shaped small hills, often resembling half an egg.
- They can be up to 1.5 km in length (although most are much smaller) and up to 50–60 m in height.
- They have a steeper 'upstream' end and a more gently sloping 'downstream' end. They are elongated in the direction of ice advance.
- They are often found in groups know as **swarms** and, given their shape, this is sometimes referred to as a 'basket of eggs' topography (Figure 4.32).
- They are formed from unsorted till, and most are believed to have little internal structure.
- They are found on lowland plains such as the central lowlands of Scotland. There is a well-known swarm at Hellifield in the Ribble Valley. Many are found at the lower end of glacial valleys.

Figure 4.32 Ribblehead drumlin field in the Yorkshire Dales

There has been considerable debate about the origin of drumlins for over 100 years and there are a few popular theories on their mode of formation:

- The first suggests that the ice moulds and shapes a pre-existing landscape of glacial drift (ground moraine) deposited in a previous glacial advance.
- Secondly, it is thought that the moving glacier deposits till around a nucleus formed by some unevenness in the floor, such as protruding rock or mass of frozen till. This material is then moulded as the ice continues to move over it.
- However, one of the most widely held views is that drumlins result from the ice becoming overloaded with debris. This reduces its capacity to carry material and deposition occurs at the base of the ice. Again, this material is shaped by further glacier advance.

Erratics

Sometimes it is possible to find a large block of rock that has been moved from one area and deposited in another which has a very different geology. These are known as **erratics** (Figure 4.33).

Figure 4.33 Norber erratics in the Yorkshire Dales

Moraines

Moraines are ridges, or a series of mounds of material, mainly running across glacial valleys. The main type is the terminal or end moraine that is found at the point where the snout of the glacier reached its furthest extent 'downstream'. **Terminal moraines** are:

- a ridge, or mounds, of material stretching across a glacial valley
- elongated at right angles to the direction of ice advance
- often steep-sided, particularly the ice-contact side, and can reach up to 50–60 m high
- often crescent-shaped, moulded to the form of the snout
- formed from unsorted ablation material (till).

Terminal moraines are formed when ice melts during a period of snout standstill, when it has reached its furthest point down the valley, and the material it has been carrying is deposited. This is why they contain a range of unsorted material, from clay to boulders.

As the glacier retreats, it is possible for a series of moraines to be formed along the length of the valley, marking points where the retreat paused for some time. These are known as **recessional moraines** (Figure 4.34).

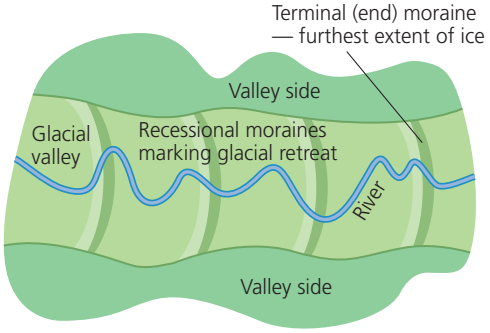


Figure 4.34 Moraines

If the climate cools for some time, leading to glacier advance, a previously deposited moraine may be shunted forwards up into a mound known as a **push moraine**. Such features are recognised by the orientation of the individual pieces of rock, which may have been pushed upwards from their original horizontal position.

Till plains

Often found behind a terminal moraine, or towards the margins of former ice sheets, in low-lying areas, **till plains** are wide areas of generally flat relief created by a till sheet. The underlying bedrock may itself be quite uneven, but it is hidden by a thick covering of glacial till, sands and gravel. The depth of the till sheet can vary greatly, being thicker where depressions exist in the underlying surface. The material tends to be quite compacted giving a poorly drained surface with bogs, lakes and slow-moving meandering streams. There are large till plains south of the Great Lakes in North America, marking the margins of the earlier Laurentide ice sheets.

Characteristic glaciated landscapes

The preceding sections illustrate the significance of the role of glaciers in shaping the landscape. Regions where glaciers have returned many times during the glacials of the Quaternary Ice Age, such as the European Alps, are often said to have a classic glaciated landscape. There are many valleys in the Alps of Austria, Italy, France and Switzerland that exhibit many of the features explored above. A good example is the Lauterbrunnen Valley close to Bern in Switzerland (Figure 4.35).

Figure 4.35 The glaciated features of an alpine valley, the Lauterbrunnen Valley, Switzerland

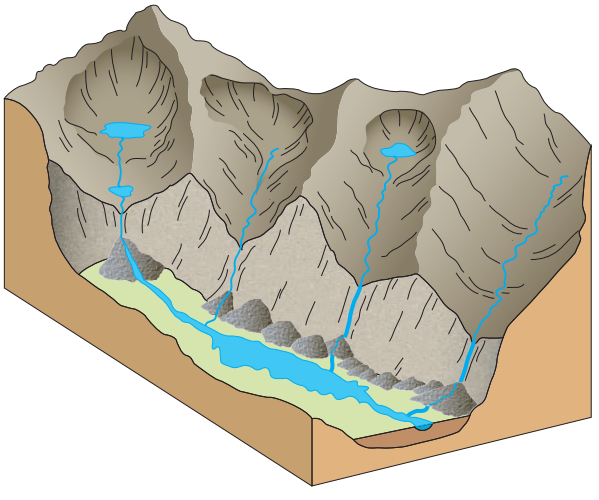


Figure 4.36 Block diagram to illustrate features of a glaciated alpine valley

Skills focus

Using Figures 4.35 and 4.36 (page 155), make a simple pencil sketch of a glaciated valley and annotate the features of a glaciated valley outlined in this chapter.

Fluvio-glacial processes

When glaciers melt huge quantities of meltwater are produced which has the capacity to transport much debris. Meltwater streams often flow under considerable pressure, either sub-glacially or in confined channels. This gives them high velocity and turbulent flow. Meltwater channels can therefore pick up and transport a larger amount of material than normal rivers of a similar size. When the meltwater discharge decreases, the resultant loss of energy causes the material to be deposited. As with all material deposited by flowing water, the heavier particles are dropped first, resulting in sorting of the material. Deposits may also be found in layers (stratified) as a result of seasonal variations in the meltwater flow. Therefore, the landscapes of regions 'downstream' of glacial or formerly glaciated areas are often dominated by **fluvio-glacial** landforms.

Fluvio-glacial landforms

Erosion

Meltwater channels are some of the main features of fluvio-glacial erosion, however understanding their origin can be complex. Some key features of meltwater channels are described below.

- They are classified according to where they flow in relation to the glacier during the glacial period (Figure 4.37).

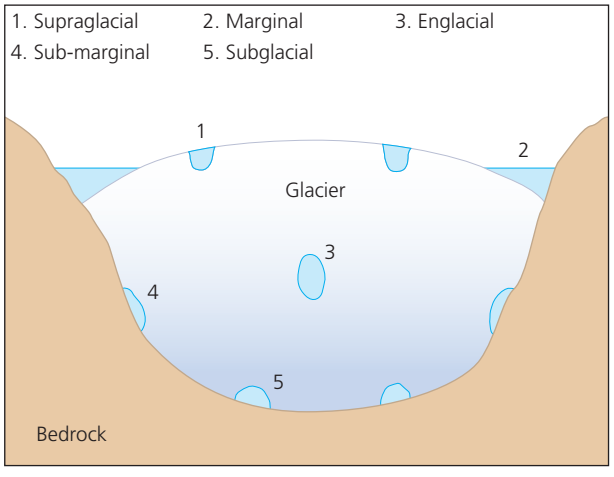


Figure 4.37 Meltwater channel types

- Sub-glacial streams are able to flow uphill due to hydrostatic pressure.
- Only meltwater flowing in channels in contact with the bedrock floor leaves any erosional trace in the post-glacial landscape.
- It is believed that beneath current glaciers two types of channel occur. Those that:
  - cut downwards into bedrock by headward erosion, known as Nye channels
  - cut upwards into the ice itself by meltwater, known as Röthlisberger channels. These only leave any postglacial trace if the channel becomes choked with sediment (see below).
- It is also believed that with high discharge and its turbulent flow, large meltwater channels can create sub-glacial valleys that are often deep and riddled with potholes.

It is only in recent decades with advances in remotely operated technology and computer-generated digital imagery that glaciologists are beginning to be able to actually see what is happening in en-glacial and sub-glacial streams.

Pro-glacial lakes and overflow channels are also features of fluvio-glacial erosion, with some well-known examples in the British Isles.

- During deglaciation, lakes develop on the edges of the ice, some occupying large areas.
- Overflows from these lakes which cross the lowest points of watersheds will create new valleys.
- When the ice damming these meltwater lakes totally melts, many of the new valleys are left dry, as drainage patterns revert to the pre-glacial stage.
- In some cases the postglacial drainage adopts them, giving rise to new drainage patterns.

The River Severn is believed to have been diverted during the last glaciation. Figure 4.38 shows the stages of this process.

- **Stage 1 Pre-glacial:** The River Severn flowed northwards to enter the Irish Sea in what is now the estuary of the River Dee. The present Lower Severn was a shorter river flowing from the Welsh borderlands to the Bristol Channel.
- **Stage 2 The last ice age:** Ice coming from the north blocked the River Severn valley to the north. The water from the blocked river formed a huge pro-glacial lake known as Lake Lapworth. The lake

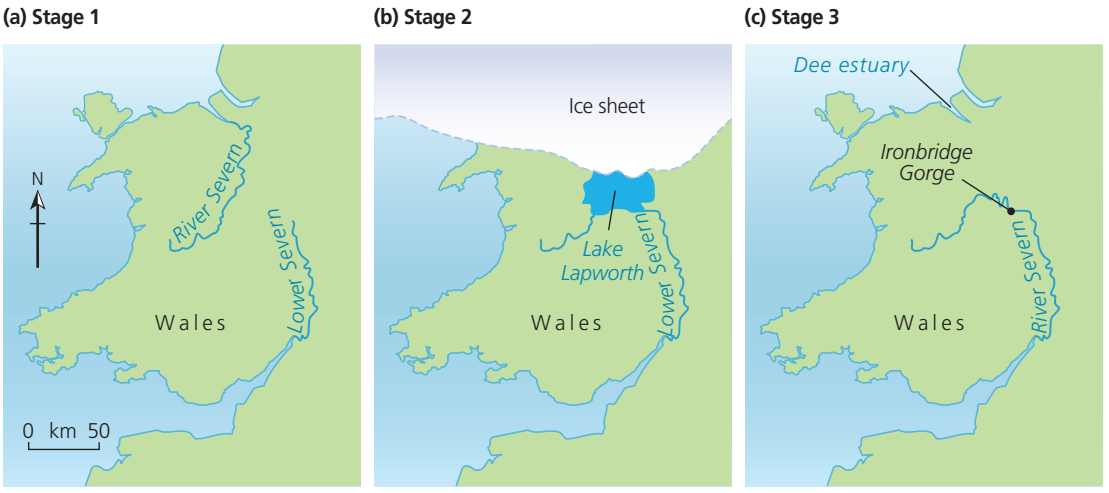


Figure 4.38 Theoretical stages in the diversion of the River Severn

eventually overflowed the watershed to the south to join the original Lower Severn. In the process it cut through a solid rock area, creating the gorge at Ironbridge.

- **Stage 3 The post-glacial period:** As the ice retreated to the north the way should have been left open for the two rivers to return to the pre-glacial situation. The route north, however, was blocked with glacial deposits, and as the Ironbridge Gorge had been cut very deep (lower than the exit to the north), the new drainage adopted this rather than its former route. The River Severn now flows from central Wales to the Bristol Channel.

Deposition

Landscapes shaped by fluvio-glacial deposition are mainly composed of the following features: **eskers**, **kames** and **outwash plains**, and pro-glacial lakes (Figure 4.39). Lakes on the outwash plain often have layered deposits in them called **varves**.

Eskers

Eskers have the following main features:

- They are long ridges of material running in the direction of ice advance.
- They have a sinuous (winding) form, 5–20 m high.
- They consist of sorted coarse material, usually coarse sands and gravel.
- They are often stratified (layered).

Eskers are believed to be deposits made by sub-glacial streams. The channel of the stream will be restricted by ice walls, so there is considerable hydrostatic pressure which enables a large load to be carried and also allows the stream to flow uphill for short distances. This accounts for the fact that some eskers run up gentle gradients. The load builds up the bed of the channel above the surrounding land or, as noted above, close to the snout of the glacier the hydrostatic pressure may drop leading to an overburdening of the channel and deposition and the

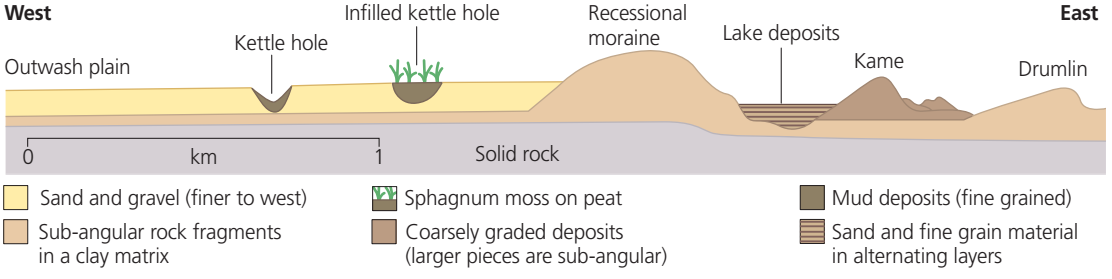


Figure 4.39 Features of a landscape modified by fluvio-glaciation and lowland glaciation

formation of a cast of the sub-glacial tunnel. A ridge is left when the glacier retreats during deglaciation. In some areas, the ridge of an esker is combined with mounds of material, possibly kames (see below). Such a feature is known as a **beaded esker**. Eskers are common in formerly glaciated areas of Europe and North America.

Figure 4.40 An esker in Northern Manitoba, Canada

Kames

Kames are mounds of fluvio-glacial material (sorted, and often stratified, coarse sands and gravel). They are deltaic deposits left when meltwater flows into a lake dammed up in front of the glacial snout by recessional moraine deposits. When the ice retreats further, the delta kame often collapses. Kame terraces are frequently found along the side of a glacial valley and are the deposits of meltwater streams flowing between the ice and the valley side.

Outwash plains (sandur)

Outwash plains are found in front of the glacier's snout and are deposited by the meltwater streams issuing from the ice. They consist of material that was brought down by the glacier and then picked up, sorted and dropped by running water beyond the position of the ice front. The coarsest material travels the shortest distance and is therefore found near to the glacier; the fine material, such as clay, is carried some distance across the plain before being deposited. The deposits are also layered vertically, which reflects the seasonal flow of meltwater streams.

Meltwater streams that cross the outwash plain are **braided**. This happens as the channels become choked with coarse material because of marked seasonal variations in discharge. On the outwash plain there is often a series of small depressions filled with lakes or marshes. These are known as **kettle holes**. It is believed that they are formed when blocks of ice, washed onto the plain, melt and leave a gap in the sediments. Such holes then fill with water to form small lakes. Aquatic plants become established in the lakes and this leads over time to the development of a marshy area and then peat.

Lakes on the fringe of the ice are filled with deposits that show a distinct layering. A layer of silt lying on top of a layer of sand represents one year's deposition in the lake and is known as a varve (Figure 4.41). The coarser, lighter-coloured layer is the spring and summer deposition when meltwater is at its peak and the meltwater streams are carrying maximum load. The thinner, darker-coloured and finer silt settles during autumn and through the winter as stream discharge decreases and the very fine sediment in the lake settles to the bottom. Varves are a good indicator of the age of lake sediments and of past climates as the thickness of each varve indicates warmer and colder periods. Also, any organic material, such as seeds and pollen, caught up in these layers can be used for paleoclimatic research.

Characteristic fluvio-glacial landscapes

The preceding sections illustrate the significance of the role of meltwater in shaping the landscape. Many areas in southern Iceland on the edges of the warm-based

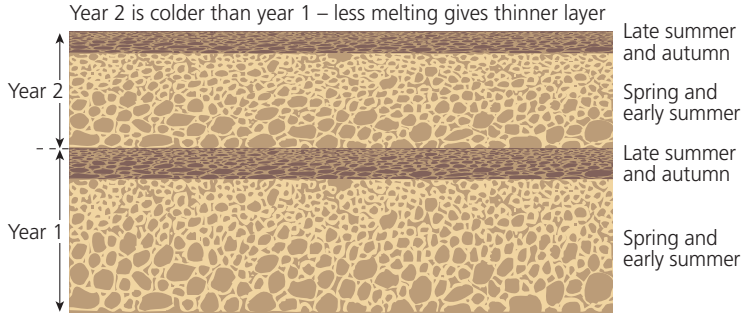


Figure 4.41 Varves

glaciers and ice sheets have fluvio-glacial landscapes composed of a range of fluvio-glacial features and landforms. The most recognisable of these are the expansive sandur or outwash plains.

Figure 4.42 The fluvio-glacial landscape of the Skeidarár Sandur in southern Iceland

The Skeidarár Sandur in southern Iceland (Figure 4.42) is a vast outwash plain from the Vatnajökull icecap. It covers 1,300 km² and exhibits the classic features of fluvio-glacial landscapes including:

- deep accumulations of black gravel, sand and silt originating from the local volcanic rock
- criss-crossed by constantly shifting braided streams; evidence of abandoned channels as the seasonally changing discharge levels affect rates of erosion and deposition
- very sparse vegetation and little evidence of human habitation. This results from the constant risk of floods or glacial burst due to volcanic eruptions in this area. There was a major eruption in 1996, and more minor events in 2004 and 2011.
- close to the glacier snout are ridges or steps that may result from kames or kettle features.

Periglacial features

Here the term **periglacial** refers to landscapes that, although not actually glaciated, are exposed to very cold conditions with intense frost action and the development of permanently frozen ground or permafrost. Some geographers apply the term more narrowly to only apply to areas affected by freeze–thaw cycles in close proximity to ice sheets.

At present, areas such as the tundra landscapes of northern Russia, Alaska and northern Canada, alongside high mountainous areas such as the European Alps, experience a periglacial climate. In the past, however, as ice sheets and glaciers spread, many areas which are currently temperate were subject to such conditions.

The climate of periglacial regions is marked by:

- persistently low temperatures, with an annual average of between –15°C and –1°C
- average annual rainfall (excluding snowfall) ranging between about 120 mm and 1,400 mm
- short summers when temperatures can sometimes reach above 15°C, but remain well below freezing in winter, with some areas occasionally reaching –50°C.

Permafrost

Permafrost generally refers to permanently frozen ground. In areas where temperatures below the ground surface remain below 0°C continuously for more than two years permafrost will occur. Currently, it is estimated that permafrost covers 20–25 per cent of the Earth's land surface. If water is present in the soil a significant amount will freeze, *cementing* the mineral and organic particles together. The ground below this frozen layer that remains unfrozen is known as **talik**. Figure 4.43 outlines how temperature changes with depth in permafrost.

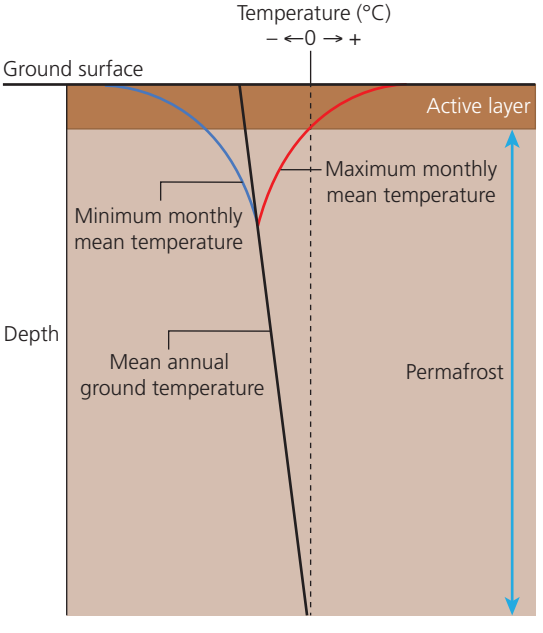


Figure 4.43 Typical temperature regime in permafrost

When summer temperatures rise above freezing, the surface layer thaws from the surface downwards to form an **active layer**. The depth of this layer varies considerably and depends upon local conditions, but may extend to 4 m. As the ice in this layer melts, large volumes of water are released.

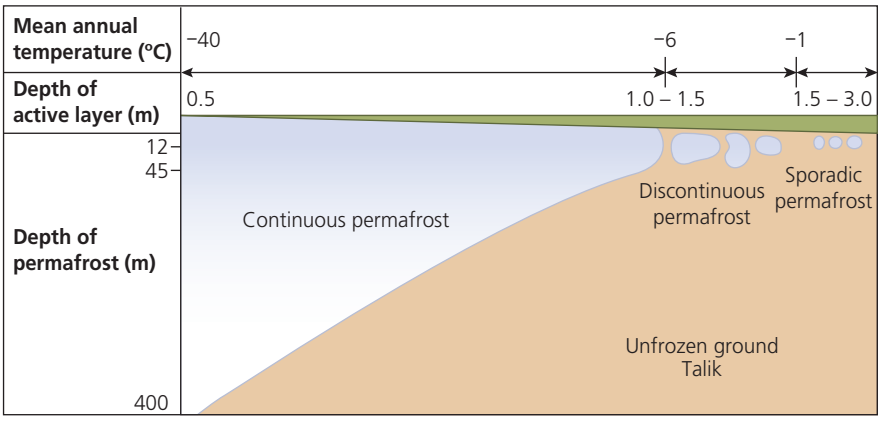


Figure 4.44 Variations in the depth of permafrost

There are three categories of permafrost (Figure 4.44):

- **Continuous permafrost** is found in the coldest regions, reaching deep into the surface layers. In Siberia today, it is estimated that the permafrost can reach down over 1,500 m. In the very coldest areas, there is hardly any melting of the uppermost layer.
- **Discontinuous permafrost** occurs in regions that are slightly warmer, where the ground is not frozen to such great depths. On average the frozen area will extend 20–30 m below the ground surface, although it can reach 45 m. There are also gaps in the permafrost under rivers, lakes and near the sea.
- **Sporadic permafrost** is found where mean annual temperatures are around or just below freezing point. In these places, permafrost occurs only in isolated spots where the local climate is cold enough to prevent complete thawing of the soil during the summer.

Periglacial mass movement processes (cryospheric slope processes)

There are a number of processes of mass movement that occur in periglacial landscapes:

- **Solifluction** occurs when summer temperatures rise enough to melt the huge amounts of water held as ice in the upper layers of the permafrost. Due to the impermeable layer of still frozen ground below, this water cannot drain away, and as there is little evaporation due to the low temperatures, this surface layer becomes very wet. Excessive lubrication from this water reduces the friction between the soil particles. On slopes as gentle as 2°, this saturated layer becomes quite mobile and soil begins to move downslope. The downslope movement of saturated soils can happen in many environments, but when related to the freezing and thawing of the active layer it is also referred to as **gelifluction**.
- Scientists also have a term that refers to any flows of earth within the still frozen permafrost. This is **congelifluction**.
- **Frost creep** refers to the gradual downslope movement of individual soil particles due to the alternating freeze–thaw cycles within the active layer. This relates to **frost heave** (see below).
- **Rock falls** result from **freeze–thaw action (frost shattering)**, and can involve the movement of large amounts of material. This process has already been described on page 148 because of its role in providing large amounts of the erosive material in glaciers. **Scree** slopes (accumulations of **talus**) at the foot of slopes resulting from frost shattering are a common feature of periglacial landscapes. On relatively flat areas, extensive spreads of angular boulders are left, which are known as **blockfield** or **Felsenmeer** (sea of rocks) (Figure 4.45).

Figure 4.45 Felsenmeer on Mount Washington, USA

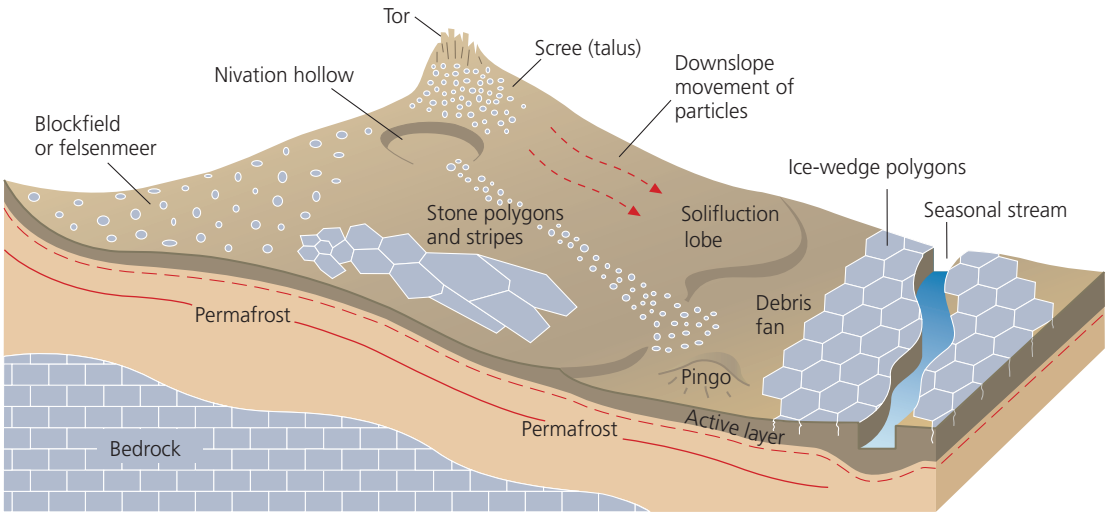


Figure 4.46 Features of a periglacial landscape

Periglacial landforms

There is a clear link between the landforms of periglacial landscapes and the processes that form them. The main landforms and the processes that form them are explored below, and illustrated in Figure 4.46.

Patterned ground

The repeated cycles of freezing and thawing of the active layer can produce a range of landforms collectively referred to as **patterned ground**. The formation of specific examples of patterned ground can be quite complex, but the main process involved is known as **frost heave**.

- As the active layer starts to freeze, ice crystals begin to develop.
- These increase the volume of the soil and cause an upward expansion of the soil surface.
- Frost heave is most significant in the fine-grained material and as it is uneven it forms small domes on the surface.
- Within the fine-grained material there are stones which, because of their lower specific heat capacity, heat up and cool down faster than the surrounding finer material.
- Cold penetrating from the surface passes through the stones faster than through the surrounding material.
- The soil immediately beneath a stone is likely to freeze and expand before other material, pushing the stone upwards until it reaches the surface.

Depending on the specific local conditions a range of features can form, summarised in Table 4.5 (page 162).

Ice wedges

Ice wedges are another feature that creates a type of **patterned ground**. Ice wedges:

- are relatively narrow cracks or fissures in the upper layers of the ground filled with ice (sometimes extending below the level of the permafrost)
- have a milky appearance due to the presence of air bubbles
- begin as small cracks, less than 5 cm across, but over hundreds of years can reach over 10 m wide and extend many metres into the ground.

There are some similarities in how they begin to form the cracked surface features in other areas like dried mud flats. However, in periglacial landscapes these result from **ground contraction** due to extreme cold, generally in areas of continuous permafrost.

- During the winter the freezing of the soil in the active layer causes the soil to contract.
- In the following summer the initial small crack is filled with meltwater (called an **ice vein** when it freezes).
- This process occurs repeatedly through the cycle of winter and summer widening and deepening the crack to eventually form the ice wedge.

Table 4.5 Characteristics of types of patterned ground

Category	Features summary	Formation summary
Circles	Circular-shaped feature Non-sorted circles – without a stone border Sorted circles – with a stone border Usually about 0.5 to 3 m diameter Slightly domed with polygonal surface cracks Occur singularly or in groups Within the circle the soil is usually fine grained silt and clay If the soil is coarser with gravels they are called stony earth circles	Usually form on flat ground Processes vary for individual types of circle Formed by seasonal frost heave
Nets	Similar in appearance to circles and polygons, though neither fully circular or polygonal in shape May be an intermediate form between circles and polygons on very gently sloping ground Creates a **hummocky** terrain	Similar to circles and polygons
Polygons	Multi-sided raised mounds of sorted material Approximately 10 m in diameter Finer particles make up the central mound surrounded by a border of larger stones The larger the border stones the larger the diameter of the polygon Not to be confused with ice-wedge polygons (Figure 4.48)	Only seen on flat ground Freezing causes domes to form The seasonal cycles of frost heave move stones to the surface Gravity acts on the larger stones making them roll downslope to form the border – thus giving these their alternate name *sorted polygons* It is also believed that ground contraction also plays a role in the sorting and shaping of the polygons (see below)
Steps	Step-like features formed on slopes More gently sloping than the slope on which they are formed They are either terraces with stones or vegetation forming a banked edge (depending on their origin) They are derived from either circles, nets or polygons Similar scale to their parent feature – up to 10–20 m wide	Created by frost-action of the permafrost Different types result from the processes that formed the parent feature
Stripes	Lines of stones running downslope Some are parallel lines of sorted material with stone borders Some are unsorted with vegetation borders Found on gradients over 6° Believed to have developed from circles, nets or polygons	Similar origins to steps Formed from the processes of frost heave Due to the slope, upon reaching the surface stones form linear features

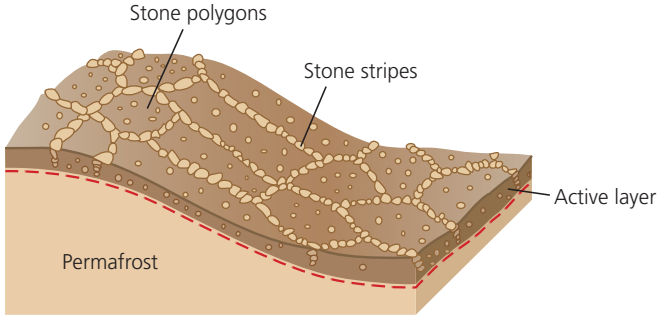


Figure 4.47 Stone polygons and stripes

The way the ground contracts creates a pattern of cracks on the surface, which when viewed from above, have a similar shape to stone polygons (Figure 4.47). These are therefore known as **ice-wedge polygons** (Figure 4.48). These are often larger than stone polygons and are generally 15–30 m in diameter, but can reach 100 m across. These are also generally only found on flat or very gently sloping surfaces.

If the climate warms enough for all the ice to melt it may be replaced by fined grained sediment that infills the void, forming an **ice-wedge cast**.

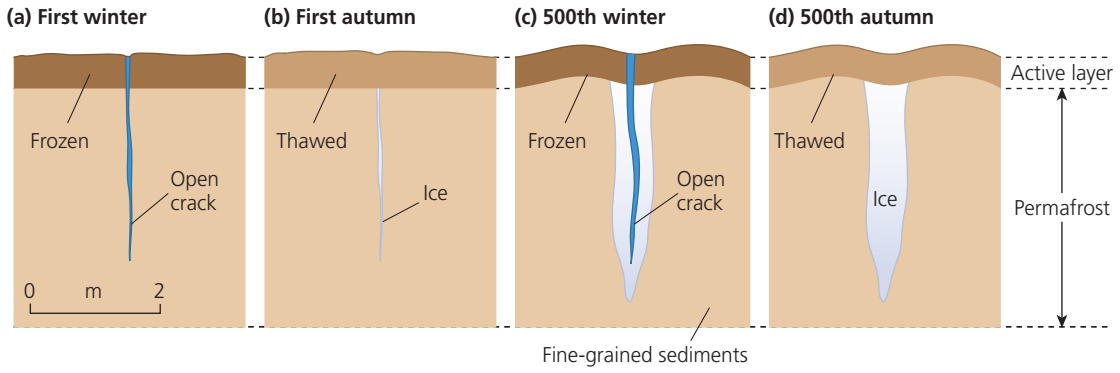


Figure 4.48 The formation of ice wedges

Pingos

Pingo is an Inuit term for a dome-shaped, ice-cored mound of earth in periglacial landscapes (Figure 4.49). Pingos range in size and age. They can be a few metres in height and diameter, but the oldest pingos, found on the MacKenzie Delta, are believed to be about 1,000 years old and are over 50 m high and over 300 m in diameter (some have been recorded over 600 m across). Some researchers suggest that they grow by 2 to 3 cm a year.

The ice core results from a slow *injection* of pressurised water into the dome. There are two types of pingo due to the different origins of this water. The specific processes underlying the formation of both types are complex. Table 4.6 provides a summary of the key features of each.

Figure 4.49 A pingo in northern Canada

Table 4.6 Key features of closed and open system pingos

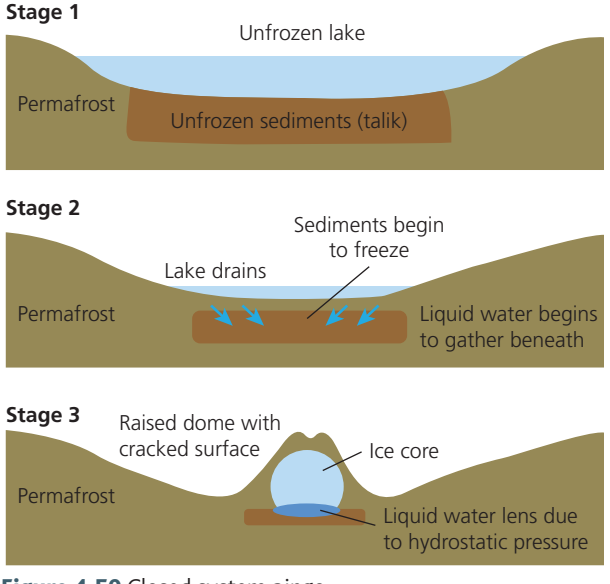
Closed system	
Famously found on the MacKenzie Delta area of the Northern Territories, Canada, so known as **MacKenzie Delta type**	
Generally found in areas of continuous permafrost	
Develop beneath lakebeds	
The growth of the ice core is *hydrostatic*	
Deep lakes (over 2 m) may remain unfrozen in winter	
The permafrost layer at the lakebed is insulated from the cold and thaws	
An area of unfrozen waterlogged ground is now sandwiched between the lake and underlying permafrost	
The lake may begin to drain, the lakebed is no longer insulated, so the waterlogged bed begins to freeze	
Due to localised differences in pressure between the lake, freezing lakebed and underlying permafrost, the newly freezing water *gathers* together to form an ice lens that expands, pushing the lakebed sediments above it up into the classic dome shape.	
The ice lens continues to grow as long as there is still unfrozen ground in the lakebed as a source of pressurised water to add to the ice core (Figure 4.50)	

Figure 4.50 Closed system pingo

Open system

Common in Greenland and Alaska, so known as **East Greenland type**

Generally found in areas of discontinuous permafrost

Found in valley bottoms

The growth of the ice core is *hydraulic*

Water is able to seep into the upper layers of the ground and flows from higher surrounding areas under artesian pressure

Water accumulates in flat low-lying areas between the upper layers of permafrost or soil and frozen ground beneath the water, then freezes

The freezing ice core expands thus doming the overlying layers into the classic pingo shape

They can grow as pressurised water continues to flow in from its surroundings (Figure 4.51)

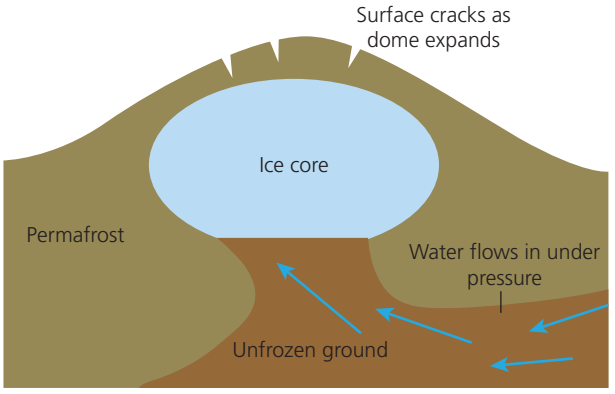


Figure 4.51 Open system pingo

As the pingo grows, increasing cracks and instability of the steep sides may expose the ice inside. As the ice melts the pingo collapses. When the core has completely melted all that remains is a doughnut-shaped mound possibly containing a small lake in the centre (Figure 4.52). Some present-day temperate regions have fossil pingos evidencing their colder past climates.

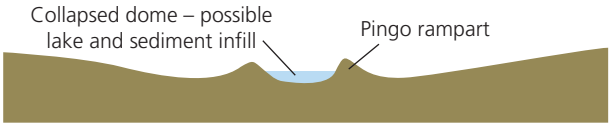


Figure 4.52 Fossil (collapsed) pingo

Solifluction (gelifluction) lobes and sheets

Where solifluction (or gelifluction) (as outlined above) occurs on slopes of around 1–20° solifluction sheets are formed. These have a smooth surface and can extend over 100 m across a slope.

Where slopes have a steeper gradient of 10–20° then solifluction lobes will form. These have a tongue-like appearance that extends downslope and can be up to 50 m in width and 5 m high. Depending on the local subsurface conditions and slope processes solifluction terraces can also form (Figure 4.53).

Terracettes

Terracettes are small periglacial features. They are often referred to as 'sheep tracks' as they have the appearance of small narrow pathways running horizontally along the side of a slope. They are narrow steps, with a *tread* surface usually only tens of centimetres wide that runs generally parallel to the contours of the slope. They are bounded by a low steep border (*riser*) that is steeper than the slope it occupies. Some geographers believe they are associated with the processes of frost-action and creep as outlined above. However, others believe they may well be formed by the trampling of animals traversing steep slopes! Many researchers have not found any in landscapes inaccessible to animals!

Figure 4.53 Solifluction lobes and solifluction terraces

Thermokarst

Thermokarst refers to depressions in the ground surface resulting from ice melting within the permafrost. This melting causes localised subsidence giving a ground surface similar to that found in areas of limestone geology known as *karst*, where similar hollows, depressions and collapsed structures result from the erosion and weathering of the limestone bedrock. As thermokarst only results from temperature changes and not erosion and weathering, features occur where human activity inadvertently warms the surface layers of the permafrost (see Figure 4.61, page 174).

Characteristic periglacial landscapes

The preceding section illustrates a range of landforms and processes typical of a periglacial landscape. Owing to the open and sparsely vegetated nature of the landscapes, rates of erosion by water and wind can also be high, continually remodelling and shaping the contemporary landscape. In regions where there are high levels of seasonal meltwater the drainage will be dominated by **braided channels** (see Figure 4.42, page 159).

Unobstructed winds blowing across periglacial landscapes can reach high velocities. These winds can shape the landscape in the way explored in Chapter 2, with both erosional features like ventifacts and depositional features formed from fine wind-blown sediments (loess).

Process, time, landforms and landscapes in glaciated settings

It is important to view each unique glaciated and periglacial landscape as an assemblage of features and landforms that combine in that place to give it its own characteristic landscape. In order to fully understand why that landscape looks as it does today it is crucial to look beyond the processes that are happening in the present and think about how different processes may have shaped its features through time.

This is especially important because the rate of development of different landscapes in different cold environments varies considerably. Some landscapes are shaped rapidly by the almost constant action of processes of erosion, weathering and deposition. For example, where warm-based glaciers are moving rapidly they will be eroding and depositing material far quicker than in a landscape being shaped by a cold-based glacier. It is also the case that many weathering processes operate at a far faster rate in more temperate cold environments. For example, cycles of freezing and thawing may happen dozens of times during a Scottish winter, on an almost daily basis, while in some of the coldest parts of the polar regions the freezing and thawing cycle may operate over a timescale of millennia. Therefore, time has an extremely important role to play in controlling the pace of the processes that shape cold environments, even at the most local level.

To fully understand the features of a characteristic glaciated landscape it must be viewed in its entirety as an assemblage of a range of landforms. So the characteristic glaciated landscape of the European Alps is not just the glacial trough, but includes arêtes, hanging valleys, ribbon lakes, striations and truncated spurs amongst many others. It also is important to remember that the landscape is not simply the result of the current processes we can witness happening today, or imagine having happened in the *recent* past, but is also the result of cycles of processes that have occurred over a very much longer timescale. This **land systems** approach to studying landscapes gives a much broader geographical context to the study of glacial landscapes.

Ice ages have come and gone on Earth, and we are currently in the Quaternary period, which spans the last 2.58 million years. In that time there have been many glacial and interglacial fluctuations, initially operating on a 41,000-year cycle and on about a 100,000-year cycle more recently. In fact there have been eight glacials in the last 740,000 years. During each of these, ice advanced from the polar regions and, in the northern hemisphere, each glacial advance slowly altered the landscape shaped by the previous ice advance. So the glaciated landscape we see today, in the current interglacial, is also the result of cycles of processes through millennia.

As periglacial landscapes often border current glacial landscapes then it is true that they too are ever-changing landscapes. In the past, during glacials, landscapes that are currently periglacial may have been glacial and covered by ice, so the characteristic landscape we see today is not just the result of recent periglacial processes, but a combination of short-term and long-term cycles of glacial and periglacial activity.

Skills focus

It is important that you are able to describe and explain the features of characteristic glaciated and periglacial landscapes. After visiting a glaciated landscape (or by using the library or internet to research contrasting glaciated or periglacial landscapes), complete the following tasks.

a Identify the different landforms in each landscape.

b Explain the processes that created the landforms.

c Try to identify the age and origin of the features; for example, do they result from present-day processes or are they relic features that are being modified today?

d Draw and annotate a sketch of the landscape to describe the characteristic features.

Review questions

6 What landforms would you associate with a:
- glaciated landscape
- fluvio-glacial landscape
- periglacial landscape?

7 Why might it be difficult to establish settlements in periglacial landscapes?

8 How does material deposited by a glacier differ to fluvio-glacial deposits?

4.4 Human occupation and development

Despite their inhospitable climates and difficult terrain, cold environments have been home to limited human populations throughout the Holocene since the Last Glacial Maximum. However, as global population continues to expand and economic development seeks to exploit untapped natural resources, some may say that the human impact on cold environments is increasing.

Environmental fragility

Scientists are increasingly concerned about the concept of **environmental fragility**. The term **fragile environment** was defined earlier. In all natural environments there is a sensitive balance between non-living (climate, geology, soils) and living (fauna and flora) components. Some natural environments appear to be quite robust and respond easily to and cope with major natural events like forest fires and volcanic eruptions. Others are very susceptible to disturbance by people, easily damaged by human activity and then slow to recover (if at all). It is said that many of the cold environments explored in this chapter would fall into this category and be classed as **fragile environments**.

Specific examples of the ways in which people affect cold environments are explored below (Figure 4.54), but reasons for their fragility and limited ability to recover include:

- short summers, long winters, cold temperatures and limited rainfall limits plant growth
- the slow rate of plant growth means that any disruption to the ecosystem takes a long time to be corrected
- some estimate that it could take over 50 years for an area of tundra, for example, to return to its former state after interference. (It could take 50 years for tyre tracks in the permafrost to be completely revegetated!)

This is a significant issue because human activity can cause very rapid change, but many natural processes in cold environments are very slow to respond. Therefore, changes and issues affecting fragile cold environments are cumulative and can quickly become permanent. This is particularly true with issues relating to climate change.

Human impacts on fragile cold environments

Despite their hostile conditions people are present in almost all cold environments; there are few places that people have not exploited. Table 4.7 gives a summary of a range of cold environments and examples of some of the activities undertaken in them.

Table 4.7 Human activities in cold environments

Cold environment	Human activities (Past and present)
Antarctica and Southern Ocean	Sealing and whaling; mining and oil extraction; tourism; scientific exploration and research; chemical, sewage and waste dumping; maritime transport
Arctic basin	Sealing and whaling; mining; oil and gas exploration; tourism; maritime transport; small-scale hunting and fur trapping; forestry; caribou herding
The Rockies, Andes, Himalayas and European Alps	Tourism; forestry; agriculture (arable and pastoral); energy production (HEP); transport routes (using accessible glaciated valleys as routeways)
Siberia	Hunting, fishing and fur trapping; mining; forestry; oil, gas and coal extraction; fresh water supply; military bases

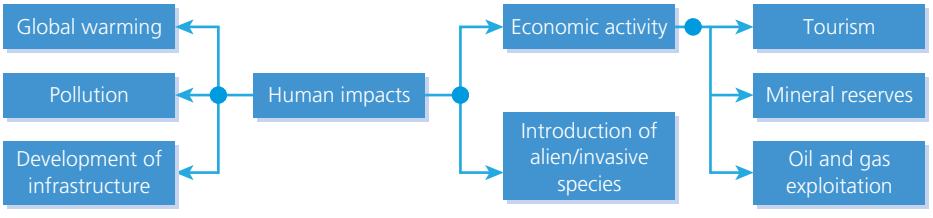


Figure 4.54 Categories of human impacts in cold environments

Links

Further information on the development of Antarctica and the Southern Ocean can be found in Chapter 7, Global systems and global governance, Antarctica as a global common, pages 334–5.

Table 4.7 (page 166) illustrates the limited range of human activities that are viable in many cold environments. In the early twenty-first century, it may not be the range of human activities that is increasing in these areas that makes them susceptible to damage, but that the scale of some of the impacts may be increasing and so too the impacts of these activities.

One of the reasons for the limited range of human activities in these landscapes is their inaccessibility and remoteness, but also the limited range of living and non-living natural resources. The limited biodiversity of most cold environments also affects their ability to be used **sustainably**. It may be that their fragility makes it almost impossible for humans to use these environments in a way that meets present needs without compromising them for future generations.

For example, the low productivity and limited species diversity mean that the plants are very specialised and any disruption causes difficulty when it comes to regeneration. In such circumstances, species have great difficulty in adapting to a changed environment.

Wide fluctuations occur in the amount of energy held in each trophic level of different food chains because population numbers change rapidly. For example, variations in the numbers of lemmings and arctic hares, both of which are liable to short-term and long-term fluctuations, have consequences for the populations of their predators, such as arctic foxes and snowy owls.

Disruption to the functioning of the biome has long-term implications. This is why there has been so much concern over the proposed exploitation of resources such as the oil reserves of north Alaska that fall within the Arctic National Wildlife Refuge.

Humans impact fragile cold environments at a wide range of scales:

- Individuals could have a direct impact on a local landscape, for example, by tourists walking on delicate vegetation.
- Individuals can contribute to global scale impacts, for example, by adding to emissions of greenhouse gases and affecting climate change.
- Communities may have impacts on their surrounding environments, for example, by changing the land use, thereby affecting the permafrost.

To a greater or lesser extent, global human activity over the last few hundred years has had, and continues to have, impacts on all fragile cold environments through resource exploitation and climate change.

Links

Links between Human impacts on fragile cold environments and Chapter 1, Water and carbon cycles, include:

- Effects of the changing carbon budget, pages 35–41, and
- Climate change mitigation, pages 41–6.

Recent and prospective impact of climate change

Climate change is a major issue in twenty-first century Geography and is explored widely elsewhere. The majority of climate scientists agree that the Earth's climate is changing, with recent decades having a pattern of accelerated warming of average global temperatures. The Earth's climate constantly changes naturally, as evidenced by the cycles of glacials and inter-glacials discussed above. However, many argue that there has been an enhanced rate of change recently that can only be explained with the inclusion of the impacts of human activity into the analysis. While the causes of climate change are not straightforward, the proposed impacts also vary widely across the globe.

Causes of climate change – natural vs. human

Table 4.8 (page 168) summarises some of the many theories that have been put forward to explain the causes of climate change.

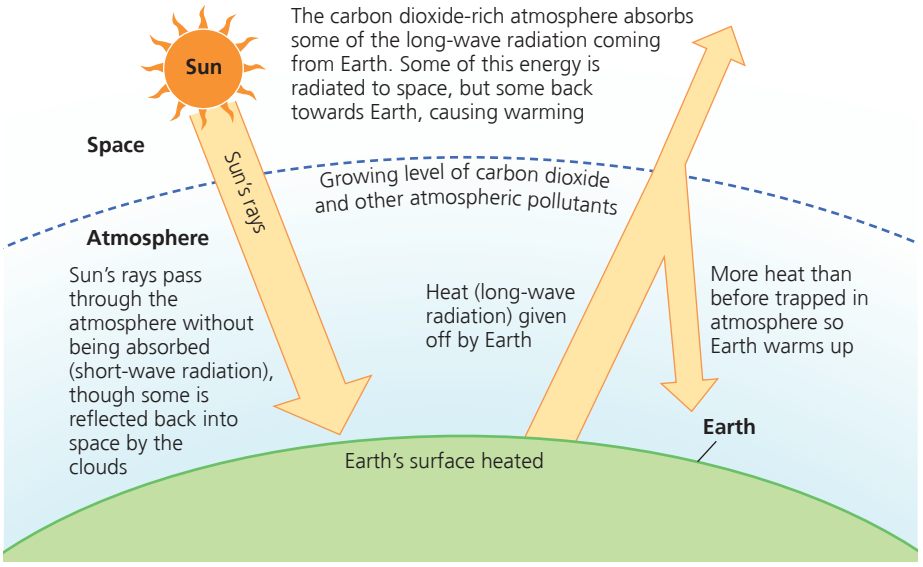


Figure 4.55 The greenhouse effect

Table 4.8 Suggested causes of climate change

Natural	Changes in solar activity (the solar constant and sunspot activity) Variations in the eccentricity of the Earth's orbit and axial tilt (affecting the amount of solar radiation reaching the surface) Meteorite impact Volcanic activity (increasing dust and altering the composition of gases in the atmosphere) Plate movement (redistribution of land masses, affecting amounts of absorption and reflection of insolation) Changes in oceanic circulation
Human	Changes in atmospheric composition, especially the build-up of carbon dioxide and other greenhouse gases from agriculture, industry, transport, energy production and waste (the enhanced greenhouse effect) Deforestation – trees act as a major carbon store; if they are felled and burnt, or allowed to rot, then the carbon returns to the atmosphere

Although there are fewer human factors listed, it is this that is seen as being the most likely explanation for the recent changes in climate. The thin layer of gases that form the Earth's atmosphere is crucial for maintaining life on Earth and without them the planet would be 30°C colder. This is known as the **greenhouse effect** (Figure 4.55).

The gases that are responsible for this greenhouse effect and trap this heat include carbon dioxide, chlorofluorocarbons (CFCs), methane, nitrous oxide and ozone. If the natural balance of atmospheric gases remains stable, and the amount of solar radiation is unchanged, then the Earth's temperature remains the same. However, Figure 4.56 suggests that there is a clear link between human activity, economic development and the balance of these gases in the atmosphere, suggesting that human activity is

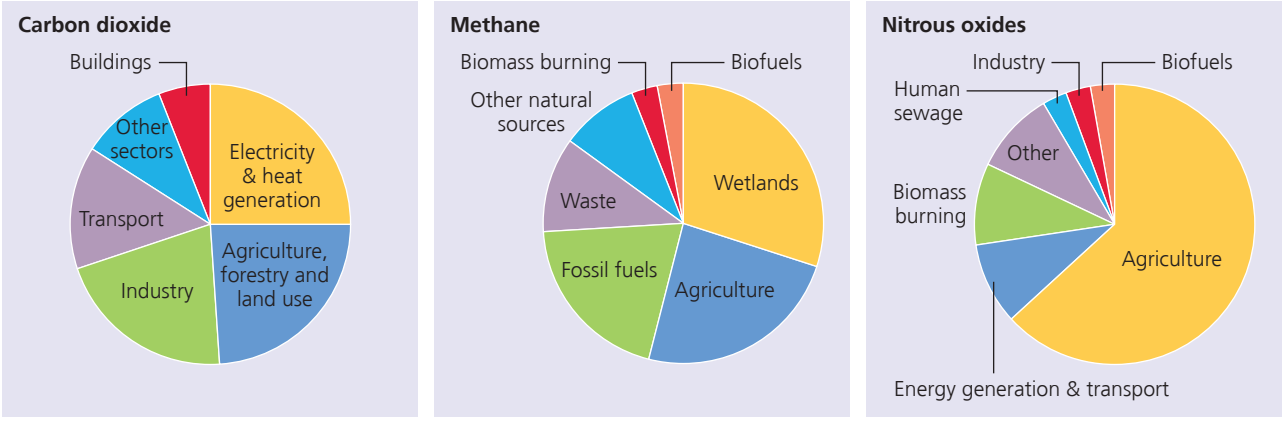


Figure 4.56 Origin of greenhouse gases

responsible for the enhanced greenhouse effect and thus current climate change.

The concentration of carbon dioxide in the atmosphere increased by about 20 per cent during the twentieth century and the current rate of increase is about 2.5 parts per million per year (Figure 4.57). Together with increases in other greenhouse gases, like methane and nitrous oxides, this has upset the natural balance and led to an increased rate of global warming in recent decades.

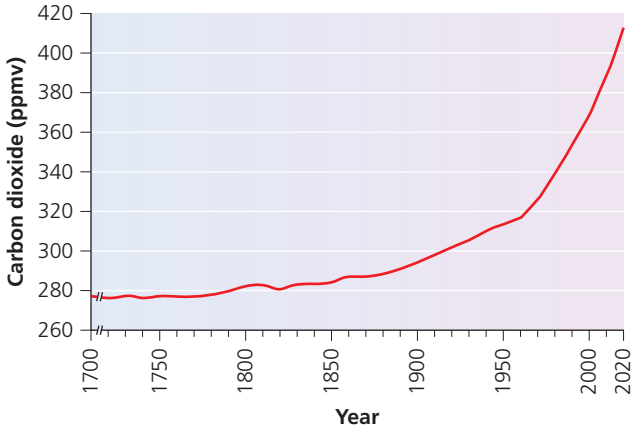


Figure 4.57 Changes in the concentration of carbon dioxide in the atmosphere, 1720–2020

Impacts of climate change in cold environments

Most geographers and climate scientists agree that the human enhanced greenhouse effect is already changing Earth's climate and already having impacts. Many of these issues are not new and have been discussed for two to three decades. In fact some impacts that were predicted 30 or 40 years ago are now occurring, with scientists making new predictions for the coming decades. Table 4.9 summarises some of these impacts.

Table 4.9 Impacts of climate change in cold environments

Type	Impacts	Evidence/details
Observable	Shrinking glaciers	Globally 90 per cent of glaciers are losing ice mass A gradual decrease in ice mass has been recorded since the 1850s, with a short period of gain around 1970, but in the decades since the rate of loss is unprecedented Millions of people rely on glacial meltwater as their freshwater supply Glaciers respond to local climate, so some glaciers in some regions do advance at certain times
	Reduced ice on rivers and lakes	In the northern USA most lakes are freezing later (about 1 day per decade) and thawing earlier (about 1.5 days per decade) than 150 years ago Many isolated communities, and the oil and gas industries, rely on frozen rivers and lakes to transport vital supplies, including drilling equipment, building materials, etc., that are too big and bulky to fly in at other times Early thaws are also increasing flood risks
	Shifting plant and animal ranges	Many predict a poleward shift in major high-latitude northern tree belts including the taiga and boreal conifer forests. These major changes will take time and so changes in treeline will be slow However, studies in the 1990s showed that the northern margins of individual species have changed markedly. For example, of 35 species of European butterflies 63 per cent had ranges that shifted north (up to 240 km in the 20th century), while only 3 per cent shifted south (Parmesan et al, 1999) Bird watchers in Britain are spotting changing patterns in the migration of certain bird species that over-winter here
Predicted and now occurring	Loss of sea ice	A small area of the Arctic Ocean has ice cover all year. It is at its smallest in September. Studies show it is shrinking. Remotely sensed images showed that in 2012 the area was the smallest on record. In 2019, satellite records showed that the 13 lowest ice extents occurred in the preceding 13 Septembers. Arctic ice is also becoming thinner, making it more vulnerable to enhanced rates of melting in the future.
	Accelerated sea-level rise	Studies show that the rate of sea level rise for most of the twentieth century was about 1.7 mm per year. Satellite monitoring then showed that mean global sea levels rose by around 3 mm per year between 1993 and 2009. In 2019, NASA reported the current average rate is around 3.4 mm per year. (Thermal expansion and the addition of meltwater contribute to the increase)
	Longer, dryer summers	By the early 2000s satellite data confirmed that longer Arctic summers were thinning the ice dramatically. Some suggest a 40 per cent thinning since the 1960s. Polar bears rely on sea ice from which to hunt seals and without it they are finding it increasingly hard to find enough food. In some places in Alaska and northern Canada polar bears have been straying into towns searching for food

Table 4.9 (Continued)

Type	Impacts	Evidence/details
Predicted future change	Highest expected temperature change to be in the highest northern latitudes	Average global temperatures could rise between 1.4 and 5.8°C by 2100 Latitudes between 40°N and 70°N could rise between 5 and 8°C by 2100, due to the loss of sea ice and snow cover reducing albedo rates (less insolation will be reflected by the surface)
	Continued contraction of area covered by snow and ice	The UN predicts mid-northern latitudes in Europe and North America may see the largest rates of decline in snow and ice cover. One estimate predicts that loss of snow and glaciers in the mountains of Asia alone will affect the lives of 40 per cent of the world's population Areas of Canada and Siberia may receive increased snowfall due to more humid climates
	Permafrost to thaw to increasing depths	A 6°C temperature increase in the Arctic by 2100 would lead to a 30 to 85 per cent loss of near-surface permafrost. This could permanently change local ecosystems and hydrology and increase the risk of wild fires and soil erosion. Local infrastructure such as buildings and roads could be affected causing economic and social problems for local communities. The impact could be global as models suggest that melting permafrost could release huge amounts of carbon dioxide and methane, thus accelerating atmospheric warming. This is a positive feedback known as the **permafrost carbon feedback**
	Further decrease in extent of sea ice	The UN predict that the Arctic may be free of sea ice by 2100, and that Antarctic sea ice may decline at similar rates, but not thin as much. Changes in sea ice could have major impacts on ocean circulation and thus the major climate patterns of the world, which is proving extremely difficult to model and predict. Increasing ice-free periods in the Arctic could have major impacts on world shipping, in addition to disrupting the lives of local populations
	Shrinking of Greenland ice sheet leading to sea level rise	The ice sheet is expected to almost completely melt by 2100 contributing to projected sea level rise by 2100 of between 18 and 59 cm above 1990s levels. Other factors like thermal expansion will be important and modelling the future is complex
	Possible growth of Antarctic ice sheet due to increased snowfall	Warmer temperatures could increase the amount of precipitation. Other theories suggest that the extent of the ice sheet may be increasing due to acceleration of glacial movement to the sea
	Possible significant impacts of small temperature increases on fragile tundra ecosystems	The IPCC suggest even a 1–2°C rise in global temperatures above 1990 levels jeopardises unique and already threatened ecosystems If temperatures rise by more than 1.5–2.5°C, 20–30 per cent of species could be at risk of extinction; if the rise is more than 3.5°C models predict extinctions of 40–70 per cent of known species
	Increased issues of invasive species in warmer tundra environments	This could increase extinction rates of indigenous species that are outcompeted for resources. Local populations' traditional ways of life may be affected

Cold environments as human habitats

The distribution of population within the cold environments

Despite the inhospitable characteristics of cold environments, many people live and work in these places. The proportion of indigenous and non-indigenous populations alters as changes in development affect the challenges and opportunities for people in these regions. Over 4 million people live and work within the Arctic Circle, but the size and structure of populations vary in different regions (Figure 4.58). All Arctic regions experience differences in natural increase and net migration.

Opportunities and risks presented by cold environments

The physical environment impacts human activities. The harsh conditions of cold environments present challenges for human occupation and development. These conditions include:

- very low temperatures
- short summers and therefore short growing seasons

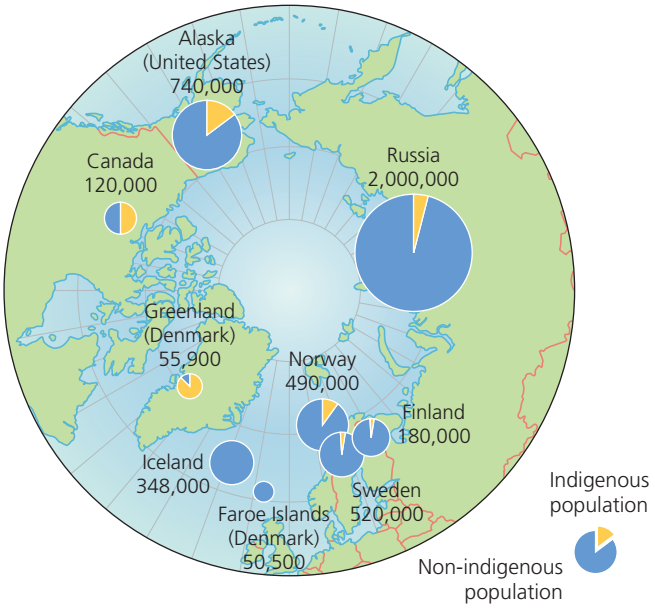


Figure 4.58 Population distribution in the Arctic

- low precipitation
- thin, stony, poorly developed soils
- permafrost
- surface thaw in summer leading to waterlogging and possible flooding
- snow lying for long periods
- blizzards.

Many indigenous people have adapted their way of life to cope with the climate and make the most of the limited resources. The traditional economic activity of indigenous populations in many tundra regions included hunting and fishing, caribou and reindeer herding, fur trading and even hunter gathering. These communities were considered to have a **sustainable** way of life, because they:

- had low population densities
- only took what resources they needed from the natural environment
- had social structures and cultural systems that actively promoted the protection of their environment.

In some cases the natural environment is revered in a religious or spiritual way. A case study of one local population is developed later in this chapter.

The peoples living in these cold environments face challenges posed by the physical environment itself, but life is made increasingly difficult by other, more variable and human issues. Such obstacles to life include:

- a shortage of labour and skills due to low population density
- increasing unpredictability of climate events – late winter freezes, or early spring melts
- remoteness and inaccessibility
- lack of permanent jobs and limited educational opportunities after secondary level
- absence of modern conveniences common in urban areas
- remoteness can lead to a feeling of economic or political neglect from national government
- impacts from outside on natural resources that are relied upon, including climate change, commercial fishing, etc.

Human adjustments to life in cold environments

Although challenging, cold environments hold many opportunities for people, so over millennia people have adapted to surviving in them.

The **Inuit**, indigenous people of areas of North America, adapted all aspects of their lives to survive in the harsh tundra landscape, including their:

- lifestyle
- shelters
- food
- transport.

Lifestyle and shelters

They lived in small communities, with small family size and low birth rates. Spread over a large geographic area at very low population density, they put little pressure on the scarce resources. They were often nomadic to limit the impact on resources in each location, but also to follow migrating animals.

Therefore, in the summer they built portable tents out of driftwood and seal skins, while in winter they needed more substantial shelter so would build snowhouses, or igloos, or in more settled communities log-and-sod houses. These log-and-sod dwellings would be built partially in holes in the ground, so the surrounding earth provided more protection from the elements.

Food

Inuit would make the most of the short summer by hunting, gathering and preserving as much food as they would need to see them through the winter. They would often dry fish or seal meat so that it would last through the year. Many indigenous communities of the north have adapted to be able to safely eat raw meat, which again limits the need for extra resources. Experience and tradition tells them how much food to hunt, and they will not take more than is needed. All parts of any prey are used: for clothing, utensils or even building materials.

Transport

In the winter many Inuit communities use dog sleds to exploit the frozen landscape, and in the summer use single person kayaks or larger open boats called umiaks. All of the above are used for transport and hunting. In many cold environments being mobile is a key to survival.

Contemporary adaptations in the Arctic tundra

Establishing permanent settlements and developing activities such as oil and gas extraction have required major technological advances, primarily because the permafrost creates a unique set of problems for construction work and engineering. Many innovations have come from efforts to overcome the problems posed by the permafrost. New methods of construction have been employed to protect the permafrost and prevent the subsidence of buildings and infrastructure. Although they are successful, these methods are more expensive than conventional construction, adding to the costs of living in the region, and continual maintenance is often necessary. Some of the methods are described below.

- Smaller buildings, such as houses, can be elevated above the ground on piles driven into the permafrost. The gap below the building allows air to circulate and remove heat that would otherwise be conducted into the ground (Figure 4.59).
- Larger structures can be built upon aggregate pads, which are layers of coarse sand and gravel, 1–2 m thick. This substitutes the insulating effect of vegetation and reduces the transfer of heat from building to ground. These pads can also be placed beneath roads, railways and landing strips.

Figure 4.59 Buildings on stilts and insulated ducting for water and power in Prudhoe Bay, Alaska

- In large settlements **utilidors** (Figure 4.60) have been built. A utilidor is an insulated box, elevated above the ground that carries water supplies, heating pipes and sewers between buildings. Pipes cannot be buried underground because of the damage that would be caused by freezing and thawing in the upper soil layers.

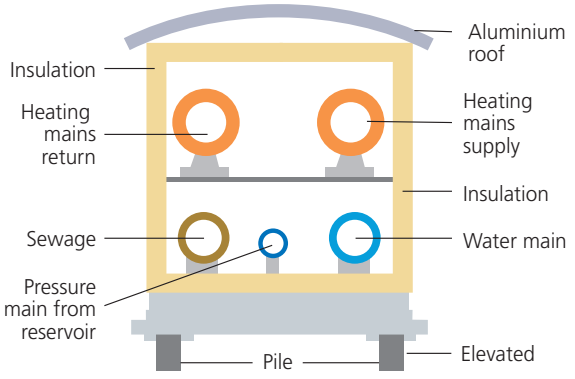


Figure 4.60 A utilidor

Typical forms of occupation and economic activity in different cold environments

Different cold environments present different challenges and opportunities, so people have learnt to occupy and exploit them in different ways. Table 4.10 summarises these activities.

Table 4.10 Economic activity in cold environments

Cold environment	Characteristics and economic activity
Alaska	With abundant natural resources, the remote wild landscapes have seen the development of oil, mining, fishing, timber and tourism
Northwest Territories, Canada	Its size and remoteness have limited much economic development, but expanding sectors include diamond mining, oil and gas, the fur trade and tourism
European Alps	Being easily accessible with steep mountain slopes, high snowfall rates and high living standards makes them popular for tourists wanting winter sports. Equally, the beautiful scenery and hot sunny weather is popular with hikers, climbers and sightseeing tourists in the summer months. The glacial troughs and narrow valleys are good locations for the production of renewable energy in the form of HEP and wind turbines, and serve as essential routeways for transport
Siberia	Its size and remoteness make it inaccessible, however its rich ore and mineral wealth have led to the development of mining of iron, silver, copper and gold, alongside coal, oil and gas extraction. The vast forests provide for a large timber industry
Iceland	Being an island, Iceland has traditionally relied on the fishing and whaling industries. Today it also exploits its geothermal energy to produce vast amounts of clean energy. In recent years, sightseeing and adventure tourist numbers have increased dramatically, with government policy actively promoting the tourism sector.
Antarctica	The extreme climate, protected status, lack of pollution and human population, make it suitable for adventure tourism and scientific research
Southern Ocean	The cold nutrient-rich waters have always been used for fishing. In the late 1990s hundreds of thousands of tonnes of krill and Patagonian toothfish were harvested, before international agreements came in to try to reduce the impact. With only limited visitor numbers allowed onto Antarctica, cruise ships have become increasingly common in the waters surrounding the continent

Links

See Chapter 7, Antarctica as a global common, pages 334–5, for further understanding of economic activity within Antarctica and the Southern Ocean.

Changing impact of human activity on fragile cold environments over time

In North America the number of Inuit was always small in terms of the vast area in which they lived, so very little pressure was put on the environment, which traditionally remained relatively undisturbed.

In the north of Europe, the Sami followed the seasonal movement of reindeer and as long as they lived at low population densities they knew their environment could sustain them, provided they managed its use sustainably.

From the seventeenth century onwards, the resources of tundra areas began to be exploited by outsiders. Major forms of economic activity, like sealing, whaling, trapping for fur and mining, particularly for gold, began to grow. Mining led to the establishment of permanent settlements whereas other activities tended to be temporary or seasonal. All activities began to bring larger numbers of people than had lived there previously. In the twentieth century, tundra regions were exploited on a much larger scale, with dramatic impacts on the lifestyle of indigenous people as well as on the physical environment.

In the tundra many of the negative human impacts occur when people try to establish permanent settlements and infrastructure. Problems are caused when vegetation is cleared from the ground surface. This reduces insulation and results, in summer, in the deepening of the active layer. Even minor disturbances, such as vehicle tracks, can greatly increase melting, because the vegetation is very slow to re-establish itself.

Buildings speed up this process by spreading heat into the ground. The thawing of ground ice leads to the development of unnatural **thermokarst**, a landscape of topographic depressions characterised by extensive areas of irregular, hummocky ground interspersed with waterlogged hollows. The damage caused by this form of ground subsidence can be seen in tilted and fractured older buildings and in damage to roads, railways and airfield runways (Figure 4.61, page 174).

Figure 4.61 Thermokarst under a car park, Fairbanks, Alaska

Apart from the production of thermokarst, human impacts on the physical environment include:

- hunting: over exploitation
- transport: risks of spillages, especially in places like Alaska where oil and gas extraction dominate; road vehicles damaging the ground
- tourism: vegetation removal, litter and waste do not easily degrade. There have been well-publicised campaigns about this issue even on the slopes of Mount Everest where huge amounts of garbage and human waste are left by climbers every season
- general air pollution
- research: even in the Antarctic, where the number of people allowed on the landmass is limited, their presence, including transport and dealing with waste, will undoubtedly have some impact.

All of these factors occur in local areas, but they are adding to the global effects of climate change, melting the snow and ice of these environments.

Management of cold environments at present and in alternative possible futures

Some suggest that it is their lack of development that gives many cold environments the potential for sustainable development in the future. Much of the Arctic tundra and Antarctica typify the common perception of wild and natural places. Their remoteness and the extreme physical processes keep them inaccessible to mass tourism and the excesses of economic development. Conservationists believe wilderness areas have intrinsic value and possess outstanding qualities that are worth conserving for the future. Areas such as these have an aesthetic value for people seeking spiritual refreshment and contemplation. Scientifically, they are important because:

- there is a need to maintain the gene pool of wild organisms to ensure that genetic variety is maintained
- animal communities can be studied in their natural environment
- wilderness is a natural laboratory for the scientific study of ecosystems
- there is a need for pure natural systems to be used as a yardstick against which managed or mismanaged systems can be compared.

There are good reasons for conserving wilderness regions, but they also often contain a range of exploitable resources. Pressure to develop these resources comes from national and transnational groups that require both energy sources and raw materials to support industrial growth. Balancing developmental pressures against the need to conserve the essential values of wilderness is the increasingly difficult task of management. Sustainable development has an important role to play here, but there is disagreement about how it may be successfully applied in many wilderness environments.

In 1964, the Wilderness Act in the USA designated a number of wilderness areas. The largest number of designated areas in any state is in Alaska, which instituted its own wilderness legislation, the National Interest Lands Conservation Act, in 1980.

Managing Antarctica at present

Links

Antarctica as a global common is explored in detail in Chapter 7, Global systems and global governance, pages 334–5.

Figure 4.62 summarises some of the key aspects of the management of Antarctica.

The whole of Antarctica is now designated as a 'reserve devoted to peace and science' by the

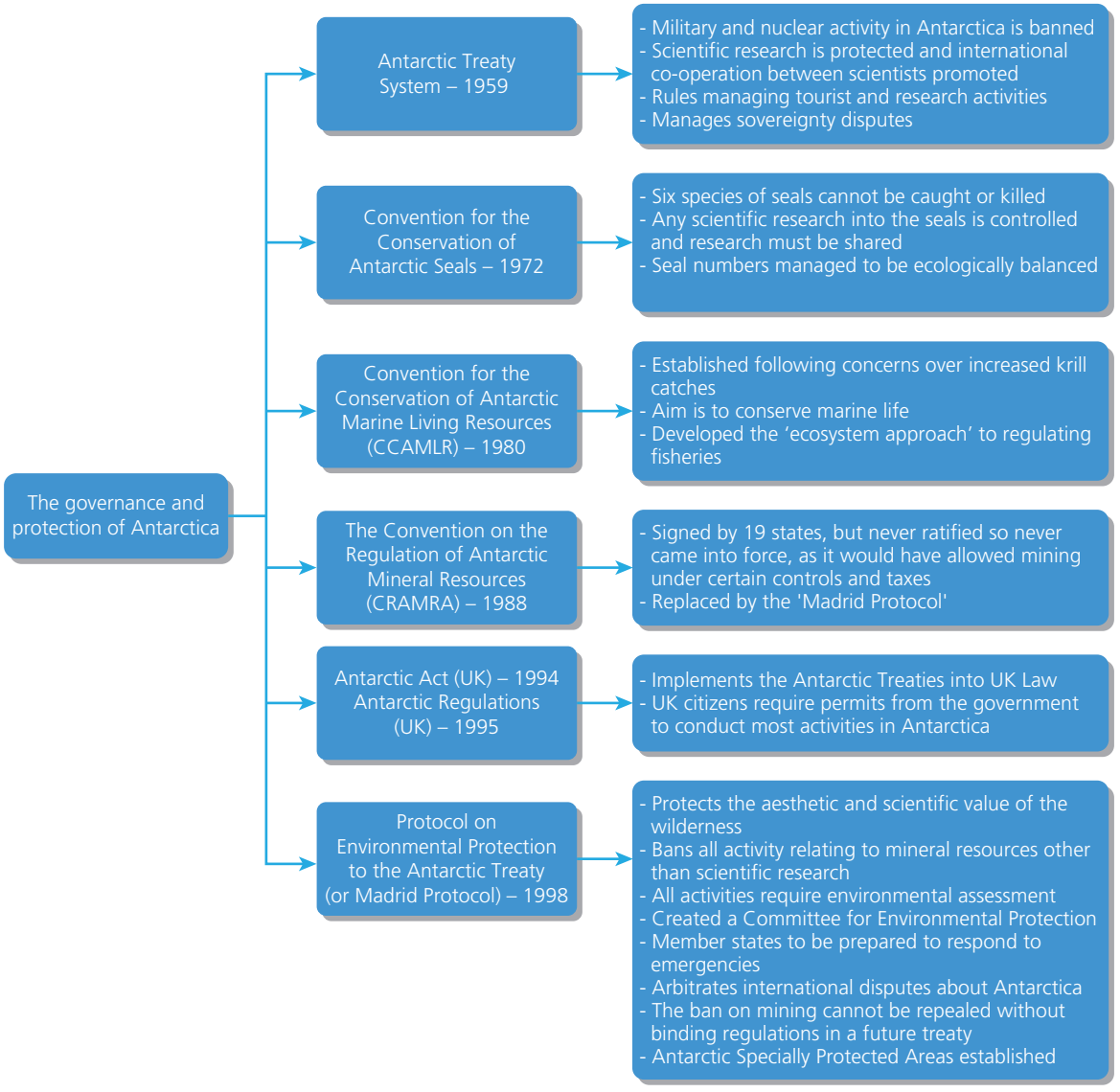


Figure 4.62 Managing Antarctica

Figure 4.63 Memorial cross at Cape Evans for Captain Scott's expedition

protocols of the Antarctic Treaties. However, it is not just the natural environment that is protected; other protocols protect the cultural and historical importance of the continent. Some sites have been designated as Historic Sites and Monuments (HSMs). Visit the website of the British Antarctic Survey to find out more about HSMs including the cross at Cape Evans on Ross Island, erected in 1913 as a memorial to Captain Scott's party that perished in 1912 (Figure 4.63).

Managing cold environments: alternative futures

Currently the Antarctic Treaties and Environmental and Madrid Protocols comprehensively protect Antarctica. However, there is a growing debate over these questions: Should current treaties be adapted or removed to encourage development in Antarctica? Or should the Treaty be maintained in its current form?

Table 4.11 (page 176) summarises some of the arguments about possible future development in Antarctica.

Table 4.11 Arguments about future possible developments in Antarctica

Observation	Arguments
Energy and mineral resources elsewhere are becoming depleted.	• Antarctica is believed to have large reserves of coal and oil, and precious metals such as gold and silver. Can we envisage Antarctica having quarries, offshore oil rigs and oil wells on land? • Mining is banned until at least 2048, and any prospecting or extraction in and around Antarctica would be prohibitively expensive due to its remoteness and climate. So, do conservationists need to be concerned?
Numbers of tourists visiting Antarctica remain high, with over 44,500 land visits and 11,000 cruise-only visitors in 2018–19.	• If tourism were to be encouraged as an economic activity could we envisage Antarctica having a commercial airport and port, with tourist hotels and shops? • The growth of adventure tourism relies on the affluent clientele being able to buy into a unique experience. If more tourists were allowed to visit current sites, operators may seek to expand operations into new sites that have not been visited by people so far. (Some controls were placed on cruise ship numbers in the early 2010s.)
In some parts of the world, commercial fishing is depleting fish stocks to near extinction levels.	• The waters of the Southern Ocean still have plentiful fish stocks. Even though fishing is heavily regulated, fishing levels around the continent are already potentially undermining conservation efforts. • If restrictions were lifted, would levels of overfishing quickly reach the same as in other areas, or would the remoteness and hostile conditions naturally limit commercial fishing activity?
Bioprospecting is a growing area of scientific research, with many keen to search for genetic and biochemical resources from Antarctica's flora and fauna.	• Antarctica's biodiversity is unique, with many companies already expressing an interest in research. • The continent's animal and plant life is of interest, because their unique adaptations to deal with the harsh conditions may be useful for future valuable commercial products, like medicines or synthetic materials. There is already great interest in chemicals found in some species that act like natural antifreeze, which could have many commercial uses. • Bioprospecting is happening in Antarctica as it is often classed as scientific research, and if profits from resulting products could be targeted towards conservation and protection of the wilderness, should bioprospecting be promoted?

This debate is too extensive to develop further here. Many geographers may instinctively fall on the side of continued protection, and maintaining and strengthening of the current protocols and treaties, but a full discussion of **sustainable development** in Antarctica would need to at least consider some of the questions above. Other cold environments will have their own unique issues raised by discussions of their future sustainable development.

Review questions

9 Why do you think people would choose to live in such harsh cold environments?

10 Can you explain why there are significant seasonal variations in the population of different cold environments?

11 Can you explain why there are debates about how to manage many cold environments in the future?

4.5 Quantitative and qualitative skills in glaciated landscapes

Quantitative and qualitative techniques in geography are defined and explored in Chapter 12 and you should read this section carefully.

There are many opportunities to visit, and undertake fieldwork, in a glaciated landscape as part of this course, and it is important to reflect on how a range of quantitative and qualitative skills could be utilised in a study of the processes and landscape outcomes in such a setting.

Quantitative data and glaciated landscapes

Glaciated landscapes provide a wealth of opportunities to employ a range of observational and measurement techniques to collect quantitative data. A good starting point is completing annotated field-sketches. Others include measuring:

- characteristics of glacial troughs and corries, such as cross-sections, depth and width
- evidence of ice movement, for example distribution and orientation of striations or glacial erratics
- till fabric analysis and the size, shape and level of sorting of glacial and fluvioglacial sediment
- characteristics of landforms such as drumlins and roche moutonnées.

Other quantitative techniques could be used to identify, measure and map post-glacial landscape features, or types of land use.

Quantitative investigations of glaciated landscapes can produce a large amount of useful data that readily lends itself to manipulation and analysis, with a range of data presentation and statistical techniques, as explored in Chapter 12. Fieldwork in glaciated landscapes also produces quantitative data that is easily mapped using a range of geo-spatial mapping techniques to help give a clear visual representation of the findings of the research.

Qualitative data and glaciated landscapes

This chapter illustrates how predicting the impact of climate change on all cold environments is complex. Even more challenging will be assessing the increased challenges this may pose for people and ecosystems in many fragile cold environments around the world. Equally, those tasked with the challenge of mitigating the worst impacts of any future climate change while improving the resilience of communities and helping them adapt to their changing circumstances, will require a more holistic assessment of many aspects of their situations. Qualitative techniques such as participant observation and in-depth interview schedules would be one way of engaging with the views of the populations involved to inform the development of bottom-up solutions to future challenges. The case study on page 180 exemplifies a region where such techniques could be employed – with Sami communities of the Sápmi region of Scandinavia.

Perhaps on a more local level, questionnaires and interviews could be used with different communities in glaciated landscapes in the British Isles to assess the views of various stakeholders about any land use conflicts, for example in the Helvellyn area of the English Lake District (page 178). The qualitative results of such surveys could then be analysed alongside more empirical data about factors such as visitor numbers, second home ownership, unemployment rates and impacts on the natural environment, to produce thorough cost-benefit or SWOT analyses of any proposed land use changes, that include the views of all stakeholders.

Fieldwork opportunities

Here is an opportunity to compile a questionnaire that could be used with stakeholders to gather information about the views of different groups involved in a current land use conflict in a glaciated landscape such as the English Lake District.

Case study of a glaciated environment at a local scale to investigate glacial processes and their landscape outcomes: The Helvellyn area of the English Lake District

Processes of landscape development

The present form of the Lake District owes much to the work of the Pleistocene glaciation modifying and deepening the pre-glacial drainage system. Ice has advanced and retreated about 11 times in the last 800,000 years. The last glacial maximum in northwest England was about 22,000 years BP.

During glacial periods, ice was the main agent at work shaping the landscape by processes of glacial erosion, with abrasion and plucking widening and deepening pre-existing valleys and hollows and steepening rock faces. As the ice retreated through this area of the Lake District, the landscape was shaped by glacial deposition and glacial meltwater. During interglacials the landscape was further shaped by periglacial processes, weathering and mass movement.

The present-day landscape is therefore an assemblage of landforms that have resulted from a range of glacial, periglacial, fluvio-glacial and postglacial processes that operated over hundreds of thousands of years. This landscape continues to be shaped by various geomorphological processes and human activity.

Evidence of many of these processes, and their outcomes in the landscape, are relatively simple to engage with through fieldwork.

Landscape outcomes

Many of the lakes – Ullswater and Windermere, for example – radiate outward from the centre of former ice accumulation. These **ribbon lakes** occupy the glacially deepened valleys. Many of the smaller lakes, or **tarns**, occupy the basins of **corries** found at higher altitudes and are in the areas most favourable for snow accumulation, so became the source of glaciers. The glaciers over-deepened the original hollows and during interglacials filled with meltwater. There are over 150 corries within the Lake District as a whole.

Using both investigations in the field and the use of secondary data, including OS maps, many aspects of these landscape features could be investigated, for example:

- orientation, of corries and ribbon lakes, within the landscape
- size and shape of features, e.g. length and width of ribbon lakes, or height of corrie headwalls.

The Helvellyn Range runs approximately north–south, including the peak of Helvellyn itself (950 m), which is part of a ridge of high ground over 600 m extending for about 11 km. Most of the corries are found on the east side of the ridge, many facing north-east, making them conducive to snow accumulation during colder periods (Figure 4.64).

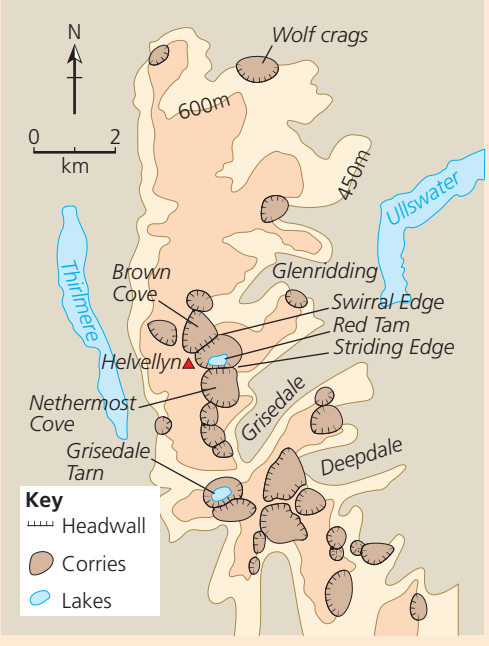


Figure 4.64 Corries, arêtes, glacial troughs and lakes in the Helvellyn area

Many of these corries are adjacent to each other, and over time erosion has created small, knife-edged ridges separating the hollows. Several such **arêtes** are found in the area, for example, Swirral Edge and the impressive Striding Edge (Figure 4.65). It may not be appropriate to investigate these features in the field, but OS maps could be used to map and identify the distribution of arêtes in the area.

Figure 4.65 Red Tarn and Striding Edge from Helvellyn summit

Grisedale and Glenridding valleys are excellent examples of glacial troughs (Figure 4.66). Both provide opportunities to investigate evidence of past processes of glacial erosion, for example:

- mapping the location and orientation of striations in exposed rock surfaces
- identifying and determining the location and scale of features such as valley sides and floors, truncated spurs, hanging valleys and possible roche moutonnée.
- identifying evidence of weathering and post-glacial features such as frost shattering and scree slopes or rock fields.

Higher up the Grizedale valley there are many depositional features including numerous moraines. In the Grisedale valley there are many moraines from various glacial phases with many exhibiting signs of subsequent erosion and weathering since their formation. Using both investigations in the field and sources, such as OS maps and aerial images, it could be possible to investigate the following features of these moraines:

- distribution
- orientation
- scale.

Figure 4.66 The glaciated trough of Grisedale valley

Wherever there is evidence of glacial deposition, it may be possible to investigate characteristics of the deposited material through till fabric analysis. (These and other techniques are outlined in the Quantitative and Qualitative Skills section on page 177).

Case study of a contrasting glaciated landscape from beyond the UK: The Sápmi region of tundra, northern Europe

Geographical context

The Sápmi region is the area stretching across the far north of Norway, Sweden, Finland and the Kola Peninsula of Russia, which is the cultural home of the Sami indigenous people (Figure 4.67).

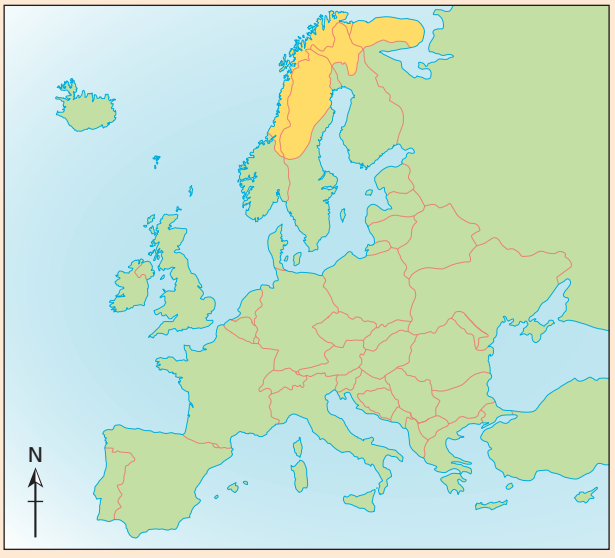


Figure 4.67 The Sápmi region or northern Europe

Approximately 2 million people live in the region with a mix of ethnic origins, including the Sami, Russians, Norwegians, Swedes and Fins. Individuals from each of these national groups would also class themselves as ethnically Sami.

A challenging place to live

Sami people have lived in this cold environment for centuries. They embrace its **challenges** and **opportunities** with **resilience**, **mitigation** and **adaptation**.

Traditional adaptations of the Sami people of Sápmi

People have bred and herded reindeer in Sápmi for over 1,000 years (Figure 4.68). Due to the diversity of environments in different parts of Sápmi different forms of reindeer breeding developed in different areas.

Figure 4.68 A Sami reindeer herder at a round-up in northern Norwegian Sápmi

Until the mid-twentieth century, many family groups remained fully nomadic, moving their herds to the shelter of the coniferous forests in the winter and the

Table 4.12 Challenges of the Sápmi region for the Sami people

Social	Economic	Environmental
• The Sami are a minority population in the region and their culture and way of life have been oppressed for a long time. In recent decades there have been efforts to protect the rights of, and promote the cultures and livelihoods of, the indigenous population. • Young Sami are attracted to the more settled comforts and education and employment opportunities of a modern urban life so are leaving their traditional communities. • Cultural traditions are threatened due to ageing populations with fewer people for traditions to be passed on to.	• For those following a more modern settled life, the Sápmi region as a whole is rich in mineral deposits, has ice-free ports for transport and has seen an increase in tourism and military installations, bringing wealth to the area. • All of the above add to the challenges experienced by those wishing to maintain traditional economies based on reindeer herding, fishing, hunting and small-scale agriculture. • Only 10 per cent of the Sami manage to remain solely as nomadic reindeer herders, as activities like commercial logging threaten their lands.	• The Sami traditionally live mostly inside the Arctic Circle, which has short milder summers, but long dark winters. • The relief is generally a low plateau, but with many lakes, rivers and marshes, making long distance travel difficult. • Vegetation is generally sparse, but the south is very densely forested. • Rivers have been dammed for hydroelectricity production; the frozen lakes and ice can no longer be used to walk the reindeer to new grazing lands, so lorries and boats have to be used.

rich pastures of the high mountains in the summer. This continual movement limited the impact on either environment. Tame reindeer were used for transport and to lead the herd to new grazing each day. The Sami survived on the reindeer themselves, consuming their meat and milk. Every part of the reindeer was processed into food, clothes or utensils, including skin, fur, sinews, bone and antlers.

With strong links between family groups and culture and traditions going back hundreds of years, knowledge about interacting with nature and understanding when to move grazing areas under changing climatic conditions was passed from generation to generation.

They also understood the importance of limiting their impact on, and protecting the natural resources around them, so many Sami communities had very low fertility levels. By spacing births a number of years apart it meant the number of young children that could not walk with the animals was at a minimum, and most waited until children could move independently with the herd before having more children. This naturally kept population numbers low thus limiting the impact on the environment and resources. The traditional Sami lifestyle could be viewed as sustainable, as it provided for their needs without negatively affecting their environment. However, during the second half of the twentieth century there have been many challenges to this traditional way of life, with some questioning the viability of sustainable development in the area in the future.

Possible sustainable development of the Sápmi region

Many view the growth of settlements and modern industry in Sápmi as unsustainable. If the traditional way of life of the Sami is seen as a sustainable use of this region then a number of issues will need to be overcome.

There are three conditions needed for the Sami and their reindeer to thrive:

- Access to old forests with good supplies of ground and tree-dwelling lichens.
  - Most Sami do not own their land.
  - Many of their grazing rights are informal and based on historical customs.
  - Commercial forestry is clearing large areas.
  - The right of the Sami to freely graze their reindeer in many areas of Sápmi is protected, but increasing numbers of landowners are making legal challenges to this.
- A family needs over 500 reindeer to live off without any other income.
  - As other land uses limit the amount of grazing land available this can put pressure on the size of the herd.
- Reindeer need to be able to move freely to find fresh sources of food.
  - Modern infrastructure, like roads and railways and HEP reservoirs, along with settlement growth, blocks traditional migratory routes.

However, many international organisations and the four national governments of the region increasingly seek to protect indigenous peoples and promote their traditions and ways of life. This seems like a difficult balancing act, with so many different Sami communities spread across areas of four separate nations who all also see the needs of others who wish to exploit the natural wealth of this region.

The Tåssåsen Sámi and their environment

An example of a Sami community is the Tåssåsen Sámi community, Östersund, Jämtland County, Sweden. The Tåssåsen Sami community lives in central parts of the Swedish mountain range west of the town of Östersund in Jämtland County and, with other Southern Sami groups, like those near Sveg to the south, has called this part of northern Europe home since prehistoric times (Figure 4.69).

Figure 4.69 Map showing the location of Jämtland County, Sweden

By the early 2000s this was one of the few remaining communities of Southern Sami living a traditional way of life.

Background

- Grazing grounds covered 111,000 hectares – with winter grazing including an area 100 km by 30 km of old forest.
- For Sami they live quite far south, so occupy mainly mountain grazing.
- Narrow valleys are used to direct the herd and bring them together to be counted and inspected.
- The reindeer herd numbered 5,500.
- The community numbered 50 individuals in families split into 13 grazing groups.
- Their culture and traditions were very strong and unlike other communities the young people had stayed and kept the population young and viable.

In the summer the Tåssåsen guide their reindeer up onto the higher mountain slopes to graze the fresh ground-growing lichens to supplement the lichens growing on the trees. As the winter approaches the herds are guided back down to the lower forested slopes. Here the trees provide shelter and the snow cover on the forest floor is much lighter and easier to dig through to reach the lichens beneath, than that outside of the forest.

The best lichen for the reindeer are those that grow hanging from trees (tree pendent lichens). These tree pendent lichens only grow in old undisturbed spruce, pine and birch forests that take 100–200 years to mature. A small area of this remains in Jämtland, used by the Tåssåsen Sami, especially in winter. If the area of forest is large and fertile enough some small herds could remain in the forest all year.

Impact on the environment

In their traditional territories, when the reindeer were able to be freely guided from grazing ground to grazing ground before they had overgrazed the lichens, the Sami had little impact on the physical environment of Jämtland.

However, in recent decades some argue that the Sami's reindeer are damaging the forest ecosystem. The debate is complicated, and the Sami would dispute this. Some claim any degradation of the forests is down to modern forestry practices.

Figure 4.70 Tåssåsen reindeer herd in winter

The Tåssåsen Sami and the future

Like all indigenous people living in the cold environments of the north, climate change is a potential threat to the Tåssåsen Sami. As outlined above, this area of the northern hemisphere is projected to have the highest rates of warming. Possible impacts:

- Lichen growth is dependent on winter snow and ice conditions. Warmer winters with less snow would reduce the amount of lichen for grazing.
- Lichen occupy a very specific ecological niche. Even slight warming will increase competition with other plant species, again reducing the amount for grazing.
- Later onset of, and warmer winters will mean fewer rivers and lakes freeze, thus limiting the movement of herds between grazing grounds. Some areas could become overgrazed.
- Reindeer and the Sami are in tune with the changing seasons. Some groups may be slow to adjust to rapid changes in seasonal patterns.
- Summer warming may increase insect populations. The reindeer could lose energy and time on dealing with insect harassment rather than foraging.

Many Sami have already made the decision to end their nomadic way of life and farm reindeer in a more settled commercial way. However, this ends hundreds of years of culture and tradition. Also, it is a more capital intensive way of farming requiring the input of more capital for fixed pasture, fences and feed, so is more expensive and requires fewer people. Those that adopt these new approaches quickly have to diversify into other modern commercial activities to make the reindeer profitable and support all members of the community.

It may be that the predicted effects of climate change mean the traditional grazing lands cannot support the reindeer herds in the future anyway. However, there is a more immediate and tangible threat to the traditional way of life of the Tåssåsen reindeer herders of Jämtland.

Much of the forest area that they have used traditionally is now owned privately, largely by commercial logging businesses. Since the 1990s there have been well-publicised court cases brought by the commercial forestry industry to attempt to ban the reindeer herders from their lands. Community leaders like Olof Johannsson say this could be the end of a culture dating back over 1,000 years.

Further research

You could use the following websites to research the Sápmi region and the Sami in more detail:

http://samenland.nl/lap_sami_si.html
www.oloft.com/casestudy.html
www.iwgia.org/en/sapmi

Further reading

The texts below develop many of the ideas set out in this chapter. Some are specialist texts and quite academic and your local library may be able to find copies should you wish to take your studies further.

Woodward, J. (2014) *The Ice Age: A Very Short Introduction*. Oxford University Press

Benn, D. & Evans, D. (2013) *Glaciers and Glaciation*. 2nd Edition. Routledge

Raw, M. and Knight, J. (2007) *Glaciation and Periglaciation Advanced Topic Master*. Philip Allan

Redfern, D. (2010) *Climate Change: Contemporary Case Studies*. Philip Allan Updates

Harvey, A. (2012) *Introducing Geomorphology: A Guide to Landforms and Processes*. Dunedin Academic Press

King, N. (2019) *A-level Geography Topic Master: Glaciated Landscapes*. Hodder Education

Knight, P. G. (2015) *A Quick Introduction to Glaciers and Glacial Landscapes*. CreateSpace Independent Publishing Platform

The following websites are also good sources of additional information on some of the topics covered.

To discover more about the characteristics of the range of types of patterned ground and more detail on the theories about the processes that formed them, and to see a range of photographic examples visit: www.erdc.usace.army.mil/CRREL/Permafrost-Tunnel-Research-Facility/Patterned-Ground/

Issues of managing Antarctica are far-reaching and complex. Suggested websites include: www.bas.ac.uk/about/antarctica/environmental-protection/

Question practice

A-level questions

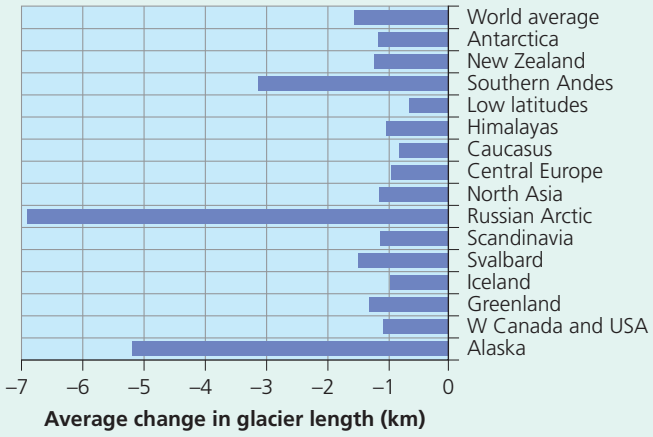


Figure 4.71 Average change in glacier length in different regions of the world

Source: The Cryosphere

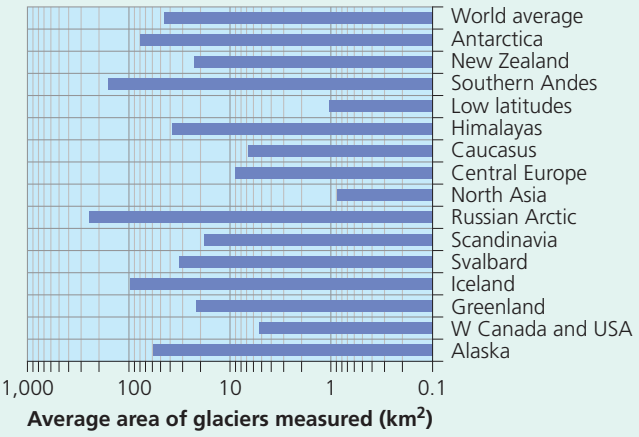


Figure 4.72 Average area of each glacier measured in the same data set as Figure 4.71

Source: The Cryosphere

1 Outline processes of ice movement. [4 marks]

2 Analyse the information shown in Figures 4.71 and 4.72. [6 marks]

3 Using Figure 4.73 and your own knowledge, assess the extent to which material is being transported in different ways by this glacier. [6 marks]

4 'Processes of glacial erosion are the most important factor in shaping glaciated landscapes.'

To what extent do you agree with this view, with reference to a local scale glaciated environment you have studied? [20 marks]

5 'Future climate change will create more challenges than opportunities for people living in glaciated landscapes.'

To what extent do you agree with this view, with reference to a glaciated landscape beyond the UK that you have studied. [20 marks]

Figure 4.73 Snout of the Sólheimajökull glacier in southern Iceland

Chapter

5 Hazards

In this chapter you will focus on the atmosphere and the lithosphere (the earth's crust and the section immediately below it – the upper mantle), which intermittently but regularly present hazards to human populations, often in a dramatic and sometimes catastrophic fashion. By studying the origin and nature of these hazards, their impacts and people's responses, you will be exploring some of the relationships between people and the environment in which they live. The study of natural hazards will also give you the opportunity to exercise and develop your observational skills, measurement and geospatial mapping skills, together with data manipulation, interpretation and presentation of statistics.

In this chapter you will study:

- the concept of a hazard in a geographical context and plate tectonics theory
- the background and examples of four types of hazard:
  - volcanic hazards
  - seismic hazards
  - storm hazards and
  - wildfires.

For each type of hazard, you should study one example of the impacts of the hazard and human responses to the hazard. For storms, you need to study a second contrasting example.

Links

There are links to Chapter 8, including The importance of place in human life experience, page 354, and Changing places: meaning and representation, page 364.

5.1 The concept of a hazard in a geographical context

The concept of a natural hazard and its potential impact

A **natural hazard** is a perceived event that threatens both life and property. Natural hazards often result in disasters that cause some loss of life and/or damage to the built environment and create severe disruption to human activities. Natural or environmental hazards include volcanic activity, seismic events (earthquakes) and tropical storms (hurricanes/typhoons/cyclones). All of these can therefore cause disruption to human systems, including death and injury, property and communication system damage and the disruption of economic activities.

These forms of hazard thus pose a risk to human populations. That risk is increased because we build shanty towns on unstable tropical slopes, urbanise volcanic zones, live in areas with active faults and on coasts susceptible to hurricanes and tsunamis. This is also exacerbated by the failure to recognise a potential hazard and act accordingly.

Natural hazards, and their effects on people, tend to have the following common characteristics:

- their origins are clear and the effects that they produce are distinctive, such as earthquakes causing buildings to collapse
- most natural hazards only allow a short warning time before the event (some hardly at all)
- exposure to the risk is involuntary, particularly for populations of less well developed countries. In developed areas, most people who occupy hazardous areas are often well aware of the risks and they choose to minimise or even ignore them
- most losses to life and damage to property occur shortly after the event although the effects of natural hazards can be felt in communities long after that time (disease, disruption to communications and economic activities)

- the scale and intensity of the event requires an emergency response.

Key terms

Adaptation – In the context of hazards, adaptation is the attempts by people or communities to live with hazard events. By adjusting their living conditions, people are able to reduce their levels of vulnerability. For example, they may avoid building on sites that are vulnerable to storm surges but stay within the same area.

Fatalism – A view of a hazard event that suggests that people cannot influence or shape the outcome, therefore nothing can be done to mitigate against it. People with such an attitude take limited or no preventative measures. In some parts of the world, the outcome of a hazard event can be said to be 'God's will'.

Natural hazards – Events which are perceived to be a threat to people and the built and natural environments. They occur in the physical environments of the atmosphere, lithosphere and the hydrosphere.

Perception – This is the way in which an individual or a group views the threat of a hazard event. This will ultimately determine the course of action taken by individuals or the response they expect from governments and other organisations.

Risk and vulnerability

Risk is the exposure of people to a hazardous event presenting a potential threat to themselves, their possessions and the built environment in which they live. People though, consciously put themselves at risk from natural hazards and the question has to be why do they do it?

Possible reasons include the following:

- **Hazard events are unpredictable:** We cannot predict the frequency, magnitude or scale of a natural hazard event.
- **Lack of alternatives:** Due to social, political, economic and cultural factors, people cannot simply uproot themselves from one place and move to another, giving up their homes, land and employment.
- **Changing the level of risk:** Places that were once relatively safe may have become, through time, far more of a risk. Deforestation, for example, could result in more flooding from torrential rain associated with tropical storms and there could also be a greater risk from landslides.
- **Cost/benefit:** There are many hazardous areas that offer advantages that in people's minds outweigh the risk that they are taking. Californian cities, for example, have a high risk from earthquakes, but people see the many advantages of living there as greater than the potential risk.
- **Perception:** see below.

Vulnerability to physical hazards means the potential for loss. Since losses vary geographically, over time and among different social groups, vulnerability therefore also varies over time and space. Researchers at the University of South Carolina have examined all the variables which link risk and vulnerability and have come up with the model shown in Figure 5.1.

This raises the question, is the risk the same for all people in an area? In other words, are some people more

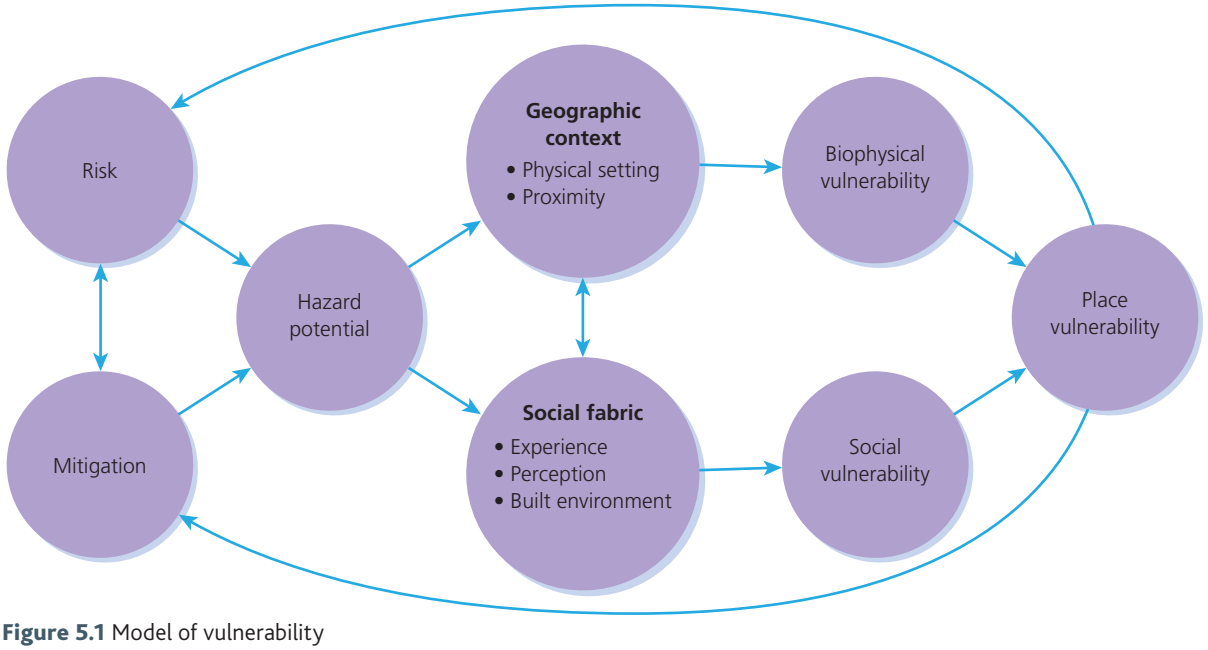


Figure 5.1 Model of vulnerability

vulnerable than others? A similar sized natural hazard event can have widely varying impacts in different parts of the world. People's wealth and the level of technology that they can apply do affect the degree to which the hazard event will impact upon them. Richer people and countries can protect themselves by building sea defences, constructing earthquake-resistant buildings, providing better emergency services, etc. They can also be better prepared by being made more aware of the risk through education. The people of cities in poorer countries are more vulnerable. As such urban areas have grown, more and more people have been forced to live in hazardous areas such as very steep hillsides that are prone to landslides, and in the lowest lying parts where they are at risk from tropical storms and tsunamis.

The perception of natural hazards

People react to the threat of hazards in different ways because of the way in which individuals receive and process information. Perception is influenced by many factors including:

- socio-economic status
- level of education
- occupation/employment status
- religion, cultural/ethnic background
- family and marital status
- past experience
- values, personality and expectations.

Perception of a hazard will ultimately determine the course of action taken by individuals in order to modify the event or the responses they expect from governments and other organisations.

There is often a great difference in the perception of a hazard between peoples of differing levels of economic development. In wealthier areas there is a sense that the better prepared you are, the more able you will be to withstand the impact of the hazard and perhaps even prevent the disaster from taking place. This is usually based upon government and community action, and is backed by capital that will fund technologically-based solutions. The sense of helplessness in the face of natural hazards tends to increase with the level of poverty and the deprivation of the people. Even in wealthier countries there are groups of disadvantaged people who tend to look upon natural hazards as part of their way of life, as they are seen as unavoidable, just as the bulk of people in poorer countries see the impacts of these events as being part of the conditions of poverty.

People may perceive natural hazards in the following ways:

- **Fatalism (acceptance):** Such hazards are natural events that are part of living in an area. Some communities would go as far as to say that they are 'God's will'. Action is therefore usually direct and concerned with safety. Losses are accepted as inevitable and people remain where they are.
- **Adaptation:** People see that they can prepare for, and therefore survive the event(s) by prediction, prevention, and/or protection, depending upon the economic and technological circumstances of the area in question.
- **Fear:** The perception of the hazard is such that people feel so vulnerable to an event that they are no longer able to face living in the area and move away to regions perceived to be unaffected by the hazard.

Management of natural hazards

Key terms

Community preparedness/risk sharing – This involves prearranged measures that aim to reduce the loss of life and property damage through public education and awareness programmes, evacuation procedures, the provision of emergency medical, food and shelter supplies and the taking out of insurance.

Integrated risk management – The process of considering the social, economic and political factors involved in risk analysis; determining the acceptability of damage/disruption; deciding on the actions to be taken to minimise damage/disruption.

Lava – Molten rock (magma) flowing onto the surface. Acid lava solidifies very quickly, but basic lava (basaltic) tends to flow some distance before solidifying (for example, on the Hawaiian Islands).

Mitigation – Long-term action taken to reduce or eliminate the risk to life and property from hazard events. Action is taken before, during and after disasters to break the cycle of damage and repair in hazardous areas.

Prediction – The ability to forecast a hazardous event and then give warnings so that action can be taken to reduce their impact. Improved monitoring, information and communications technology have meant that this has become more important and accurate in recent years.

Resilience – The sustained ability of individuals or communities to be able to utilise available resources to respond to, withstand and recover from the effects of natural hazard events. Communities that are resilient are able to minimise the effects of the event, enabling them to return to normal life as soon as possible.

People respond to natural hazards and the threats that they pose by seeking ways to reduce the risk. Responses

can come from individuals, the local community, national governments and international agencies.

Community **resilience** is the sustained ability of a community to utilise available resources to respond to, withstand and recover from the effects of natural hazards. Communities that are resilient are able to minimise the effects of a hazard, making the return to normal life as effortless as possible.

A key feature of the modern approach is that hazards are best combated by efficient management. Modern management techniques, with their gathering of information, careful analysis and deliberate planning, aim to make the most efficient use of the money available to confront natural hazards. A process known as **integrated risk management** is often used. This incorporates:

- identification of the hazard
- analysis of the risks
- establishing priorities
- treating the risk and implementing a risk reduction plan
- developing public awareness and a communication strategy
- monitoring and reviewing the whole process.

The governments of many countries use such schemes. An example developed by the New Zealand government is shown in Figure 5.2.

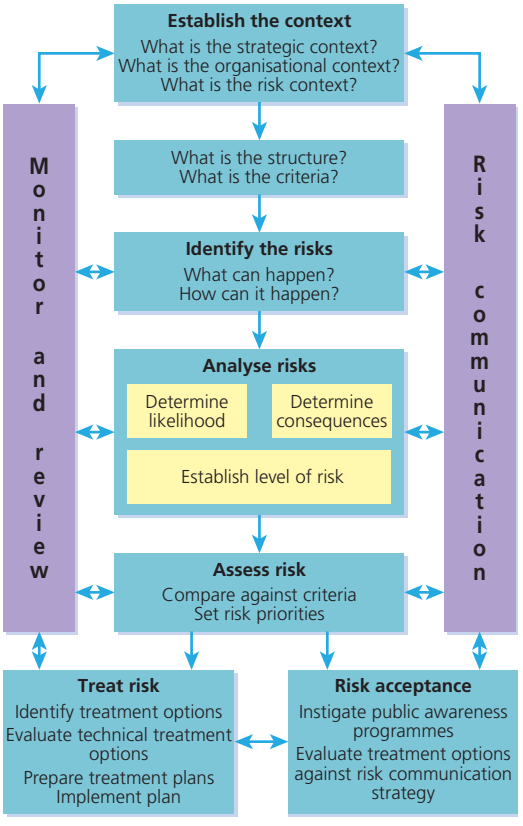


Figure 5.2 The process of risk management

People and organisations therefore try to manage natural hazards in the following ways:

- **Prediction:** It may be possible to give warnings that will enable action to be taken. The key to this is improved **monitoring** in order to give predictions which means that warnings can be issued. The National Hurricane Centre in Florida is a good example of an agency demonstrating how prediction can depend upon monitoring, through the use of information from satellites and land-, sea- and air-based recordings.
- **Prevention:** For natural hazards this is probably unrealistic although there have been ideas and even schemes such as seeding clouds in potential tropical storms in order to cause more precipitation, which in theory would result in a weakening of the system as it approached land.
- **Protection:** The aim is to protect people, their possessions and the built environment from the impact of the event. This usually involves modifications to the built environment such as improved sea walls and earthquake-proof buildings. One way in which governments can act, and people react, is to try to change attitudes and behaviour to natural hazards which will reduce people's vulnerability. **Community preparedness (or risk sharing)** involves prearranged measures that aim to reduce the loss of life and property damage through public education and awareness programmes, evacuation procedures and provision of emergency medical and food supplies and shelters. There can also be attempts to modify losses through insurance (richer areas) and international aid (in poorer regions).

All attempts at management must be evaluated in terms of their success. Successful schemes include the use of dynamite to divert lava flows on Mt Etna and pouring sea water on **lava** flows in Iceland. On the other hand, the well-prepared Japanese government was caught out in 2011 by an earthquake in north-eastern Japan that triggered a tsunami. 15,894 people were killed and 2,500 were reported missing. Hundreds of thousands of buildings were destroyed or badly damaged. According to the Japanese government the final cost was estimated at US $200 million.

Disaster/risk management cycle: This illustrates the ongoing process by which governments, businesses and society plan for and reduce the impact of disasters, react during and immediately following an event, and take steps to recover after an event has occurred. Appropriate actions at all points in the cycle lead to greater preparedness, better warnings and reduced vulnerability or the prevention of hazard events during the next cycle. The complete cycle includes the shaping of public policies and plans that either modify the causes of the hazard events or mitigate their effects on people, property and infrastructure.

One of the main goals of disaster management, and one of its strongest links with development, is the promotion of sustainable livelihoods and their protection and recovery during such events. Where this goal is achieved, people have a greater capacity to deal with disasters and their recovery is more rapid and long lasting.

The Federal Emergency Management Agency (FEMA) was created in the USA in 1978. The agency's primary purpose is to co-ordinate the response to a disaster that has occurred in the United States and that has overwhelmed the resources of local and state authorities. They have created major analysis programmes for floods, hurricanes and earthquakes. Their operations are carried out very much along the lines shown by the disaster/risk management cycle model shown in Figure 5.3.

Disaster/response curve: To show that hazard events can have varying impacts over time, Park (1991) devised his impact/response model (Figure 5.4). This model shows an early stage, before the disaster strikes, where the quality of life is normal for the

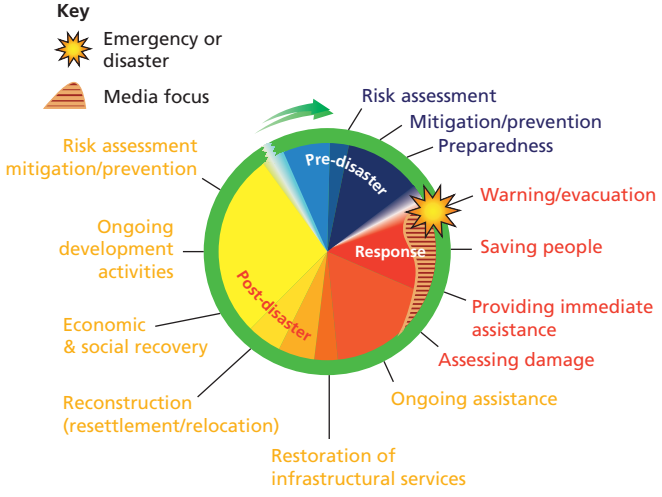


Figure 5.3 The disaster/risk management cycle

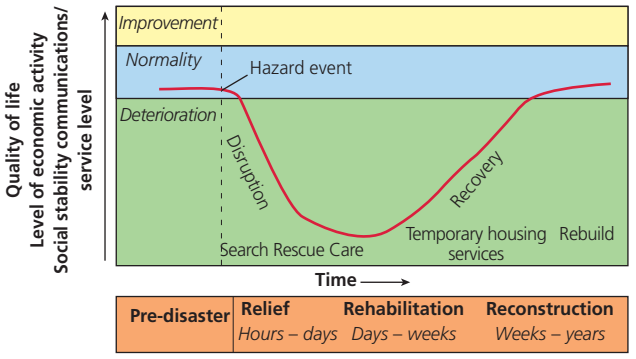


Figure 5.4 The Park impact/response model

area. Here people try their best to prevent such events and prepare in case they should happen. When the event happens, the quality of life suddenly drops with people taking immediate action to preserve life and, if possible, the built environment.

The next stage Park called **relief**, where medical attention, rescue services and overall care are delivered. This can last from a few hours to several days if the event has been very damaging. From this point the quality of life of the people of the area starts to slowly increase.

Next comes **rehabilitation**, where people try to return to normal by providing food, water and shelter for those most affected. This period can last anything from a few days to weeks.

Finally comes **reconstruction**, where the infrastructure and property are repaired or rebuilt and crops regrown. At this time people use the experience of the event to try to learn how to better respond to the next one. This period can take from weeks to several years.

Key terms

Frequency – The distribution of a hazard through time.

Magnitude – The assessment of the size of the impact of a hazard event.

Primary effects – The effects of a hazard event that result directly from that event. For a volcanic eruption these could include lava and pyroclastic flows. In an earthquake, ground shaking and rupturing are primary effects.

Secondary effects – These are the effects that result from the primary impact of the hazard event. In volcanic eruptions these include flooding (from melting ice caps and glaciers) and lahars. In an earthquake, tsunamis and fires (from ruptured gas pipes) are secondary effects.

Soil liquefaction – The process by which saturated, unconsolidated soil or sand is converted into a suspension during an earthquake. It is then able to act like a liquid and flow, particularly when under pressure.

Park also showed how different events can have different impacts. This is shown in Figure 5.5 (page 190) by the speed of the drop in the quality of life, the duration of the decline and the speed and nature of recovery. The difference in the three lines could be related to the type of hazard, the degree of preparedness or the speed of the relief effort, and the nature of recovery and rebuilding.

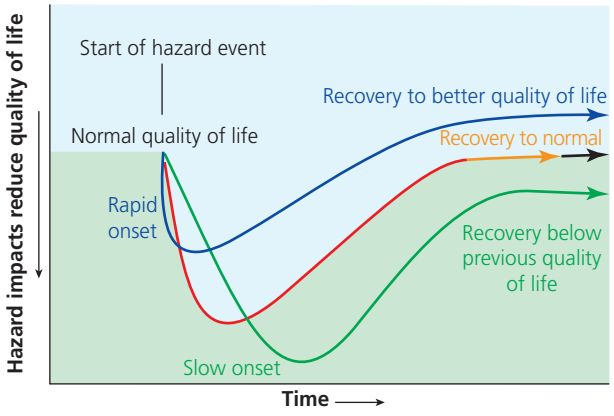


Figure 5.5 Variations within the Park model

Distribution, frequency and magnitude

Distribution refers to the spatial coverage of the hazard. This can refer to the area affected by a single event, some of which can have a very localised effect, while others have a much wider effect, such as tsunamis which can cross large oceans. Volcanic eruptions have been known to have a global effect with the spread of dust in the upper atmosphere and the consequent short-term climatic change. Distribution of a hazard can also refer to the areas where the particular hazard is likely to occur. Earthquakes and volcanoes, for example, are generally associated with tectonic plate boundaries whereas tropical cyclones usually occur between 5° and 25° north and south of the Equator.

Frequency refers to the distribution of the hazard through time whereas **magnitude** assesses the size of the impact. At its most basic level, the frequency-magnitude principle leads us to expect many small insignificant events and fewer large-scale ones.

Skills focus

Crowdsourcing and 'big data' (see page 662) can help in the recovery from a hazard event. Research an example where they have been used and assess their effectiveness.

The effects of natural hazard events

The effects are the impact that an event has upon both the physical and human environments. Some commentators differentiate between primary and secondary effects when considering hazard impacts. For an earthquake, for example, the **primary effect** would be the ground shaking and cracking followed by the **secondary effects** of **soil liquefaction**, landslides, tsunamis and the effects on people and the built environment such as collapsing buildings, fires, flooding and the knock-on effects which could be with the population for a long time. Communication systems could be out of order, the ability to produce food crops may take some time to be restored and the economy of a region may be so damaged that the legacy of the hazard event will be around for years.

Review questions

1 Assess the validity of the statement, 'The richer you are the greater the chance of surviving a natural hazard'.

2 For one example of a hazard, assess the extent to which the Park model can be applied to the impacts and responses to that hazard.

5.2 Plate tectonics

Earth structure

The interior of the Earth is divided into three main components: the crust, the mantle and the core (Figure 5.6, page 192).

The core

The **core** is made up of dense rocks containing iron and nickel alloys. It is subdivided into a solid inner core and a very hot molten outer one, with a temperature of over 5,000 °C. This heat is produced mainly as the result of two processes: **primordial** heat left over from the Earth's formation and **radiogenic** heat produced by the decay of radioactive isotopes, particularly Uranium-238, Thorium-232 and Potassium-40.

The mantle

The **mantle**, the thickest layer, is made of molten and semi-molten rock rich in iron and magnesium (such

as peridotite). At a depth of between approximately 700 km and 100 km, the upper part of the mantle becomes hotter and more fluid. This is called the **asthenosphere**. Above the asthenosphere the mantle becomes solid.

The crust

The outermost layer is called the **crust**. There are two types of crust; a relatively thin layer of dense basalt is found under the oceans and a thicker layer of less dense granitic rock makes up the continents. Table 5.1 summarises the differences between the two types of crust.

Table 5.1 Differences between continental and oceanic crust

	Continental crust	Oceanic crust
Thickness	30–70 km	6–10 km
Age	Over 1,500 million years	Less than 200 million years
Density	2.6 (lighter)	3.0 (heavier)
Composition	Mainly granite; silicon, aluminium, oxygen	Mainly basalt; silicon, magnesium, oxygen

The theory of plate tectonics

Plate tectonic theory was developed in the 1960s. It completely revolutionised the study of Earth Science. The theory emerged from earlier work by the German researcher, Alfred Wegener, who gathered evidence (Table 5.2, page 192) to suggest that just one giant continent (Pangaea) existed about 300 million years ago. He believed that it later split into two smaller continents (Laurasia and Gondwanaland). Today's continents were then formed from further splitting of these two masses. Wegener called his ideas, the theory of **continental drift**.

Key terms

Asthenosphere – Part of the Earth's mantle that lies below the lithosphere, at depths between about 100 and 700 kilometres.

Conservative plate margin – A plate margin where two tectonic plates are moving past one another with no addition or destruction of plate material.

Constructive plate margin – A plate margin where new crust is generated as the plates pull away from each other. These are found at mid-oceanic ridges.

Destructive plate boundary – A plate margin where crust is destroyed as two plates converge. These are usually associated with island arcs or young fold mountains.

Palaeomagnetism – A record of the history of the Earth's magnetic field, preserved in magnetic minerals in volcanic rocks.

Plate tectonic theory – The theory that states that the Earth's crust is made up of several rigid plates moving relative to one another.

Plume – A hot column of magma which rises from deep within the Earth.

Rift valley – A long, deep valley found in the centre of a spreading ridge. It is formed between parallel faults where a block of the crust has sunk down.

Sea-floor spreading – The theory that the ocean floor is moving away from the mid oceanic ridge and across the deep ocean basin, to disappear beneath continents and island arcs.

Tectonic plate – One of a series of rigid sections of the Earth's crust. They float on the upper mantle and move relative to one another.

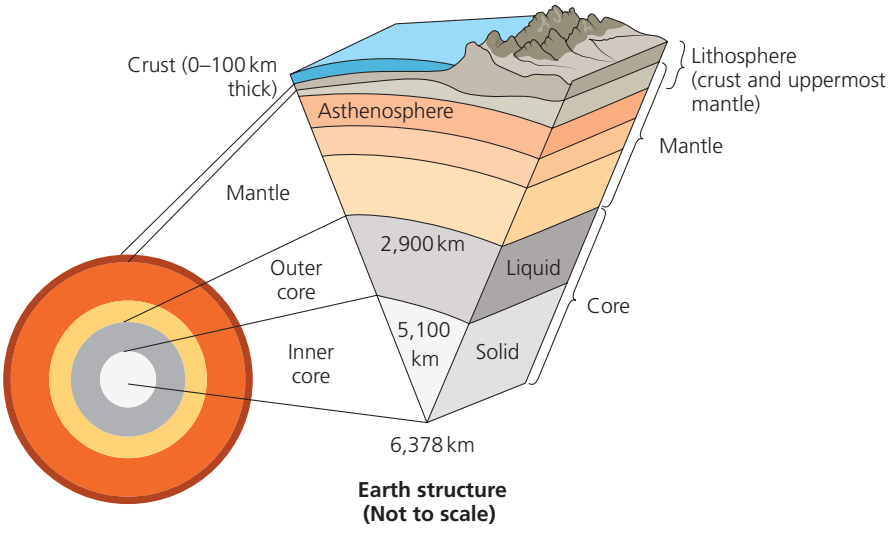


Figure 5.6 The structure of the Earth

Table 5.2 Evidence for the theory of continental drift

Geological evidence	• The 'jigsaw fit' of South America and west Africa • Evidence of similar ancient glacial deposits found in South America, Antarctica and India. They must have been formed together and then moved • Structural faults in rocks in Brazil and west Africa match when the two are compared • Similar rock sequences in northern Scotland and eastern Canada, indicate that they were formed under the same conditions in one location
Biological evidence	• Fossils found in India are comparable with fossils in Australia • Fossil remains of the reptile *Mesosaurus* are found in both South America and southern Africa indicating they were once joined together • Identical plant fossils have been found in the coal deposits of both India and Antarctica

Crustal evolution

Little is known about the early development of the Earth's crust; much of the geological record has disappeared over time as elements of the crust have been recycled. The oldest rocks are approximately four billion years old. It is thought, however, that about three billion years ago there was an increase in the growth of continental crust; rocks of that age are found in the middle of the large continental plates.

Plate movement, as we think of it today, developed about 750 million years ago. It was at this time that Pangaea split up and life on Earth began to evolve rapidly. Deep mountain belts, surface erosion and the recycling of continental crust back into the mantle also developed.

Twenty-first century developments in the theory of plate tectonics suggest that the crust and rigid mantle collectively make up the **lithosphere** and that the lithosphere is divided into seven large continental and ocean **tectonic plates** and several smaller ones (Figure 5.7).

The theory states that the lithosphere is able to slide over the asthenosphere and this allows plate movement. As the lithospheric plates move across the Earth's surface (typically at five to ten centimetres per year), driven by forces that are not yet fully understood, the plates interact along their boundaries where they converge, diverge or slip past one another. Such interactions are thought to be responsible for most of the Earth's seismic and volcanic activity, although earthquakes and volcanoes can also occur in plate interiors. Where plates push together (converge), mountains are formed. Where plates pull apart (diverge), continents break apart and oceans form. The continents within the plates, and the ancient hearts of the continents (cratons), drift with the plates. Over millions of years these movements result in significant changes to the Earth's geography.

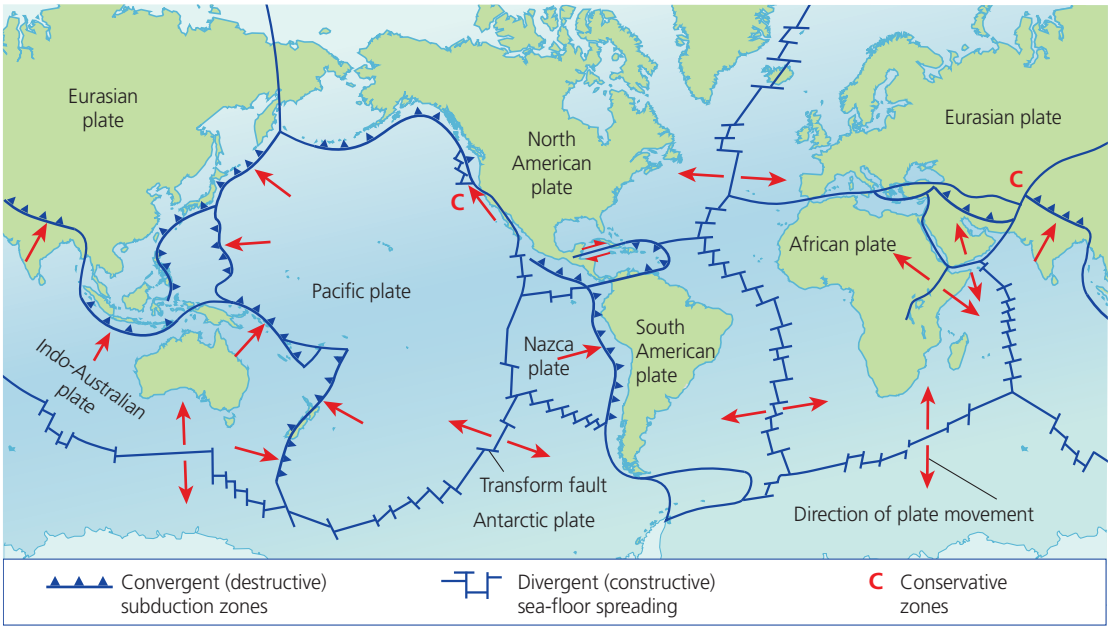


Figure 5.7 Tectonic plates and their margins

Evidence supporting the main theory of plate tectonics was starting to emerge in the 1940s. The mid-Atlantic ridge was discovered and studied along with a similar feature in the Pacific Ocean. A study of the polarity of ancient magnetic particles in the rock (**palaeomagnetism**) found that:

- the polarity of the rock either side of the mid-Atlantic ridge alternated in a striped pattern – that was mirrored on either side of the ridge (Figure 5.8)
- the oceanic crust was slowly moving away from the plate boundary
- the oceanic crust got older with distance from the mid-oceanic ridge (although geologically speaking it is very young, nowhere older than 200 million years).

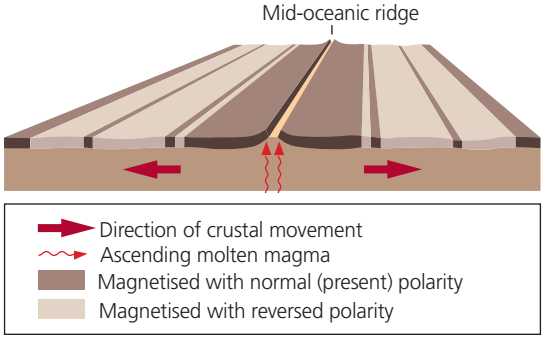


Figure 5.8 Magnetic 'stripes' on the Atlantic Ocean floor

Sea-floor spreading: If the sea floor is spreading, it implies that the Earth must be getting bigger. This is not the case. To accommodate the new crust being created at mid-oceanic ridges, crust must be destroyed elsewhere. Evidence of crust being destroyed was found in huge oceanic trenches where large areas of ocean floor were being pulled downward in a process known as **subduction**.

Plate movement

Convection currents

Lithospheric plates are massive and require huge forces to move them. It was previously thought that **convection currents** within the mantle drove plate movement. Uneven distribution of temperatures towards the base of the mantle caused convection cells which dragged the lithospheric plates with them when the moving mantle reached the surface. This theory has been partly discounted because most of the mantle is not fluid. Scientists now believe that convection currents play only a supporting role in the current dynamic models (ridge push and slab pull) in which plates move as part of a **gravity-driven system** (see Figure 5.9, page 194).

Ridge push or gravitational sliding

At constructive boundaries, less dense, hot **magma** wells up and produces an ocean ridge standing some

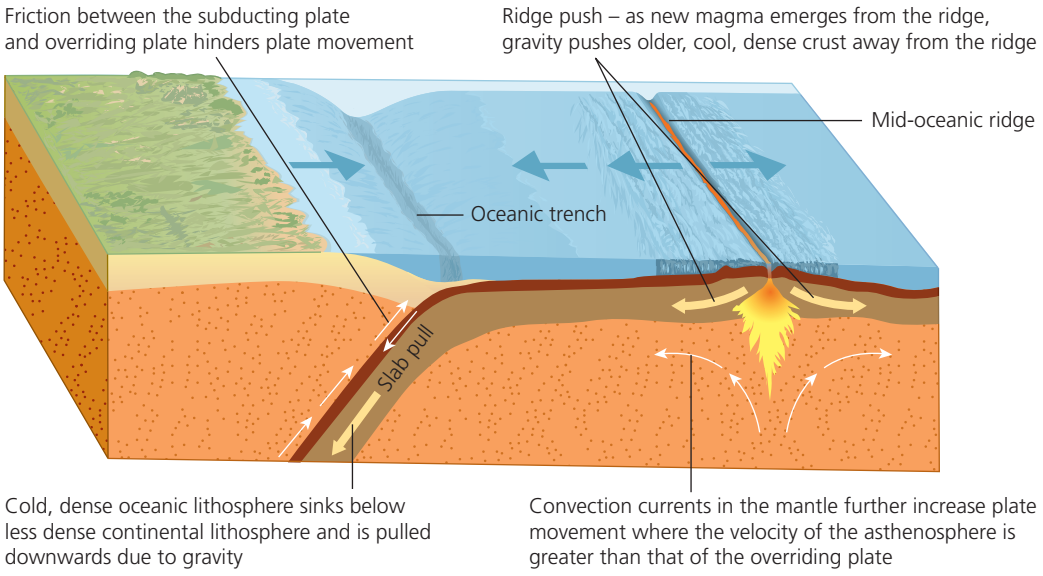


Figure 5.9 Mechanisms contributing to plate movement

two to three kilometres above the ocean floor. As this newly formed rock gets older, it cools becoming denser. Gravity acting on this older, denser lithosphere causes it to slide away from the ridge, down the sloping, semi-molten asthenosphere below. The occurrence of shallow earthquakes, resulting from the repeated tearing apart of the newly formed crust, indicates that there is also some frictional resistance to this force. Some experts prefer to call this process **gravitational sliding**.

Slab pull

At a destructive (subduction) boundary, older, colder oceanic plates are denser than the underlying mantle. As the subducting plate is much colder and heavier than the hotter mantle, it sinks into the mantle due to the downward gravitational force acting on it and pulls the whole oceanic plate down into the mantle. The force that the sinking edge exerts on the rest of the plate is called slab pull.

Currently, although many scientists consider slab pull to be a stronger factor in driving plate movements than ridge push or mantle convection, there are others who disagree and believe ridge pull to be more significant. Each plate moves at its own rate, which suggests that the relative importance of the driving and retarding forces must vary from plate to plate. It therefore seems unlikely that any single mechanism is the sole driving mechanism of plate motion.

Plate margins

As mentioned previously, the lithospheric plates interact with one another at their margins. It is along these margins that most volcanic and seismic activity occurs. The margins are also where distinctive landforms are located. There are three types of plate margin:

- **constructive (divergent) plate margin**
- **destructive (convergent) plate margin**
- **conservative (passive) plate margin**.

Constructive (divergent) plate margins

Ocean ridge system

At constructive plate margins, plates are moving away from one another and new lithosphere is created. Under the oceans, this has produced an extensive **ocean ridge** system, comprising underwater mountains and volcanoes, that stretches for nearly 65,000 kilometres. Over 90 per cent of the mountain range lies at an average depth of 2,500 metres.

The ocean ridge system has different names depending on location (for example, the Mid-Atlantic Ridge (MAR), the East-Pacific Rise, the Juan de Fuca Ridge, and the Galapagos Rise). The MAR is about three kilometres in height above the ocean floor and between 1,000 and 1,500 km wide. Along its length, there are numerous transform faults and

an axial rift valley where the crust has partially collapsed into the low-pressure zone left by the erupting magma. Where the magma reaches the surface, volcanoes can grow into islands (such as in Iceland).

Rift valleys

Within continents, constructive margins produce **rift valleys** (Figure 5.10), such as the Great Rift Valley of East Africa which is up to 120 km wide and 1,500 m deep. This is thought to have developed over a hot spot (pages 196–7). Rising magma caused the whole region to be uplifted. This created weaknesses in the crust, through which low viscosity basalts emerged and flooded the area. As with the oceanic ridge, there was a partial collapse of the crust and the development of a deep, steep-sided rift valley.

The two sides of the rift valley are slowly moving apart, and it is believed that it will eventually flood with sea water and become a new ocean separating Africa into two parts.

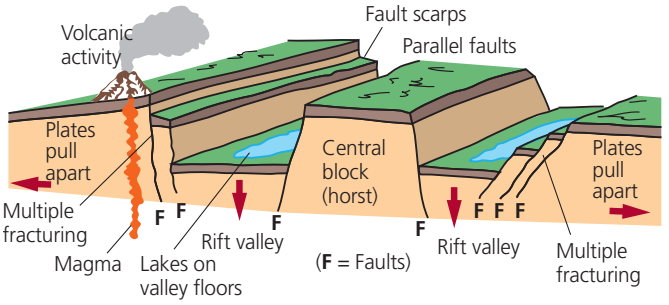


Figure 5.10 Cross section of a rift valley

Destructive (convergent) plate margins

These occur where lithospheric plates move towards one another. The nature of the boundary depends upon the combination of plate types.

Oceanic/continental convergence

The denser oceanic plate is forced under the lighter continental plate in the process of **subduction** (Figure 5.11). The downward displacement (downwarping) of the oceanic plate forms a **deep-sea trench** that runs parallel to the plate boundary. The Challenger Deep, in the Marianas Trench, is the deepest part of the ocean at just over 10,900 metres.

During the subduction process, sediments that have accumulated on the edge of the continental plate are deformed by folding and faulting and then uplifted to form **young fold mountains** (such as the Andes). Fold mountains are parallel chains of high volcanic mountains with an inter-montane plateau between them.

Ocean/ocean convergence

Where oceanic crust meets oceanic crust (Figure 5.12, page 196), the more dense oceanic crust subducts beneath the less dense oceanic crust and a line of volcanic islands known as an **island arc** can appear. On the western side of the Pacific Ocean, where the Pacific plate is being subducted beneath the smaller Philippines plate, a line of volcanic islands, including Guam and the Marianas, has been formed from magma upwelling from the Benioff zone.

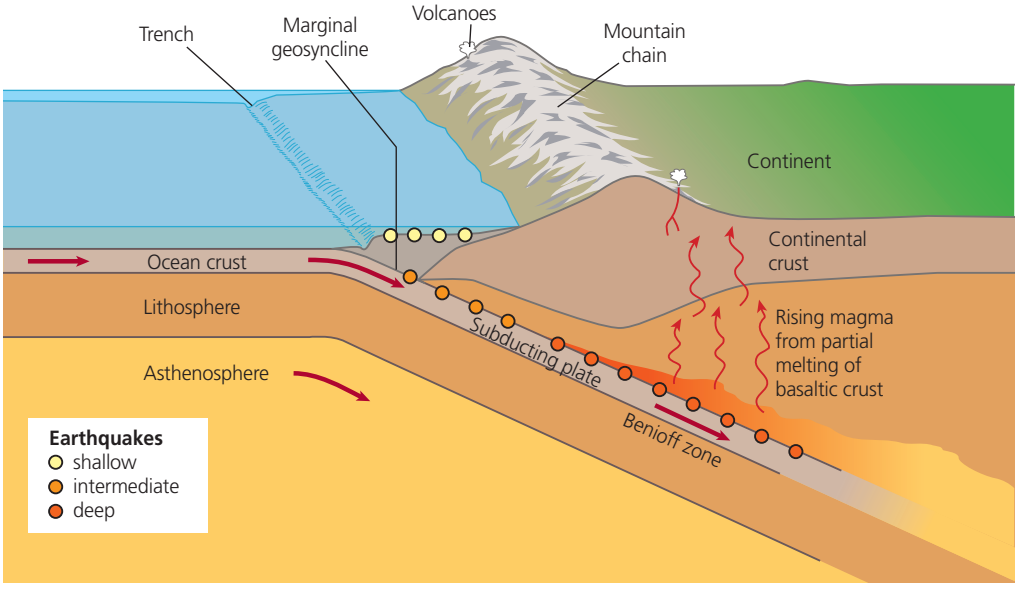


Figure 5.11 Cross section of oceanic/continental plate convergence at a destructive margin

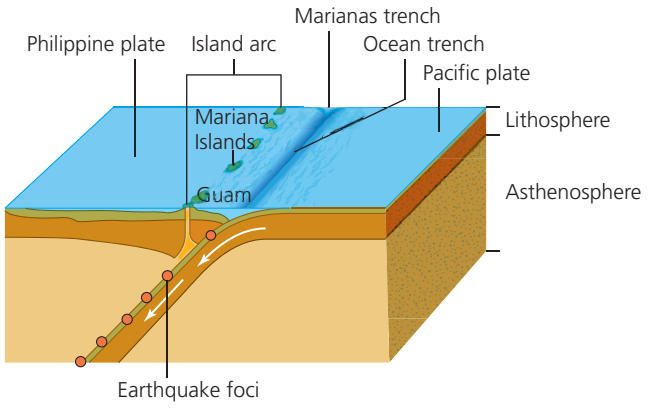


Figure 5.12 Cross section of oceanic/oceanic plate convergence at a destructive margin

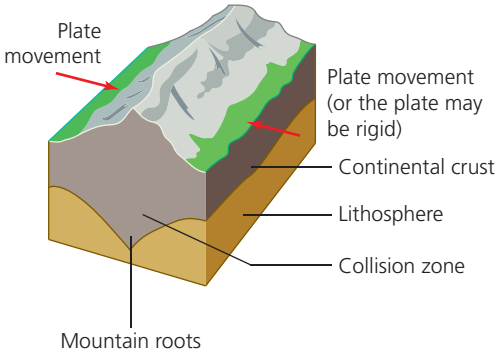


Figure 5.13 Cross section of continental/continental plate convergence (collision boundary) at a destructive margin

Continental/continental convergence (collision boundary)

If two low density continental plates converge, subduction does not occur. Sediments between them are forced up into young fold mountains. This is what is happening where the Indo-Australian plate is being forced northwards into the Eurasian plate, forming the Himalayas. There is no volcanic activity, but the movement can trigger shallow-focus earthquakes (Figure 5.13).

Conservative (passive) plate margins

Where two plates slide past one another, parallel to the plate margins, there is no subduction and therefore no volcanic activity. The movement of the plates does create stresses between the plate margins, particularly when they get stuck. Sudden movements of the plates trigger shallow focus earthquakes. The most well-known example of a conservative plate margin is the San Andreas Fault complex in California, where the Pacific and North American plates move in the same direction but at different speeds. Stresses set up by this movement cause transform faults to develop running at right angles to the main fault (Figure 5.14).

Mantle (magma) plumes

In the 1970s, a theory was developed to explain the presence of volcanic activity away from plate boundaries. It proposed that localised heating at the core/mantle boundary caused a **plume** of magma to rise through the mantle and 'eat' into the plate above the so-called 'hot spot'. Where the lava broke through to the surface, active volcanoes formed above the

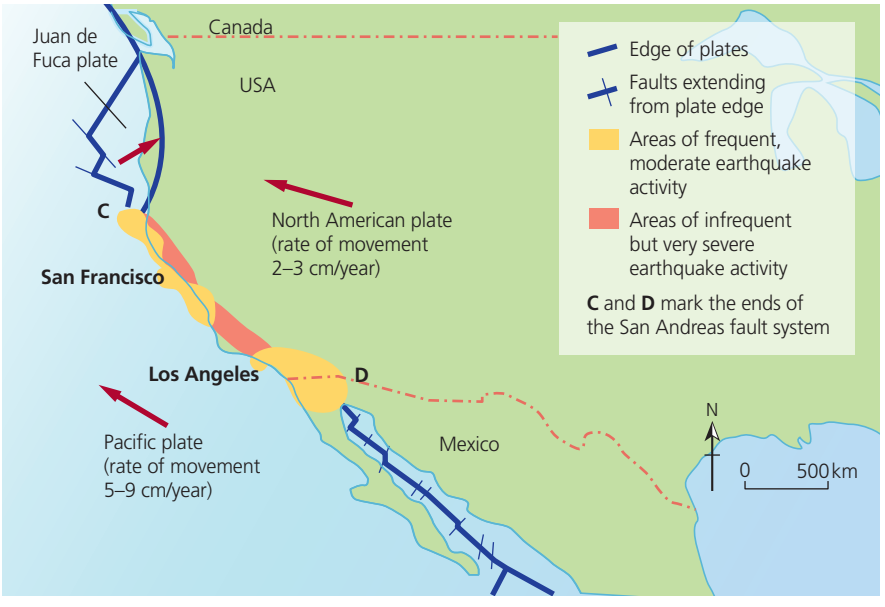


Figure 5.14 The San Andreas Fault system – a conservative plate margin

hot spot. While the hot spots are thought to be fairly stationary, the tectonic plates continue to move. As a plate continues to move away from the hot spot, it takes with it the volcanoes that have formed. These volcanoes cool and subside, and over millions of years a chain of islands, atolls and seamounts (known as hotspot tracks) can form. The volcanoes get progressively older (and become extinct) the further they move away from the hot spot. The youngest and most active volcanoes are found directly over the plume. The Hawaiian Island chain is an example of hot spot activity (Figure 5.15). The Loihi Seamount is a submarine volcano to the south east of Hawaii, formed due to hot spot activity, that will eventually form a new island in the chain.

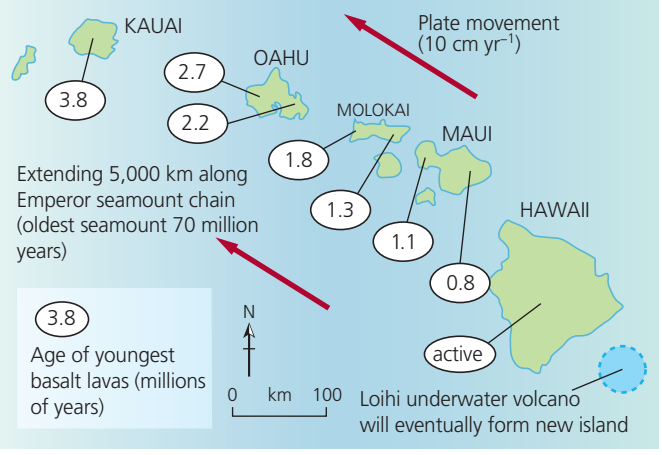


Figure 5.15 The Hawaiian hot spot

The Hawaiian hot spot has been active for around 70 million years. It has created a 6,000 kilometre long chain of volcanic islands in the north west Pacific Ocean. Scientists can use hot spot tracks to study how tectonic plates have moved over time compared to mantle sources.

Whilst the hot spot theory can explain the relationship between age and distance from plate margins for some volcanic chains, it can't be applied to all volcanic activity. There has been some doubt expressed by scientists in relation to hot spot theory about whether hot spots are fixed in position, leading some to believe that other, as yet unknown, processes are at work.

Review question

3 For a named tectonic boundary, outline the links between the nature of the boundary and the landforms found along it.

5.3 Volcanic hazards

Key terms

Ash – Dust-sized particles of rock produced by the explosive eruption of some volcanoes. This material may be carried in the air for long distances from the volcano which formed it.

Composite volcano – Large, steep-sided, symmetrical cone-shaped volcano formed from alternating layers of lava flows, volcanic ash, cinders, blocks, and bombs.

Lahars – Mudflows composed of pyroclastic material and water that flows down from a volcano, usually along a river valley.

Magma – Molten rock that is found beneath the surface of the Earth.

Pyroclastic flow – A high-density mass of gases, hot ash and larger material that flows rapidly down the sides of the volcano.

Tephra – Any type of rock fragment that is forcibly ejected from a volcano during an eruption.

Tsunamis – Giant sea waves generated by shallow-focus underwater earthquakes, violent volcanic eruptions, underwater debris slides and landslides into the sea.

Volcanic bombs – Rocks that are more than 5 mm in diameter that are thrown into the air by a volcanic eruption.

Volcanic explosive index – A scale used to measure the explosiveness of volcanoes.

Nature and distribution

Most volcanic activity is associated with plate tectonic processes and is mainly located along plate margins (Figure 5.17, page 199).

Volcanic activity is therefore found at the following sites:

- Along **constructive plate margins** where plates are moving apart and magma is forcing its way to the surface. In some locations along oceanic ridges, where the lava builds up to the ocean surface, volcanic islands form, for example, Iceland. Volcanoes also form within continental rift valleys. The Great Rift Valley in east Africa has created many volcanoes, such as Kilimanjaro, even though they are just outside the valley (Figure 5.10, page 195). Volcanoes formed here have fairly gentle sides because of the low viscosity of the basaltic lava. Eruptions are frequent but relatively gentle (effusive).

- On or near **subduction zones**: The line of volcanoes, or 'ring of fire' that surrounds the Pacific Ocean is associated with plate subduction. The deeper the oceanic plate descends, the hotter the surroundings become. This, together with the heat generated from friction, begins to melt the oceanic plate into magma in a part of the subduction zone known as the Benioff zone. As it is less dense than the surrounding material, this molten magma begins to exploit weaknesses in the crust and rises as columns of magma. The magma collects in huge sub-surface reservoirs called plutons. Eventually, some of the magma reaches the surface and form volcanoes (Figure 5.11, page 195). The andesitic lava, which has a viscous nature (flows less easily), creates complex, composite and explosive volcanoes (compared with the basaltic emissions on constructive margins). If the eruptions take place offshore, a line of volcanic islands known as an **island arc** can appear (Figure 5.15, page 197).
- Over **hot spots**: The Hawaiian Islands (Figure 5.15) in the north Pacific Ocean are examples of shield volcanoes that are located away from plate boundaries over a hot spot. They have low-angled slopes made from low viscosity, basaltic lava that flows great distances from the volcanic vent itself. Two of the world's most active volcanoes – Kilauea and Mauna Loa – can be found on the island of Hawaii. Mauna Loa last erupted in 1984, and Kilauea's last eruption lasted from 1983–2018.

Skills focus

There is no volcanic activity taking place in the British Isles at present but this was not always the case. Consult the geological record to find evidence of past volcanic activity and map this on an outline of the British Isles.

Magnitude and frequency of events

Like all natural phenomena, volcanic events only become hazardous when they impact upon people and the built environment, for example, by killing and injuring people, burying and destroying buildings and infrastructure and bringing agricultural and other economic activities to a halt. The Kamchatka peninsula in the far east of Russia is one of the most geologically active zones on the planet. Despite the fact that there are 300 volcanoes (including 29 that are active) on the peninsula, they are not considered a hazard because so few people live there.

Volcanic eruptions show an enormous variation. The type of volcano and volcanic activity depends upon the nature of the lava, which in turn depends upon the location of the volcano with regard to tectonic plate margins. If the lava is not viscous (a thin fluid), the gases may escape easily. If the lava is highly viscous (thick and pasty), the gases will not move freely and can build up tremendous pressure within the volcano. The resulting explosive eruptions can blow volcanic dust into the high atmosphere, reducing the amount of incoming solar radiation and causing short-term global climate change. The main method of measurement of **magnitude** has been the **volcanic explosivity index (VEI)**, a logarithmic scale running from 0 to 8 (Figure 5.16).

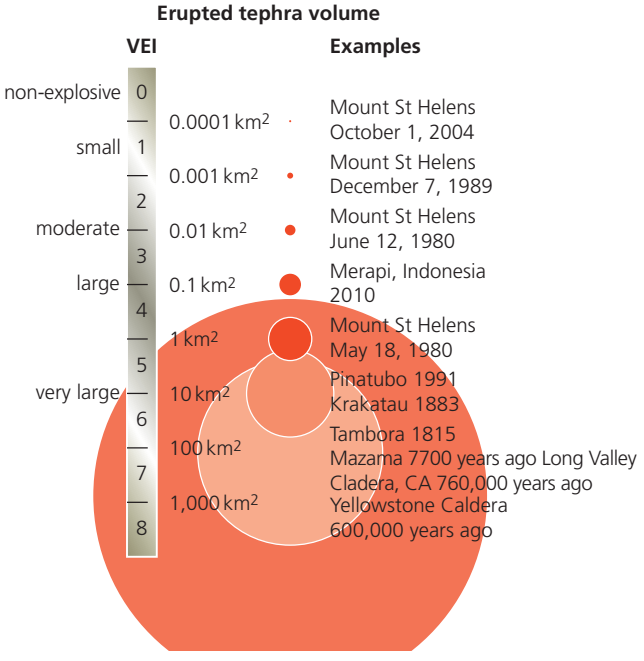


Figure 5.16 The Volcanic Explosivity Index (VEI).
Source: USGS

To determine the **frequency** of eruption of any volcano, its previous history of activity can be interpreted by volcanologists using the deposits associated with the volcano itself and those within the wider region it can effect.

The impacts of volcanic activity

A volcanic event can produce a variety of effects, the impact of which can range from the area immediately

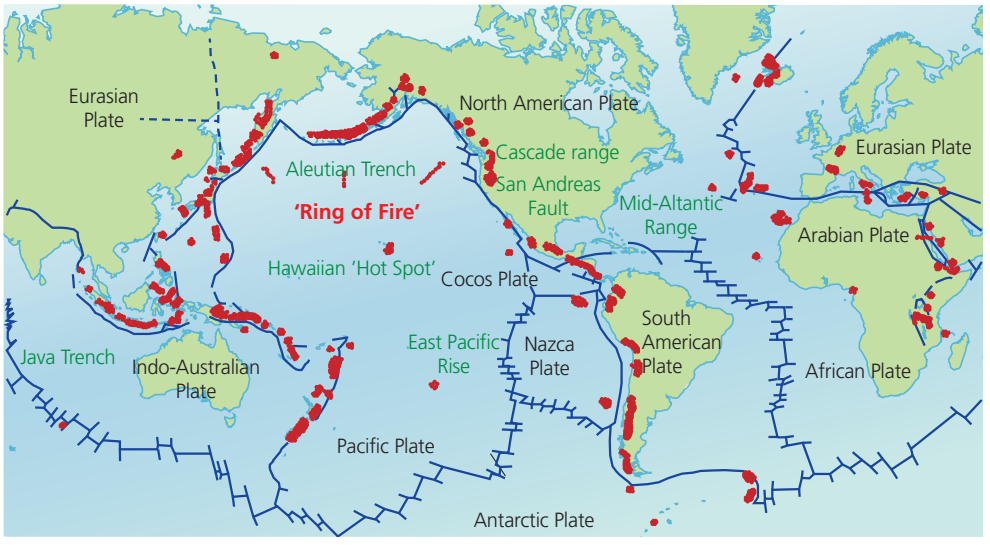


Figure 5.17 Global distribution of active volcanoes

around the volcano to the whole planet. The impact of a volcanic eruption can be categorised into primary and secondary effects.

Primary effects

Primary effects are brought about by material ejected from the volcano and described below:

- **Tephra:** Solid material of varying grain size ranging from **volcanic bombs** to **ash**, all ejected into the atmosphere. Generally, the larger the tephra particles, the shorter the distance of travel from their source. Volcanic bombs as large as a bus are very dangerous, but few people get in their way. At the other end of the scale, fine ash can be thrown high into the atmosphere where strong winds can blow it around the Earth, reducing incoming radiation and cooling the Earth.
- **Pyroclastic flows (also known as nuées ardentes):** Very hot (over 800°C), gas charged, high-velocity flows made up of a mixture of gas and tephra. These usually hug the ground and flow down the sides of the volcano with speeds of up to 700 km per hour. The Roman city of Pompeii (Italy) was destroyed in 79 AD by such flows from Mt Vesuvius.

  Lava flows rarely cause injury to people due to their relatively low velocity. They are, however, often unstoppable and can therefore damage crops and buildings and block roads.
- **Volcanic gases:** These include carbon dioxide, carbon monoxide, hydrogen sulphide, sulphur dioxide and chlorine. In 1986, carbon dioxide emissions from the lake in the crater of Nyos (Cameroon) killed 1,700 people.

Secondary effects

- **Lahars (volcanic mudflows):** Unconsolidated ash from a recent eruption combined with water may be swept down river valleys in the form of a hot, dense, fast-moving mudflow. The water can come from heavy rain, for example, Mount Pinatubo, Philippines in 1991, or from melting snow and ice, for example, Nevada del Ruiz, Colombia in 1985. The mudflow from the Nevada del Ruiz completely destroyed the town of Armero and only a quarter of the 28,700 population survived.
- **Flooding:** When an eruption melts glaciers and ice caps, serious flooding can result. This happened in Iceland in 1996 when the Grimsvotn volcano erupted.
- **Volcanic landslides.** These range in size from less than 1 km^3 to more than 100 km^3. The high velocity and great momentum of landslides allows them to cross between valleys and run up slopes several hundred metres high. For example, the landslide at Mount St Helens in 1980 had a volume of 2.5 km^3, reached speeds of 50–80 m/s and surged up and over a 400 m high ridge located about 5 km from the volcano.
- **Tsunamis:** Sea waves generated by violent volcanic eruptions such as those formed after the eruption of Krakatoa (Indonesia) in 1883. Tsunamis from this eruption are estimated to have killed 36,000 people. (For details on tsunamis see the section on seismic hazards, pages 206–11.)
- **Acid rain:** Volcanoes emit gases which include sulphur. When this combines with atmospheric moisture, acid rain results.

- **Climatic change:** The ejection of huge amounts of volcanic debris into the atmosphere can reduce global temperatures and is believed to have been an agent in past climatic change.

Responses to volcanic hazards

Preparedness

In AD 79, the Roman cities of Pompeii and Herculaneum were overwhelmed by a pyroclastic flow and ash fall resulting from the eruption of nearby Mount Vesuvius. It is thought thousands of people died. Despite being part of a sophisticated civilisation, they had no understanding of the danger that existed on their doorstep. Even in the twenty-first century, people can be caught out (Mt Nyiragongo, 2002, Democratic Republic of the Congo). Thankfully, this is rare as scientists are now able to predict with increasing certainty when a volcanic eruption will occur. Early **prediction** means that people can be evacuated from the danger zone and lives can be saved.

Many earthquakes are now monitored continually, either by local observatories or by satellite. Signs that a volcano may erupt include:

- an increase in the release of various gases, particularly sulphur dioxide and carbon dioxide
- a rise in the level of lava lakes in volcanic craters
- the bulging upwards of surrounding land due to pressure from below
- an increasing number of relatively small earthquakes caused by the rising magma.

A study of the previous eruption history of any volcano is important, along with an understanding of the type of activity produced. Ash, lahars and pyroclastic flows leave characteristic deposits around volcanoes. We can assume that if this type of hazard appeared once, it will be repeated. It is also possible to identify areas at greatest risk and the frequency of eruptions from these layers of deposits.

Mitigation

There are a variety of actions that can be taken before, during and after an eruption to help reduce or eliminate long-term risks caused by volcanoes. **Risk assessments** are one form of mitigation. Governments of countries at risk of volcanic hazards, such as the Philippines, carry out risk assessments and produce a series of alert levels

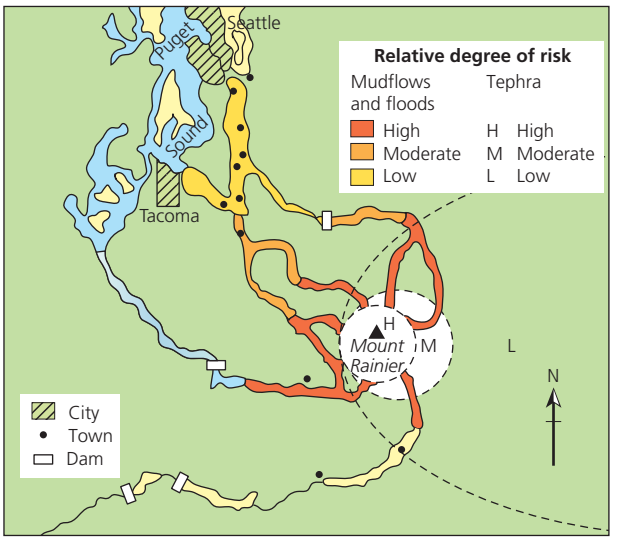


Figure 5.18 Risk assessment of Mt Rainier, Cascade Range USA

(Table 5.4, page 203) to warn the public of the threat. Figure 5.18 shows the risk assessment made for the area around Mt Rainier (part of the Cascade Range in the USA), which is not surprising as 3.5 million people live and work close to it.

During an eruption, some physical actions can be undertaken to reduce the impacts:

- Once viscous lava has started to flow, it may be possible to divert it from the built environment. On Mt Etna in Sicily, digging trenches, dropping blocks into the lava stream and using explosives have been successful in slowing down the flow and, in some cases, diverting it.
- In 1973, the inhabitants of Haeimaey (Iceland) were able to pour seawater on the front of a lava flow to solidify it before it cut off their vital fishing port from the open sea.
- In parts of the Hawaiian Islands, barriers have been built across valleys to protect settlements from lava flows and lahars.

The most common way to reduce the impact on people is to evacuate vulnerable areas when risk becomes intolerable. This comes with its own problems. If an evacuation is carried out needlessly, future evacuations are more difficult to manage. Evacuations have been carried out successfully several times in the Philippines (for example, Pinatubo 1991, Mayon 2018). Sometimes permanent evacuation may be necessary. For example, two-thirds of the Caribbean island of Montserrat is designated an exclusion zone following an eruption of the Soufrière Hills volcano in 1995.

Prevention

There is no way to prevent a volcanic eruption. Disasters can be prevented or reduced by mitigation (see above), particularly evacuation in advance of an eruption.

Adaptation

Perhaps the best response is to move away from the area surrounding the volcano. In the case of Montserrat, this move proved to be permanent for many people and they had to change their way of life, often in a new country. In the east of the Democratic Republic of the Congo, farmers returned to their fields after the eruption of Nyiragongo (2002) and within a short time they were able to change the crops they cultivated, moving from maize to cash crops such as cabbage and bananas. In some cases, their income doubled from before the eruption.

Volcanic event

A recent volcanic event: Mount Mayon, Philippines (2018)

Mount Mayon, at 2,462 m above sea level, is the most active volcano in the Philippines. It is a composite volcano with records of eruptions dating back to 1616. It covers an area of 250 km^2 and, although within a national park, it is surrounded by the city of Legaspi and eight other towns (Figure 5.39, page 225). It is a popular tourist spot and is renowned for its symmetrical conical shape. Its activity is regularly monitored by PHIVOLCS, a Philippine government department dedicated to providing information on the activities of volcanoes, earthquakes and tsunamis.

Mayon is a result of the subduction of the Philippine Sea Plate underneath the Sunda Plate as part of the Pacific Ring of Fire.

Risk and vulnerability

- Large ash emissions from Mount Mayon generally settle within a few kilometres of the volcano and affect the nearby city of Legaspi – the capital of the province of Albay, with a population of almost 200,000 people and a large centre for tourism, education, health, services, commerce and transportation for the region. This ash, made up of fine grains of rock, settles on roofs, often causing them to collapse.
- Ash emissions also settle in the gullies (Figure 5.19) on the flanks of the volcano. The high rainfall amounts in the area (3,432 mm, with January/February being the wettest season), mean there is a high risk of lahars, which destroy anything in their path. In the short term, ash emissions destroy local agriculture, but in the long term they enrich the soil.
- Nuées ardentes (a type of pyroclastic flow) are also concentrated in the gullies. These too destroy anything in their path.
- Lava flows destroy local agriculture

The 2018 event

- Historical data indicates that an eruption event cycle begins with ash emissions and basaltic eruptions, followed by longer term, andesitic lava flows from the central crater. These lava flows travel far down the flanks of the mountain. Pyroclastic flows and mud flows have commonly swept down many of the approximately 40 ravines that radiate from the summit and have often devastated populated lowland areas.
- Mount Mayon has erupted regularly throughout the first 20 years of this century, but the most notable eruption was in 2018. Table 5.3 (page 202) gives a timetable of events.

Figure 5.19 Natural-colour image of Mount Mayon, 15 December 2009. A small plume of ash and/or steam is blowing west from the summit. Dark-coloured lava or debris flows from previous eruptions streak the flanks of the mountain. A ravine on the south-east slope is occupied by a particularly prominent lava or debris flow

Table 5.3 Timetable of events during the 2018 eruption of Mount Mayon

Date	Volcanic activity	Immediate reaction
13 January 2018	• A phreatic (high-pressure steam driven) eruption began that blasted a steam and ash plume approximately 2,500 m into the air for just over an hour and a quarter • Ash fell in many of the surrounding townships and a sulphurous smell was noted in the town of Camalig (Figure 5.20)	• PHIVOLCS raised the alert level of Mayon Volcano from 1 to 2 (Table 5.4) • Approximately 40,000 residents evacuated from a six-kilometre radius
14 January 2018	• Three more phreatic eruptions and 158 rockfall events were recorded • The summit crater began to glow brightly, indicating the growth of a new lava dome and the start of lava flows towards its slopes	• The alert status was upgraded to Level 3
16 January 2018	• Lava flows reached the limits of the six-kilometre evacuation zone	• The province of Albay declared a state of calamity
22 January 2018	• A three-kilometre tall ash column developed • Later that day, lava fountains were observed along with pyroclastic flows, ash plumes, lava bombs and rockfalls	• The alert status was raised to Level 4
23 January 2018	• Intermittent 300–500-metre-high lava fountains and ash plumes developed • Lava bombs and rockfalls could also be observed, along with the sound of explosions	• Schools and colleges were shut • The danger-zone was expanded up to nine kilometres, despite the alert status remaining at Level 4 • Further evacuations took place (eventually totalling 74,000) • Warnings were issued of the possibility of rock falls, landslides and sudden explosions or dome collapse that could generate hazardous volcanic flows • Heavy rains were reported, bringing concerns over the development of lahars
Early February 2018	• The eruption slowly subsides	• Some schools and colleges reopened
March 2018	• Further fall in volcanic activity	• The alert level was dropped to Level 3 and then to Level 2 at the end of the month

Figure 5.20 Ash cloud rising over Mount Mayon on 22 January 2018, as seen from Camalig, Albay

Table 5.4 PHIVOLCS alert levels for Mayon volcano

Alert levels	Main criteria	Interpretation
0: No alert	**Quiet**	**No eruption in foreseeable future**
1: Abnormal	**Low-level unrest**. Slight increase in seismicity	**No eruption imminent**
2: Increasing unrest	**Moderate unrest**. Low to moderate level of seismic activity	**Unrest, probably of magmatic origin;** could eventually lead to eruption
3: Increased tendency towards hazardous eruption	**Relatively high unrest**. Volcanic quakes and tremors may become more frequent	**Magma is close or at the crater**. If the trend is increasing unrest, an eruption is possible within weeks
4: Hazardous eruption imminent	**Intense unrest**. Persistent tremors, many 'low frequency' type earthquakes may occur	**Hazardous eruption is possible within days**
Level 5: Hazardous eruption	**Hazardous eruption ongoing**	Pyroclastic flows (e.g. lava, hot ash or gases) may sweep down along gullies and channels, especially along those fronting the low part(s) of the crater rim. Additional danger areas may be identified as the eruption progresses. Aircraft should stay away lest they encounter potentially dangerous ash clouds

Source: Philippine Institute of Volcanology and Seismology (PHIVOLCS)

Impacts

- Fortunately, nobody was killed during the eruption. This is a testament to the well-organised emergency procedures that had been set up in the Philippines. By 5 February 2018, a total of 86,052 persons had been affected, of which, 64,895 stayed in 58 designated evacuation centres.
- A total of US $3.4 million worth of damages to agriculture (rice, corn and abaca) occurred, affecting more than 10,000 farmers. Some roads were made impassable by landslides and ash falls. Some, but not all, flights in and out of Legaspi were cancelled.

Responses to the eruption

Local and national response

- The Filipino government committed US $1 million for a 'Cash for Work' programme as well as providing hygiene packages for 50,000 families for ten days. As well as this, it provided food packs to evacuees for 100 days.
- The army was called in to help – to enforce the evacuation and to be present if there was any civil disturbance.
- The Philippine Red Cross set up first-aid stations and welfare desks to provide psychosocial support to affected individuals.
- The National Council of Churches in the Philippines provided immediate relief assistance to 3,446 families.

International response

Following the declaration of a Level 3 alert for the volcano, the United States government advised its nationals against travelling to Mayon. Canada and the United Kingdom also discouraged their nationals from visiting the volcano.

The United States government, through the United States Agency for International Development (USAID), committed over US $100,000 which was used to distribute vouchers for families to buy essential items and hygiene kits to help reduce illness in evacuation centres. They also constructed latrines, bathing cubicles and handwashing stations, and promoted safe water and hygiene practices through educational materials.

Volcanic activity has continued on Mayon with further episodes in 2020. It is interesting to note that whilst nobody was killed in the 2018 eruption which reached an alert Level of 4, five people were killed in the eruption in 2013 despite that eruption only being classified at alert Level 2.

Review question

4 Study Figure 5.1 (page 186), the model of vulnerability. Using the information contained in this study of Mayon and, replace the generalities of the model with the real details of the eruption. What does the completed figure tell you about the extent of the risk to people and the environment (and their vulnerability) in living in the area around Mount Mayon?

Fieldwork opportunities

There is the possibility of doing some work (along with other physical geography aspects, such as glaciation) in Iceland. You will have a chance to observe active volcanoes, lava flows, ash deserts, geysers and the rift which is opening up between the North American and Eurasian plates.

Other volcanic events worth studying

Etna (Sicily, Italy) 1991–99

Montserrat (West Indies) 1995–97

Merapi (Indonesia) 2010

Mt Ontake (Japan) 2014

Review question

5 Why have people sought to manage volcanic hazards through prediction rather than mitigation?

5.4 Seismic hazards

Key term

Earthquake – As the crust of the Earth is constantly moving, there tends to be a slow build-up of stress within the rocks. When this pressure is released, parts of the surface experience, for a short period, an intense shaking motion.

Nature and distribution

As the crust of the Earth is mobile, there tends to be a slow build-up of stress between masses of rocks moving in different directions. When this stress is suddenly released, parts of the surface experience an intense shaking motion that lasts for from just a few seconds up to 30 seconds. The Boxing Day earthquake of 2004, however, lasted for nearly six minutes. This is an **earthquake**. The point at which this pressure release occurs within the crust is known as the **focus**, and the point immediately above that on the Earth's surface is called the **epicentre** (Figure 5.21). The depth of the focus is significant in terms of the effects on the surface and three broad categories of earthquake are recognised:

- **shallow focus (0–70 km deep):** these tend to cause the greatest damage and account for 75 per cent of all the earthquake energy released
- **intermediate focus (70–300 km deep)**
- **deep focus (300–700 km deep)**

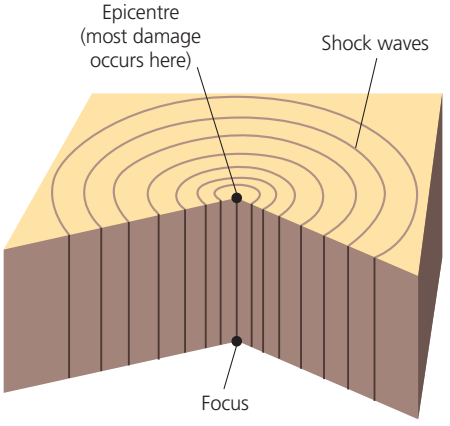


Figure 5.21 The focus and epicentre of an earthquake

The vast majority of earthquakes occur along plate boundaries (Figure 5.22), the most powerful being associated with destructive margins. At conservative margins, the boundary is marked by transform faults and sudden differential movement along these faults produces the earthquake. Perhaps the most famous of these is the San Andreas Fault in California, which represents the boundary between the North American and Pacific plates. In reality, the San Andreas system consists of a broad complex zone in which there are a number of fractures of the crust (Figure 5.14, page 196).

Some minor earthquakes occur away from plate boundaries. In February 2018, an earthquake of intensity 4.6 on the Mercalli Scale was felt over much of Wales and south-west Britain. An earthquake of this size occurs somewhere on mainland Britain roughly every four years.

This compares to Japan where, according to the Meteorological Agency of Japan (JMA), there are approximately 2,000 earthquakes every year.

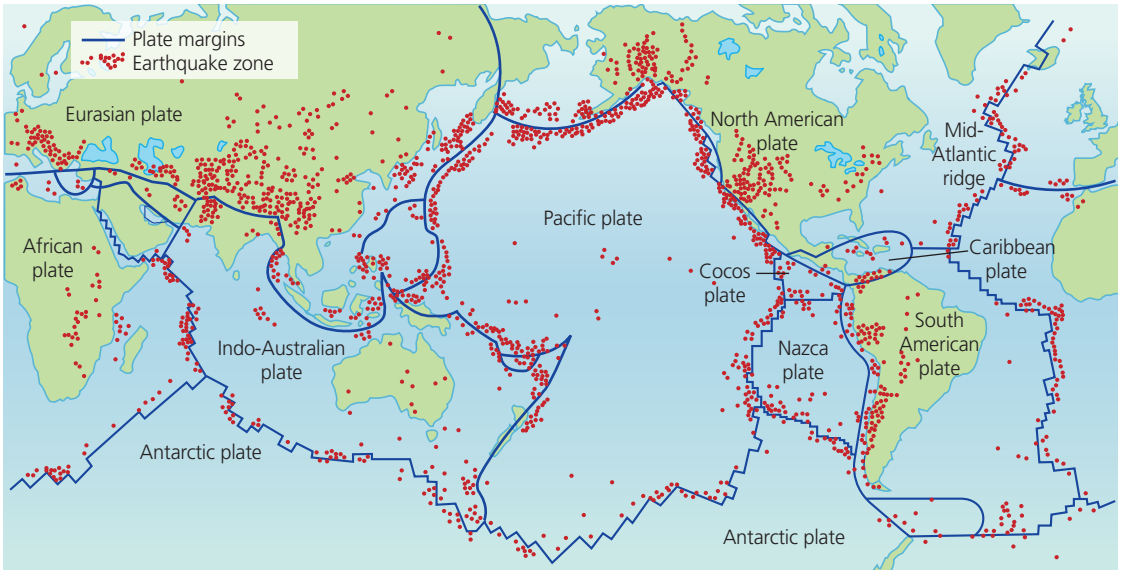


Figure 5.22 Global distribution of earthquakes

It has been suggested that human activity could be the cause of some minor earthquakes, through building large reservoirs which puts pressure on the underlying rocks, or subsidence of deep mine workings. In recent times, people have become increasingly worried about the process of fracking (hydraulic fracturing of rock in order to release gas). A series of earthquakes, the largest with a magnitude of 2.9, were recorded near a fracking site near Blackpool, Lancashire. This caused the drilling to be stopped, pending a further review into the safety of the activity.

Magnitude and frequency of events

The **magnitude** of earthquakes is measured in a variety of ways. The **moment magnitude scale** measures the size of an earthquake in terms of energy released (Mw). It is a logarithmic scale, which means that an increase of one step on the scale corresponds to about 32 times the amount of energy released. Two steps denote an increase in energy of 1,000.

It is calculated using a formula that includes the rigidity of the rock affected, the distance moved and the size of the area where movement takes place.

While magnitude is a quantitative measure of the energy released at the source of an earthquake, intensity is the qualitative measurement of the strength of shaking produced by an earthquake. The **Mercalli scale** (Table 5.5, page 206) measures the intensity of an event and its impact.

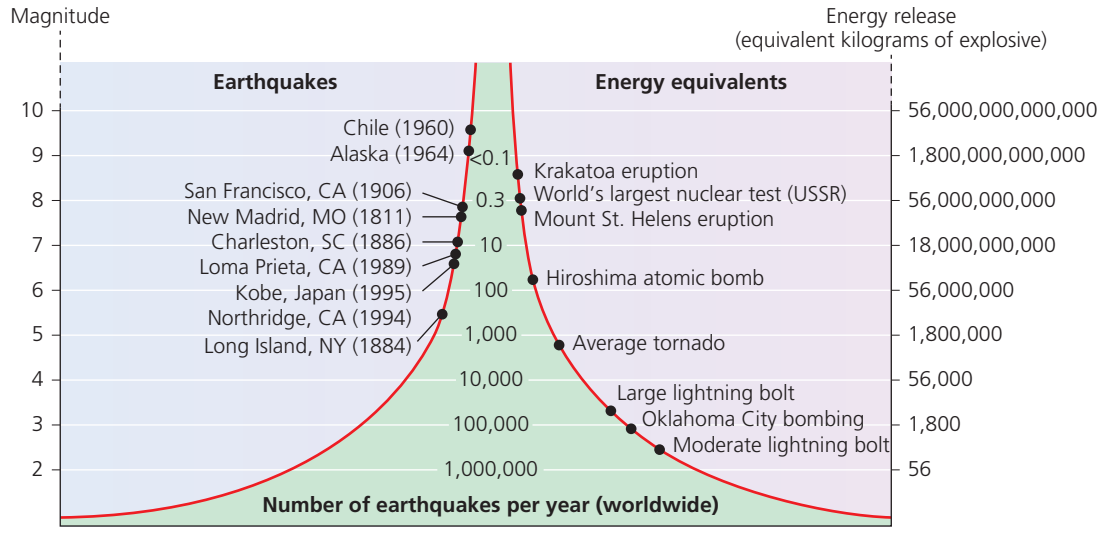


Figure 5.23 The moment magnitude scale (Mw)

Source: USGS

Table 5.5 The Mercalli scale

I. Not felt	Not felt except by a very few under especially favourable conditions
II. Weak	Felt only by a few persons at rest, especially on upper floors of buildings
III. Weak	Felt quite noticeably by persons indoors, especially on upper floors of buildings. Many people do not recognise it as an earthquake. Standing motor cars may rock slightly. Vibrations similar to the passing of a truck. Duration estimated
IV. Light	Felt indoors by many, outdoors by few during the day. At night, some awakened. Dishes, windows, doors disturbed; walls make cracking sound. Sensation like heavy truck striking building. Standing motor cars rocked noticeably
V. Moderate	Felt by nearly everyone; many awakened. Some dishes, windows broken. Unstable objects overturned. Pendulum clocks may stop
VI. Strong	Felt by all, many frightened. Some heavy furniture moved; a few instances of fallen plaster. Damage slight
VII. Very Strong	Damage negligible in buildings of good design and construction; slight to moderate in well-built ordinary structures; considerable damage in poorly built or badly designed structures; some chimneys broken
VIII. Severe	Damage slight in specially designed structures; considerable damage in ordinary substantial buildings, with partial collapse. Damage great in poorly built structures. Fall of chimneys, factory stacks, columns, monuments, walls. Heavy furniture overturned
IX. Violent	Damage considerable in specially designed structures; well-designed frame structures thrown out of plumb. Damage great in substantial buildings, with partial collapse. Buildings shifted off foundations
X. Extreme	Some well-built wooden structures destroyed; most masonry and frame structures destroyed with foundations. Rails bent
XII. Extreme	Few, if any (masonry) structures remain standing. Bridges destroyed. Broad fissures in ground. Underground pipelines completely out of service. Earth slumps and land slips in soft ground. Rails bent greatly
XIII. Extreme	Damage total. Waves seen on ground surfaces. Lines of sight and level distorted. Objects thrown upward into the air

Source: USGS

Seismic records enable earthquake frequency to be observed, but these records only date back to 1848 when an instrument capable of recording shock waves was first developed. Before that we have to consult historical records as to the date and the effects of the identified event.

The impacts of seismic events

Very few people are killed by earthquakes if they are standing in the middle of an open area. People are killed by what the earthquakes do to buildings (roof collapse), the sea (tsunamis) and slopes (landslides). Some of the effects on the human environment are short term; others occur over a longer period and will depend to a large extent on the ability of the area to recover.

Primary effects

The initial, or **primary** hazard (impact), of an earthquake is **ground shaking**, caused by shock waves travelling through the crust from the focus up to the surface and then radiating outwards. In extreme cases waves, rather like sea waves, can be observed. The severity of the shaking depends upon the magnitude of the earthquake, its depth, the distance from the epicentre and the local geological conditions. In the Mexico City earthquake of 1985, for example, the shock waves that devastated the city were amplified several times by the ancient lake sediments upon which the city is built. The ground motion in the city measured five times that in the outlying districts, which had different soil foundations.

Another primary hazard is **ground rupture**; the visible breaking and displacement of the Earth's surface, probably along the line of a fault. Ground rupture poses a major risk for large engineered structures such as dams, bridges and nuclear power stations.

Secondary effects

- **Soil liquefaction:** when violently shaken, soils with a high water content lose their mechanical strength and start to behave like a fluid. The Palu (Indonesia) earthquake of 2018 caused major soil liquefaction within and close to the city of Palu. This led to mudflows that submerged many buildings and in one case a whole village.
- **Land slides/avalanches:** slope failure as a result of ground shaking. Following the Gorkha earthquake in Nepal in 2015, scientists from NASA found that 4,312 landslides occurred between the time of the earthquake in April and the onset of the monsoon rains in June. Most of these landslides were comparatively small rockfalls and debris avalanches but some were large enough to block whole valleys. In the Langtang valley, the village of Langtang was almost completely buried.
- **Tsunamis**: giant sea waves generated by shallow-focus underwater earthquakes involving movements of the sea bed, or landslides into the sea (see below)

- **Fires:** resulting from broken gas pipes and collapsed electricity transmission systems.

Tsunamis

A tsunami is a dangerous series of waves that can be triggered by any event that disrupts the sea floor. Tsunamis are most often generated by shallow-focus underwater earthquakes but also by volcanic eruptions, underwater debris slides and large landslides into the sea. Tsunamis have a very long wavelength (sometimes over 100 km) and a low wave height (under one metre) in the open ocean, and they travel at speeds of over 700 km per hour. When they reach shallow water bordering land, they slow down and increase in height (Figure 5.24).

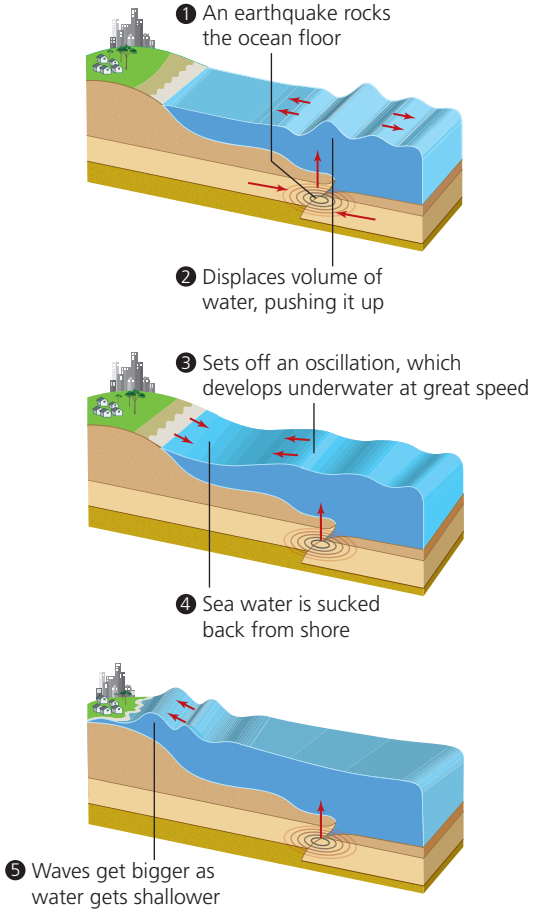


Figure 5.24 Anatomy of a tsunami

Quite often, the first warning given to coastal populations is the wave trough in front of the tsunami which results in a reduction in sea level, known as a drawdown. Behind this comes the tsunami itself, which can reach heights in excess of 25 m. The event usually consists of a number of waves, the largest not necessarily being the first. When a tsunami reaches land, its effects will depend upon:

- the height of the waves and the distance they have travelled
- the length of the event (at source)
- the extent to which warnings could be given
- coastal physical geography, both offshore and on the coastal area
- coastal land use and population density.

The wave will wash boats and wooden coastal structures inland, and the backwash may carry them out to sea. People are drowned or injured by the tsunami as both the water itself and the debris that it contains are hazards. Buildings, roads, bridges, harbour structures, trees and even soil are washed away, most tsunamis having an effect at least 500–600 m inland, depending upon the coastal geography. Tsunamis generated by the explosion of the volcano Krakatoa in 1883 have been estimated to have drowned over 35,000 people and produced waves that travelled around the world, the highest being over 40 m.

Around 90 per cent of all tsunamis are generated within the Pacific Basin, associated with the tectonic activity taking place around its edges. Most are generated at convergent plate boundaries where subduction is taking place, particularly on the western side of the Pacific (25 per cent of all events). Some recent events which have taken place in that area are described below:

- December 2004, **east Indian Ocean off Sumatra, 'the Boxing Day' tsunami:** Generated by a powerful magnitude 9.0 earthquake some 25 km below the Indian Ocean floor off northwest Sumatra (one of the biggest earthquakes ever recorded). The tsunamis which spread across the Indian Ocean killed an estimated 300,000 people in bordering countries and devastated hundreds of communities. Extensive damage was caused to coastal communications, particularly bridges and railway lines and damage to local economies left people unable to feed themselves. A positive result was that a warning system was set up among the countries that border the Indian Ocean. Despite the enormous human cost, the insurance industry estimated that the disaster cost less than $5 billion.
- March 2011, off the Pacific coast of the **Tohoku region of Japan** (Figure 5.25, page 208). Earthquake occurred 70 km offshore and of a magnitude of 9.0: the most powerful earthquake ever recorded in Japan (it was estimated that it moved the main island of Japan, Honshu, 2.4 m east).

Tsunamis generated reached over 40 m in height and in some places penetrated over 10 km inland. These tsunamis devastated part of Japan's Pacific coast, causing nearly 16,000 deaths with over 6,000 people injured and 2,500 missing. It was estimated that over 300,000 people were also displaced from their homes. The built environment suffered with 127,000 buildings destroyed, 277,000 'half collapsed' and 750,000 partially damaged. In addition, the Fukushima nuclear power plant was seriously affected resulting in the evacuation of people living within 20 km of the facility. The assessed cost of the damage by insurers was over $30 billion while it has also been estimated that the economic cost to Japan was $235 billion. Tsunami waves ran right across the Pacific to North and South America; in Chile the tsunami created waves 2 m high.

- In September 2018, following a series of small tremors, a shallow earthquake of magnitude Mw 7.5 hit the island of **Sulawesi in Indonesia**. The epicentre was 77 km away from the city of Palu. A localised tsunami with an estimated height of 4–7 m struck **Palu** shortly afterwards.
- The earthquake was caused by a mostly horizontal movement of the crust (most large tsunamis are caused by earthquakes with vertical movement). It is thought that this triggered underwater rockslides, causing the tsunami. As Palu was located at the end of a narrow bay, its force became more concentrated. The waves that struck Palu reached a height of nearly six metres. This particularly affected a festival that was taking place on Palu's main beach at the time, killing or injuring many of the participants.

Figure 5.25 Devastation caused by the Tohoku tsunami, Japan 2011

Skills focus

Find photographs on the internet of two urban areas in very different parts of the world that have suffered from devastating earthquakes. Use your observational skills to describe the damage shown in each photograph, in particular focusing on the differences that may be seen. If you identify any major differences, try to explain how they have come about.

Responses to seismic hazards

Preparedness

Earthquakes can occur with no warning whatsoever so reactions to them must become second nature. The simple mantra of 'Drop, cover, and hold on' should be instilled in children at a young age and practised regularly. There are other actions that can be taken in advance of a seismic event:

- Within the home, any heavy items (such as televisions, refrigerators and bookcases) should be secured and breakable items stored at a low level.
- Families should create a communication plan (for example, a WhatsApp group) and establish an emergency meeting place.
- If possible, households should have an emergency supply kit to last a few days containing essential items, such as tinned food and water, clothing, bedding, a first-aid kit, toilet paper, a torch, a fire extinguisher, a whistle and a radio.
- Homes should be made structurally sound. This applies most in those countries where building codes are lax or non-existent.
- If financially possible, people should take out a specialised earthquake insurance policy.

The prediction of earthquakes is difficult. Regions at risk can be identified through plate tectonics. Some attempts have been made to predict exactly when and where earthquakes might strike based upon monitoring groundwater levels, release of radon gas and unusual animal behaviour. Fault lines, such as the San Andreas, are also monitored for signs of movement. Despite this, the ability to predict an earthquake a few hours before the event is unrealistic for the foreseeable future.

Close studies of fault lines can sometimes indicate the point along the fault where the next earthquake could be due. A study of the pattern of the events along the

San Andreas fault between 1969 and 1988 indicated the existence of a '**seismic gap**' (an area that had not had any real seismic activity for the past twenty years) in the area of Loma Prieta. This area suffered an earthquake in October 1989 that measured 6.9 Mw and was the worst to hit the San Francisco region since 1906. As a result of the seismic survey, this event was not entirely unexpected, but like all earthquakes, it was not possible to predict it precisely. This system is limited in its use; the Northridge earthquake (1994) took place on a previously unknown fault line.

Mitigation

- **Early warning systems**: The damaging surface waves of earthquakes take time to travel from the epicentre to where they may impact upon people. Early warning systems, such as that in Japan, give people a short time to make themselves safe. In Japan, the 'Earthquake Early Warning' system aims to reduce earthquake-related damage by slowing down trains immediately, controlling lifts and enabling people to protect themselves quickly.
- **Hazard-resistant structures**: Buildings can be designed to be earthquake resistant. There are three main ways in which this can be achieved.
  - By putting a large concrete weight on top of the building that will move, with the aid of a computer program, in the opposite direction to the force of the earthquake in order to counteract stress.
  - Putting large rubber shock absorbers in the foundations that will allow some movement of the building without structural failure.
  - By adding cross-bracing to buildings to hold them together better when they shake. Older buildings and structures, such as elevated motorways, can be **retrofitted** with such devices to make them more earthquake proof.
- A comparison between the 1989 Loma Prieta earthquake (6.9 Mw) and the 1988 event in Armenia (6.8 Mw) shows the effects of different types of buildings. In California, where there were more earthquake-proof buildings, there were 63 deaths whereas in Armenia over 25,000 people died, many inside buildings that collapsed due to weak foundations and a lack of earthquake-proofing features. In the town of Leninakan, for example, over 90 per cent of more modern 9–12 storey pre-cast concrete frame buildings were destroyed.
- **Tsunami protection**: Tsunamis cannot be predicted entirely, even if the magnitude and location of an earthquake is known. Certain automated systems can be installed to give warnings, the best of which uses sea-bed pressure sensors, attached to buoys, which constantly measure the pressure of the overlying water column. Regions with a high tsunami risk use warning systems (such as a klaxon) to warn the population before the wave reaches land. The Pacific Tsunami Warning and Mitigation System (ICG/PTWS) is co-ordinated by UNESCO and is based on Hawaii. It monitors earthquake activity and issues warnings to countries around the Pacific Edge if tsunamis are likely. Some countries have built prevention walls up to 12 m in height. These have not proved very effective, as large tsunamis are likely to overwhelm them.

Prevention

Trying to prevent an earthquake is thought to be almost impossible. This, however, has not stopped studies into the feasibility of schemes to lubricate active faults with water or oil to stop them becoming stuck. Some people have even gone as far as to suggest nuclear explosions at depth.

Adaptation

Long-term change in people's behaviour depends on levels of economic development, education and national and regional priorities. People in more economically developed countries, such as the USA and Japan, are generally more able to adapt their environment than those in low income countries. Land-use planning measures may include:

- identifying areas most at risk from a seismic event and regulating land-use planning for those areas or limiting the way in which the land can be used
- putting key buildings, such as schools and hospitals, in low-risk areas and open spaces, such as parks, in more high-risk areas
- including open spaces in plans to allow for safe areas away from fires and aftershock damage to buildings.

Emergency services must adapt their organisation and planning to deal with seismic events. For example, heavy lifting gear needs to be available and emergency service personnel should be first-aid trained to deal with any casualties following a seismic event as it could be some time after the event before trained medical personnel arrive.

Seismic event

A recent seismic event: Lombok, Indonesia (2018)

The island of Lombok in Indonesia is part of the chain of the Lesser Sunda Islands, along with Bali to the west and Sumbawa to the east (Figure 5.26).

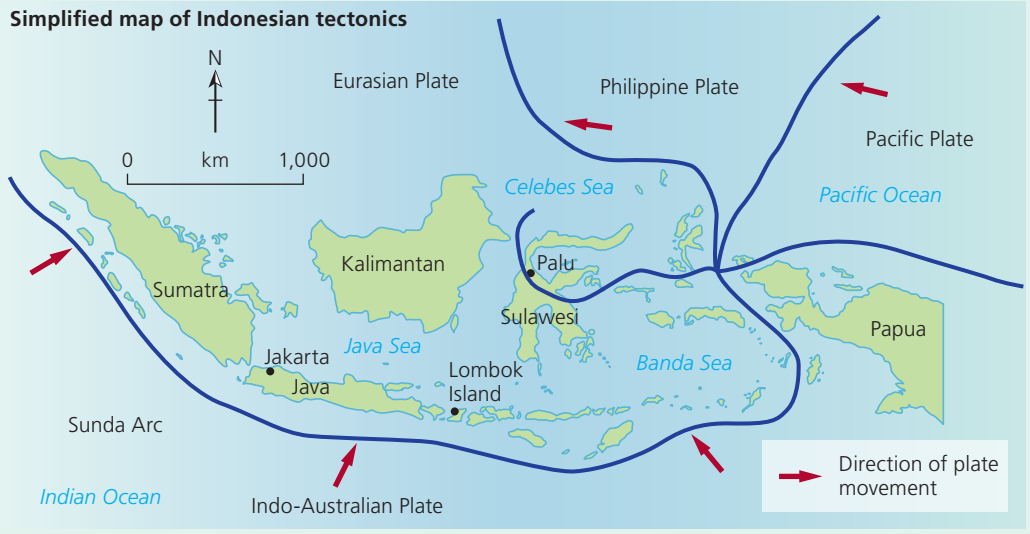


Figure 5.26 The location of Lombok and Palu

Risk and vulnerability

Indonesia is a meeting point of several tectonic plates and is one of the most seismically active places on Earth (Figure 5.27). It is a typical island arc structure where the Indo-Australian plate is moving northward towards, and subducting under, the south-easterly moving Eurasian plate at a rate of approximately six centimetres per year. In addition, the Philippines and Pacific plates are moving westwards, also subducting under the Eurasian plate. The Sunda Arc subduction zone can be traced from the northern tip of Sumatra to beyond Lombok; it creates a deep submarine trench parallel to the coastline. Most of the earthquakes affecting Indonesia, as well as many volcanoes, occur along this subduction zone.

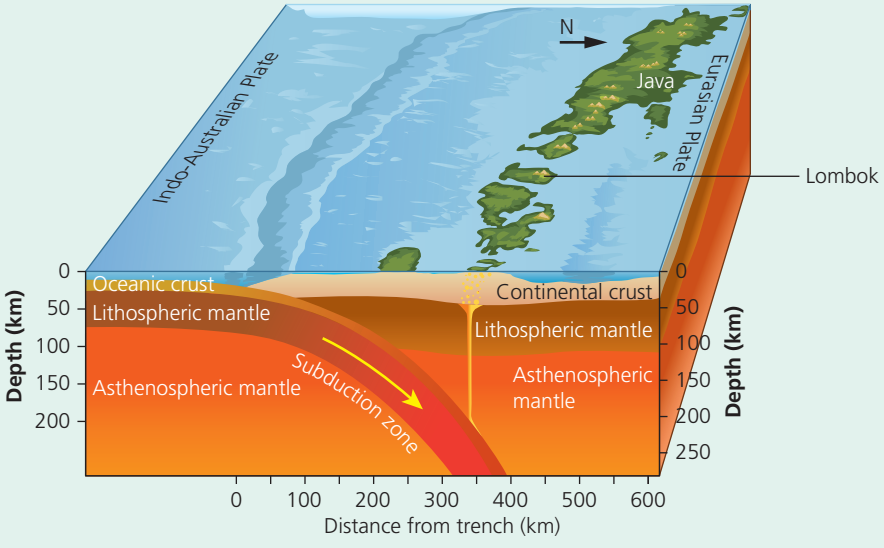


Figure 5.27 A cross section of the Sunda Arc

Source: Earth Observatory of Singapore

The Lombok earthquake is a collective name for a series of seismic events that took place between 29 July and 19 August 2018. On 29 July, the Indonesian Meteorological, Climatological and Geophysical Agency (BMKG) recorded an earthquake of 6.4 Mw. It was the first of a series of shallow earthquakes, with a further 1,000 smaller aftershocks. There were two further major earthquakes: the main shock on 5 August (7 Mw) and an aftershock on 19 August.

Impacts

The Indonesian National Board for Disaster Management (BNPB) reported that:

- 564 had died
- 1,584 had been injured
- 445,343 people were in makeshift shelters
- 129,715 houses were damaged
- 432,015 persons had been displaced and were living in 2,700 displacement sites
- 458 schools had been damaged
- the total cost of the damage was estimated to be around $600 million.

Responses to the earthquakes

The emergency plan of action was continually updated when necessary. Table 5.6 shows the response timetable.

Table 5.6 The emergency plan of action: Lombok earthquakes

Date	Response
29 July 2018: A 6.4 magnitude earthquake strikes off Lombok	• The Indonesian Red Cross allocates money from their emergency fund to enable them to meet the humanitarian needs of 1,000 households (4,000 people)
30 July 2018	• The local state governor declared a three-day state of emergency • People were advised not to return to their houses for fear of aftershocks. Assurances were made to help rebuild housing and infrastructure in Lombok using the Indonesian armed forces • Financial help was also promised
5 August 2018: A second stronger earthquake, of 7.0 magnitude and depth of 15 km, hits Lombok	• The Indonesian Red Cross launched an emergency appeal for funds to assist 80,000 people affected by earthquakes in Lombok for 18 months
9 and 18 August 2018: New 5.9 and 6.4 magnitude earthquakes strike Lombok	
26 August 2018	• The transition process from emergency to recovery phase starts following an announcement by the Indonesian National Board for Disaster Management (BNPB)
September 2018	• An earthquake hits elsewhere in Indonesia (Palu) and funds and help are diverted away from Lombok
8 November 2018	• The emergency appeal asks for further funds to deliver assistance to 20,000 households in Lombok for 30 months • It includes mid to longer-term recovery needs in the affected areas as well investing in increased preparedness and resilience

Source: International Federation of Red Cross and Red Crescent Societies (IFRC)

Most of the cost of US $600 million damage was met by the Indonesian government and Indonesian NGOs. A hospital ship was dispatched, ferry journeys were made free and the postal service helped move large quantities of aid.

It is interesting to note that the subduction zone along the west coast of Sumatra 'breaks' approximately once every 200 years. With this 200-year period approaching, it might mean that a powerful earthquake of magnitude 8.5 to 9 could happen by 2040.

Review questions

6 Study Figure 5.1 (page 186), the model of vulnerability. Using the information contained in this study of an earthquake, replace the generalities of the model with the real details of the Lombok event. What does the completed figure tell you about the extent of the risk to people and the environment of Lombok?

7 Outline the links between plate tectonics and volcanic and seismic hazards.

8 Explain why tsunamis are potentially the most damaging consequence of both volcanic and seismic events.

Other seismic events worth studying

Haiti 2010

Christchurch (New Zealand) 2011

Tohoku (Japan) 2011

Bohol (Philippines) 2013

Gorkha (Nepal) 2015

Kathmandu (Nepal) 2015

5.5 Storm (typhoons, hurricanes) hazards

Key terms

Storm surge – A rapid rise in sea level in which water is piled up against a coastline to a level far in excess of the normal conditions at high tide. They are produced during the passage of a tropical storm when wind-driven waves pile up water against a coastline combined with the ocean heaving upwards as a result of much lower air pressure.

Nature, distribution and magnitude

Nature

Tropical revolving storms (hurricanes, cyclones, typhoons) are intense low-pressure weather systems that develop in the tropics. These violent storms usually measure some 200–700 km in diameter. Studies of atmospheric pressure patterns show that there is a broad area of low pressure that encircles the globe either side of the Equator. The winds on the north side of this zone blow from the north-east (the north-east trade winds) and on the southern side winds blow from the south-east (south-east trade winds).

Within this area of low pressure, air is heated over the warm (27°C+) and deep (70 m+) tropical ocean. This air rises in discrete parcels, causing thundery showers to form that may group together into larger thunderstorms. Low pressure develops at the surface. If the storms are at a sufficient distance from the Equator (5°+ latitude) then the converging winds are acted upon by the Coriolis force and the incoming air spirals into and accelerates around the depression. As it does so, the air also rises and cools adiabatically (by expansion), leading to condensation of the moisture vapour in the air. Latent heat is released which acts as feedback, fuelling further uplift. These tropical depressions travel westwards over the ocean; some never develop into full-blown tropical storms or ever make landfall. Those that do can be very intense and destructive.

The tropical revolving storm exists while there is a supply of latent heat and moisture to provide energy and low frictional drag on the ocean surface. Once the system reaches maturity, a central eye develops. This is an area 10–15 km in diameter in which there are calm conditions, clear skies, higher temperatures and descending air. Wind speeds of more than 300 km/hour have been observed around the eye. Figure 5.28 shows the structure of a typical mature tropical revolving storm. Once the system reaches land or colder waters polewards, it will decline as the source of heat and moisture is removed.

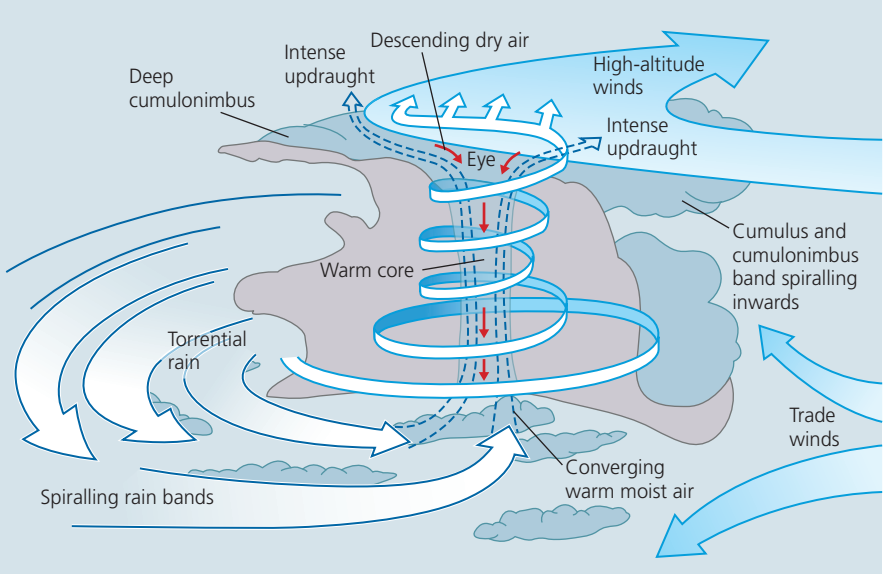


Figure 5.28 The structure of a tropical revolving storm

Distribution

Such storms occur between 5° and 20° degrees north and south of the Equator (Figure 5.29). Once generated, they tend to move westwards and are at their most destructive:

- in the Caribbean Sea/Gulf of Mexico area where they are known as **hurricanes** (11 per cent of all tropical revolving storms)
- on the western side of Central America (eastern Pacific) (17 per cent)
- in the Arabian Sea/Bay of Bengal area where they are known as **cyclones** (8 per cent)
- off southeast Asia where they are known as **typhoons** (this is the major area for such storms with one-third of all storms every year occurring here)
- off Madagascar (southeast Africa) (11 per cent)
- off north-western and north-eastern Australia (20 per cent). Off north-western Australia they are given the local name, **willy-willies**.

Magnitude

Tropical revolving storms are measured on the **Saffir-Simpson scale**, a five-point scale based solely upon wind speed (Table 5.7).

Table 5.7 The Saffir-Simpson hurricane scale

Category	Wind speed
Five	≥ 252kph
Four	209–251 kph
Three	178–208 kph
Two	154–177 kph
One	119–153 kph

There is a school of thought that thinks that the Saffir-Simpson scale should be scrapped because it only predicts damage from wind. In recent years, nearly 90 per cent of tropical storm deaths in the United States were not the result of wind, but water – storm surges, extreme rains and inland flooding.

The average lifespan of a tropical storm is 7–14 days, although it is believed that they may be lasting longer as a result of climate change. Every year, around 80–100 such storms develop around the world, with approximately 80 per cent of these going on to become tropical revolving storms.

The impacts of climate change

In an article published in the journal *Science* in 2005, climatologists examined the changes in tropical cyclone number, duration and intensity since 1970, the period when satellite data became available. They found that although the overall number of cyclones across the world had decreased, with the exception of those that occurred in the North Atlantic Ocean, there was a great increase in the number and proportion of very strong cyclones. One explanation for these changes is that the increase in ocean temperatures leads to an increase in:

- the air temperature above the warmer water, resulting in enhanced uplift and local lowering of atmospheric pressure
- oceanic evaporation and a corresponding increase in atmospheric water vapour content.

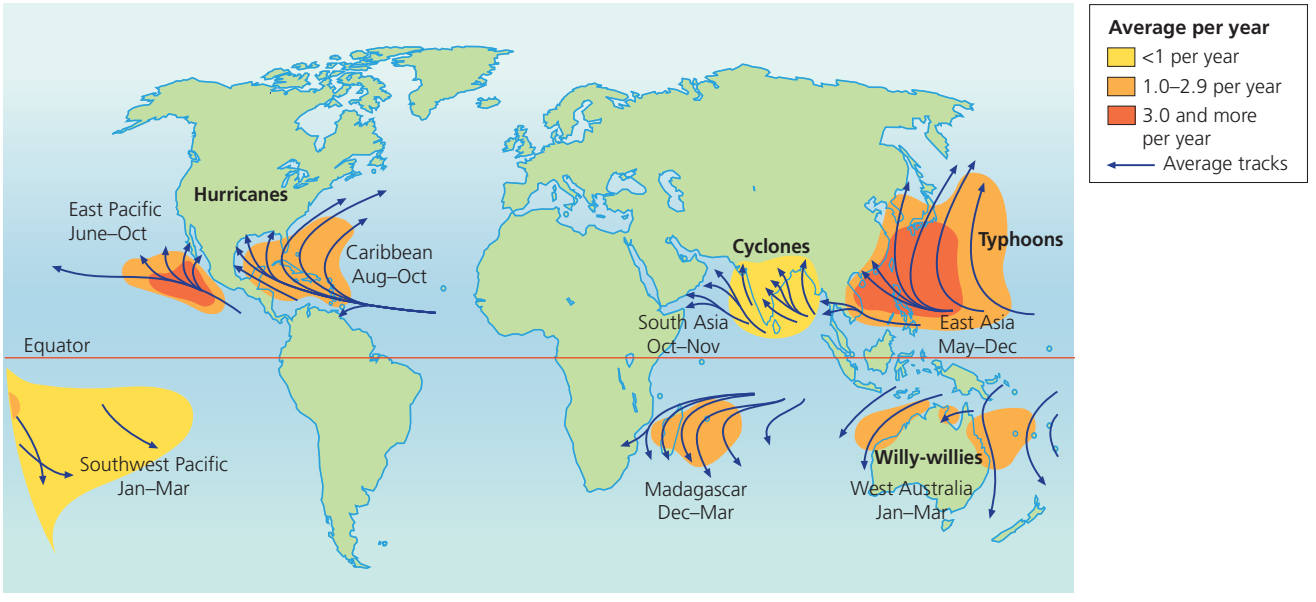


Figure 5.29 Global distribution and seasons of tropical revolving storms

These two factors then result in parcels of very moist rising air which, upon cooling with altitude, releases huge amounts of latent heat in the form of condensation, further enhancing uplift.

The predictions for changes by 2100 vary but can be summarised by:

- an increase of 2–11 per cent in the average intensity of storms
- a decrease of between 6 and 34 per cent in the total number of tropical storms
- a substantial increase in the frequency of the most intense storms.

Another study published in the journal *Nature* in 2017 suggested that slow-moving tropical storms are becoming more common. As polar areas get warmer, there is increasingly less difference in atmospheric pressure between the poles and the tropics. This reduces the speed of the winds between the two areas. Since it is these winds that carry the tropical cyclones within them, the tropical cyclones are slowing down as well.

Slower moving storms produce more rainfall. This is further enhanced by the greater amount of water carried by the storm due to higher ocean and air temperatures. There is predicted to be a 20 per cent increase in the amount of precipitation that will fall within 100 km of the centre of a storm. Wind damage from slower moving storms will increase simply because there will be more time for it to cause damage. Slower storms will also create a larger wall of seawater in front of them, increasing the damage potential from storm surges.

The Category 1 Hurricane Florence in September 2018 stalled for several days after making landfall in North Carolina, USA. High rainfall (760 mm+) lasting five days caused widespread inland flooding. Coastal areas also suffered a 1.7 m storm surge. The storm caused at least 55 deaths and costs were estimated at $17.9 billion. Research showed that the rainfall associated with Hurricane Florence was made worse by around 50 per cent due to human-influenced climate warming.

The impacts of tropical revolving storms

The vulnerability of people to storm events depends upon a range of factors, both physical and human:

- the intensity of the storm (scale 1–5)
- speed of movement, that is, the length of time over the area
- distance from the sea
- physical geography of the coastal area – width of coastal plain or size of delta, and location of any mountain ranges relative to the coast
- the preparations made by a community
- warnings and community response.

The major impacts that a tropical revolving storm can have on an area are:

- **Winds:** They can often exceed 150 km/hr (over 250 km/hour in a scale 5 event). Such high winds can cause structural damage to buildings (even collapse) and roads, bridges, etc. They can bring down electricity transmission lines and devastate agricultural areas. The huge amounts of debris that are flung about are a serious threat to peoples' lives.
- **Heavy rainfall:** It is not unknown for rainfall to exceed 200 to 300 mm, bringing about severe flooding, landslides and mudslides. If there is high relief near the coastal area, rainfall could increase to over 500 mm/day.
- **Storm surges:** High sea levels result when the wind-driven waves pile up and the ocean heaves upwards as a result of the lower atmospheric pressure. These storm surges can have a devastating effect on low-lying coastal areas such as river deltas where the flooding can extend a long way inland. Storm surges cause the majority of deaths in such events and agricultural areas can suffer for a long time as soil is contaminated by sea water.

Responses to storm hazards

Preparedness

Modern data gathering techniques mean that the prediction of tropical storms is increasingly sophisticated.

- **Hurricanes**: Weather bureaux (for example, the National Hurricane Centre in Florida, USA) are able to access data from weather satellites that give a real-time view of the location, track, size and strength of a hurricane. They also investigate potential hurricanes and gather data with specially adapted aircraft and weather balloons. Powerful supercomputers analyse the data from these different

sources and compare it to historical data from previous storms. The future path and strength of storms can then be predicted and people can be warned to **evacuate** vulnerable areas. It is essential that such warnings are correct, as there is a high economic cost associated with evacuation; the State of Georgia in the USA estimates that evacuation would cost $2.3 million per kilometre of coastline. Also, if there are false alarms, people might become complacent and refuse future advice.

Hurricane drills can be practised. In Florida, for example, 'Project Safeside' is a hurricane awareness programme with precautionary drills for use in schools.

- **Cyclones**: As cyclones tend to follow an erratic path, it is not always possible to give more than 12–18 hours warning. This means that in less developed regions, where communications are poor, there is insufficient time for a proper evacuation. Bangladesh's 'Cyclone Preparedness Program', an initiative devised by the Bangladesh government, the United Nations, the International Red Cross and the Bangladesh Red Crescent Society, includes:
  - an awareness campaign about cyclone warning signals, using a variety of ways to share information, such as public discussions, posters, leaflets, film and theatre
  - an early warning system based on modern weather radar stations that transmit weather updates and identify potentially dangerous storms while they are still at sea
  - the construction of storm shelters to reduce storm related deaths – in 2020, there were 4,000 shelters, with plans for 1,400 more
  - an ongoing programme by the Bangladeshi government of replanting former mangrove forest areas to protect areas just inland from the coast
  - encouraging people to have emergency packs containing dried food and a portable stove so that they are able to live in the storm shelters for some time.
- There has been a dramatic fall in storm-related deaths in Bangladesh. In 1970 over 300,000 people were killed by just one storm. In 2019, 2.1 million people were relocated to storm shelters as a result of the approaching Storm Bulbul. Despite widespread damage only eight people were killed.

Mitigation

In the USA, the Federal Emergency Management Agency (FEMA) has a handbook with suggestions on how to reduce damage to infrastructure, for example, by making roads, bridges, public buildings, parks and beaches as resilient as possible. For homeowners, it has produced a factsheet with advice on how people can protect their homes from flood and wind damage, from waterproofing all external walls to securely fixing down roofs to the rest of the building. They highly recommend taking out specialist insurance.

Prevention

Like most other natural hazards, tropical cyclones cannot be prevented. A number of prevention techniques have been unsuccessfully considered over the last 80 years, including:

- seeding clouds with dry ice or silver iodide so that the storm loses more water over the ocean and therefore has less latent heat energy to be released on land
- cooling the ocean with icebergs to reduce evaporation
- changing the radiation balance in the storm environment by blowing black soot into the storm
- exploding the storm apart with hydrogen bombs and blowing the storm away from land with giant fans.

All these ideas fail to take into account the sheer size and power of tropical storms. NASA has estimated that during its lifetime an average Atlantic hurricane can expend as much energy as 10,000 nuclear bombs.

Adaptation

With the growing frequency of extreme events (such as tropical storms) due to climate change, humans will increasingly have to adapt both their lifestyle and environment to live with these storms. Adaptation may include:

- **land-use planning** to identify the areas at greatest risk, for example, from storm surges, and to limit certain types of development in such areas
- building **sea walls**, **breakwaters** and **flood barriers** and putting houses on stilts. The sea wall that was built to protect Galveston (Texas) from further flooding, after a deadly storm surge of 1900, was so expensive that it is unlikely to be repeated and other measures would need to be considered elsewhere

- limiting expenditure on developments in high-risk areas and directing population away from them, as well as having plans to reduce evacuation times and for post-disaster redevelopment, such as has been done by local authorities in the USA
- **retrofitting** structures so they are wind resistant and safer during an event. In poorer areas, the need for land usually outweighs planning considerations to avoid high-risk areas so retrofitting buildings in high-risk areas is important. In 1994 in Dominica (West Indies), some homes were retrofitted by the Organization of American States and the government of Dominica. The value of the project was shown in the following year when all the retrofitted buildings withstood the impact of Hurricane Marilyn.

Finally, as with other natural phenomena, people in richer areas are urged to take out insurance. In poorer countries it is important that following a cyclone event, aid is available, both in the short and long term, as damage to the economy of the affected area is likely to last for many years. In the past, poorer regions lost more lives than richer ones. Richer regions lost more property, of a higher monetary value. This has changed. Globally, many fewer lives are lost to tropical storms but the difference in the value of property lost (in monetary terms at least) is still as big as ever.

Tropical storm events

A recent hurricane event: Hurricane Michael, northern Florida, USA (2018)

Hurricane Michael in October 2018 was the strongest storm to make landfall in the USA since Hurricane Andrew in 1992. Initially, it developed as a tropical depression in the northern Caribbean, between Cuba and Mexico, during the weekend of 6 and 7 October 2018. It had already hit parts of Mexico and Cuba with heavy winds and torrential downpours on the Sunday night and Monday morning. By 9 October, having become a category 2 storm, it started moving north through the Gulf of Mexico. It made landfall early in the afternoon of 10 October having briefly been a category 5 storm (Figure 5.30).

Risk and vulnerability

The storm came ashore on the coast of Florida near Mexico Beach, Bay County, Florida, a small coastal town with a population of 1,000. It then reduced in strength as it headed inland to hit parts of Georgia and North and South Carolina, which were already suffering from the effects of Hurricane Florence earlier that year.

Figure 5.30 The path of Hurricane Michael, October 2018

Source: NOAA

Initially categorised as a category 4 hurricane, a later report published by the National Hurricane Center in Florida said the storm was stronger than that, with winds at landfall estimated to be over 250 kph. Hurricane Michael was one of only four category 5 hurricanes to ever reach land in the USA.

Impacts

- Hurricane Michael caused a trail of destruction stretching from Florida (particularly the cities of Mexico Bay and Panama City) to Georgia, North Carolina and South Carolina. It caused flash flooding that turned roads into rivers.
- Wind, storm surge and rain (floods) were directly responsible for 16 deaths across the states of Florida, Virginia, North Carolina and Georgia. Additionally, the storm was responsible for an additional 43 deaths caused by falls during the post-storm clean-up, traffic accidents and medical issues made worse by the hurricane.
- Winds and storm surge caused devastating damage in Bay County, Florida, with more than 45,000 structures damaged and more than 1,500 destroyed, including two hospitals with severe damage. Trees were uprooted and power lines were brought down. Roofs were ripped off and swept into the air and homes were ripped open by fallen trees.
- In the town of Mexico Beach, 1,584 of the town's 1,692 buildings were damaged and more than 800 were destroyed.
- Power cuts, at their greatest extent, affected nearly 400,000 electricity customers in Florida, about four per cent of the whole state.

- $18.4 billion of the estimated total $25 billion damage from Hurricane Michael occurred in Florida. Agricultural and forestry losses totalled about $3.87 billion, primarily in Florida and Georgia.
- In the United States, there were at least $6.23 billion in insurance claims related to the hurricane.

Responses to the hurricane

Before the storm

Table 5.8 Responses to Hurricane Michael

Date	Response
7 October 2018	• In advance of the storm, residents were advised by the governor of Florida to prepare for the storm • The governor declared a state of emergency for 26 counties
8 October 2018	• The governor declares a state of emergency for an additional nine counties • A mandatory evacuation is issued for three coastal counties, including Bay County • Schools were closed • The mayor of Panama City asked two-thirds of the county (120,000) residents to evacuate
9 October 2018	• President Donald Trump approves a major disaster declaration request for Florida, giving access to federal resources and money for relief and recovery efforts • A state of emergency is declared in Georgia and plans for evacuation are made
	• As the storm approached, more than 20 million people in five states remained under either a hurricane or tropical storm warning, while areas in North and South Carolina flooded by Florence braced for more rainfall • Emergency Preparedness organisations, like Direct Relief, provided emergency medical packs to health facilities that were in the storm's path

After the storm

- Bay County, Florida had to borrow more than $250 million to pay for the clean-up whilst it waited for reimbursement from state and federal government.
- 1,500 national guard personnel were activated in the state of Georgia.
- On 11 October 2018, President Trump declared a major disaster in five counties. Residents were able to receive grants for house repairs, temporary shelter, loans for uninsured property losses and business loans. He also signed an emergency declaration for Georgia, authorising funding for 75 per cent of the cost of emergency protective measures and the removal of storm debris.
- The Federal Emergency Management Agency (FEMA) and several private and not-for-profit organisations established the Hurricane Michael Relief Network to provide direct relief to residents that were affected by the disaster.

Figure 5.31 Damage in the city of Mexico Beach a day after Hurricane Michael

Further information and advice on hurricane safety can be found on the US National Weather Service website at www.weather.gov/safety/hurricane-plan.

A recent typhoon event: Typhoon Haiyan, Philippines (2013)

Typhoon Haiyan (also known as Yolanda) was one of the strongest tropical cyclones ever recorded and the deadliest Philippine typhoon in modern history. It originated from an area of low pressure in the Federated States of Micronesia (western Pacific Ocean) on 2 November. Tracking westward it gradually developed into a tropical storm by 4 November (Figure 5.32, page 218). It then began a period of rapid intensification that brought it to typhoon intensity by 5 November. By 6 September, the system was assessed as a Category 5 or super-typhoon with wind speeds already reaching over 250 km/hr, with gusts reaching over 300 km/hr.

On 7 November Haiyan reached the Philippines where it caused great destruction in the area around Tacloban, which was struck by the northern eyewall, the most powerful part of the storm. By this time, central pressure had dropped to 895 mb and the highest wind speeds reached 315 km/hr (calculated as sustained for one minute). This meant that Haiyan was the most powerful storm ever to strike land. After crossing the Philippines, Haiyan continued westward towards southeast Asia, eventually reaching the north of Vietnam on 10 November as a severe tropical storm. It eventually dissipated as a tropical depression over southern China on 11 November.

Figure 5.32 The track of Typhoon Haiyan (wider area and central Philippines)

Risk and vulnerability

Tacloban lies in the centre of the Philippines where tropical storms cross from the Pacific Ocean towards southeast Asia, which makes the region the most affected in the world by this type of hazard.

In recent years many Filipinos have moved in large numbers from countryside areas to cities such as Tacloban looking for opportunities to work and make money. This has led to high densities of population and often overcrowded residential areas, which makes the people more vulnerable when events such as Haiyan occur. Many of the people have built poorly constructed houses on the flat land nearest the sea and such locations make people there vulnerable to storm surges and flooding. Others have built houses on steep hillsides, which again makes them vulnerable to the high rainfall of tropical cyclones which, on slopes where vegetation has been removed, will cause landslides that can devastate whole communities. Poor people are especially vulnerable to the aftermath of such events. Unable to provide for themselves, they look to regional and national government agencies who often do not have the resources to cope in such emergencies. This makes people, such as those in Tacloban, dependent on foreign government help and on NGOs, such as charitable organisations.

Impacts

Apart from the severe winds, Haiyan also caused:

- very heavy rainfall. In one area of the Philippines, nearly 300 mm was recorded, much of which fell in under 12 hours
- huge storm surges, measuring up to six metres, which were responsible for most of the damage and loss of life.

Major effects of the storm:

- There were over 6,000 **fatalities**, with over 1,000 people declared missing. Bodies were being discovered well into 2019. Some estimates have put the death toll as high as 10,000.
- There was widespread **damage to buildings and infrastructure** (Figure 5.33). The American army officer in charge of relief flew over Tacloban and observed 'I do not believe there is a single structure that is not destroyed or severely damaged in some way.' Some journalists described the devastation as 'off the scale' and 'apocalyptic'.
- It was estimated that nearly 2 million people were made **homeless** and more than 6 million people **displaced**, with at least 20,000 fleeing to the capital, Manila.
- The low-lying areas on the eastern side of Tacloban were particularly vulnerable with some areas completely **washed away**.
- Large areas were **cut off**, with communication and power systems destroyed.
- With **flooding and landslides**, major roads were blocked (also with fallen trees) which made relief work almost impossible except from the air.
- Tacloban city's **airport** was damaged by the storm surge (the terminal building destroyed by a 5.2 m storm surge) but facilities had to be repaired quickly as it became the centre of the international relief operation.
- Widespread **looting** was reported.
- **Local government** in many areas collapsed as so many local officials were killed during the storm.
- The **total damage** has been estimated at $2.9 billion.

Responses to the typhoon

After the event it was feared that there would be substantial outbreaks of diseases such as cholera, leading to an even higher death toll. Prompt action by the WHO and other relief agencies ensured that outbreaks, such as that of dysentery, were kept isolated and to a minimum.

Immediately after the storm, the United Nations stated that 'access remains a key challenge as some areas are still cut off from relief operations. Unknown numbers of survivors do not have basic necessities such as food, water and medicines and remain inaccessible for relief operations, as roads, airports and bridges were destroyed or covered in wreckage.' Many countries contributed to the relief effort through cash donations or practical help. The UK government deployed two navy ships to the Tacloban area supplying over 200,000 tonnes of aid, while the Americans sent numerous ships and aircraft and over 13,000 service personnel.

Review questions

9 Both south-eastern USA and Tacloban (Haiyan) were hit by very powerful tropical storms but the estimated cost of the damage to each area was completely different. Why was this?

10 Why are tropical revolving storms potentially more damaging than any tectonic hazard?

Figure 5.33 Devastation caused by Typhoon Haiyan in Tacloban, Philippines, November 2013

Other tropical storms worth studying

Cyclone Nargis (Myanmar) 2008

Hurricane Sandy (USA) 2012

Typhoon Hagibis (Japan) 2019

5.6 Fires in nature

The nature of wildfires

Wildfire is a natural process in many ecosystems and can be a necessary and even beneficial process in some of them. It is also true to say that fire can be a very destructive process in the natural world, one which has both natural and human causes.

Fires can be major events, occurring on a large scale and causing widespread destruction and killing much wildlife. Surface fires sweep rapidly over the ground, consuming plant litter, grasses and herbs, and scorching trees, and it is possible for ground temperatures to rise to over 1,000°C. At higher levels there are crown fires that spread through the canopy of trees.

The nature of the fire will depend upon the types of plants involved, strength of the winds, topography of the area in question and the behaviour of the fire itself. Once vegetation has dried out, the nature of the fire will depend largely upon the wind, as the largest fires occur in dry windy weather with low humidity. Wind drives the fire forwards and burning embers that ignite more vegetation are more easily spread in windy weather. The key factors therefore in any fire are the climate and the nature of the plants involved.

Key terms

Pyrophytic (vegetation) – Pyrophytes are plants adapted to tolerate fire. Methods of survival include thick bark, tissue with a high moisture content and underground storage structures.

Retardants – Chemicals sprayed on to fires in order to slow them down. They are composed of nitrates, ammonia, phosphates, sulphates and thickening agents.

Causes of wildfires and their spread

For a natural fire to occur and to spread, two things are needed:

- **An ignition source:** in the case of natural fires, lightning is by far and away the main cause. Climate will affect the frequency of electrical storms, particularly one in which there is hardly any rainfall. Increasingly, fires are the result of human intervention, particularly those which occur

in and around settlements. Such fires are started by falling power lines, carelessly discarded cigarettes, children playing with matches, camp fires and barbecues and agricultural fires (controlled burning) which get out of hand. In some cases, there is evidence that devastating fires of this kind have been deliberately started by arsonists. In the USA and Australia, greater access to wild areas by tourists has increased the danger of fires through some of the ways listed above.

- **Fuel:** the fuel has to be of sufficient quantity and dry enough to burn. Climate affects the frequency and duration of droughts during which the vegetation and plant litter has an opportunity to accumulate and dry out. Climate also affects the type of vegetation that will grow in an area and the rate at which plant litter can be produced.

Distribution

Wildfires are essentially a rural hazard and can occur in most environments, although with the continued expansion of human habitation, they now occur within the boundaries of even substantial settlements in California, Australia and the countries of southern Europe. Areas that are most susceptible are those with a Mediterranean and tropical wet season/dry season (savanna) climates. In Mediterranean climates, winter rainfall encourages vegetation growth but hot dry summers, with occasional thunder storms, result in a fire hazard. A similar pattern occurs in savanna regions although it is the summers that are wet and the winters that are dry. Such regions include:

- Mediterranean climates:
  - parts of Australia – south west, south east (Victoria and southern New South Wales) and parts of coastal southern Australia
  - California
  - southern Europe – including southern France, Italy, Greece, Turkey and Mediterranean islands
- Savanna:
  - Northern Australia
  - Tropical regions of Africa and South America
  - North-eastern India
  - Northern Florida.

Fires and natural ecosystems are closely linked, particularly in the areas named above. Fire can clear vegetation and aid new seed germination, stimulate the growth of certain plants and rid an area of insects and some parasites. Some species are **pyrophytic** in that they can withstand fire through some mechanism such as resistant bark, for example, the baobab tree. In Australia, plants such as banksia only release their seeds after a fire; this ensures that the seeds fall on well-cleaned, brightly-lit ground recently fertilised with ash.

Fire has not been a real hazard of tropical rainforest areas due to the humid climate, but recent burning for forest clearance has often got out of hand, resulting in widespread fires that burned out of control for long periods. Both the Amazon Basin (Brazil) and south-east Asia (Indonesia) have suffered from such fires in recent years.

A good example of an area where wildfires are a continuing problem is the Los Angeles Basin in California. Wildfires that occur here often get a great deal of media coverage (in some cases globally) as it is likely that a wildfire event could threaten the homes of famous film and TV personalities. Wildfires are common in the Los Angeles area and pose a serious threat for the following reasons:

- much of the area is covered by drought-resistant chaparral, which is a scrub vegetation, as the area is too dry for much tree growth. This vegetation, after the summer drought, can be tinder dry.
- the dry Santa Ana wind descends from the local mountains, increasing the dryness of the vegetation to the point where a spark, lightning or a carelessly discarded cigarette can cause a major fire. This wind also allows for the easy spread of fire, and makes it very difficult to control.
- much of the area, outside the centre of Los Angeles, consists of low-density building where the natural vegetation has been allowed to remain between properties, which exposes a large number of them to the fire risk.

The impacts of wildfires

Primary effects

- **Loss of crops, timber and livestock:** Forest fires can have a huge impact in timber producing areas, with the loss of trees that will take many years to replace.

- **Loss of life:** Although many fires are events from which people can get out of the way, some fires move so fast that people can be trapped, although this is not usual. In the Victoria (Australia) bushfires of 2009, 173 people lost their lives.
- **Loss of property:** At one time, only a few rural communities were at risk from fires, but with urban expansion, the fringes are now susceptible, for example, Sydney (Australia) and Los Angeles (see above). The cost of damage and of fighting the fires can run into hundreds of millions of dollars. After such events, large numbers of people can be left homeless.
- **Release of toxic gases and particulates:** The Southeast Asia 'haze' is fire-related large-scale air pollution that occurs regularly and is still present in parts of the region. It is largely caused by illegal agricultural fires due to slash-and-burn practices in Indonesia. The resulting pollution covers a number of countries in the region including Singapore, Malaysia, Thailand and Vietnam. This introduces a political dimension, as the governments of these countries have demanded that the Indonesian government take action against companies responsible for illegal forest fires. The size of the haze produced by the Indonesian fires of 1997–98 can be seen in Figure 5.34. These fires were described by one commentator as 'probably amongst the two or three, if not the largest forest fires, in the last two centuries of recorded history'.
- **Loss of wildlife:** In Indonesia, the fires of 1997 destroyed the Wein River orangutan sanctuary on Kalimantan.
- **Damage to soil structure and nutrient content:** With the intense heat generated at ground level, wildfires can destroy many soil nutrients and lead to an alteration in the soil's structure. AccuWeather estimates the total damage and economic loss (both insured and uninsured) caused by the California wildfires in 2019 was $80 billion. This was based on an analysis of all direct and indirect impacts of the fires.

Secondary effects

- **Evacuation:** Many people will flee from the area of the fire. Such people will not be allowed back into the affected area, often for a long period, if not forever. Emergency shelters/accommodation will have to be found along with food, etc.
- **Increased flood risk:** In certain environments, where rain comes in heavy bursts, the loss of so much vegetation, and the consequent decrease in interception, can lead to increased flooding.

Figure 5.34 The extent of the haze produced by the Indonesian fires of 1997–98

Responses to wildfire hazards

Preparedness

- **Education**: In areas affected by fires, governments often produce comprehensive guidance to help authorities and people prepare for fire hazards. For example, in California (USA), the Department of Homeland Security recommends:
  - signing up and being aware of fire warning systems produced by the state of California and the federal weather service
  - familiarisation with evacuation routes and roads
  - putting together a bag containing basic essentials for life away from home
  - designating a room in houses (and other buildings) that can be shut off from outside air when smoky conditions prevail and using a portable air filter to reduce indoor pollution
  - creating a fireproof safe place at home for all important paper documents (passports, driving licences and so on) and storing all digital documents on cloud storage
  - using fire-resistant materials to build house extensions or outbuildings
  - clearing flammable materials and vegetation within 10 metres of buildings to create a fire-resistant zone

  - taking out sufficient insurance cover to replace everything that could be lost, including the house itself.
- **Community action**: In Victoria (Australia), a community education programme, known as 'Community Fireguard', has been established to help individuals and communities develop a high level of fire preparedness and resilience. The programme focuses on identifying the most vulnerable areas in fire-prone communities and then makes residents aware of how they can be responsible for their own safety. The residents develop their own fire survival strategies and techniques, such as local warning systems, ensuring that buffer zones are maintained, carrying out clean-ups of the local area, running equipment training sessions and preparing emergency plans.
- **Computer modelling**: This has been employed to understand and predict fire behaviour. It involves studying the ways in which fires behave to comprehend and predict fire behaviour. Figure 5.35 shows a fire propagation model.

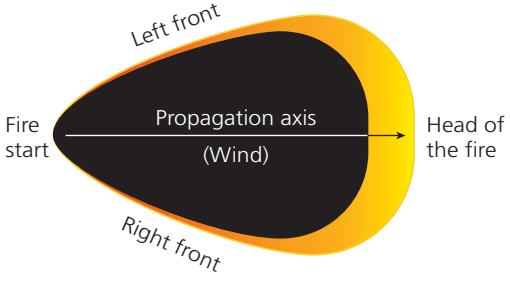


Figure 5.35 Fire propagation model

Mitigation

To reduce the risk of fire damage to both property and forest, long-term mitigation action must be undertaken. In the USA, The Colorado State Forest Service and Colorado State University encourage the creation of 'fire-adapted communities' to increase forest resilience. They provide homeowners and communities with advice on how to reduce and even prevent damage to property by fire, such as through fire-resistant landscaping and planting and incorporating fire protection measures within homes.

Prevention

- **Education**: Wildfires, unlike the other hazards in this chapter, are preventable to a certain extent. In regions prone to wildfires, education programmes, signs and posters are used to urge people to be careful and advise them on how to avoid starting fires by:
  - not leaving a fire or barbecue unattended and always extinguishing it fully with soil and water
  - not discarding a lighted cigarette out of car windows
  - removing litter, particularly any broken glass.
- Of course, this does not prevent fires being started deliberately.
- **Managing the vegetation**: Clearing areas between trees and other vegetation ahead of wildfire events to create firebreaks helps prevent the spread of a wildfire.

Dealing with the event as it happens

Once a fire has started it has to be dealt with. This involves fighting the fires as they happen, which can be a potentially very dangerous operation and, although rare, firefighters can suffer fatalities when they are involved in this action. Six firefighters died either directly or indirectly in the Australian fires of 2019–20. Spraying with water and chemicals both from the ground and in the air is costly so firefighters also work on the ground beating out the flames and bulldozing large breaks in the vegetation, preventing the fires jumping in certain directions. **Retardants**, to slow the pace of the wildfire, are often sprayed on fires in areas which are inaccessible or when human safety and structures are endangered.

Adaptation

Wildfires have always existed. Some species of vegetation have even evolved to become dependent on fire to reproduce. Unfortunately, wildfires have become more common. Climate is changing and as such, fire seasons are hotter and longer than ever before. Also, human activity and settlement have encroached upon previously untouched areas. Communities need to adapt to live with wildfires.

According to the Centre for Climate and Energy Solutions, wildfires cost the USA $24 billion in 2018. Cities and areas throughout the USA prone to fires, such as Montecito in California, have put in place measures to enable residents to adapt to living with wildfire hazards. For example, houses must maintain a mandatory 35-metre space around the home and ember-resistant and heat-resistant building materials, such as tempered glass and cement render to cover wooden walls, must be used. Despite these measures, fires still pose a major hazard in Montecito.

Wildfire event

A recent wildfire event: Australian bushfire season 2019–20

Risk and vulnerability

The 2019–20 Australian bushfire season (known locally as the Black Summer) began in June 2019 with several significant uncontrolled fires. It continued throughout the summer, mainly in the south-east of the country, peaking during December and January (Figure 5.36).

The nation's capital, Canberra, was covered in a blanket of cloud, which spread out across the Pacific Ocean (Figure 5.37).

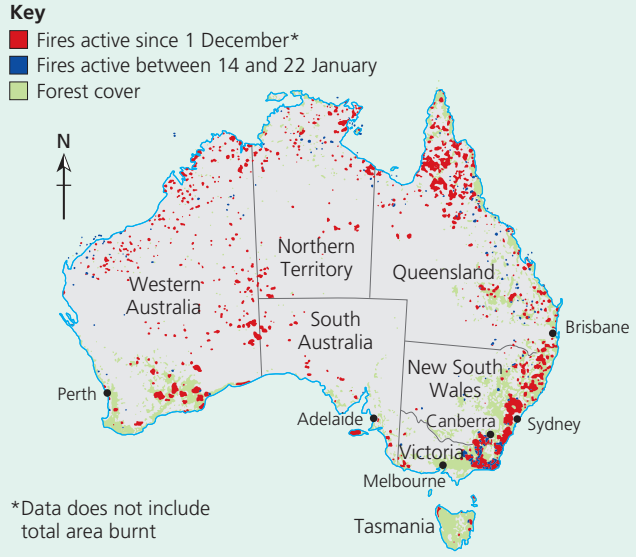


Figure 5.36 Map of location of wildfires in Australia from December 2019 to January 2020 using infrared data

Source: Springer Nature

An extremely dry and warm spring and early summer provided optimal conditions for the fires, which began in Queensland and spread southward (Figure 5.36). According to the Australia Bureau of Meteorology, it was the driest spring on record for the continent, and the temperatures for spring were above average across nearly all of Australia. In the summer, there was a record heatwave with a country average of 40.9°C.

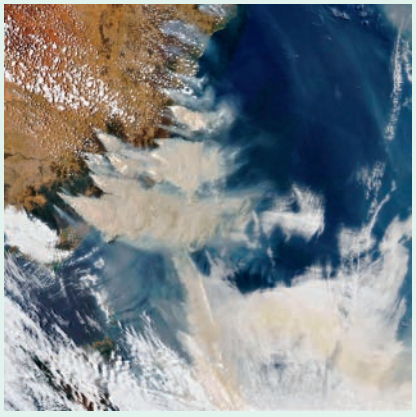


Figure 5.37 A satellite image of smoke from bushfires in Victoria and New South Wales, Australia on 2 January 2020

Impacts

- More than 20 per cent of Australia's forests burned during the devastating bushfires – an unprecedented proportion globally.
- By the beginning of March 2020, an estimated 18.6 million hectares of forest and over 5,900 buildings (including 2,779 homes) had been destroyed by the fires, with at least 34 people killed.
- It is estimated that one billion animals were killed and some endangered species may have become extinct.
- Analysts projected the total costs of the fires to be in the neighbourhood of US$70 billion
- The smoke from the fires travelled approximately 11,000 kilometres across the South Pacific Ocean to Chile and Argentina.
- NASA estimated that 306 million tonnes of CO_2 had been emitted.
- In the course of firefighting operations, an air tanker and two helicopters crashed resulting in the deaths of the three air tanker crew. Two fire trucks were involved in fatal incidents caused directly by fire conditions, killing three fire fighters.

Responses to the wildfires

- A number of states of emergency were declared across New South Wales, Victoria and the Australian Capital Territory.
- Reinforcements from all over Australia were called in to assist fighting the fires and relieve exhausted local fire crews in New South Wales.
- The Australian Defence Force was mobilised to provide air support to the firefighting effort and to provide manpower and logistical support.
- Firefighters, supplies and equipment from Canada, New Zealand, Singapore and the United States, among others, helped fight the fires, especially in New South Wales.

Kangaroo Island

Figure 5.38 Vegetation on Kangaroo Island after the forest fires

One location that gained global attention was Kangaroo Island in the state of South Australia (Figure 5.38, page 223). Located 112 km south-west of Adelaide, it is just 13.5 km from the mainland at its closest point. Over a third of the island is protected by nature reserves, which are home to native wildlife such as sea lions, koalas and diverse bird species. In the west, Flinders Chase National Park is known for penguin colonies.

More than half the island (215,000 hectares) was damaged. Following the fire there were reports of so many corpses of wild animals caught in the fire that they had to be buried in huge pits. Injured animals were taken in and fostered by local people or sent to the mainland for treatment.

Not only has the fire taken away homes and people's livelihoods, but there is also concern that some of the wildlife on the island has now become extinct.

Review questions

11 Some natural environments are more likely than others to experience fire as a hazard. Identify such environments and describe their global situation.

12 To what extent can the Model of vulnerability (Figure 5.1, page 186) be applied to the Australian wildfires of 2019/20?

13 'Resources are stretched. Even before we could recover from one disaster, here is the next one.' Does this really sum up the attitude of people living in a multi-hazard environment?

Case study of a multi-hazard environment beyond the UK: The Philippines

The Philippines is a country off the mainland of southeast Asia consisting of an archipelago of over 7,600 islands (Figure 5.39). It is also home to over 100 million people.

The Philippines has the fourth richest economy in south-east Asia and is considered to be a newly industrialised country. Despite this, there is a large imbalance in the distribution of wealth with 21.6 per cent of the population (mainly in rural areas) below the poverty line.

By virtue of its geographical position it is prone to a range of natural disasters. Its position on the western rim of the 'Pacific Ring of Fire' brings both **earthquakes** and **volcanic activity**, with secondary impacts such as **lahars**. It is also situated in an area where tropical storms (**typhoons**) cross from the Pacific towards southeast Asia making the Philippines the most exposed country in the world to this particular hazard. These storms bring a great deal of **flooding** and the threat of **landslides**, which again add to the hazardous nature of life in these islands. All of this hinders the attempts by government to reduce the incidence of poverty and to reduce the number of people and assets vulnerable to these hazards. The cost to the Philippine economy has been estimated at 0.5 per cent of national GDP and indirect and secondary impacts further increase this cost. In addition to the significant economic cost, there are substantial social and environmental impacts. In 2018, for example, the country was subject to two severe earthquakes. A magnitude 6.1 Mw hit Luzon in April 2018, damaging buildings and killing five people. A more powerful, but less damaging, earthquake of magnitude 7.1 Mw struck off southern Mindanao in December 2018. In addition, Mount Mayon erupted in February of that year (see pages 201–3). The islands were also touched or crossed by nine typhoons and tropical storms, including a category 5 'super' typhoon Mangkhut in September. Mangkhut was responsible for at least 127 deaths, including those of miners and their families caught in a landslide caused by heavy rain on bare soil and rock.

Major threats to the population of the Philippines are described below:

Earthquakes: These are a fairly common occurrence in the Philippines. In the twenty-first century, between 2000 and 2017, there were 17 earthquakes with a magnitude of at least 6.0 Mw, which have caused numerous deaths and damaged a large number of buildings. The most devastating in recent years have been:

- **1976 off Mindanao (Moro Gulf):** this earthquake (magnitude 7.9 Mw) caused tsunamis which inundated the western shore of Mindanao killing up to 8,000 people

Figure 5.39 Map of the Philippines

- **1990 Luzon:** (magnitude 7.8 Mw) killed over 1,600 people and caused widespread ground rupturing and soil liquefaction
- **2013 Bohol:** a magnitude 7.2 Mw earthquake which killed over 200, injured 800, and caused damage to tens of thousands of buildings.

Volcanic activity: The Philippines has a number of active volcanoes on its islands (opinions vary, but 23 would seem to be a reasonable figure). Among these are three volcanoes which have a history of violent eruptions:

- **Mayon:** This is the most active, having erupted nearly 50 times in the last 400 years, with the most destructive event coming in 1814. The last eruption was in 2018. Mayon is noted for its lahars, with volcanic ash mixing with heavy rainfall from tropical storms to produce rivers of mud.
- **Pinatubo:** Erupted violently in 1991, being the second largest eruption on the planet in the twentieth century. Pinatubo can be deadly as large numbers of people live in the vicinity of the volcano (500,000 people within 40 km). Early warnings and evacuations probably saved at least 5,000 people in 1991 (Figure 5.41, page 226).

  The explosion was so violent that it hurled gases, ash and steam into the upper atmosphere where they continued to affect global temperatures for at least two years (average global temperature down 0.5°C).

  This was followed by Typhoon Yunya, with very heavy rainfall combining with volcanic ash to produce deadly lahars. Final death toll was around 850.
- **Taal:** One of the most active volcanoes in the country, having 33 eruptions noted on the historical record. All of these eruptions have come from a volcano that sits in the middle of a crater lake formed when a huge explosion in prehistoric times formed the caldera of Taal. In the historical record it is thought that Taal has been responsible for 5,000–6,000 deaths as it lies close to populated areas.

Tropical storms: Numerous storms cross the country every year bringing with them the risk of severe flooding, landslides and lahars if combined with a volcanic eruption. Such disturbances account for the highest numbers of deaths from natural hazards in the Philippines, most fatalities being the result of storm surges.

Typhoon Mangkhut was an extremely powerful and catastrophic tropical cyclone that caused extensive damage in September 2018. It was the strongest typhoon to make landfall anywhere in the Philippines since Haiyan in 2013 and was the third-strongest tropical cyclone worldwide in 2018. It caused 127 deaths in the Philippines and cost US $627 million.

According to the Red Cross, about three million people were affected in the worst-hit parts of the Philippines and another 11 million further south faced flooding that made it very difficult for them to return to their rice farms.

In a multi-hazard environment such as the Philippines, people are often at a greater risk than elsewhere. Not only are they vulnerable to the main hazards of earthquakes, volcanic activity and typhoons, but these hazards can be linked to others, in some cases being the direct cause of those hazards. These hazards – tsunamis, landslides/mudslides (Figure 5.40), flooding, fires and lahars – are sometimes seen simply as secondary effects of the main hazard, but they are often recognised in their own right as a hazard. Earthquakes can lead on to tsunamis, fires and landslides and in a human context, the hazard of crime (looting). Heavy rainfall associated with typhoons can cause serious flooding and landslides. The combination of volcanic activity (ash) and the same heavy rainfall can initiate lahars which can devastate whole communities.

Figure 5.40 Landslides in northern Philippines caused by excessive rainfall nearly block a road

Figure 5.41 The Pinatubo Volcano eruption in the Philippines, 1991

In the past two decades, the Philippines has experienced over 300 natural disasters and Filipino officials insist that the typhoon hazard is becoming more severe, in part because of climatic change. Over the past decades, Filipinos have flocked to risky low-lying areas, havens for cheap, crowded housing where there is a clear link between poverty and vulnerability to natural disasters. Rapid urbanisation has exacerbated the problem, with its tightly packed, flimsily constructed housing and the environmental degradation such as deforestation that has only added to the problem (steep, unprotected slopes leading to rapid run-off, flooding and the potential for landslides).

The Philippines institutional arrangements and disaster management systems have tended to rely on a response or reactive approach, when they should be, in contrast, taking a more effective proactive approach in which disasters can be mitigated, even avoided, by appropriate land-use planning, construction and other preventative measures which avoid the creation of disaster-prone conditions. There has been widespread emphasis on post-disaster relief and short-term preparedness (forecasting/evacuation), rather than on mitigation and post-disaster support for economic recovery. The current system tends to be more of a centralised top-down administrative system than one which is community based and gives rise to local initiatives. What is required is initiatives which emphasise the bottom-up approach that will help the most vulnerable communities cope with the hazards when they occur.

Some programmes have been started along these lines. The Philippines National Red Cross (PNRC), in conjunction with foreign Red Cross organisations, has started programmes at community level. The aim of these is to reduce the impacts of natural disasters by encouraging people to collaborate in protecting their lives and the resources upon which they depend. Such plans address the following areas:

- co-operation and partnership with governmental bodies in order to gain financial support for mitigation measures and to ensure that the programmes have long-term sustainability
- training local volunteers in disaster management
- identifying risk through land-use mapping and determining which mitigation measures might be possible
- initiating mitigation measures which could be physical (sea walls, dykes), health related (clean water supplies) or planning tools (land-use plans, evacuation plans)
- dissemination of information to the whole community.

Many commentators have pointed to the great resilience of the Filipino people in the face of the problems in living in a multi-hazard environment. Others have simply described the attitude of the people as fatalistic; they accept that these events are part of living in such an area (even 'God's will') and that losses are inevitable from time to time.

The central government, however, is trying to do something about the situation. The United Nations claims that the Philippines already has some of the best risk-reduction laws in the world, but unfortunately most of these are still on paper. With a country of over 7,600 islands, too much of the responsibility for reducing disaster risk falls to local government where money is not always wisely spent. Recent government legislation, however, calls for 70 per cent of disaster spending to be used on long-term plans with only 30 per cent going on emergency aid. The main problem of living in a multi-hazard environment remains, though, as one government official put it, 'resources are stretched; even before we could recover from one disaster, here is the next one.' An infographic summarising the hazardous events in the Philippines in 2018 can be found at:

https://reliefweb.int/report/philippines/philippines-2018-highlights-events

Review question

14 'Resources are stretched. Even before we could recover from one disaster, here is the next one.' Does this really sum up the attitude of people living in a multi-hazard environment?

Case study at a local scale of a place in a hazardous setting: Gili Trawangan, Indonesia

On 5 August 2018, the island of Gili Trawangan was left devastated after an Mw 7.0 earthquake hit the Lombok area of Indonesia. Most of the island's buildings were destroyed.

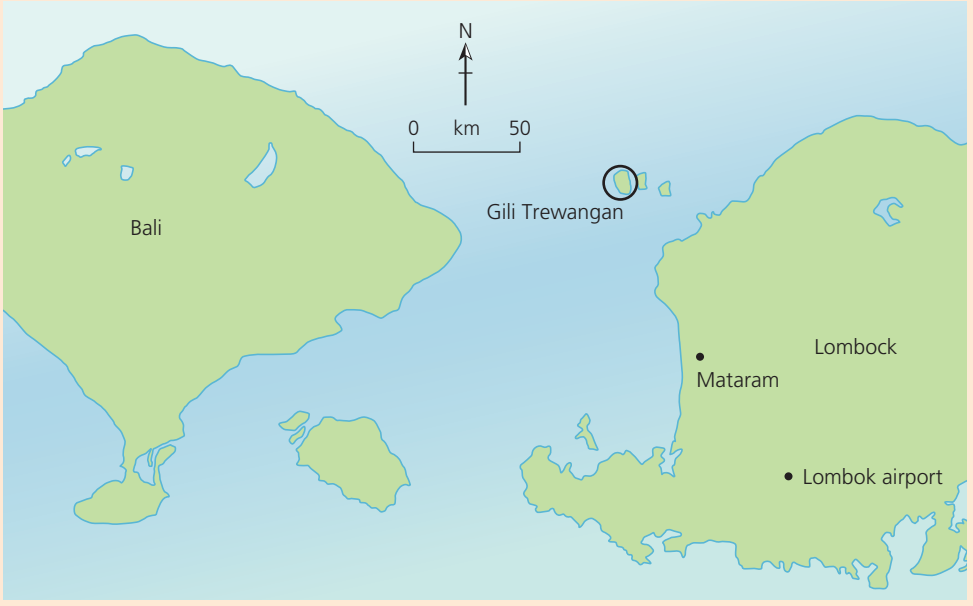


Figure 5.42 The location of Gili Trawangan

Gili Trawangan is a tourist spot in Indonesia, popular with divers, party lovers and honeymooners. It is the largest in the chain of three Gili islands just off the north-west coast of Lombok (Figure 5.42). It is the only island in the chain to rise significantly (30 m) above sea level. It is three kilometres long by two kilometres wide and has a permanent population of around 1,500 people. There are no motorised vehicles on the island. The main means of transportation are bicycles (rented by locals to tourists) and Cidomo (a small horse-drawn carriage (Figure 5.43)). Locals usually use motor boats to travel between the Gili islands and there are links to Lombok, Bali and other islands.

Risk and vulnerability

As with the bigger islands of Lombok and Bali, Gili Trawangan is located within an island arc structure where the Indo-Australian Plate is moving northward towards and subducting under the south-east moving Eurasian Plate. The friction between the two plates is immense, causing the subducting plate to get 'stuck' with a resulting build-up of stress in the rocks. When the plates become 'unstuck', huge amounts of energy are released resulting in earthquakes.

The development of the island has been haphazard. Originally, it was a stop-off location for fishermen, but by 1971 coconut plantations had been established and 350 inmates from the overcrowded Mataram prison were sent to help with the first harvests. When the coconut harvests failed, many of these workers decided to stay, building makeshift homes and businesses on the private, abandoned plantation land. In the 1980s, backpackers started discovering the island and it rapidly developed cheap hotels and bunkhouses.

From the 1980s to the present day, there have been land disputes between the settlers and private companies who own the land on the island. This has resulted in further chaotic cycles of building, demolition and then rebuilding. Since the 1980s, most of the tourist developments have been created by westerners moving to the island. As a result of this haphazard development and the island's location:

- many of the buildings on the island are structurally unsound
- the only links to the mainland are by boat
- there is no permanent police force on the island
- there is no hospital (and ambulances and so on), though many of the dive centres have first-aid trained staff.

Figure 5.43 Cidomo – horse-drawn carriage, Gili Trawangan

Impacts of the earthquake

- At least eight people died.
- About 30 per cent of the buildings were destroyed and another 40 per cent were damaged.
- The water and electricity supplies were cut off.
- Horses used to pull the carts on the car-free island were running loose.
- There were reports of looting and it took days for the military to arrive to make the island secure.
- The number of tourists fell dramatically. According to the general manager for Aston Sunset Beach in Gili Trawangan, the average number of tourists arriving on the island each day before the earthquake could reach 3,000 people. Three months after the earthquake it was still only around 900. Many of the absent travellers were Indonesian rather than foreign visitors.

Responses

- Doctors, dive instructors and tourists set up a basic camp with oxygen and medical supplies and organised people into groups from minor injuries to the most severe.
- Many people camped out overnight, as they were too afraid to sleep indoors.
- Thousands of panicked tourists left the islands the day after the earthquake (Figure 5.44). Most headed straight for the international airport, where some slept on the floor as they waited for a flight out.

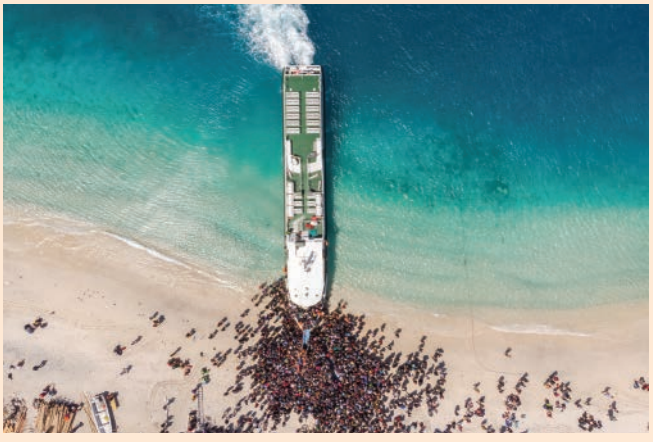


Figure 5.44 People crowd the beach as they wait to be evacuated from Gili Trawangan island to neighbouring Lombok island the day after a 7.0 magnitude earthquake struck

- The Indonesian National Agency for Disaster Management responded slowly. Complaints were made that despite the risk of earthquakes, there had not been any disaster mitigation training from the central government.
- The Indonesian Financial Services Authority (OJK) agreed to reduce electricity and water bills, as well as helping businesses secure bank loans.
- By February 2019, around 50 per cent of the 500 hotels and restaurants on Gili Trawangan had returned to normal operations.

In September 2019, the government inspected the island and categorised the amenities according to the level of repair needed, from those that were lightly damaged, those with average need for repair and those that were badly damaged and needed rebuilding. After completing the inspection, the government announced that:

- Gili Trawangan was making a rapid recovery and that a number of hotels, cafes and restaurants were already receiving visitors
- the Cidomo, the traditional horse-drawn carriage transport on Gili Trawangan, needed to be retained and better organised
- the Ministry of Public Works would implement road improvements with work starting on paving the roads from November 2019
- the jetty at Gili Trawangan would be given priority and rebuilt to meet international standards, so that it could better meet the needs of the thousands of tourists who visit the Gili Islands. It was scheduled for completion by November 2019.

The inspection concluded that there were no serious problems as some hotels had already begun operating. However, there was a problem with the poor management of rubbish and inadequate space for a rubbish dump for its disposal. They recommended the introduction of a new rubbish management system to keep the Gili islands clean.

Conclusion

The island's rather chaotic and haphazard development left it vulnerable to tectonic hazards. The earthquake meant that a spotlight had been shone on the island resulting in a greater scrutiny of Trawangan and its tourist industry. As a result, not only is it being made more resilient, but it is also being modernised, possibly taking away from what attracted the first visitors back in the 1980s.

Review question

15 To what extent are the developments taking place in Gili Trewangan after the earthquake sustainable?

Further reading

Bankoff, G. (2015) *Culture of Disaster: Society and Natural Hazard in the Philippines* (Routledge)

Bryant E. (2004) *Natural Hazards* (Cambridge University Press)

Dunn C. and Degg M. (2012) *Top Spec Geography: Tectonic hazards* (Geographical Association)

Fury S. (2019) *The Disaster Survival Handbook: The Disaster Preparedness Handbook for Man-Made and Natural Disasters (Escape, Evasion and Survival)* (Survival Fitness Plan)

Keller E.A. and DeVecchio D.E. (2019) *Natural Hazards: Earth's Processes as Hazards, Disasters, and Catastrophes* (Routledge)

Livermore R. (2018) *The Tectonic Plates are Moving!* (Oxford University Press)

Rothery D. (2010) *Volcanoes, Earthquakes and Tsunamis* (Teach Yourself Series)

Skinner M. (2003) *Hazards* (part of the *Access to Geography* series) (Hodder & Stoughton)

Warn S. & Holmes D. (2008) *AS/A2 Geography Contemporary Case Studies: Natural Hazards and Disasters* (Philip Allan)

Eyjafjallajökull (Iceland) and Mount Ontake (Japan) – Resources on both are available from the Geographical Association (GA)

Question practice

A-level questions

1 In the context of natural hazards, outline the differences between primary and secondary impacts. [4 marks]

2 Using Figure 5.29 (page 213) only, describe the global distribution and seasonality of tropical revolving storms. [6 marks]

3 To what extent is the changing nature of storm hazards a result of the changing carbon cycle? [9 marks]

4 Study Figure 5.3 (page 189) showing the disaster/risk management cycle. Using examples, analyse the extent to which this model can be applied to all natural hazards. [9 marks]

5 'The mitigation of earthquake risk requires the knowledge of hazards to be delivered to the people that need it.'

Using evidence, assess the extent to which you agree with this statement. [20 marks]

Chapter 6

Ecosystems under stress

This section focuses on the biosphere and in particular the nature and functioning of ecosystems and their relationships to the nature and intensity of human activities. You have to study the impact of population growth and economic development on ecosystems at various scales; this affords you the opportunity to engage with fundamental contemporary people–environment issues including those relating to biodiversity and sustainability.

There are also opportunities to consider the synoptic links between parts of this topic and many other aspects of geography.

In this chapter you will study:

- ecosystems and processes, including ecosystems in the British Isles over time and local ecosystems
- ecosystems and sustainability
- marine ecosystems
- biomes.

Key terms

Biodiversity – The variation of life forms within a given ecosystem, biome or for the entire Earth. Biodiversity is often used as a measure of the health of biological systems.

Biomass – The mass of living biological organisms in a given area or ecosystem at a given time. Biomass can refer to species biomass, which is the mass of one or more species, or to community biomass, which is the mass of all species in the community.

Biome – A major habitat category, based on distinct plant assemblages which depend on particular temperature and rainfall patterns, for example, tundra, temperate forest and rainforest.

Climatic climax – A biological community of plants and animals which has reached a state of dynamic equilibrium through ecological succession.

Ecosystem – A system in which organisms interact with each other and the environment.

Fauna – The animals of a particular region, habitat or geological period.

Flora – The plants of a particular region, habitat or geological period.

Food chain – An arrangement of the organisms of an ecological community according to the order in which they eat each other and transfer energy.

Food web – A scheme of feeding relationships, resembling a web, which unites the member species of a biological community.

Net primary production (NPP) – The rate at which an ecosystem accumulates energy or biomass, minus the energy it uses for the process of respiration. This typically corresponds to the rate of photosynthesis minus respiration by the photosynthesisers.

Nutrient cycling – The movement of nutrients in the ecosystem between the three major stores of soil, biomass and litter.

Plagioclimax – The plant community that exists when human interference prevents the climatic climax vegetation being reached.

Seral stage – An individual stage within a sere, for example, colonisation or stabilisation.

Sere – The entire sequence of stages in a plant succession.

Sub-climax – The development of an ecological community to a stage short of the expected climax because of some factor, such as repeated fires in a forest, which arrests the normal succession.

Succession – the series of changes in an ecological community that occur over time.

Trophic level – An organism's position in the food chain. Level 1 is formed of autotrophs which produce their own food. Level 2 consume level 1 and level 3 consume level 2, etc.

6.1 Ecosystems and processes

An **ecosystem** is a community of living and non-living things that work together (Figure 6.1). Ecosystems have no particular size; they can be as large as a rainforest or a lake or as small as a tree or a puddle. The water, water temperature, plants, animals, air, light and soil all work together consuming and emitting energy as they feed, reproduce and die. Because the flow of energy in an ecosystem is uni-directional, they can be considered **open systems**. All the parts of an ecosystem work together in a balanced way.

Nature of ecosystems

Structure

Each ecosystem has two main components:

- abiotic
- biotic.

Abiotic components

Abiotic factors are all of the non-living things in an ecosystem. They influence the structure, distribution, behaviour and inter-relationships of all the organisms in that ecosystem. Abiotic factors vary among different ecosystems. For example, abiotic factors found in aquatic ecosystems may be things like water depth, pH, sunlight, turbidity (amount of water cloudiness), salinity (salt concentration), available nutrients (nitrogen, phosphorous, etc.) and dissolved oxygen. Abiotic variables found in terrestrial ecosystems include things like rain, wind, temperature, altitude, soil, pollution, nutrients, pH, types of soil and sunlight.

- **Climatic factors:** These include precipitation, temperature, sunlight, wind, humidity, etc. Precipitation varies by latitude, continentality, topographical position (for example, rain shadow) and altitude. Some places have year-round precipitation while others have seasonal drought or very little precipitation at all. Temperature varies by latitude: locations near the Equator are warmer than locations near the poles or the temperate zones. Continental areas have a wider range of temperatures than coastal (maritime) locations. Humidity influences the amount of water and moisture in the air and soil, which, in turn, affects rainfall. Wind can bring precipitation or can have a drying effect depending on the humidity of the air.
- **Topography** is the layout of the land in terms of elevation, gradient and aspect. For example, south-facing slopes in the northern hemisphere will receive more sunlight (energy) than those facing north.
- **Altitude** is the height above sea level. The characteristic life forms in mountains depend on elevation because of the changes in climate with altitude.

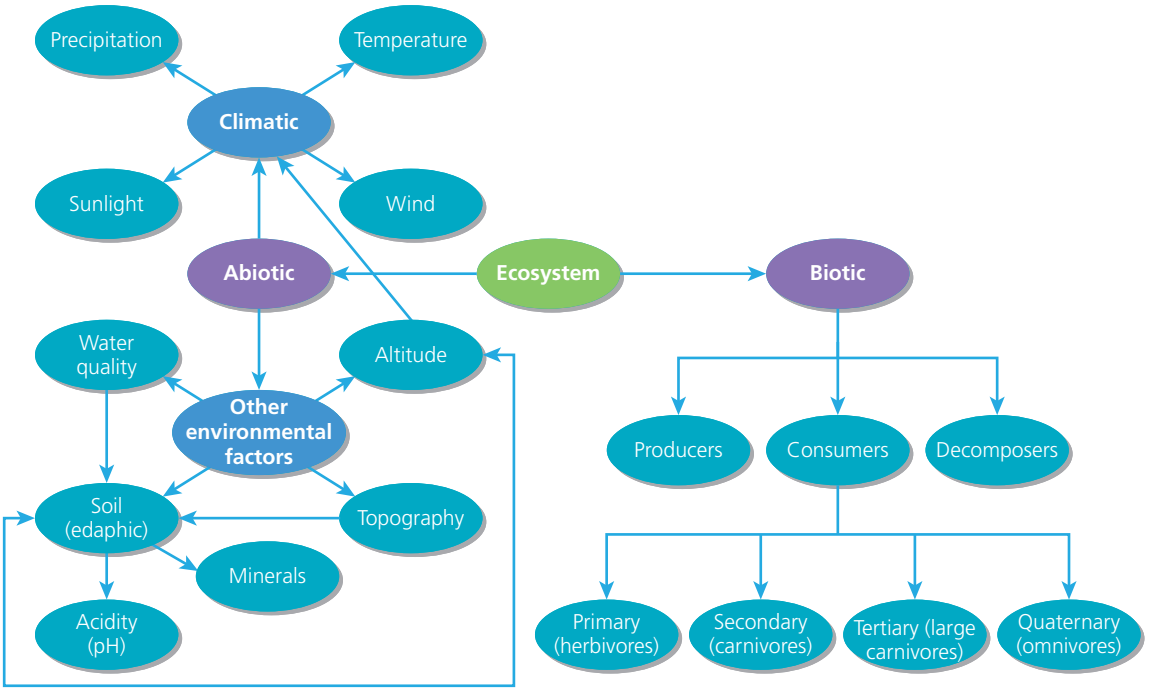


Figure 6.1 The components of an ecosystem

– Temperatures drop with altitude (approximately 1°C/100 m). Winters can be severe.
– Mountains force air masses to rise, cooling the air and eventually causing precipitation. In the lee of the mountains there is a rain shadow.
– Wind speed increases with height. Stronger winds cause physical damage; they discourage tree growth above certain elevations (the tree line). Only specifically adapted ground-hugging plants, such as grasses, sedges, mosses, shrubs and lichens can survive the extreme cold and winds of alpine zones.

This dependency causes life zones to form: bands of similar ecosystems at similar altitude.

Links

See Chapter 10, Population and the environment, pages 477–81, to understand the relationship between soils and population.

- **Soils** are rich ecosystems in themselves. They are composed of both living and non-living matter. Recent research combining above-ground and below-ground approaches has shown that vegetation–soil interactions are the driving force for many ecosystem processes and their responses to global change.

Biotic components

The biotic components are the living organisms including plants, animals and micro-organisms (bacteria and fungi) that are present in an ecosystem.

Biotic components can be classified into three groups:

- **Producers:** Green plants have chlorophyll that enables them to convert solar energy into chemical energy by photosynthesis. As the green plants manufacture their own food they are known as **autotrophs**.

  The chemical energy stored by the producers is utilised partly for their own growth and survival with the remaining stored in the plant for future use.

- **Consumers (heterotrophs):** Animals lack chlorophyll and are unable to synthesise their own food. Therefore they depend on the producers for their food.

  There are four types of consumer:

Table 6.1 The four types of consumer

Type of consumer	Characteristics	Food source
Primary	• Feed on producers • Herbivores	Rabbit, deer, goat, cattle
Secondary	• Feed on herbivores • Primary carnivores	Cats, foxes, snakes
Tertiary	• Large carnivores • Feed on secondary consumers	Lions, wolves
Quaternary	• Largest carnivores • Feed on tertiary consumers • Not eaten by other animals Omnivores	Lions, tigers

- **Decomposers** (such as bacteria and fungi): These break down the dead organic materials of producers (plants) and consumers (animals) for their food and release into the environment the simple inorganic and organic substances produced as by-products of their metabolisms.

  These simple substances are re-used by the producers, resulting in a cyclic exchange of materials between the biotic community and the abiotic environment of the ecosystem.

Energy flows

An ecosystem is an **open system**. The inputs and outputs are shown in Figure 6.2.

Organisms require energy for growth, maintenance, reproduction, locomotion, etc. Hence, for all organisms, there must be both a source and a corresponding loss of energy.

All the energy used by living things comes ultimately from the **Sun**. Energy enters living systems as a result of **photosynthesis** by plants and some bacteria. Often, less than two per cent of the incoming sunlight is captured; of this more than half of the energy captured by plants is used in respiration and is lost as heat (Figure 6.3).

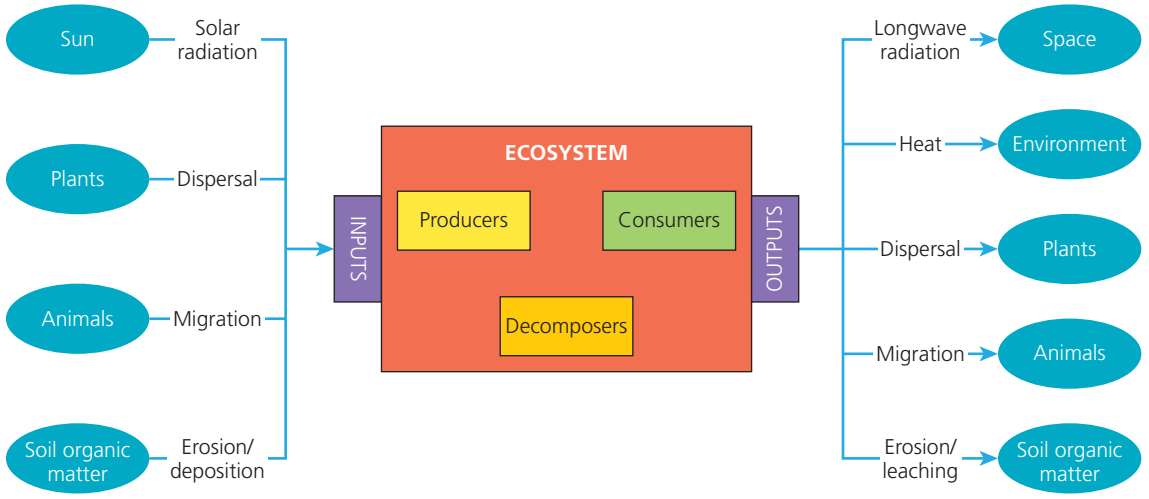


Figure 6.2 Inputs and outputs of a typical ecosystem

There are two types of organisms that have direct access to the energy in plant tissues:

- Herbivores feed on the plant while it is alive.
- Decomposers feed on the plant after it is dead.

Much of the energy in herbivore **biomass** is taken by:

- **carnivores**: these meat eaters survive mainly by eating herbivores
- decomposers.

In most ecosystems, the majority of the energy goes to the decomposers. For example, in an area of grassland only ten per cent of the energy in plants is taken by grazing animals (such as antelope). Herbivores then use almost all of their energy intake on respiration and maintaining their bodies (**maintenance**); the rest goes to herbivore biomass (the flesh and blood of the animal).

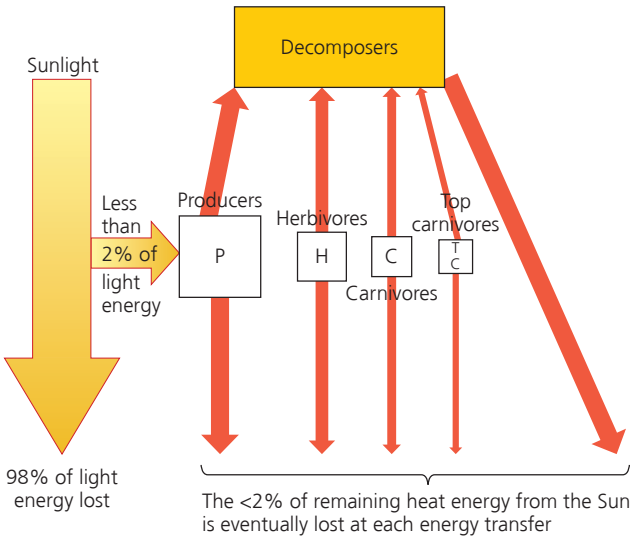


Figure 6.3 Generalised energy flow and heat loss through an ecosystem

Almost all of the energy taken in by carnivores goes to maintenance. The decomposers receive most of the plant energy and use up over half of that in maintenance. The rest may be locked up in **soil organic material** or taken by organisms that feed on decomposers. Ultimately, all of the energy originally captured by plants is transformed and lost as heat; energy is not recycled.

Biomass and net primary productivity

Biomass is a term in ecology for the mass of living organisms in a given ecosystem. Biomass can refer to:

- the mass of all living species in a given ecosystem (community biomass)
- the mass of one species within an ecosystem
- the mass of decaying material. This includes leaf litter and decaying organic material on a forest floor. It can be a substantial component of, for example, a forest ecosystem.

Phytoplankton, along with some larger rooted plants such as kelp, make up most of the plant biomass of marine aquatic systems. In the animal kingdom, small aquatic fauna such as krill comprise a disproportionate amount of the animal biomass.

In forest ecosystems trees, shrubs and forest floor detritus may comprise a considerable amount of the plant biomass. Faunal statistics show that arthropods (such as insects) comprise a very high percentage of animal biomass, with large mammals making up a relatively small percentage. In the case of organisms other than animals and plants, fungi dominate the biomass.

Plants characteristically comprise the greatest part of the biomass of terrestrial ecosystems. It is important to note that the contribution of individual species within an ecosystem to the community biomass may change considerably from season to season and even from year to year.

Biomass can be expressed as the average mass per unit area (for example, tonnes/ha or g/m^2) or simply as the total mass in the community. It can also be expressed in units of energy (for example, joules/m^2).

The product of photosynthesis is a carbohydrate, such as glucose (a sugar), and oxygen which is released into the atmosphere. The **primary productivity** of a community is the amount of biomass produced through photosynthesis per unit area and time by plants. Primary productivity is usually expressed in units of energy (for example, joules/m^2/day) or in units of dry organic matter (for example, kg/m^2/year). The total energy fixed by plants in a community through photosynthesis is referred to as **gross primary productivity** (GPP). A proportion of the energy of gross primary productivity is used by plants in a process called **respiration**. This provides a plant with the energy needed for various plant physiological and morphological activities.

Subtracting respiration from gross primary production gives **net primary productivity (NPP)**, which represents the rate of production of biomass that is available for consumption by **heterotrophic** organisms (bacteria, fungi and animals). Estimates of global terrestrial NPP range from 48–69 billion tonnes per year. This almost doubles when oceanic NPP is taken into account.

Food chains, trophic levels and food webs

As has been outlined above, the source of all food is the activity of producers or autotrophs; mainly **photosynthesis** by plants. The producers feed the primary consumers (herbivores) who in turn feed the secondary consumers (carnivores). These can be a source of food for tertiary consumers.

Such a path of food consumption is called a **food chain**. This chain follows a direct, linear pathway from one species to another (Figure 6.4). Examples of simple food chains include:

- Phytoplankton → zooplankton → fish → seal → killer whale
- Grass → grasshopper → frog → heron

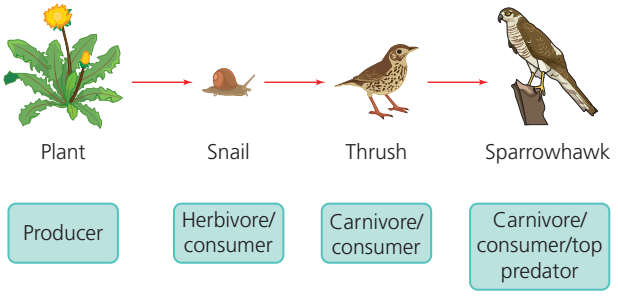


Figure 6.4 A typical food chain for a suburban garden in the UK

Each level of consumption in a food chain is called a **trophic level**.

The trophic level of an organism is the position it occupies in a food chain. The number of steps an organism is from the start of the chain is a measure of its trophic level. Food chains start at trophic level 1 with primary producers such as plants. They then move to herbivores at level 2, predators at level 3 and typically finish with carnivores or apex predators at level 4 or 5. Figure 6.5 is called a number pyramid and shows the number of organisms in each trophic level. It does not take into consideration the size of the organisms and it over-emphasises the importance of small organisms. In a pyramid of numbers the higher the trophic level the fewer organisms there are. If the biomass of the organisms at each level were taken into account rather than numbers then the mass of larger (but fewer) organisms (for example, a tree) would be better represented. This is called a **pyramid of biomass**.

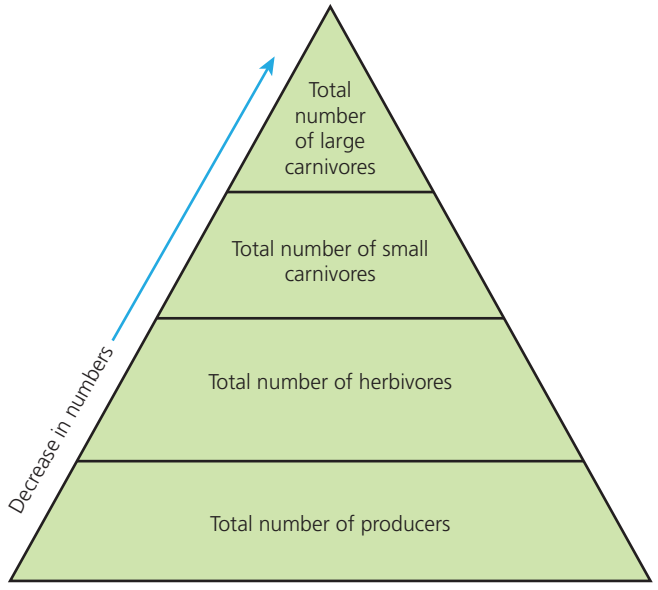


Figure 6.5 A number pyramid showing trophic levels

Most food chains are interconnected. Animals typically consume a varied diet and, in turn, serve as food for a variety of other creatures that prey on them. These interconnections create **food webs**.

The food web in Figure 6.6 only contains 25 species. Most food webs are much more complicated than this, numbering in the hundreds to thousands of species.

Review question

1 How can a systems approach to the study of ecosystems explain the role organisms play in the environment?

Skills focus

Describe the food web shown in Figure 6.6.

Succession

Ecological succession is the gradual process by which ecosystems change and develop over time. Nothing remains the same and habitats are constantly changing. There are two main types of succession:

- **Primary succession** is the series of plant community changes which occurs on an entirely new habitat which has never been colonised before. For example, a recent lava flow.
- **Secondary succession** is the series of plant community changes which takes place on a previously colonised but disturbed or damaged habitat. For example, after felling trees in a woodland, land clearance or a fire.

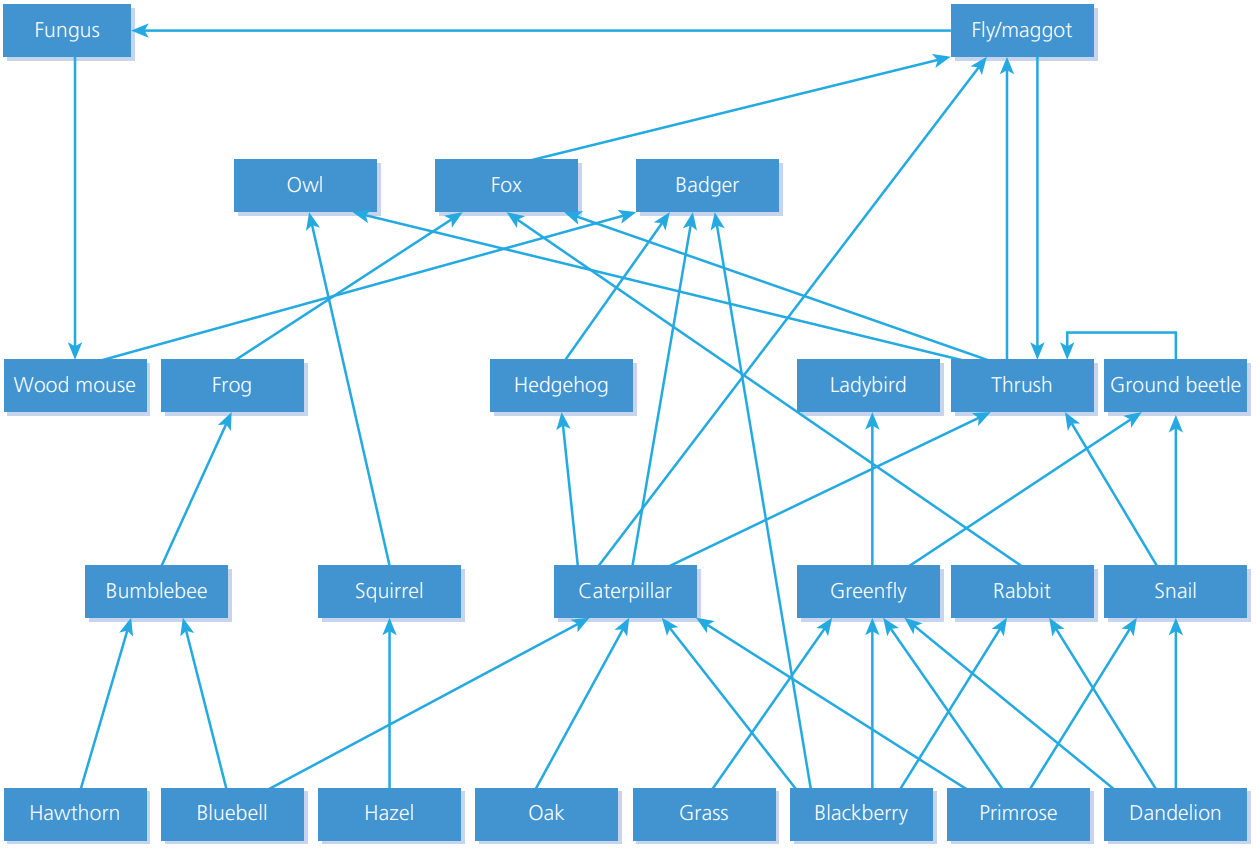


Figure 6.6 A simplified food web for an English woodland

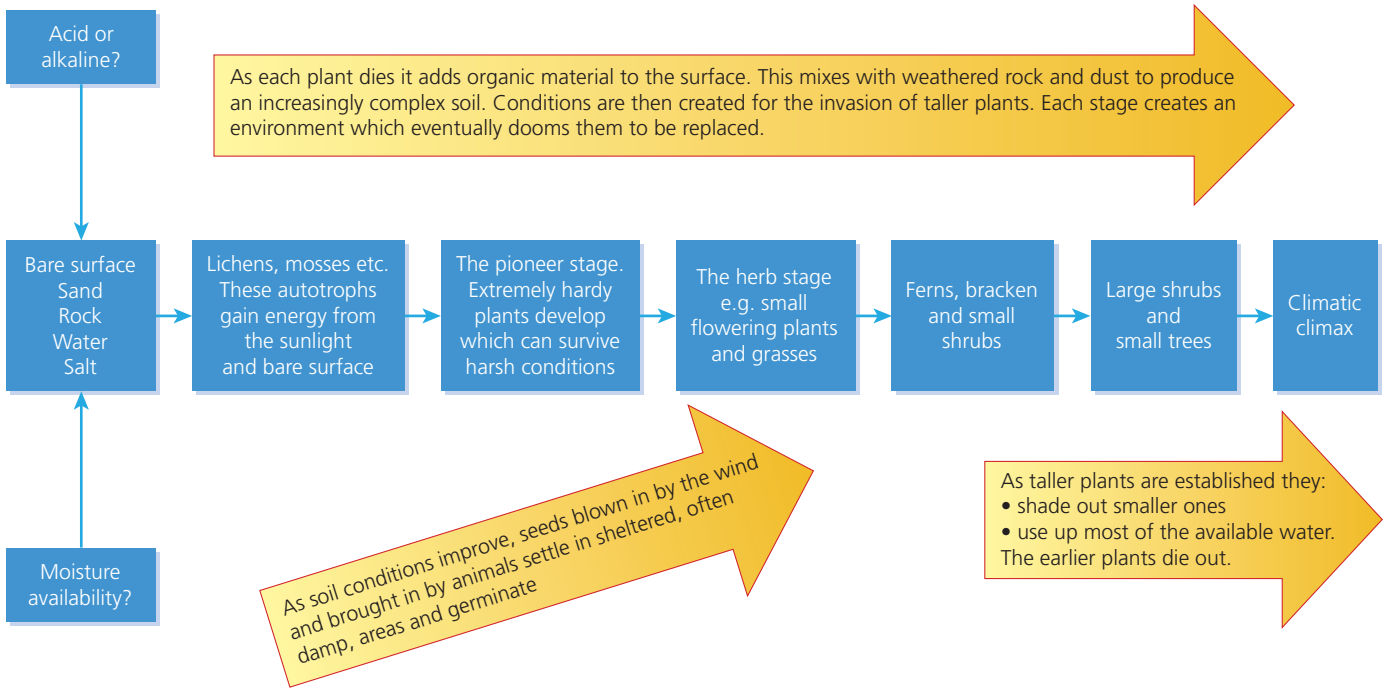


Figure 6.7 A generalised model of succession

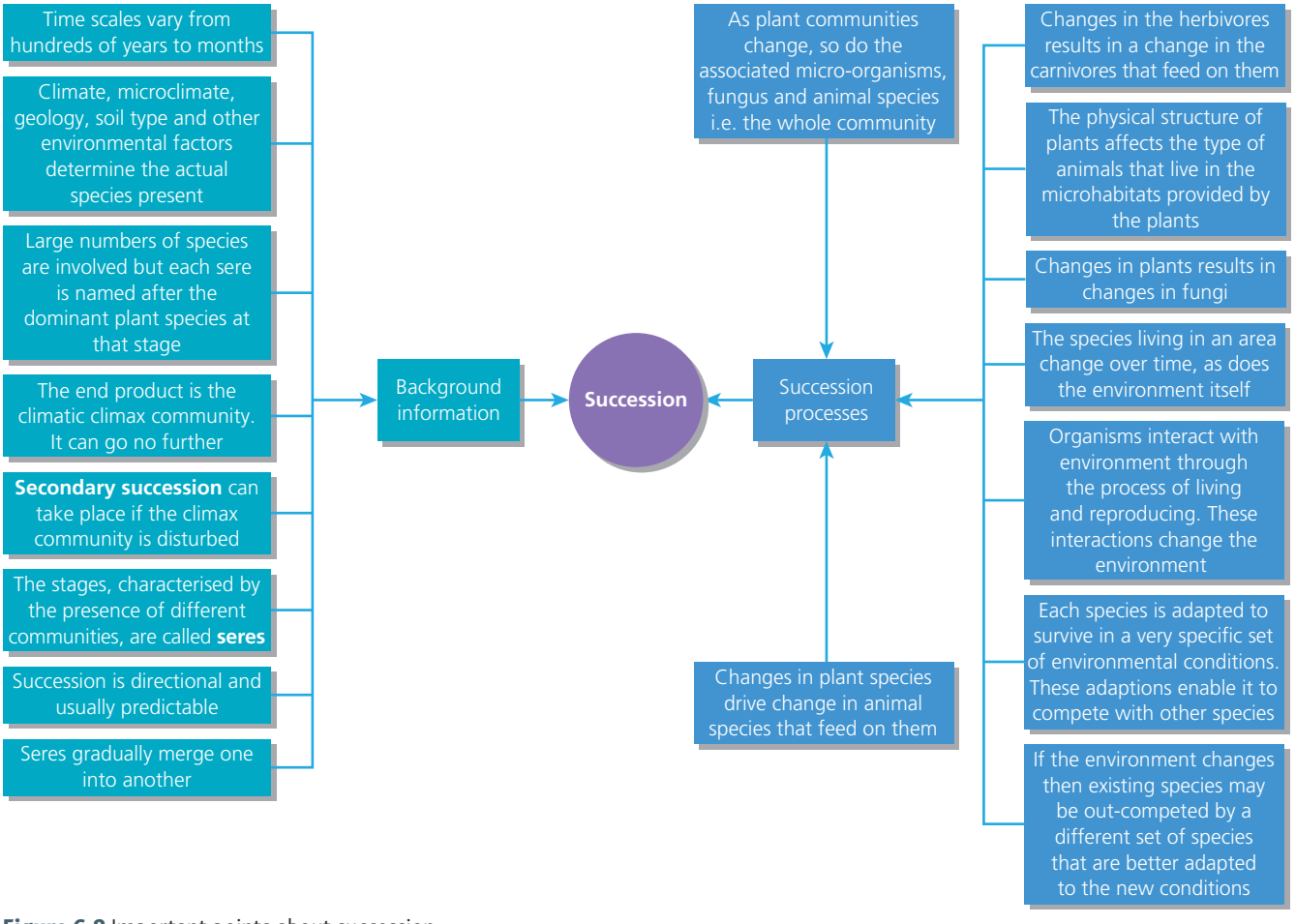


Figure 6.8 Important points about succession

Succession is the term used to signify the changes in the composition of a community of plants and animals over time (Figure 6.7 and Figure 6.8, page 236). These changes result in an increase in the complexity of the structure and species composition (**species diversity**) of such a community. Plants invade an area when the conditions are suitable and then die off when the conditions are unfavourable local conditions. This means that there is a change in the dominant plant species with time.

Successions can start on a variety of surfaces and their names reflect this:

- **Lithosere**: A succession that starts on a newly exposed bare rocky surface, for example, newly erupted volcanic lava or rock exposed as an ice sheet retreats.
- **Psammosere**: A succession that starts on bare sand, for example, coastal sand dunes.
- **Hydrosere**: A succession that starts in fresh water, for example, a pond.
- **Halosere**: A succession that starts in salt water conditions, for example, a salt marsh.

Primary succession or **prisere** develops by the gradual colonisation of a lifeless abiotic surface. A **succession** goes through several stages; the entire sequence of stages is called a **sere** and the stages are referred to as **seral stages.** As succession occurs each new stage:

- is structurally more complex than the one it replaces
- contains more biomass
- has a change in the number of species
- has greater net primary productivity
- has greater flows of nutrients and energy.

The species living in a particular place gradually change over time but so does the physical and chemical environment within that area.

An example of a lithosere is shown in Table 6.2.

Table 6.2 A lithosere in the UK

Bare rock surface	This is initially colonised by bacteria and single-celled photosynthesisers that are able to survive on few nutrients and get most of their energy from the Sun. The surface conditions are often dry and the soil little more than particles of weathered rock
Seral stage 1 Colonisation	The first plant species to colonise an area are called pioneers. These are lichens that are adapted to the severe (dry, windy, soil-free) conditions (Figure 6.9). They begin to break up the rock to form a thin layer of proto-soil. As they die they add dead organic matter to weathered rock and windblown dust. This creates a simple soil which improves water retention. Mosses are then able to develop
Seral stage 2 Establishment	As the soil develops further, ferns and small herbaceous plants and grasses begin to grow. **Species diversity** increases. There are more invertebrates living in the soil, which increases the organic content. This enables the soil to hold more water
Seral stage 3 Competition	Larger plants begin to establish themselves. These include shrubs and small trees. They use up a lot of available water and shade the ground. Some of the earlier colonisers are unable to compete and die out. Herbivores become established and predators begin to move in
Seral stage 4 Stabilisation	Fewer new species colonise. Complex food webs develop. This stage is dominated by larger, fast-growing trees such as birch and rowan. Top predators are found at this stage
Seral climax	This is the final seral stage. It represents the maximum possible development that a community can reach under the prevailing climatic (temperature, light and rainfall) conditions. This is called a **climatic climax** community. The total number of larger plant species falls as a few large species dominate the area. In the case of southern England, these are broad-leaved deciduous trees. For example, oak and ash

Figure 6.9 Lichen colonising bare rock in sub-arctic Canada

Fieldwork opportunities

Succession studies provide an opportunity for fieldwork. This could be carried out on a piece of wasteland in an urban area.

Climatic climax, sub-climax and plagioclimax

A convenient concept to describe the end product of succession is the idea of a **climatic climax** community (Figure 6.10, page 238). This is the final stage in succession. It is an association of plants and

animals which has attained a state of equilibrium within a given area.

Many ecologists believe however that the time required to achieve this state is unrealistically long. In most cases, external disturbances and environmental change occur so frequently that the realisation of a climax community is unlikely and therefore it has come to be regarded as a less useful concept. Nevertheless many authors and nature-enthusiasts continue to use the term 'climax' to refer to what might otherwise be called mature or old-growth communities.

The British Isles is in a temperate deciduous forest **biome**, so broad-leaved deciduous forest, dominated by oak and ash is the *normal* **climatic climax community for much of England and Wales**. At a more local scale however, the situation is far more complex. Heather moorland, coniferous forest, marshland and alpine (montane) communities are widely distributed, occurring wherever local **soil**, drainage or relief factors restrict the growth of deciduous forest.

It should also be noted that human activity over thousands of years has cleared almost all of the UK's **primary** forest, replacing it with agriculture, plantations and other **secondary growth**.

At any point in the development of an ecosystem, succession can be halted by an arresting factor. This can occur suddenly or gradually and can have a permanent or temporary effect. The result is a **sub-climax** community where the climatic climax has not been reached. Table 6.3 summarises these arresting factors.

Table 6.3 Arresting factors in ecological succession

Arresting factor	Possible cause(s)
Topoclimax	A change in topography. This could be a landslide, a volcanic eruption (lava or ash) or deposition of mud following a river flood
Hydroclimax	A change in drainage. This could be caused by a raised water table following increased precipitation
Biotic climax	The introduction of an alien species to an area. For example, rabbits are the most significant known factor in species loss in Australia
Plagioclimax	An environment maintained by management. This could be forest clearance in Amazonia for ranching, for example. It includes all agriculture, sports fields, parks, gardens, etc.

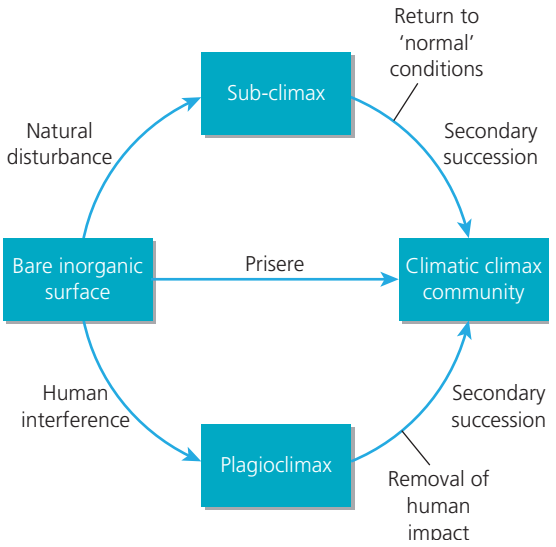


Figure 6.10 Relationship between sub-climax, plagioclimax and secondary succession

Plagioclimax communities are determined by planned or unplanned human activity (Figure 6.11). The process of succession has either been halted before reaching climatic climax or it has been deflected to a different climax. This could be caused by:

- cutting down existing vegetation
- burning (for example, to remove forest)
- planting trees/crops
- grazing (for example, goats in the Mediterranean region)
- draining wetland.

The plagioclimax will remain as long as the human activity continues. If this activity stops then the succession will revert to its original path unless there has been a permanent change to other environmental conditions (Figure 6.12).

Figure 6.11 A golf course: an example of a plagioclimax

Figure 6.12 An abandoned golf course: an example of secondary succession

Mineral nutrient cycling

Nutrients are the chemical elements and compounds needed for organisms to grow and function. Energy flow and **nutrient cycling** are interdependent. The rate of nutrient cycling may affect the rate that energy can be trapped. For example, plants cannot make new cells (grow) if essential nutrients are absent.

Nutrients are stored in three compartments within an ecosystem:

- Soil: A mixture of weathered rock, air, water and decomposed organic matter on the surface of the Earth.
- Litter: The amount of dead organic matter on top of soil.
- Biomass: The total of plant and animal life in an ecosystem. These nutrients are cycled from one store to another as shown in Figure 6.13.

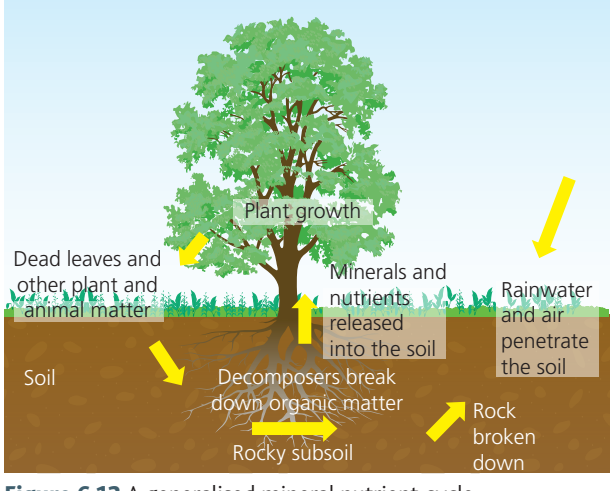


Figure 6.13 A generalised mineral nutrient cycle

All nutrient cycles can be summed up in a Gersmehl diagram (see Figure 6.14). In this model the size of the circles is in proportion to the amount of nutrients they store and the thickness of the lines is proportional to the amount of nutrient flow.

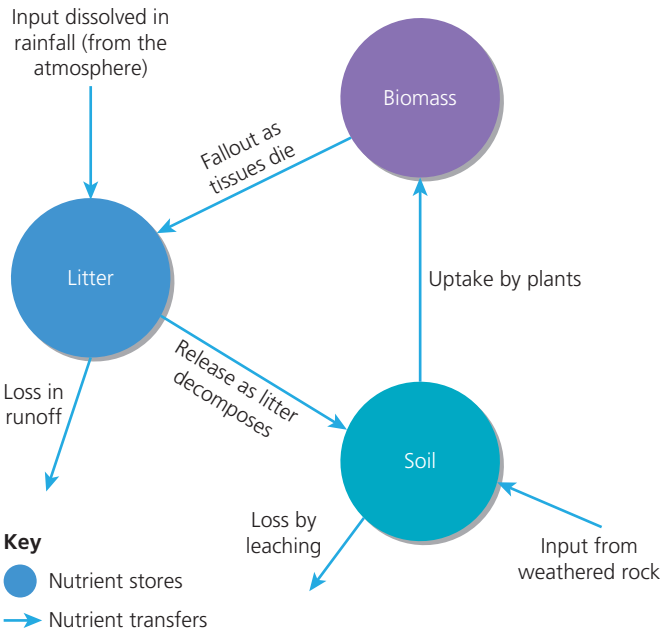


Figure 6.14 A model of the mineral nutrient cycle

Review questions

2 Why might there be seasonal changes in biomass components?

3 Under what circumstances would a lithosere develop in the UK today?

4 How might succession on an area of urban dereliction vary from that of a lithosere?

5 Why might Kinder Scout (Figure 6.33, page 255) and its blanket bog be considered a plagioclimax?

6 How will the sizes of the three nutrient stores and the nutrient transfers shown in Figure 6.14 vary between winter and summer for an area of deciduous woodland in southern Britain? Draw Gersmehl diagrams to show your answer.

Terrestrial ecosystems

The term **ecosystem** (a contraction of ecological system) is generally understood to be the entire assemblage of organisms (plant, animal and other living beings) living together in a certain space with their environment, functioning as a loose unit. Together, these components and their interactions with, and relationships to, each other form a dynamic and complex new whole, functioning as an 'ecological unit', with additional characteristics that cannot be found in the individual components. Nor could any organism live completely on its own without involving any other species of organism. A terrestrial ecosystem is one found only on landforms, that is, all ecosystems that are not aquatic

or marine. They range in scale from sand grains to continents, though the latter tend to be called **biomes.**

Example of a terrestrial ecosystem in the UK: The chalk downlands of the South Downs

The South Downs is a chalk ridge that extends from near Winchester to the high sea cliffs of Beachy Head 100 km to the east. The ridge is an escarpment that reaches heights of over 200 m. The most dramatic feature of the South Downs is the steep, mostly northerly-facing scarp slope (Figure 6.15) that is broken only by the gaps created by the rivers Meon, Arun, Adur, Ouse and Cuckmere. The gently sloping south-facing dip slope supports well-drained, easily worked soils.

The South Downs is comprised of tilted layers of relatively soft chalk containing bands of flint nodules. Both the chalk and the flint are the remains of animals and plants in the Cretaceous sea over 100 million years ago. The soils derived from the chalk of the South Downs are mostly thin, well-drained and poor in several minerals and nutrients.

The South Downs supports significant areas of unimproved calcareous grassland, with extensive areas occurring along the north-facing escarpment, and more widely scattered fragments on the south-facing, shallow sloping, predominantly arable dip slope.

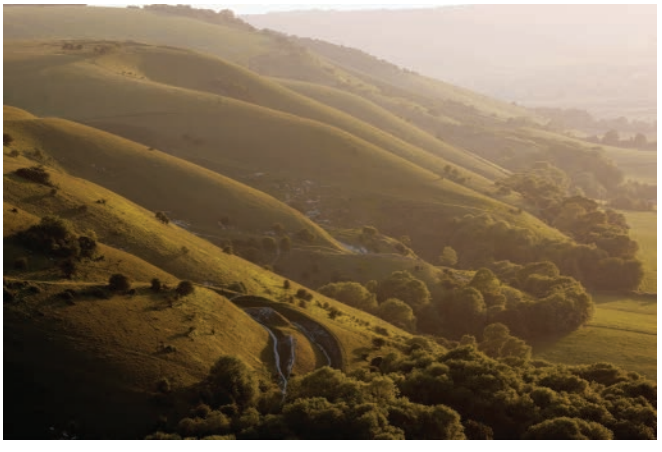


Figure 6.15 Looking west along the north-facing scarp slope of the South Downs at Fulking

Abiotic factors

Climate

This area of southern England combines the highest average daytime temperatures found in the British Isles with the highest sunshine averages on the British mainland. Rainfall is below the UK average with 950 mm/yr. The warm, dry dip slope and cooler, damper scarp provide a range of conditions for wildlife. Species at the northern extent of their distribution occur in the South Downs.

Topography

Steep slopes on the scarp prevented ploughing and today support most of the relict ancient woodland and chalk grassland. Gently sloping, south-facing dip slopes have hot, dry conditions which favour growing cereal crops. Very little semi-natural habitat remains on the dip slope. The gaps in the ridge created by southward flowing rivers are narrow, steep sided and flat floored.

Geology, soils and drainage

The underlying chalk gives rise to thin, infertile, well-drained soils. These lack many minerals, especially potassium and phosphates, though they are of course rich in calcium. This soil type, known as **rendzina**, only allows slow rates of plant growth. It is largely responsible for the diversity of small, low-growing herbs able to co-exist in downland turf.

Pockets of deeper soils are present on the plateau, dip slope and valley bottoms. These more fertile areas are either cultivated or support woodland. In a few places a thin covering of wind-blown soil called 'loess' remains over the chalk. This is thought to derive from dry conditions during the last Ice Age. It produces a very unusual plant community known as 'chalk heath'. The soils of the river floodplains that cut through the chalk downs derive from water-borne silt and are highly fertile.

The poorly drained areas of the floodplains were traditionally used as grazing marsh, but with increased drainage has come conversion to arable use.

Biotic factors

There are many different **habitats** (that is, places where living things live) and these are summarised in Figure 6.16.

Two of these habitats are described below:

- **Chalk heath** is an extremely rare habitat found on the South Downs in a few, scattered locations. It occurs where deposits of acidic soil created by windblown deposits overlie the chalk. Characterised by acid-loving species such as heathers growing alongside typical chalk-loving plants, chalk heath was probably much more widespread in the past. Its vulnerability to disturbance of the soil and scrub encroachment has

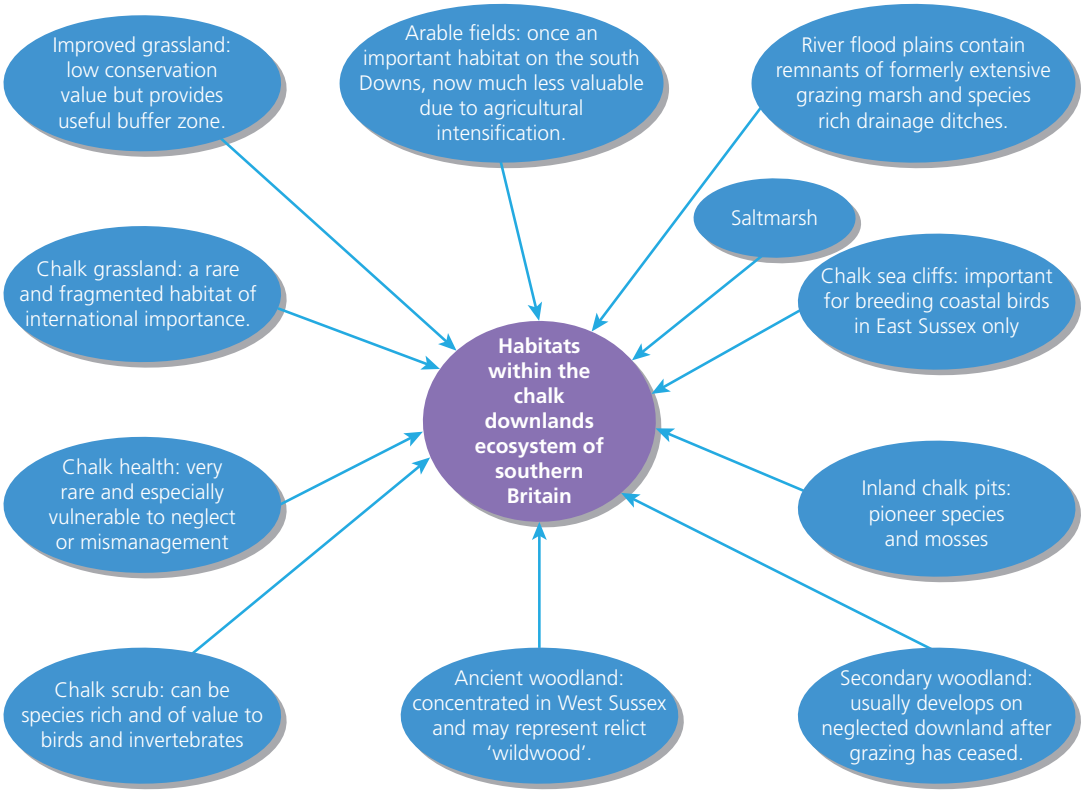


Figure 6.16 The range of habitats found in the chalk downlands of southern Britain

allowed this habitat to dwindle in size to present day minute levels. Lullington Heath National Nature Reserve has what is regarded as the best example of chalk heath in Britain, but there are other chalk heath sites within the South Downs.

- **Unimproved chalk grassland** is typically species rich (with up to 56 species in one square metre) and is maintained by sheep or rabbit grazing. Once covering extensive areas of the South Downs, mainly in East Sussex, this habitat is now scarce, fragmented and under threat from changing patterns of land use. Legal protection for the remaining areas of unimproved chalk grassland is increasing by designating some locations as **Sites of Special Scientific Interest (SSSI)** and others to Annex 1 of the EC Habitats Directive. The best examples are now under consideration as **Special Areas of Conservation**.

Factors influencing change in the downland area

- Many of the habitats within the South Downs National Park have become highly **fragmented** in recent years. In particular, chalk grassland, chalk heath and unimproved grazing marsh mostly occur as isolated patches within agricultural and/or unmanaged land. The implications for wildlife of fragmentation are numerous and include increased risk of local extinctions. On small, isolated sites species may not be able to sustain themselves without appropriate management. Chalk grassland fragments, for example, are vulnerable to scrub invasion.
- Increased edge to area ratio. When a habitat is next to intensively worked farmland, the edge effects of spray drift (pesticide or fertiliser) and run-off can soon cause an erosion of the high value habitat from the edges inwards.
- Species movement between small, isolated patches of available habitat can be severely restricted, especially for many of the characteristic chalk grassland plants which do not produce wind-dispersed seeds. The same applies to the many invertebrates which will not willingly cross open arable fields, like molluscs and some butterflies.
- The myxomatosis epidemic of the mid 1950s onwards showed that rabbit grazing is vital for the maintenance of many chalk habitats. Although some fluctuation still occurs in rabbit populations, they have largely recovered to the extent that they are once again a major source of grazing pressure on the South Downs.

- Leisure use of the South Downs is increasing with activities such as hang gliding, mountain biking and four-wheel drive rallies, all of which can disturb rare species and damage the grass cover, leading to erosion of the thin soil.

Review question

7 The South Downs have recently been designated as a national park. How does this help in protecting the downland ecosystem?

Global changes influencing ecosystems

Climate change

> The natural environment underpins all aspects of our lives. It will be affected by climate change, and yet we will be increasingly reliant on it to help us manage the impacts that a changing climate will bring.
>
> Defra, 2010

Ecosystems are highly vulnerable to climate change; this can aggravate existing stress factors such as pollution, land conversion and invasive non-native species. While some species could benefit from climate change, far more are set to lose out, according to UK government estimates.

The UK may see changing patterns of wildlife and plants as species try to adapt by moving northwards, or have to compete with new non-native species. Habitats may come under increasing pressure – from salt marshes threatened by sea-level rise to beech woodland susceptible to summer droughts. Species could also experience reduced food supply if earlier breeding periods are at odds with the food available at the time.

Figure 6.17 summarises the UK government view of what the impacts could be.

On a global scale, atmospheric warming is likely to be the greatest cause of species extinctions for millions

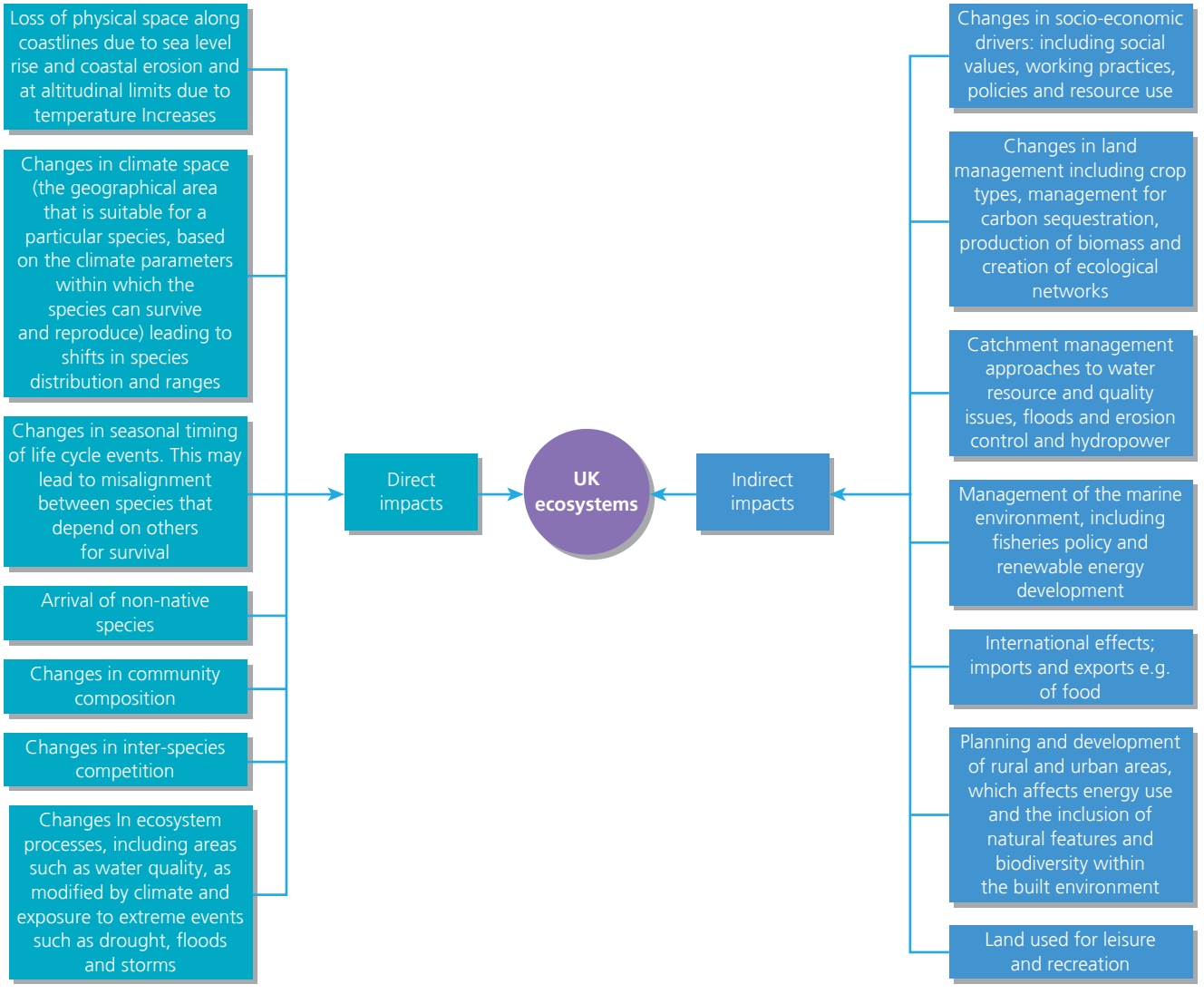


Figure 6.17 Direct and indirect impacts of climate change on UK ecosystems

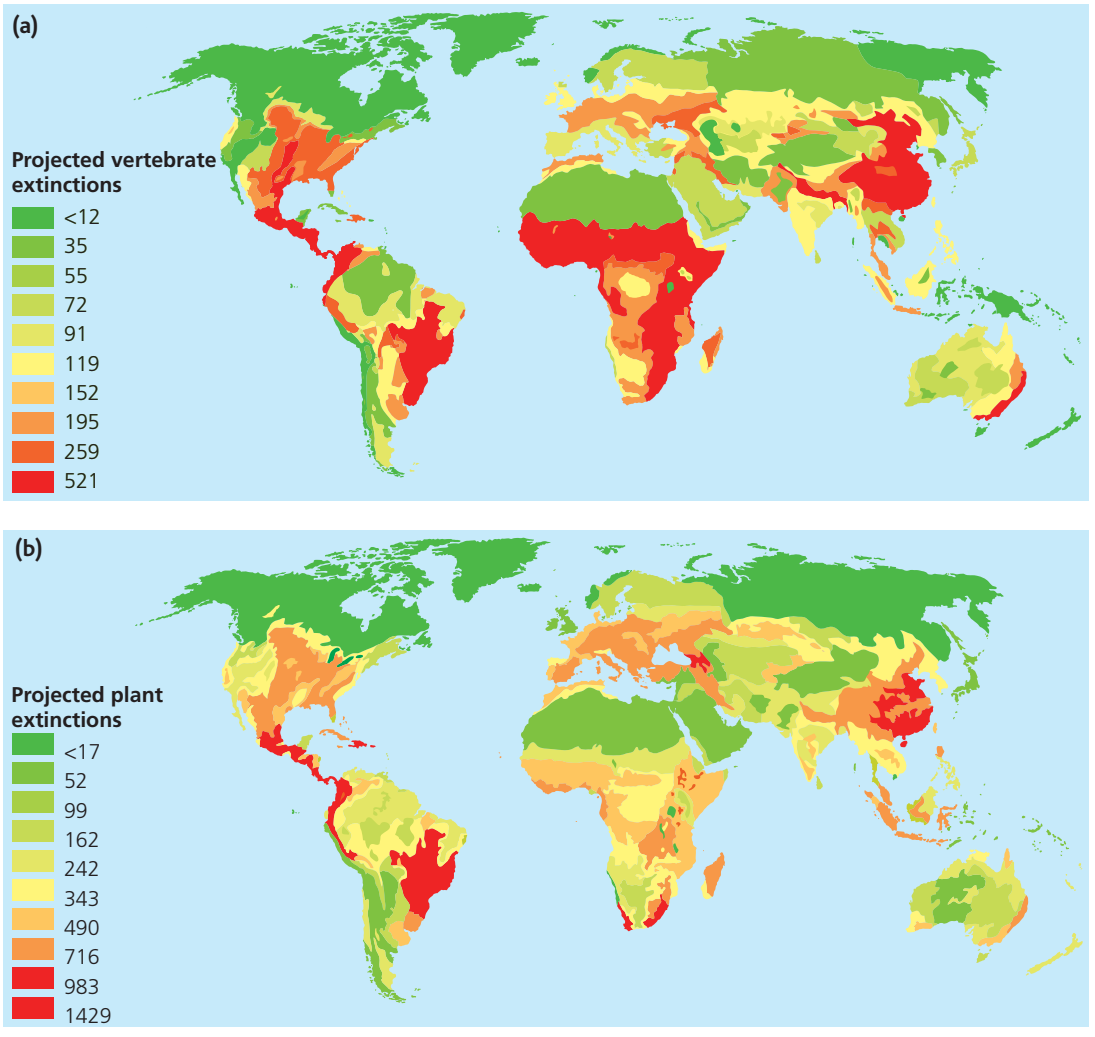


Figure 6.18 **(a)** Projected combined species loss for mammals, birds, amphibians and reptiles; **(b)** Projected plant species loss

Source: Chaudhary & Brooks (2018) Environmental Science and Technology

of years. The **Intergovernmental Panel on Climate Change (IPCC)** says that rising temperatures (along with habitat destruction and pollution) will put many species at risk (Figure 6.18). Predictions of species extinction vary widely. One conclusion, based on the results of many smaller published studies, is that extinction risks will accelerate as global temperatures rise, threatening up to one in six species.

These projected rates of species extinctions are 10 times greater than recently observed global average rates and 10,000 times greater than rates observed in the distant past (as recorded in fossils). Examples of species that are particularly climate sensitive and could be at risk of significant losses include animals that are adapted to mountain environments, such as the pika (Figure 6.19) and animals that are dependent on sea ice habitats, such as polar bears.

Review question

8 Study Figure 6.18. Why are the numbers of species lost in polar regions so low when the animals and plants that are there are so vulnerable to change?

Figure 6.19 A pika in the Canadian Rockies

Human exploitation of the global environment

According to the UN Department of Economic and Social Affairs, from 1960 to 2020 the global population has increased by over 150 per cent, going from 3 billion in 1960 to nearly 7.7 billion in 2020. It is expected to reach 10 billion by 2050. Low-income countries have accounted for most population growth in the past 60 years, but this is beginning to change. Some high-income countries such as the USA are still experiencing high rates of population growth, while some developing countries such as China, Thailand, North Korea and South Korea have very low rates.

According to the World Bank, Global economic activity increased by just over 800 per cent between 1960 and 2020, despite the downturn in 2006. Average income per person has quadrupled during this period. As per capita income grows consumption changes, with wide-ranging potential for effects on ecosystems. The share of additional income spent on food declines and the consumption of goods and services rises. Dietary composition changes, with less consumption of starchy staples (rice, wheat, potatoes) and more of fat, meat, fish, fruits and vegetables.

The impact of science and technology on ecosystems is most evident in the case of food production. Much of the increase in agricultural output over the past 40 years has come from an increase in yields per hectare rather than an expansion of the area under cultivation. For instance, wheat yields rose over 200 per cent, rice yields rose over 100 per cent and maize yields rose nearly 160 per cent in the 40 years leading to 2015. At the same time there were some unintended effects, such as **eutrophication** (see Figure 6.22, page 246) of freshwater systems and hypoxia in coastal marine ecosystems, both of which result from excessive application of inorganic fertilisers. Advances in fishing technologies have contributed significantly to the depletion of marine fish stocks.

For terrestrial ecosystems, the most important direct drivers of change in ecosystems have been:

- land-use change, mainly conversion of natural vegetation to growing crops. At present, some 11 per cent of the globe's land surface is used in crop production – approximately 36 per cent of the land estimated to be suitable for crop production.
- the use of new technologies which have contributed significantly to farmland and timber stock exploitation.
- deforestation and forest degradation affect almost 10 per cent of the world's remaining forests, amounting to an annual loss of 7.6 million hectares. This loss is of particular concern in the tropical rainforests (17 per cent of the Amazon rainforest has been lost in the last 50 years). Data on the loss of boreal forests is limited, but it is believed to be significant. Approximately 10 per cent of the arid lands of the world are considered degraded, with the majority of these areas being in Asia.

Freshwater ecosystems have been affected by water abstraction, invasive species and pollution, particularly high levels of fertiliser. The introduction of non-native invasive species is one of the major causes of species extinction in freshwater systems. For example, after the American signal crayfish was introduced to Britain, it reduced the now-endangered native white-clawed crayfish population by more than 50 per cent.

Local ecosystems

Ponds

Figure 6.20 A pond in rural Lancashire

A pond is an example of a distinctive local ecosystem (Figure 6.20). In the UK, the Department for Environment, Food and Rural Affairs' (Defra) **Biodiversity Action Plan (BAP)** defined them as small, permanent or seasonal water bodies of up to two hectares in size. They are depressions where water collects, and are often quiet environments with a shallow depth. This shallowness allows the sun to penetrate to the bottom, which in turn allows plants to grow. Pond plants (**flora**) either grow entirely underwater or partially on the surface. Some plants such as reeds and bulrushes will also grow along the

pond's edge. Ponds in the UK support a large variety of animal (**fauna**) and plant life, such as birds, small fish, amphibians (frogs, toads and newts), insects, protozoa, algae, fungi and lily pads.

A pond's ecosystem, like all others, consists of abiotic environmental factors and biotic communities of organisms. Abiotic environmental factors of a pond's ecosystem include:

- temperature
- water flow
- salinity
- the percentage of dissolved oxygen.

A water body's salinity may determine the different species present. For instance, marine organisms tolerate salinity, while freshwater organisms will not thrive when exposed to salt. In fact, freshwater ecosystems often have plant species present which will absorb salts that are dangerous for freshwater organisms.

A pond ecosystem consists of four **habitats**:

- **The shore**: The nature of the shore determines the type of organisms found there. Rocky shorelines may not allow plants to grow, while muddy or sandy shorelines attract grasses, algae, earthworms, snails, protozoa, insects, small fish and micro-organisms.
- **The surface film**: The pond's surface is inhabited by pond skaters and free-floating organisms.
- **Open water**: An open-water habitat permits sizeable fish (carp, pike), plankton, phytoplankton and zooplankton to grow. Phytoplankton includes a large variety of algae, while zooplankton refers to insect larvae, etc. Fish feed on plankton or tiny organisms. Deep water provides space for many creatures, such as the great diving beetle, a voracious predator that feeds on smaller invertebrates, tadpoles and even small fish hiding among the pond weed. Snails graze on the algae that stick to the leaves of larger plants that grow in deeper ponds.
- **The pond bottom**: This habitat varies depending upon the pond's depth. Shallow ponds (less than 30 cm) with sandy bottoms provide a nesting environment for earthworms, snails, and insects. Deeper ponds (over 30 cm) often have muddy bottoms, which allow various micro-organisms, such as flatworms and dragonfly nymphs, to reproduce and survive.

A pond's **food chain** (see Figure 6.21) has three basic **trophic levels**:

- The first trophic level represents the producers and autotrophs, such as phytoplankton and plants. Producers prepare their own food with the energy emitted from the sun through a process known as photosynthesis.
- The second trophic level consists of herbivores, such as insects, crustaceans and invertebrates that inhabit the pond and consume the plants.
- The third trophic level comprises of carnivores, such as various sizes of fish, which feed on both the plants and herbivores of the first and second trophic levels.

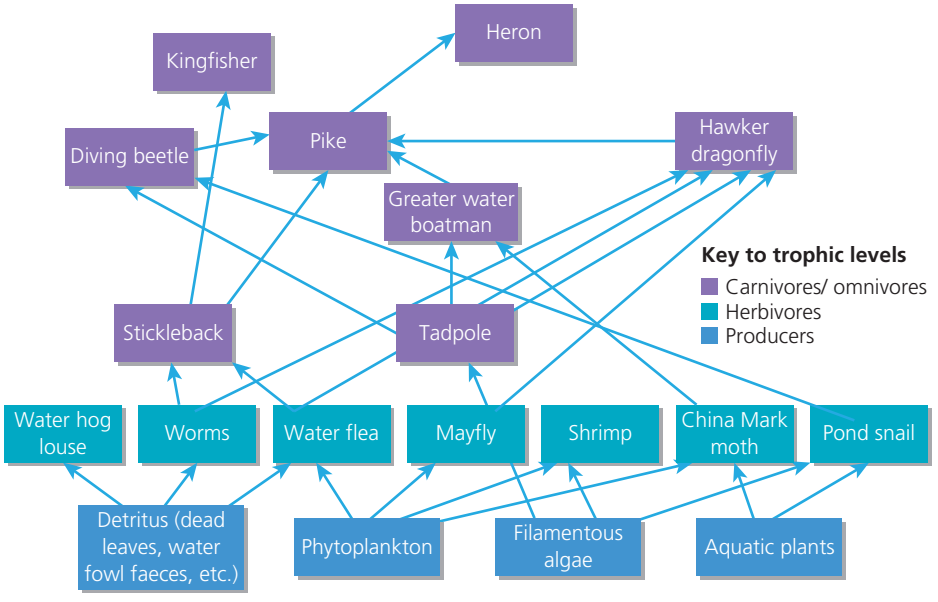


Figure 6.21 Generalised food web of a pond

Decomposers located at the bottom of the food chain help decompose dead organic matter, which further breaks down into carbon dioxide and essential nutrients, such as nitrogen, phosphorus and magnesium. These nutrients supply the necessary life force for the autotrophs, etc., to produce food for the second trophic level organisms; this results in the perpetual flow of energy in the pond's ecosystem.

The importance of ponds

Ponds support many species of plants and animals – more than 150 UK priority species are associated with them. They are particularly good for insects, such as damselflies and dragonflies.

Ponds provide important homes for amphibians including the protected great crested newt and common toad. They are also home to water voles, grass snakes and Daubenton's bats. Waterbirds such as swans, moorhens and tufted ducks rely on ponds for feeding and nesting, while waders like lapwing, redshank and snipe probe the muddy margins for invertebrates. Ponds provide stepping stones between isolated patches of habitat, linking up the countryside and allowing species to move about freely.

Recent research shows that 80 per cent of wildlife ponds in the UK are in a 'poor' or 'very poor' state and we have lost almost 500,000 in the last 100 years. This loss has a major effect on wildlife as they provide many species with a suitable breeding and feeding habitat. Ponds have been, and continue to be, lost to urban development and land-use change, agricultural drainage and in-filling, fragmentation and through poor management.

Fresh waters are adversely affected by a number of human pressures, including:

- excessive groundwater and surface water abstraction, which has caused many spring-fed ponds to dry up
- pollution, especially eutrophication. This is explained in Figure 6.22
- the effects of inappropriate land use or poor management, for example, afforestation, land drainage and overgrazing
- introduction of invasive plant and animal species, for example, the floating pennywort. This is a North American plant which was introduced to the UK in the 1980s by the aquatic nursery trade. It roots in shallow margins of slow-flowing water bodies, particularly in ditches, slow-flowing dykes and lakes. It causes de-oxygenation of the water, which in turn affects fish and invertebrate populations and causes a choking of ponds and drainage systems
- urban, industrial, and agricultural development within catchments
- inappropriate development of recreation and navigation
- inter-basin water transfer schemes and the construction of dams and reservoirs.

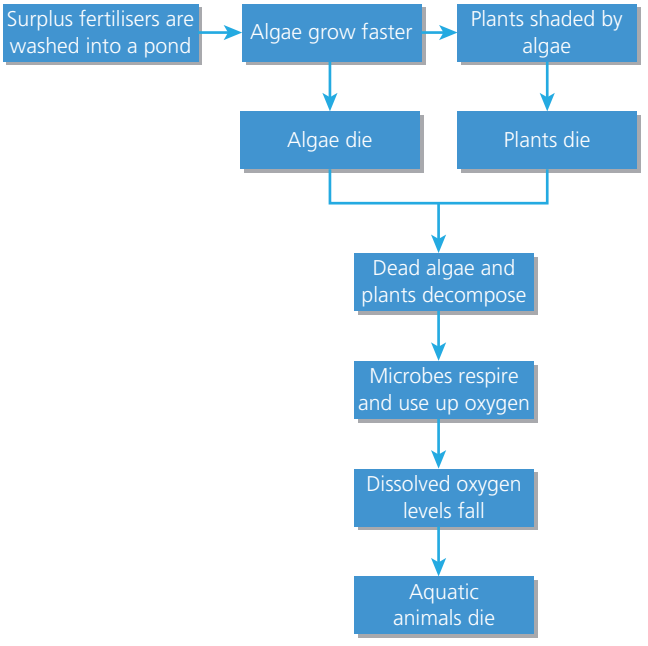


Figure 6.22 Eutrophication explained

Succession in a pond: a hydrosere

In time, an area of open freshwater such as a pond will silt up and dry out. If left alone for long enough (over 200 years) it will eventually become woodland. During this process, a range of different habitats, such as swamp and marsh, will succeed each other (Figure 6.23).

Stage 1: Open water

Deep freshwater will not support rooted, submerged plants because there is not enough light for photosynthesis at the bottom of the pond. There will be micro-organisms and plankton floating in the water.

Stage 2: Rooted plants

Over time, sediments will be transported into the pond by streams or rainwater. Large amounts of sediment can be deposited in this way. The water depth will gradually decrease, allowing rooted, submerged plants, such as starwort and pondweed, to grow. Waterlilies, which are rooted, but with floating leaves, may also become established. The added plant life is accompanied by a variety of invertebrates and fish. The

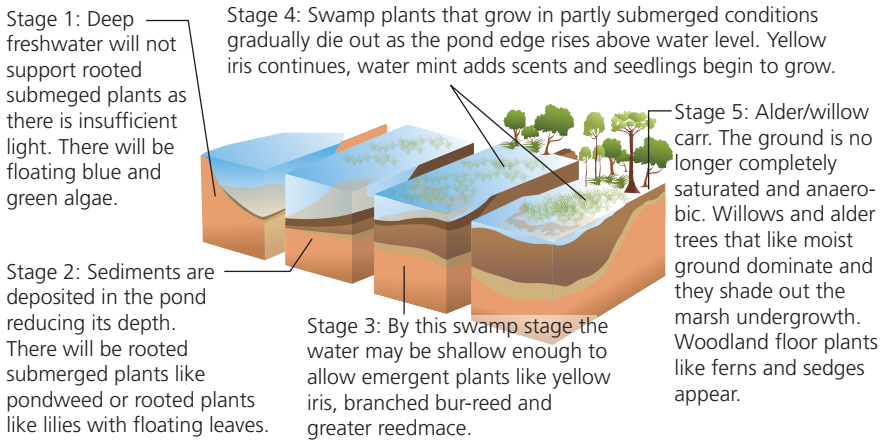


Figure 6.23 A hydrosere

large numbers of different species present give rise to complex food webs.

The vegetation traps and holds more and more of the incoming sediment so that the water becomes shallower. Decomposing dead plant and animal matter provides food for detritivores and increases the nutrients in the water. This promotes plant growth.

Stage 3: Swamp stage

By this stage, the water may be too shallow to support fully submerged plants. Instead, emergent plants, such as yellow iris grow partly in and partly out of the water. These plants tend to have tall, flexible, spear-shaped leaves. This allows them to cope with large differences in water level, always having some portion of the leaves above the water for photosynthesis.

The water in swamp areas teems with invertebrate life and provides an ideal place for amphibians to breed.

Stage 4: Marsh stage

Swamp plants which are adapted to grow in partially submerged conditions will gradually die out as the marsh floor progressively rises above the water level.

Some plants such as yellow iris, which grow equally well in swamp or marsh conditions, will continue to grow, while other marsh plants such as water mint become established. Tree seedlings, such as willow, which favour wet soil conditions, will become well established and begin to grow up.

Willow has a very high transpiration rate, transferring large quantities of water from the sediment into the atmosphere. Together with the silt-trapping effect of the marsh plants, this greatly increases the rate at which the marsh dries out.

Stage 5: Carr and woodland stages

The soil is still wet, but no longer completely waterlogged and air is able to enter along root channels. Willow and alder dominate the ground and shade out the lower-lying marsh vegetation. This is replaced by a variety of woodland floor plants including sedges, rushes, ferns and small flowering herbs which are adapted to low light levels and which flourish in wet conditions. Willows are able to support a wide range of invertebrates (more than 450 different species). This means that there is plenty of food for insectivorous birds.

In drier areas, increasingly aerobic decomposition will accelerate nutrient recycling, increasing the organic and nutrient content of the soil. Tree transpiration will continue to dry out the soil to the point where climax tree species such as oak, beech or ash can become established.

Pond conservation

One national example of pond conservation is the Million Ponds Project. This is a 50-year initiative set up by various charities and led and co-ordinated by Freshwater Habitats Trust. It aims to create an extensive network of new ponds across the UK, to reverse a century of pond loss and ensure that once again the UK has over one million countryside ponds.

Its first phase (2008–2012) aimed to create networks of clean water ponds across England and Wales. It also sought to change attitudes, so that pond creation is embedded as a routine activity in land management practices. The team worked with landowners and managers to create over 1,000 ponds for around half (49) of the 105 rare and declining pond species that are a national priority for conservation action under the UK government's Biodiversity Action Plan (BAP).

To achieve its aims, the project:

- funded new ponds and pond complexes for threatened freshwater plants and animals
- provided technical and on-the-ground support for pond creation
- engaged with and trained a wide range of stakeholders
- published best practice information in the extensive Pond Creation Toolkit, including the BAP species map
- raised the profile of ponds with policy makers and the media.

One critical element of the project is that these new ponds must have clean water. This is important because most countryside ponds are now badly damaged by pollution, and evidence shows that pond wildlife is declining across the UK.

Phase II of the Million Ponds Project is now underway in England and Wales, and was scheduled to run from 2012 to 2020, including:

- identifying Important Areas for Ponds (IAPs) nationally – these are core areas in which to protect existing ponds, and from which stepping stones can be created into the wider countryside
- developing targeted landscape-scale pond creation projects in areas like the New Forest and in more intensively managed landscapes, such as catchments in Oxfordshire and Leicestershire
- investigating the potential to further embed pond creation into national policy initiatives, such as conservation credits, which could potentially create many thousands of clean water ponds in the future
- working with academic partners to develop applied research projects.

In 2014 and 2015 the project aimed to expand the Million Ponds Project to Northern Ireland and Scotland. Discussions with Natural England indicated that clean water pond creation would be an important part of the England Biodiversity Strategy objective to ensure that 90 per cent of Priority Habitats were in good condition by 2020 – with thousands of new high-quality ponds created each year.

Case study of a specified ecosystem at a local scale: The Sefton Coast sand dunes

The Sefton Coast area is dominated by recreational beaches and one of the largest coastal dune systems in Great Britain; it extends 17 km along shore, 4 km inland, covers an area of 2,100 ha and shows dune heights reaching up to 30 m (Figure 6.24).

The sand dune area is well protected. It is a designated SSSI (Special Site of Scientific Interest), SAC (Special Area of Conservation), SPA (Special Protection Area) and under the RAMSAR convention on wetlands. There are three NNRs (National Nature Reserves) and three LNRs (Local Nature Reserves) within the borough of Sefton. Key species of this area include the sand lizard, natterjack toad, great crested newt, red squirrel, dune helleborine and pendulous flowered helleborine. There are 460 species of flowering plants recorded, including 33 locally or regionally rare species.

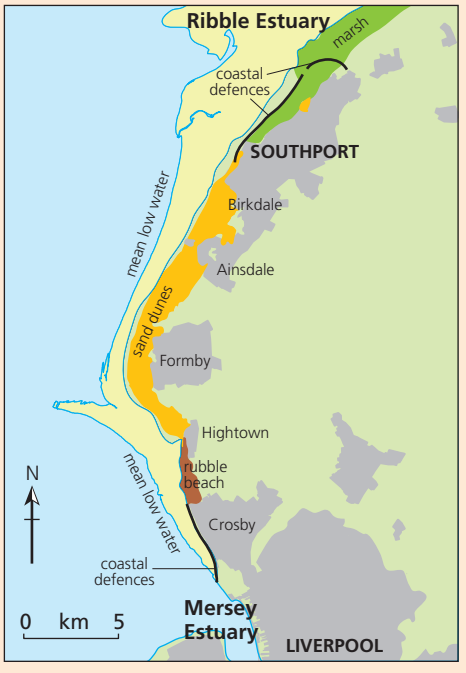


Figure 6.24 The location of the Sefton Coast sand dunes

Figure 6.25 Formby Point dune system

Succession on sand dunes

Sand dunes form above the highest high-tide mark behind a beach (Figure 6.25). As one moves from the bare sand of the beach to the dune heath and forest inland, the location is older and vegetation and animal life have had longer to colonise and adapt to the environmental conditions. This sand dune succession is known as a **psammosere**. Each stage of a psammosere is known by its location on the coastline (Figure 6.26).

- **The strand line**: This is bare sand and it is inhospitable for plant growth. It is:
  - dry and salty, lacking in plant nutrients
  - has a high pH from fragments of sea shells
  - unstable.

  There are a few highly specialised plants that have evolved strategies to grow in such a location. These are the **primary colonisers** of the psammosere and are restricted in their distribution to this very stressful type of habitat. Examples are:
  - sea rocket
  - prickly saltwort.

  These annual plants may form miniature sand dunes as sand accumulates around the plant body and the plant is able to grow upwards a little through the accumulating sand. Alternatively, they may be washed away by the waves brought in on the next high tide.
- **Embryo dunes**: Sand accumulation which persists above the high-tide line is colonised by the first perennial plants in dune succession. These are specialised grasses such as:
  - sand couch
  - lyme grass.

  Both of these grasses are able to grow upwards through accumulating wind-blown sand and as a result low, hummocky dunes are formed.
- **Mobile dunes**: Upward growth of the embryo dunes allows the surface to be raised so it is out of the reach of all but the highest storm tides (Figure 6.27, page 250). This, together with the washing by fresh rainwater, results in a slightly less salty and unstable substrate. Under these conditions marram grass is able to colonise. Marram grass is the major dune building grass. It is tall and robust (but flexible in the wind) and very

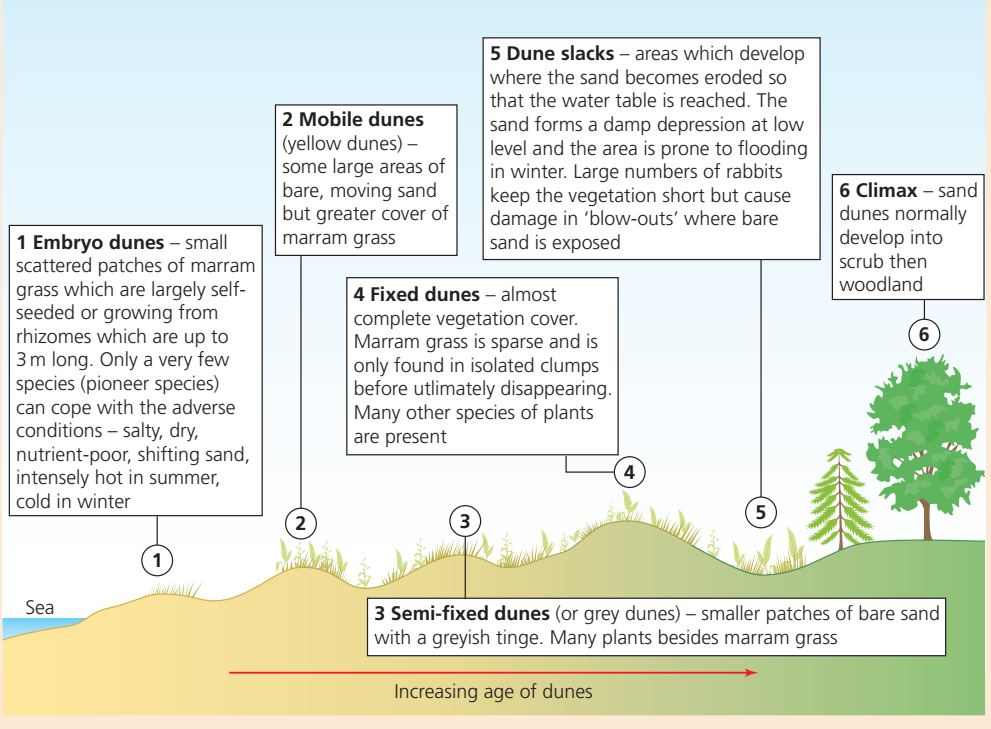


Figure 6.26 Plant succession on the Sefton Coast sand dunes

effective at trapping sand by reducing the windspeed at the surface. It is able to grow upwards through accumulating sand at rates of up to one metre a year. The environment is still extremely stressful; marram is the dominant and only species present.

Figure 6.27 Erosion of mobile dunes at Formby during a high spring tide

On the landward side of the frontal dunes the surface is more sheltered from the onshore winds and the effects of sea spray. More (still highly adapted) plants are able to colonise and the species diversity starts to increase. Good examples of these plants are:
- sea holly
- sea spurge.

As the conditions for plant growth improve with increasing distance from the sea, the species diversity continues to increase. A much smaller and very fine-leaved grass starts to fill in the bare spaces between the clumps of marram. This is:
- red fescue.

If conditions remain stable at the sand surface red fescue and mosses will continue to cover the bare sand between the patches of marram.

- **Fixed dunes:** When the vegetation has developed so that it forms a more or less complete cover, the dunes are said to be 'fixed'. These fixed dunes still represent a stressful environment. The pH is still very high, drought is a problem and nutrients are still in very short supply. In addition to these abiotic factors, the dunes may be affected by grazing or trampling. A thin, brown, organic layer has, however, started to form at the surface of the soil. If the dunes are grazed, for example, by rabbits or sheep, a **fixed dune grassland** will develop.

  On the Sefton Coast the rabbit population was severely affected by the introduction of a disease called myxomatosis in the 1950s. The absence of grazing pressure caused by the loss of rabbits allowed a different type of vegetation, large coarse grasses and a woody plant called creeping willow, to develop. In recent years grazing by domestic stock has been re-introduced to the Ainsdale area and rabbit populations are now healthy again.

- **Dune slacks:** There are a number of wet and sometimes very large depressions in the dunes. These are called dune slacks. The water table fluctuates annually because of the differences in rainfall and evapotranspiration between the summer and winter months. Because of this, many dune slacks are flooded during the winter period. The wet sand is colonised by plants and a succession occurs. This succession on wet sand, however, is very different to that on the dry dunes. A range of wetland plants are important and early vegetation can be extremely species-rich with plants such as sedges and horsetails together with a number of mosses. These slacks are the breeding ground for the natterjack toad (Figure 6.28).

- **Scrub and woodland:** Since the rabbit population was almost destroyed in the 1950s, tall woody plants that invaded naturally have become widely established on the dunes. These include:
  - birch
  - hawthorn.

  In addition, pines from the plantations seed into the open dunes.

Figure 6.28 A protected damp slack designed to encourage natterjack toads

Skills focus

How might you survey the succession in a location such as Formby?

The human impact on the dunes of the Sefton Coast

From the middle of the nineteenth century, after the Southport to Liverpool Railway was built, vast amounts of 'night soil' (human manure) from the backstreets

of Liverpool were dumped in the dunes in Freshfield, Formby. Dunes were flattened and about 80 ha of duneland was converted to growing asparagus. This continues on National Trust land today, though most of the flattened areas have become car parks, caravan sites or just open fields.

Uncontrolled public access in the 1960s and 1970s caused severe dune erosion at Formby Point. Dogs and people caused disturbance to ground-nesting birds. Off-road vehicles damaged large areas of sand dunes in the 1970s and 1980s until controlled by the introduction of countryside wardens.

Over 25 per cent of the Sefton Coast dunes are managed as golf courses. Management techniques such as drainage, irrigation, mowing, fertilising and reseeding have damaged habitats. However, much golf management is sympathetic to the conservation of the dune–links landscape and its characteristic species and habitat types. Approximately 70 per cent of each golf course area is out of play and direct golf management.

Management

In the early twentieth century the area was planted with pine trees. The aim was to stabilise the dunes and turn the 'wasteland' into a more productive estate providing a timber crop, woodland products and improved opportunities for agriculture and game. The first pines were planted around 1900 and areas continued to be planted until the 1960s. Today, these woodlands support an important population of the red squirrel, now rare in England. The slacks were generally not planted, although drainage ditches were put in to lower the water table. This drainage has led to a loss of wet slacks, an important habitat for protected species such as the natterjack toad.

Once the woodland canopy was formed, trees shaded out the area beneath and so no other plants were able to grow. Huge areas of natural dune landscape with specialised sand dune plants and animals were lost in this way. The trees also cause the water levels in the soil to drop; pine needles fall on the surface and alter the structure and chemistry of the soil. Together these create very different conditions from the original dunes. The needles are very acid and there is a very large accumulation of plant litter on the soil surface. In these pine woodlands, the pine trees are often the only plant species present (Figure 6.29).

In one part of the area, the Ainsdale National Nature Reserve, the seaward plantations have been removed. This has allowed the recolonisation by specialised plants such as yellow bartsia, and animals such as the sand lizard and natterjack toad. Monitoring of the project shows that this is already a success. The landward plantations have been allowed to stay because of their importance to the red squirrel. The National Trust manages 414 hectares of land along the Formby coast. Its aim is to safeguard the resource and the wildlife that lives there for future generations. One of the management challenges is the loss of mobile and fixed dunes due to increased marine erosion.

Figure 6.29 Pine woodland on the dunes at Formby

Conservation of fixed dunes

The fixed dunes and associated species suffer from a number of factors threatening to reduce their nature conservation value. The spread of scrub and rank vegetation, in particular, leads to associated problems of soil development and the drying out of slacks. In order to manage this there has been:

- **scrub cutting and clearance:** this allows light and warmth to reach the surface so colonisers can begin to grow
- **mowing** to control the height and density of creeping willow in dune slacks where it was rapidly overwhelming the short, botanically diverse dune turf
- **turf-stripping and excavation:** some of the slacks have been scraped to provide breeding pools for the natterjack toad.
- **grazing**: domestic grazing was re-introduced onto the Ainsdale dunes by English Nature. It has controlled target species such as creeping willow; there has been an increase in species diversity with a corresponding return to a mix of vegetation with bare sand patches.

6.2 Ecosystems and sustainability

Biodiversity

> The preservation of biodiversity is not just a job for governments. International and non-governmental organisations, the private sector and each and every individual have a role to play in changing entrenched outlooks and ending destructive patterns of behaviour.
>
> *Kofi Annan*, former Secretary General of the United Nations

Biodiversity, which is short for **biological diversity**, is the term used to describe the whole variety of life on Earth. It encompasses the diversity of all living things, from human beings to micro-organisms, the diversity of all the habitats in which they live and the genetic diversity of individuals within a species.

All of these are part of the global ecosystem – the Earth's entire collection of living things and the environment in which they live. Ecosystems form the basis for all forms of life support, including provision of food, clean air and water, climate and disease regulation and many others. Biodiversity underpins the health of ecosystems.

Measuring biodiversity

From the very broad definition given above it is clear that no single measure of biodiversity will be adequate. Biodiversity cannot be captured in a single number because measurements have two components:

- the number of entities, for example, the number of individuals, the number of species, the number of different habitats, etc.
- the degree of difference (dissimilarity) between those entities.

Measuring biodiversity is commonly used to provide evidence for the need for conservation or for planning more generally. Often **indicator species** are used as a way of measuring biodiversity. For example, stoneflies are found in clean, fast-moving streams with a gravel or stone bottom. The presence of stonefly nymphs indicates highly oxygenated (healthy) water. Mussels are used to measure the health of coastal environments worldwide.

This method can be very useful but it introduces an aspect of how we value different components of biodiversity. For example, we are more likely to use the abundance of birds or butterflies on a farm as a measure of biodiversity than the richness of microbes in the soil.

Species richness is one of the most commonly used measures of biodiversity and it refers to the number of different species; it tells us how many different species there are but does not convey information on the frequency of individuals of each of those species.

The advantages of this:

- Species richness has proven to be measurable in practice. It is doable.
- A lot of information already exists on patterns in species richness, and this has been made available in scientific literature. This allows comparisons over time.
- Species richness acts as a proxy measure for many other kinds of variation in biodiversity. In general, greater numbers of species tend to indicate more genetic diversity (in the form of a greater diversity of genes through to populations), more species diversity and greater ecological diversity.
- It is commonly seen as the unit of practical management and is most easily applied to legislation and political debate. For a wide range of people, variation in biodiversity is pictured as variation in species richness.

The disadvantages of this:

- There is a lack of agreement as to what constitutes a species. For example, it is not an adequate method of defining single cell organisms.
- Some assemblages of organisms may be very closely related (for example, two species of rat) while others are diverse (for example, a species of rat and a species of snail).

A more recently developed measure of biodiversity is the **Living Planet Index (LPI).** The LPI measures trends in thousands of vertebrate species populations and provides information on changes in the abundance of the world's vertebrates. It can quickly convey information on which habitats or ecosystems have species that are declining most rapidly. This information can be used to define the impact humans are having on the planet and for guiding actions to address biodiversity loss.

Trends in biodiversity

Global trends in biodiversity

Through geological time, global biodiversity has steadily increased from when life as we know it started about 550 million years ago (mya). In that time there have been five episodes of **mass extinctions** (for example, the end of the dinosaurs about 65 mya). Some modern scientists believe that we are now in another period of mass extinction – caused by man.

Certainly, population sizes of animal species measured by the LPI have declined overall by 60 per cent between 1970 and 2014. This decline has continued rapidly since then.

Biodiversity is declining in both temperate and tropical regions, but the decline is greater in the tropics – particularly in South America (with a fall of 89 per cent). The tropical LPI showed a 56 per cent reduction in 3,811 populations of 1,638 species from 1970 to 2010. The 6,569 populations of 1,606 species in the temperate LPI declined by 36 per cent over the same period. Latin America shows the most dramatic decline of 83 per cent.

Exploitation through hunting and fishing, and habitat degradation and loss are the main causes of decline (Figure 6.30). Climate change is the next most common primary threat, and is likely to put more pressure on populations in the future.

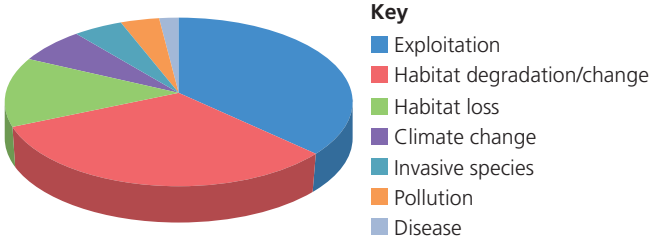


Figure 6.30 Primary threats to LPI populations according to WWF

According to the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (IPBES), since 1900, the average abundance of native land-band species has declined by at least 20 percent, a trend that shows no sign of slowing down. The loss of habitat for human land use – particularly for agriculture, urban development and energy production – continues to be a major threat, made worse by hunting.

The LPI for freshwater species shows an average decline of 83 per cent. The main threats to freshwater species are habitat loss and fragmentation, pollution and invasive species. Changes to water levels and freshwater system connectivity – for example, through irrigation and hydropower dams – have a major impact on freshwater habitats.

Marine species declined by more than half between 1975 and 2015. The period from 1970 through to the mid-1980s experienced the steepest decline, after which there was some stability, before another period of decline. The steepest declines can be seen in the tropics and the Southern Ocean – species in decline include marine turtles, many sharks and large migratory seabirds like the wandering albatross.

Many species have completely disappeared from areas dominated by human influences (Figure 6.31, page 254). Even in protected areas, native species are often out-competed or consumed by organisms introduced from elsewhere. Extinction is a natural process, but it is occurring at an unnaturally rapid rate as a consequence of human activities. Already we have caused the extinction of 5–20 per cent of the species in many groups of organisms, and current rates of extinction are estimated to be 100–1,000 times greater than pre-human rates.

Example of a local area subject to biodiversity change

One of the most extensive, semi-natural habitats in the UK is **blanket bog.** It ranges from Dartmoor in the south to Shetland in the north. It is a globally restricted peatland habitat confined to cool, wet, typically oceanic climates. Blanket bogs comprise of generally upland **peat** bogs that are fed by rainfall alone. Peat is an accumulation of dead organic material, mainly plant remains, that is unable to break down because of the wet, acidic and cold conditions in which it is found. These peat bogs take thousands of years to form and support mosses, insects, rare birds and even carnivorous plants. Peat depth is very variable, with an average of 0.5–3 m being fairly typical.

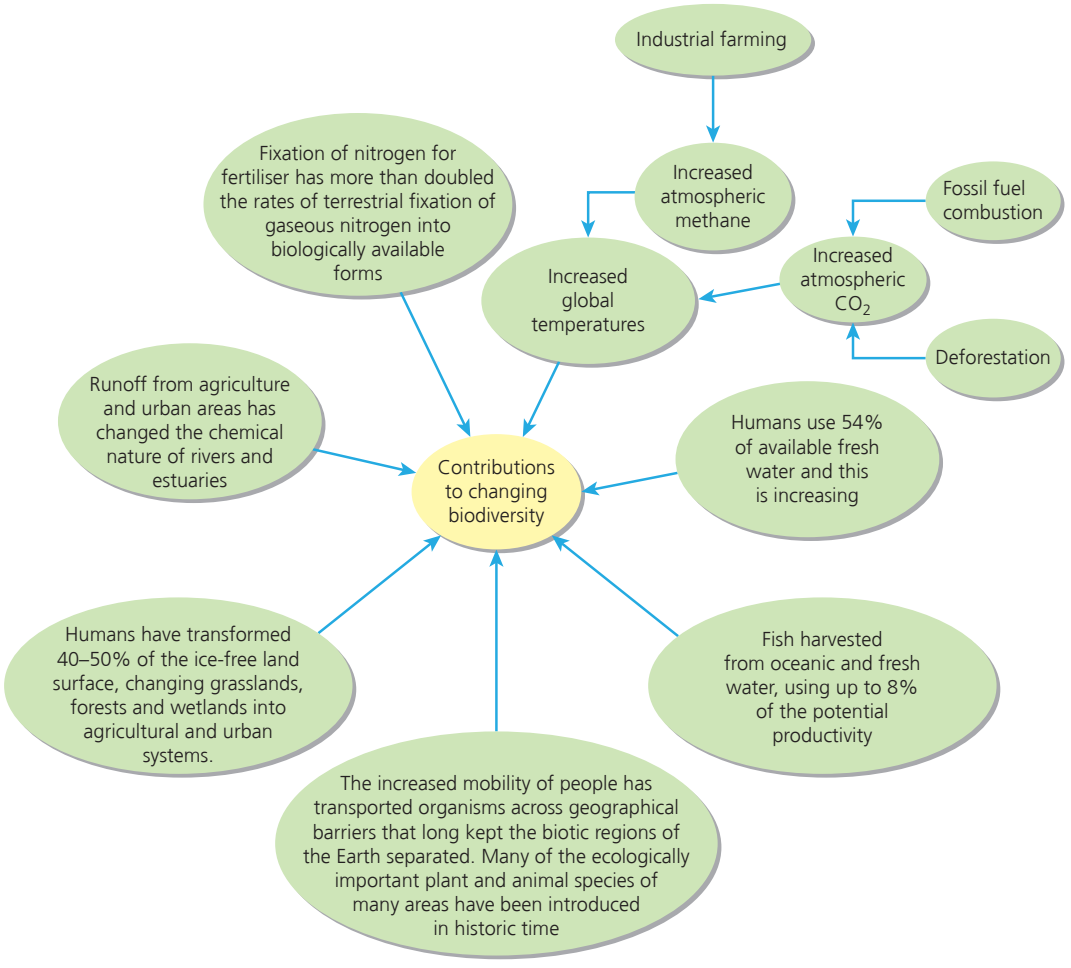


Figure 6.31 Contributions to changing biodiversity

Because peat is essentially semi-decomposed plant matter, it represents a huge carbon store: more carbon is stored in UK peat than in all the forests in Britain and France combined! When these habitats are healthy, and peat is actively forming, they are building up carbon, so they have an important role to play in helping to combat climate change.

Kinder Scout

One area of blanket bog is **Kinder Scout, Derbyshire**. Kinder Scout is an area of moorland and National Nature Reserve in the Dark Peak of the Derbyshire Peak District in England. Part of the moor, at 636 m above sea level, is the highest point in the Peak District. The 7,000-year-old blanket peat bog sits on top of a plateau (Figure 6.32 and Figure 6.33) formed from horizontal layers of an impermeable sandstone called Millstone Grit. The land is owned and managed by the National Trust.

Figure 6.32 The plateau of Kinder Scout

Figure 6.33 Location of Kinder Scout, Derbyshire. © Crown copyright and database rights, 2021, Hodder Education, under licence to Ordnance Survey. Licence number 100036470

Kinder Scout is also one of the most damaged areas of moorland in the UK and its future is in jeopardy as a result of catastrophic wild fires, a long history of overgrazing by sheep, air pollution and the routes that thousands of visitors have taken.

In spring 2011, the National Trust started a five-year project to restore the peat bogs of Kinder Scout.

- They started by fencing off much of the moorland from sheep. The fencing should be in place for 15 years, at which point the Trust and other bodies will take stock of the impact.
- They planted nearly half a million pods of cotton grass in 2011 (they are dropped by helicopter but then are planted by hand); these help to bind the loose peat together and combat erosion.
- This is part of a longer-term reseeding programme. Following on from that, heather brash was spread across large areas of Kinder to encourage heather to re-seed. Heather is cut and harvested from nearby local moors in winter and is spread over the worst eroding areas of blanket bog. It has four purposes:
  - acting as a skin, thus reducing the effects of erosion
  - providing a microclimate for seeds to grow in as they are protected from the harsh weather
  - providing seeds – heather is cut in winter when the seeds are ripe and thus they are transported onto the moors
  - providing a fungi that helps moorland plants to thrive.
- Rangers have built over 6,000 small dams in erosion gulleys, rewetting the moor, reducing peat erosion and helping heather to flourish. These dams also act as a flood prevention measure for settlements in the valley by slowing the rate of water release. In addition, they have planted 10,000 trees in the surrounding cloughs.

Table 6.4 Causes and impacts of biodiversity change on blanket bog on Kinder Scout

Cause	Consequence	Impacts and potential impacts
Climate change		
Increased mean temperatures	Longer growing season	Bracken has become invasive at higher altitudes Potentially increased nitrogen deposition leading to changes in plant species Mire vegetation (that is, plant types that thrive in wet conditions) may become less dominant
Hotter summers	Increased evapotranspiration and lowering of water table	Changes in species composition. Increase in the release of dissolved organic carbon leading to declining water quality
Drier summers	Drought Drier ground conditions Wildfire	Shift in the dominance of species Possible wind erosion of dried peat Possible increased agricultural potential Changes in red grouse populations
Wetter winters	Increased overland flow	Loss in peat stability; increased slides
Storm events	Increased rainfall intensity	Gullying erosion
Other influences		
A long period of over-grazing	Overgrazing	Loss of biodiversity
Access to walkers from nearby industrial towns and cities since the 1940s	Footpath erosion. Damage to surface vegetation that holds the peat in place	Large areas of peat lost Exposed dried-out peat subject to wind and gullying erosion
Increased wildfires	Loss of peat	Diminution of species numbers
Draining of peat bog	Lowering of water table	Sphagnum mosses have been replaced by other species such as heather and moor grass Species of moorland birds that nest and feed in the heather, such as golden plover and curlew, as well as the mountain hare are threatened
Industrial pollution		Reduction in number of species

Evaluation of the scheme:

Within two years of starting the restoration project, the percentage vegetation cover, runoff and sediment yield all show large changes on bare peat sites. While near complete vegetation cover can be achieved relatively quickly (six years) the change in species richness can take up to 12 years. The water table has risen and continues to rise slowly. Rare wildlife like the Bilberry Bumblebee has started to return.

Skills focus

a Using Figure 6.32 (page 254) and Figure 6.33 (page 255), describe the landscape of Kinder Scout and the surrounding area.

b Describe the pattern of drainage for the Kinder Scout area. How has that helped Kinder Scout to develop into a blanket bog?

Human population and ecosystems

An ecosystem is a dynamic complex of plant, animal, and micro-organism communities and their non-living environment interacting as a functional unit. Humans are an integral part of ecosystems.

Ecosystem services are the benefits that people obtain from ecosystems. The **Millennium Ecosystem Assessment (MEA)** analysed 24 ecosystem services and found that 15 were being degraded or used unsustainably. The decline in services affects the world's disadvantaged people most strongly, impedes sustainable development globally and, in developing countries, represents a considerable barrier to achieving the UN's Millennium Development Goals of reducing poverty and hunger.

The MEA grouped ecosystem services into four categories:

- **provisioning services,** such as the supply of food and water
- **regulating services**, which help to stabilise ecosystem processes such as climate and water storage and purification
- **supporting services**, including soil formation and nutrient cycling; and
- **cultural services**, such as recreational, spiritual, religious and other non-material benefits.

The MEA also considers human well-being to consist of five main components:

- the basic material needs for a good life
- health
- good social relations
- security
- freedom of choice and action.

Figure 6.34 (page 258) shows one attempt to highlight the importance of ecosystems and biodiversity to human populations. Human well-being is the result of many factors, many are directly or indirectly linked to biodiversity and ecosystem services while others are independent of these.

Many people have benefited over the last century from the conversion of natural ecosystems to human-dominated ecosystems and the exploitation of biodiversity. At the same time, however, these losses in biodiversity and changes in ecosystem services have caused some people to experience declining well-being, with poverty in some social groups being made worse.

Although biodiversity and ecosystem services experience change due to natural causes, there are also five indirect drivers of change:

- demographic
- economic
- sociopolitical
- cultural and religious
- scientific and technological.

Current change is driven by these **anthropogenic** factors, in particular growing consumption of ecosystem services (as well as the growing use of fossil fuels), which results from growing populations and growing per capita consumption. Figure 6.35 (page 258) is an illustration of this. Global economic activity increased rapidly from 1950 onwards and despite occasional economic downturns, is expected to continue to rise. At the same time the global population reached 6 billion in 2000 and is projected to reach 8.1–9.6 billion by 2050.

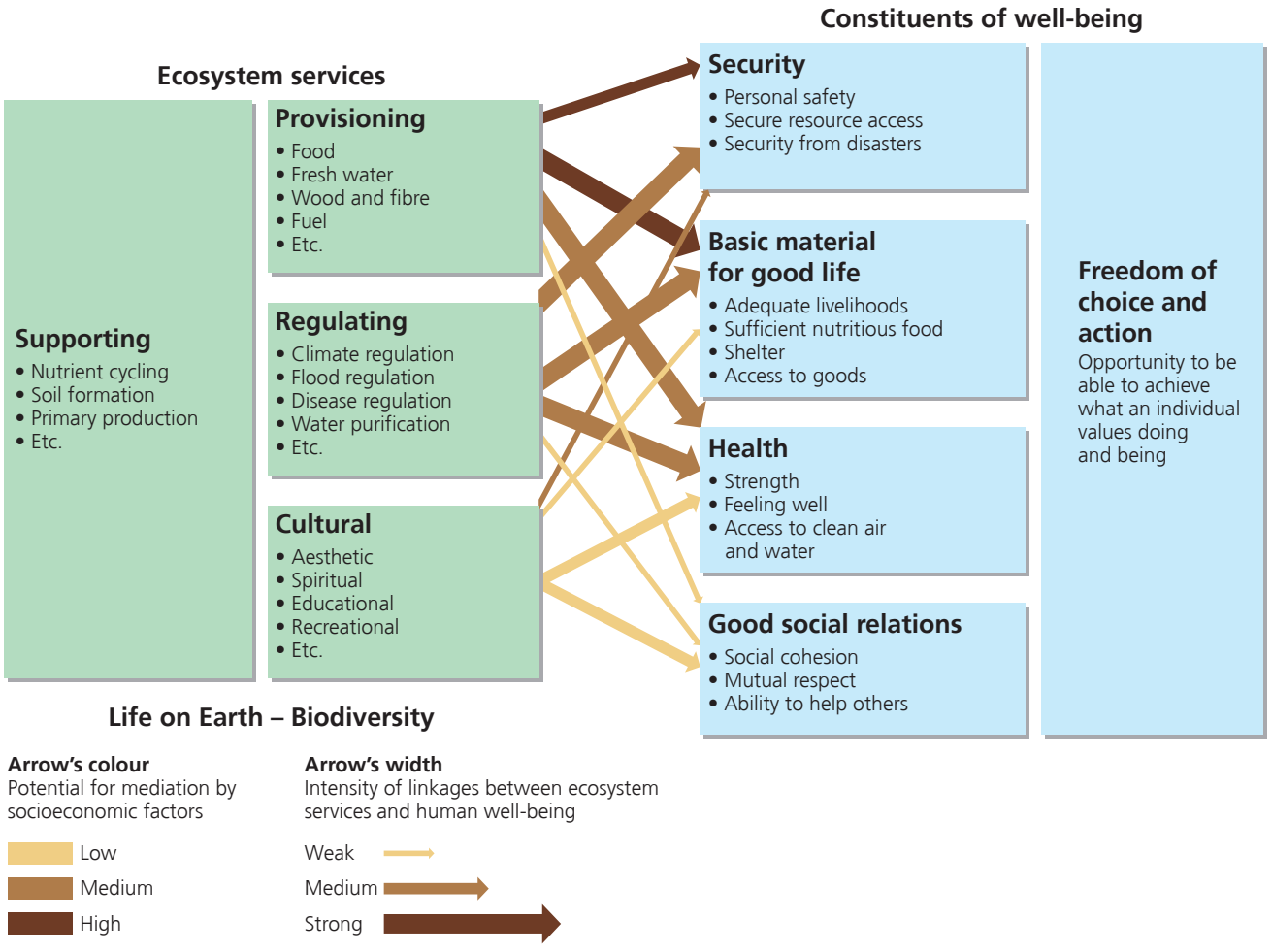


Figure 6.34 Linkages among biodiversity, ecosystem services and human well-being

Source: UN Millennium Ecosystem Assessment: http://pdf.wri.org/mea_synthesis_030105.pdf

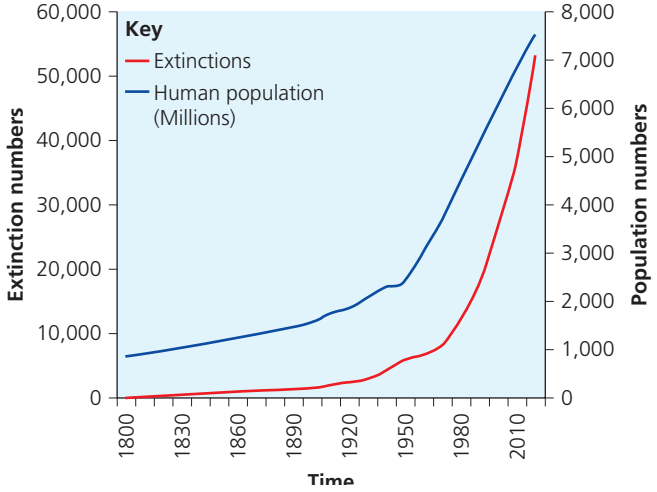


Figure 6.35 Species extinction and human population

Source: USGS

The UK ecosystem services

In 2011, a major UK government-sponsored report, the National Ecosystem Assessment (NEA) attempted to put a value on the UK's ecosystems. The conclusion was that the parks, lakes, forests and wildlife are worth billions of pounds to the economy. It concluded that the health benefits of merely living close to a green space are worth up to £300 per person per year.

> The natural world is vital to our existence, providing us with essentials such as food, water and clean air – but also cultural and health benefits not always fully appreciated because we get them for free.
>
> *Caroline Spelman,* former Environment Secretary

The economic benefits of nature are seen most clearly in food production, which depends on organisms such as soil microbes, earthworms and pollinating insects. If their health declines – as is currently happening in the UK with bees – either farmers produce less food, or have to spend more to produce the same amount.

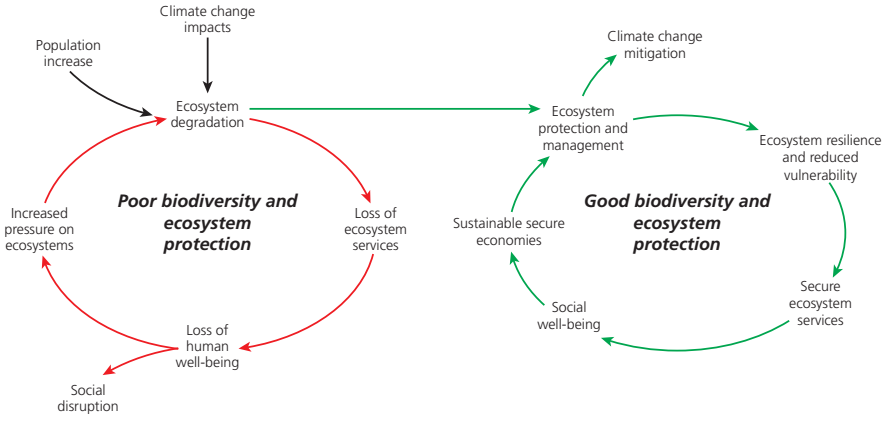


Figure 6.36 Beating the vicious cycle of poverty, ecosystem degradation and climate change

The report also concluded that:

> Humans rely on the way ecosystem services control our climate – pollution, water quality, pollination – and we're finding out that many of these regulating services are degrading.
>
> About 30 per cent of the key ecosystem services that the UK relies on are degrading and about 20 per cent are getting better. Our air quality, however, has improved a lot.

Ecosystem development and sustainability

Figure 6.36 shows two different scenarios that could result from human population growth and climate change. One way, the positive feedback loop, shows how poor ecosystem management leads to decline in human well-being and eventual further decline in ecosystems. Good ecosystem management, on the other hand, causes there to be a negative feedback loop which leads to climate mitigation and social well-being.

Review questions

9 How can a systems approach to the study of ecosystems explain the role organisms play in the environment?

10 For four discrete environments, name a species that is at risk and the primary threat to that species.

11 What type of activities might be considered 'good ecosystem management'?

12 What are the similarities and differences between the ways different aboriginal people have adapted to changing circumstances within their biome?

13 Explain why climate change appears to be the biggest threat to many ecosystems.

14 How do ecosystem services affect you?

15 Why is there a wide range of forecasts for the future of ecosystems? What does this tell us about our fundamental understanding of how the planet works?

6.3 Marine ecosystems

The marine ecosystem is the largest ecosystem on Earth because oceans and seas account for almost 70 per cent of its surface. It can be divided into smaller ecosystems, such as rocky shores, salt marshes, submarine canyons, hydrothermal vents, deep basins, seamounts and coral reefs.

Coral reef ecosystems

There are two broad groups of corals:

- deep-sea cold-water coral reefs
- tropical coral reefs.

Deep-sea cold-water coral reefs

These are a recent discovery on Europe's margin. They are found in the deeper, nutrient-rich high latitude waters where ocean currents accelerate and prevent sedimentation. Colonial stone corals, such as Lophelia pertusa and Madrepora oculata, occur along much of the northwest European continental margin and in some Scandinavian fjords, and have recently been discovered in the Mediterranean. They grow slowly (5–25 mm/year), but over time they form extensive reefs. The largest reef yet discovered, off the coast of Norway's Røst Island, is 40 km long and 2–3 km wide.

Tropical coral reefs

Tropical coral reefs are found between 30° north and south of the Equator, in areas where surface water temperatures do not drop below 16 °C (Figure 6.37, page 260). The total area of the world's tropical coral reefs is around 284,300 km^2. These reefs are able to grow upwards at rates of 1–100 cm/year. They can form huge structures over incredibly long periods of time, making

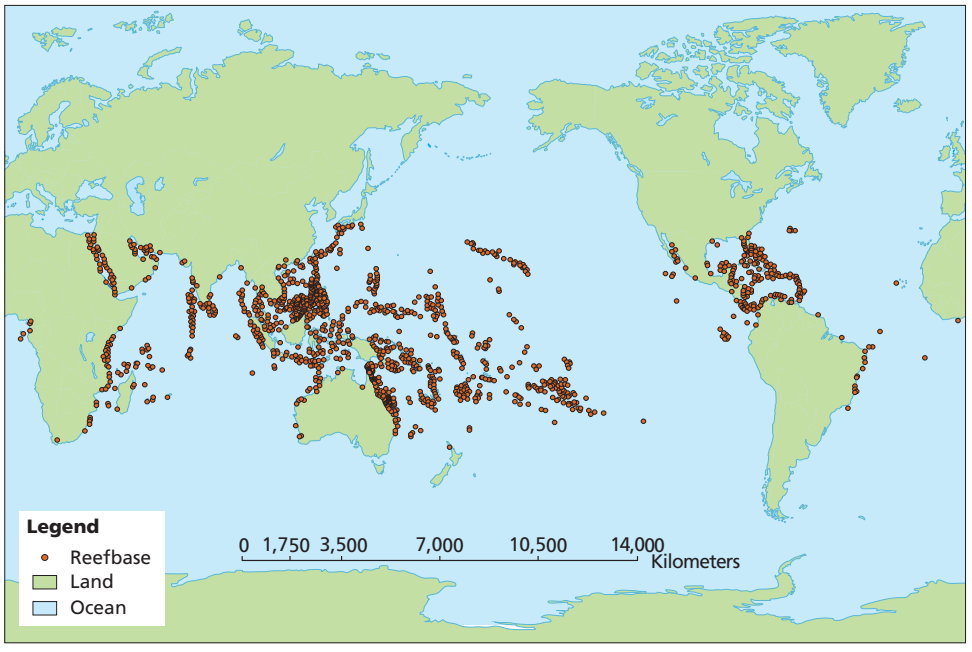


Figure 6.37 The location of the major warm water stony coral reefs of the world

them the largest and oldest living systems on Earth. For example, Australia's 2,000 km-long Great Barrier Reef was formed over the course of five million years.

A coral reef is composed of calcium carbonate, or limestone. This is absorbed from the water by colonies of coral polyps and coralline algae. Most of the underlying foundation of the reef is dead, made up of layer upon layer of coral skeletons. The living reef is built over the top of this by tiny coral polyps adding new limestone to the massive base structure. Tropical coral polyps have small, symbiotic algae, or zooxanthellae, growing inside them. The algae get shelter and food (in the form of nutrients from captured plankton) from the polyp, while the polyp gets some food from the algae via photosynthesis. This photosynthesis means that the algae (and so the coral) need sunlight to live, explaining why tropical corals only grow where the sea is shallow, clear of sediment and away from river mouths.

The algae also give corals their colour. If they become stressed, such as if the water temperature becomes too high, they leave the polyp. This exposes the white, calcium carbonate skeletons of the coral and is called coral bleaching.

Coral reefs are divided into three main types:

- Fringing reefs grow directly from a shore with no area of open water between the beach and the inshore edge of the reef. This is the most common type of reef. Because there is no intervening lagoon to act as a buffer between the land and the reef, they are particularly sensitive to pollution and sedimentation from rivers.
- Barrier reefs are extensive linear reef complexes that develop parallel to the shore, separated from it by a lagoon (a wide band of sea water of varying depth).
- Atolls are roughly circular reef systems surrounding a central lagoon.

Examples of tropical coral reef systems

Although most of the world's coral reefs develop under, more or less, the same conditions, they face a variety of threats from human activity. These are covered in the following three examples.

Belize Barrier Reef Reserve System

The 300-kilometre-long Belize Barrier Reef Reserve System (Figure 6.38) is the second largest reef system in the world. Within it are seven protected areas that have been designated a UNESCO World Heritage Site.

The unique array of reef types within one self-contained area distinguishes it from other reef systems. As well as the barrier reef complex there are three atolls. The whole reef complex is 960 km^2 and comprises approximately 450 sand and mangrove cayes (small, sandy islands).

Approximately 70 hard and 36 soft coral species have been identified on the barrier reef and over 500 species of fish and invertebrates live in the area. However, it is believed that only about 10 per cent of all the species on the reef have been discovered.

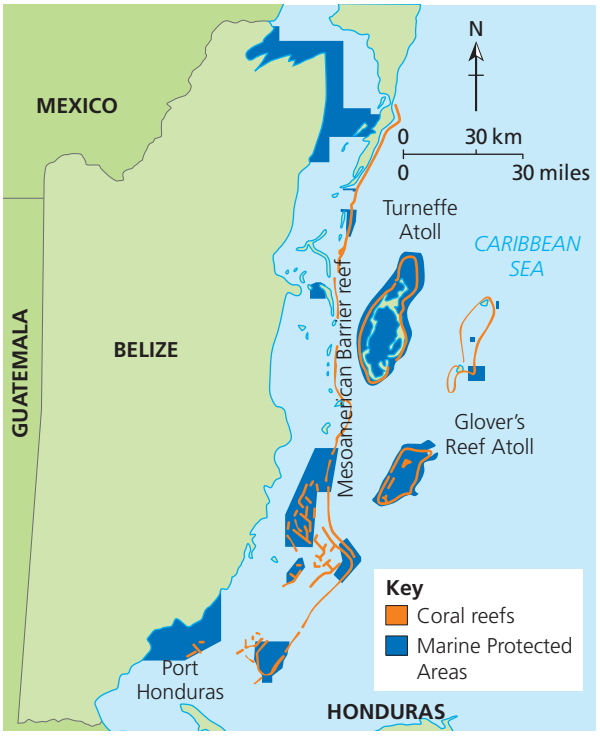


Figure 6.38 The Belize Barrier Reef Reserve System

Physical influences

- **Coral bleaching** is the term used to describe a phenomenon where coral loses its zooxanthellae. This can cause the coral to die if the algae are not replaced within weeks. Coral bleaching is caused by unusually warm sea surface temperatures. Coral bleaching events worldwide have been attributed to sea surface temperatures (SSTs) rising and staying as little as 1°C higher than the usual average monthly maximum SST during the hottest months of the year. In Belize, bleaching along the reef has become more common as the oceans become warmer. In 2015 and 2016, almost a quarter of the corals off the Belizean coast were affected by bleaching.
- Coral reefs develop best where the salinity of the sea water is between 34 and 37 parts per 1,000. Where rivers flow into the sea there are gaps in the reef because of the lowering of the salinity locally. This can be seen where the Belize River flows into the lagoon between the mainland and the barrier reef.
- Studies regarding links between the acidification of sea water and coral development have been inconclusive. Although scientists think that coral's ability to form an exoskeleton of calcium carbonate would be reduced, this has not always been found to be the case. In fact during the summers of 2010 and 2011, scientists in the Caribbean observed rapid acidification events driven by enhanced coral growth and calcification, which were likely caused by an increase in food supply thanks to phytoplankton blooms.

  Not only does the formation of $CaCO_3$ produce CO_2, well-fed corals respire more CO_2 than their symbiotic algae can take up through photosynthesis – making the water around them more acidic.
- A disease epidemic swept through populations of an important reef grazer – the sea urchin, Diadema antillarum, virtually killing it off across the Caribbean. Over-fishing and the loss of these urchins meant that there was rapid algal growth (an algal bloom). These blooms limit the sunlight available and cause lack of oxygen in water. When oxygen levels decline, marine animals, coral reefs, seagrass beds and other vital habitats suffer and may die.
- Major hurricanes such as Hurricane Earl (2017) have damaged the Belizean reefs. The violence of the waves fragment and kill branching coral in shallow water. Even the soft corals are damaged by objects being tossed around in the waves. Sometimes the scars that result are rapidly overgrown by fast growing opportunistic algae. The heavy rainfall on land and the increased river run-off increases cloudiness and reduces water salinity. Fortunately for Belize it is rarely affected by hurricanes, with just eight major hurricanes in the last 100 years.

Human influences

In 1996, UNESCO designated the Belize Barrier Reef Reserve System a World Heritage Site, with Belize responsible for its protection. It is something that the country has struggled with. By 2009, the site was on UNESCO's 'danger' list, with the organisation saying that the country needed to enact better management and safeguards.

- Mangrove cutting and excessive development were cited as the main concerns.
  - Development of the tourist industry has meant land clearance, resulting in increased amounts of sediment being washed into the sea. Mangroves act as nurseries for small reef fish and as filters, removing much of the sediment being washed off the land.
  - Coastal developments, such as tourist hotels, are creating problems for the satisfactory treatment of sewage. Each hotel has its own disposal system, but a lot of waste is discharged into the sea after little treatment.

- UNESCO also expressed concern about oil exploration. In 2017, the Belize government was the first in the world to impose a moratorium (temporary halt) on oil exploration and exploitation in the maritime zone of Belize.
- The government has introduced special environmental taxes that go directly to conservation and a booming eco-tourism market.
- The fish population on Belize's reefs has been steadily declining as a result of over-fishing. One study indicated that fish biomass had already been reduced by up to 80 per cent by the late 1970s. The government has tried to stop trap fishing as well as visits from foreign commercial trawlers, although this can be difficult to manage. In 2008, when the International Society for Reef Studies recommended that countries along the reef must protect parrot fish, Belize was the first to respond with a law restricting their catch. In 2018, Belize announced plans to more than triple the size of its 'no-take' fishing zones, from 3 to 10 per cent, to help ocean wildlife to recover.
- They also announced plans to ban single-use plastic and Styrofoam products (bags, utensils and so on) by Earth day, 2019.

The Jamaican coral reef

Jamaica, the third largest Caribbean island is 230 km long by 80 km wide with 891 km of coastline and a coral reef area of 1,240 km^2. Well-developed fringing reefs occur along most of the north and east coasts, while patchy fringing reefs occur on the broader shelf of the south coast (Figure 6.39). In addition to the reefs surrounding mainland Jamaica, there are reefs and corals on the neighbouring banks and shoals within Jamaica's Exclusive Economic Zone.

After a series of natural and manmade disasters in the 1980s and 1990s, Jamaica has lost 85 per cent of its coral reefs. The physical influences are very similar to those experienced by Belize, but the human influences differ.

Human influences

Pollution comes in many forms including:

- Onshore development such as bauxite mining in Cockpit Country is particularly important for the economy of Jamaica. This activity also releases large quantities of particulate matter into local rivers. This sediment load is carried to the sea, creating cloudy water and reducing the necessary sunlight penetration coral needs in order to grow. The corals within Montego Bay Marine Park are under threat.
- Coastal developments such as tourist hotels in Jamaica are creating problems for the satisfactory treatment of sewage. Coastal hotels each have their own disposal systems, but much waste is discharged to sea after little treatment. Although they are now beginning to install treatment plants, many are not used efficiently.
- Excess nutrients are discharged to the marine environment through fertilisers from agriculture. These nutrients may lead to eutrophication (an excessive growth of marine plant life and decay). This leads to algal blooms.
- In recent years there has been a water shortage in Jamaica; in 2012 water was rationed. There have been calls for the building of desalination plants on the island. These plants remove all the unwanted chemicals from sea water and produce a fresh water supply. The unwanted chemicals are then flushed back out to sea causing localised increases in salinity and temperature. Studies have found effects ranging from no significant impacts through to widespread alterations to community structure in seagrass, coral reef, and soft-sediment ecosystems when discharges are released in areas where there is little mixing of sea water.
- The fish population on Jamaican reefs has been steadily declining as a result of over-fishing. One study indicated that fish biomass had already been reduced up to 80 per cent by the late 1970s.

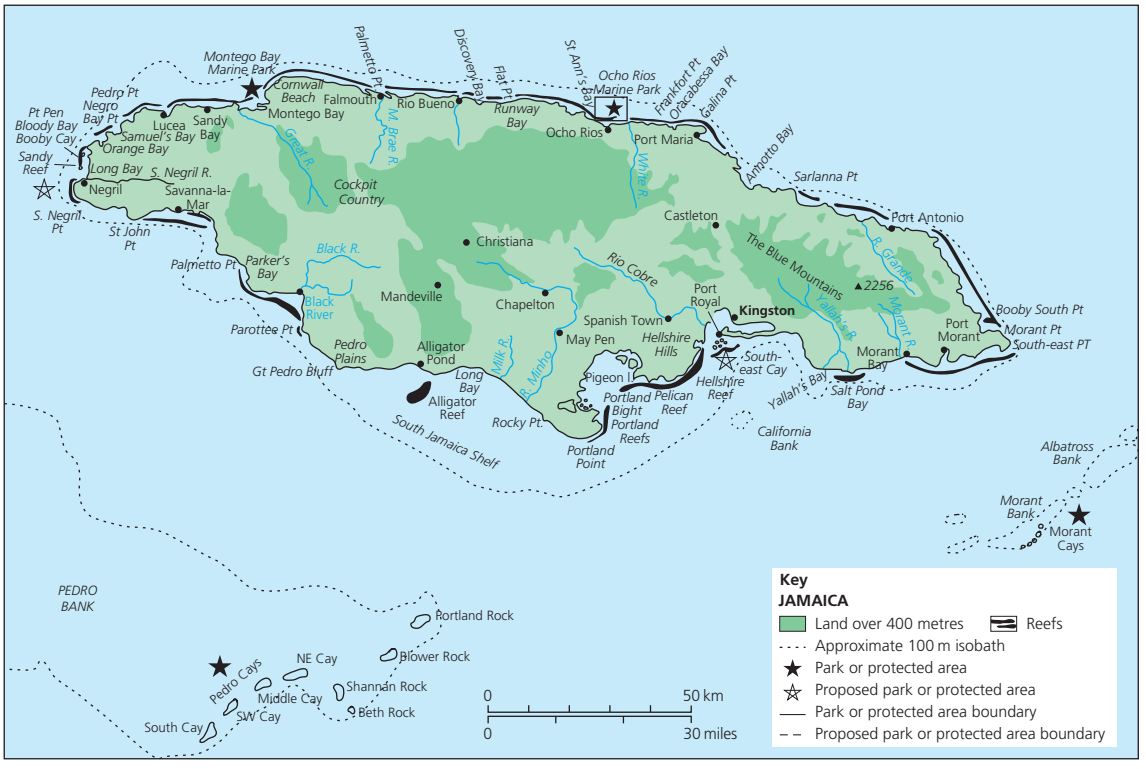


Figure 6.39 The Jamaican coral reef system

The Great Barrier Reef of Australia and the Fitzroy drainage basin scheme

The Fitzroy River basin in Central Queensland is the second biggest river basin in Australia and the largest on the eastern seaboard. The catchment area is approximately 145,000 km^2 in area (almost the size of England). It contains important freshwater and estuarine wetlands. The basin contains six major rivers which produce a total annual discharge of 5,900 gigalitres.

In 2010–2011, there were extensive floods in the drainage basin resulting in large amounts of sediment being carried downstream, emerging from the river mouth. Here, the Fitzroy River flows into the lagoon separating the southern part of the Great Barrier Reef from the Australian mainland. This part of the reef is a tourist hotspot as well as a centre of the fishing industry. The increase in sediment had the potential to cause the coral lasting damage.

The drainage basin management scheme includes:

- A series of weirs, dams and barrages along the Fitzroy and its tributaries designed to store water for use in agriculture (cotton, peanuts, corn, melons etc.), industry, power generation and domestic use (for example in the city of Rockhampton). These dams and weirs slow down the river flow and reduce its ability to carry sediment.
- The Fitzroy water quality programme (https://www.barrierreef.org/what-we-do/reef-trust-partnership/water-quality-improvement/regional-actions/fitzroy) aims to reduce the amount of sediment carried by the Fitzroy River by 50,000 tonnes per year by repairing damaged agricultural land and introducing more environmentally friendly management land use practices.

The future prospects for coral reefs

In the Caribbean as a whole, many coral reefs have collapsed and been replaced by beds of algae. The proportion of the reef covered by live coral has fallen from 50 per cent in the 1970s to just 8 per cent in 2013, changing the fish communities dramatically. In fact a third of reef-building corals are in danger of extinction and reefs the world over are in serious decline.

To lose the reefs would be to lose the planet's most diverse ecosystems. People would also suffer. More than 450 million people live close to coral reefs and rely on them as sources of tourism revenue and protein and as buffers that dampen the energy of incoming storms.

Climate change is the biggest threat (Figure 6.40). It leads to warmer and increasingly acidified oceans. Some scientists fear that the reefs are doomed to extinction. Opponents of this viewpoint state that not all corals, or all reefs, are the same. Some, like those in American Samoa, have genetic advantages that allow them to thrive in shallower warmer waters and some can recruit strains of algae that tolerate higher temperatures. Others grow in waters that are naturally acidic, where carbon dioxide seeps from the ocean floor. Corals can acclimatise to changing conditions, and there is some evidence that reefs which bleach extensively for one year are better able to handle warmer waters a decade later. Reefs can also change at the community level, shifting from sensitive species like the elkhorns to sturdy, robust ones like big boulder corals.

Humans can help by setting up marine protected areas – underwater national parks – where fishing is forbidden. Not only do they allow local reefs a chance to recover, but they can seed nearby areas with coral larvae. This will slow down the decline but if predictions of ocean warming in the next 100 years are correct then coral reefs will die.

Review question

16 Outline the threats faced by coral reefs. Categorise them into:

- local threats
- global threats.

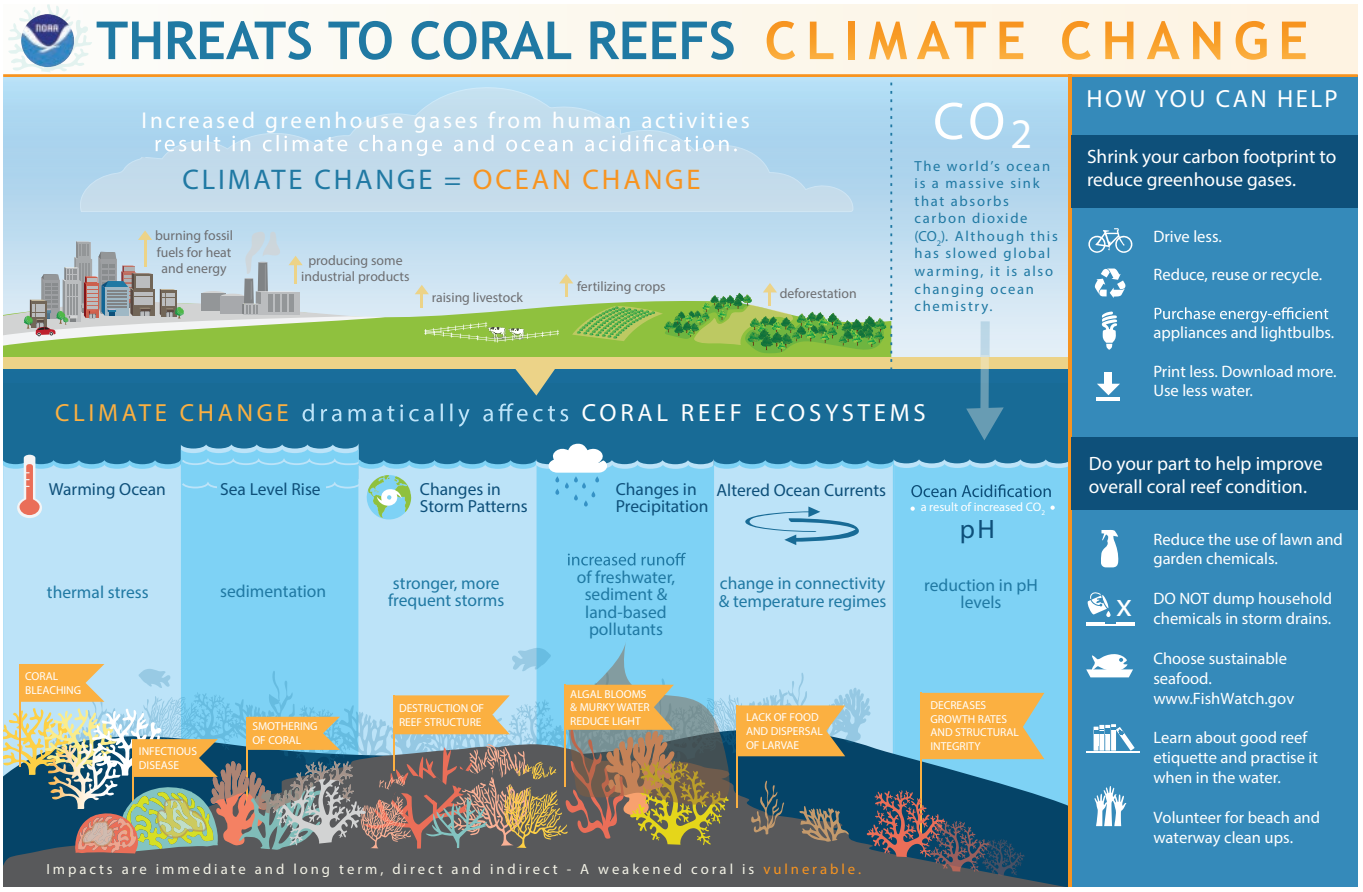


Figure 6.40 Threats to coral reefs posed by climate change

6.4 Biomes

Concept of a biome

Climate, topography and soil determine the changing character of plant and animal life, as well as ecosystem functioning, over the surface of the Earth. Although no two locations contain exactly the same assemblage of species, we can group biological communities and ecosystems into categories based on climate and dominant plant form, which give them their overall character. These categories are referred to as biomes.

A biome, therefore, is a large geographical area of distinctive plant and animal groups which are in the climatic climax stage of succession; a global-scale ecosystem. Each biome gets its name from the dominant vegetation and climate found there. There are 11 major terrestrial biomes plus several marine and aquatic ones. Each biome consists of many ecosystems whose communities have adapted to the differences in climate and the environment within that biome.

Ecosystems belonging to the same biome but in different parts of the world develop a similar vegetation structure and ecosystem functioning (including productivity and nutrient cycling). An example of this is the way in which the woodland/scrubland ecosystems characteristics of Mediterranean climates (cool, wet winters and hot, dry summers) look similar and function in the same way whether in southern California, southern France, Chile, South Africa or Australia.

The major world biomes are based upon the underlying properties of temperature and soil moisture regimes; these are largely determined by latitude, climate, topography and the native vegetation that is adapted to these local conditions. Comparing maps of the world's vegetation (Figure 6.41) and climate (Figure 6.42) zones reveals marked similarities. A third map showing zonal soil distribution at the same scale would also bear comparison.

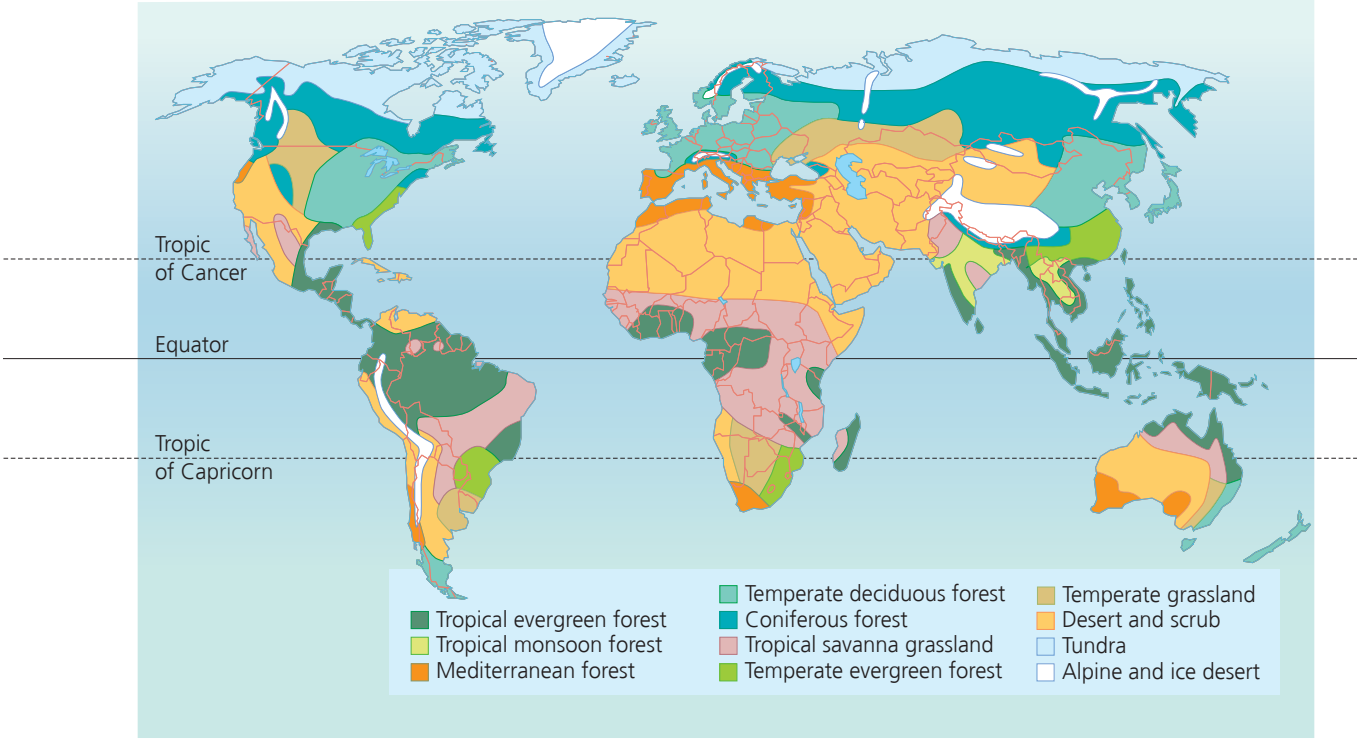


Figure 6.41 Vegetation zones of the world

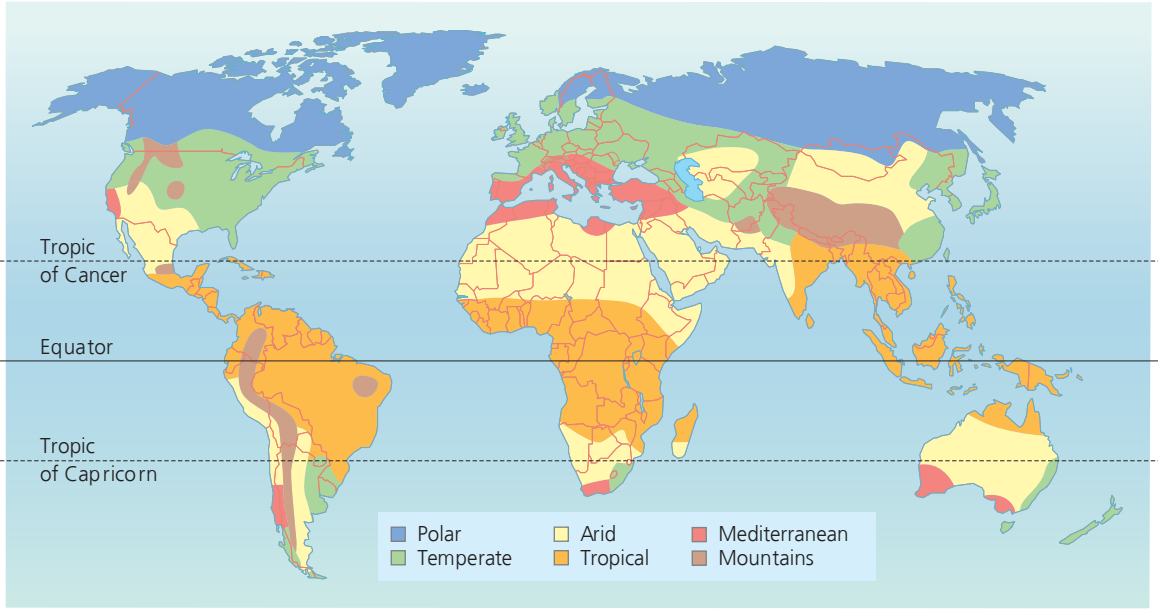


Figure 6.42 Climate zones of the world

Skills focus

Analyse the extent to which the distribution of the world's vegetation zones (Figure 6.41, page 265) mirrors that of the world's climate zones (Figure 6.42).

Robert Whittaker, an American ecologist, plotted rainfall vs temperature for points all over the globe on a single graph. He then looked at what biomes had developed at those sites, and was able to group the different biomes according to mean annual temperature and precipitation. The results are summarised in Figure 6.43.

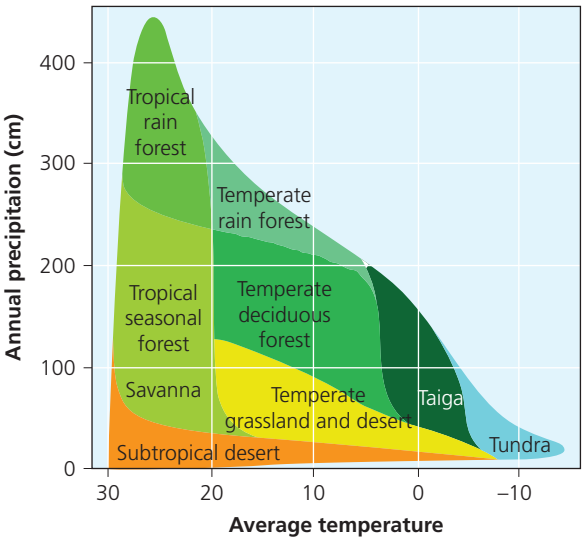


Figure 6.43 Whittaker biome diagram

Contrasting biomes–Biome 1

Tropical evergreen forest

These are mostly distributed between the latitudes 10° north and 10° south of the Equator. They are found in:

- the Amazon river basin of South America, with more than half of it in Brazil; this area holds about one-third of the world's remaining equatorial forests
- the equatorial part of Africa including large parts of the Republic of Congo, the Central African Republic and Gabon
- the Guinea coast of Africa in isolated pockets stretching from Liberia to Cameroon
- SE Asia and Oceania, for example, Malaysia, Indonesia and Papua New Guinea.

Main characteristics of the tropical evergreen forest

The climate

The equatorial climate as shown in Figure 6.44 has no seasons. It is hot and wet all year round. The average temperatures rarely go below 24°C or rise above 28°C. There is a low diurnal range of temperature; during the daytime it might reach above 30°C, but the nights remain very warm. The high temperatures are because the Sun is high in the sky all year round. They are, though, moderated by the presence of cloud.

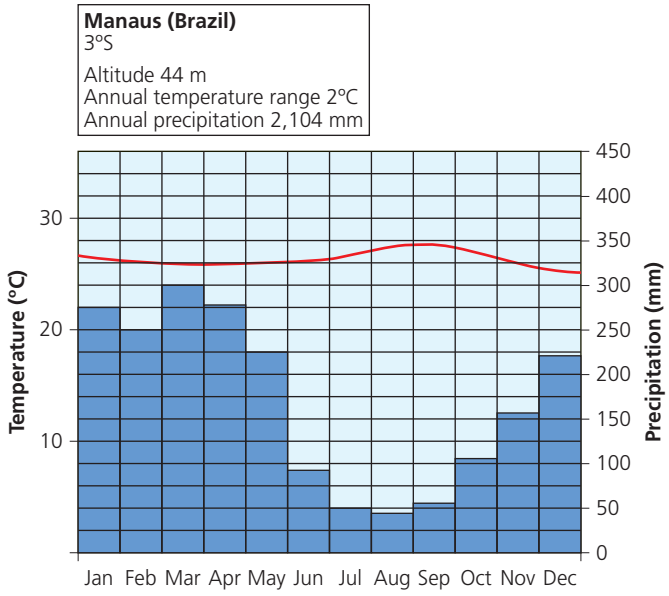


Figure 6.44 Climate graph for Manaus, Brazil

Tropical forests are to be found where the trade winds from the northeast and the southeast come together in the inter-tropical convergence zone (ITCZ, a low-pressure zone). At these low latitudes, intense solar energy from the overhead sun induces convection and rising air that loses its moisture through frequent rainstorms. Rainforests are subject to heavy rainfall, often over 2,000 mm/year. In these equatorial regions, rainfall may be all year round without apparent 'wet' or 'dry' seasons, although forests on the northern and southern fringes of the region do have a short dry season. Even in seasonal forests, the period between rains is usually not long enough for the leaf litter to dry out completely. During the parts of the year when less rain falls, the constant cloud cover is enough to keep the humidity high and prevent plants from drying out. The rainforest has been likened to a natural greenhouse.

The pattern of rainfall in the rainforest varies during the day. In the morning, skies are generally clear, though there may be a low mist. Evapotranspiration is rapid as the sun beats down on the humid forest and the low-pressure conditions allow this air to be rapidly uplifted. As the air rises it cools and water vapour condenses into clouds. These clouds continue to build until they become towering, grey, cumulonimbus clouds in the early afternoon. In the middle of the afternoon, heavy rain, often with thunder and lightning, returns the previously uplifted moisture back to ground level. The cycle begins again and the day ends as it started with clear skies.

Day and night are more or less the same length at the Equator. Dawn arrives at around 6 a.m. and night falls quickly at 6 p.m. There is little twilight. Twelve hours of sunlight every day allows photosynthesis to take place all year.

On the forest floor there is little breeze, though this can increase 40 or 50 m up. Right on the Equator there is a distinct lack of wind because this is where the trade winds converge.

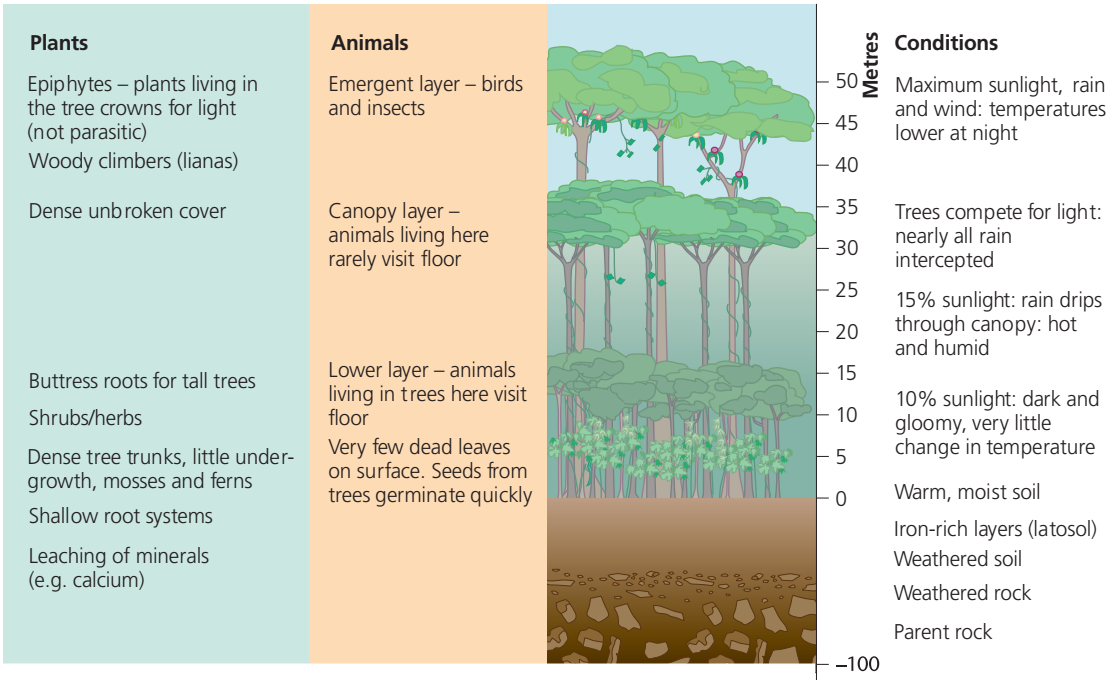


Figure 6.45 The layered structure of tropical rainforest

The vegetation

The stable climate, with evenly spread rainfall and warmth, allows most rainforest trees to be evergreen, constantly shedding and replacing leaves all year round. The moisture in the rainforest from rainfall, constant cloud cover and transpiration creates intense local humidity. Each canopy tree transpires some 760 litres of water annually; this translates to roughly 190,000 litres of water transpired into the atmosphere for every hectare of canopy trees. Large rainforests (and their humidity) contribute to the formation of rain clouds and generate as much as 75 per cent of their own rain.

Primary rainforests are very productive. The net primary productivity is 2,200 g/m²/year. This means that for every square metre of land, the forest produces 2,220 g of living matter. They are characterised by a unique vegetative structure consisting of several layers (Figure 6.45, page 267). The **canopy** refers to the dense ceiling of leaves and tree branches formed by closely spaced crowns of forest trees. The canopy is 20–40 m above the forest floor and is penetrated by scattered **emergent** trees that can reach above 50 m. Below the canopy ceiling are multiple leaf and branch levels known collectively as the **understorey.** The lowest part of the understorey, 1.5–6 m above the floor, is known as the **shrub layer,** made up of shrubby plants and tree saplings. The canopy screens light from the forest floor so that less than 5 per cent of the incident sunlight reaches it; this makes the ground layer quite open. When a tree dies naturally it brings down others as it falls, creating a small clearing in the forest. Saplings grow quickly, taking advantage of the light. The fallen tree is rapidly decomposed, assisted by detrivores, and by the hot, humid conditions.

Fungi are very common at all levels of the rainforest. They assist the detrivores in decomposing the litter.

It is estimated that 90 per cent of the animal and bird species that exist in the equatorial forest biome reside in the canopy. Since the tropical rainforests are estimated to hold 50 per cent of the planet's species, the canopy of rainforests worldwide may hold 45 per cent of life on Earth.

The soil

Rainforest soils are called **latosols** (Figure 6.46). These can be more than 40 m deep. The constant hot wet climate of the rainforest provides perfect conditions for chemical weathering of the bedrock and there is a constant supply of minerals from the parent rock to the soil. **Ferrallitisation** is the name for the process by which the bedrock is broken down by chemical weathering into clay minerals and sesquioxides (hydrated oxides of iron and aluminium).

The red colour of the soil is partly the result of the presence of iron and aluminium minerals. As there is a moisture surplus in the equatorial climate (because rainfall exceeds evapotranspiration), there is downward movement of water through the soil. Silica minerals are washed out of the horizon (A) and transported downwards by this water in a process known as **leaching** or **eluviation**. Iron and aluminium compounds are less soluble and are left behind. The iron compounds give the soil its rich red colour.

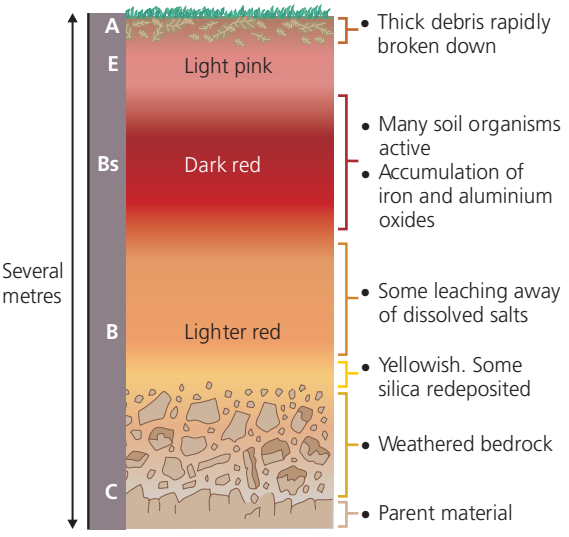


Figure 6.46 Profile of a latosol

In the rainforest areas precipitation is greater than potential evapotranspiration for much of the year. During this time there is a **soil moisture surplus**; the soil is saturated. Water cannot infiltrate into the soil and so leads to surface run-off. Rivers regularly flood during the wettest months.

Nutrient cycling

In the rainforest, nutrient cycling is very fast (Figure 6.47). The hot, damp conditions on the forest floor allow for the rapid decomposition of dead plant material, which in turn provides plentiful nutrients that are easily absorbed by plant roots. However, as these nutrients are in high demand from the rainforest's many fast-growing plants, they do not remain in the soil for long and stay close to the surface of the soil. If vegetation is removed, the soils quickly become infertile and vulnerable to erosion.

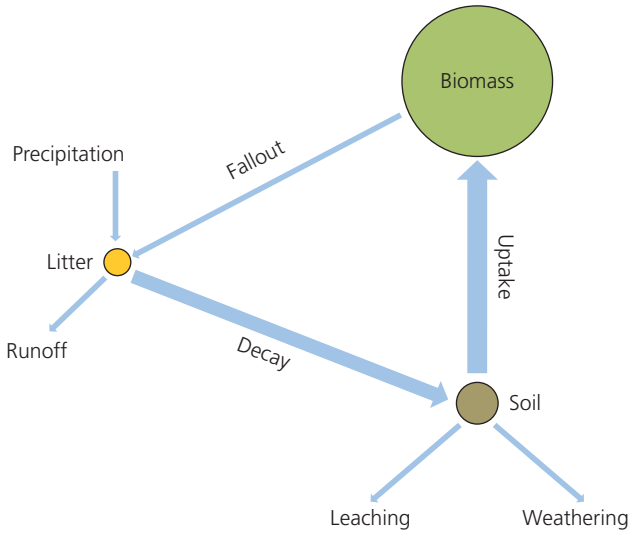


Figure 6.47 The mineral nutrient cycle for a tropical rainforest

Ecological responses

The vegetation has developed and adapted to the physical conditions of the rainforest in a number of ways:

- Leaves have drip tips and waxy surfaces. Others have jagged edges and holes in the leaves. All these allow the excess rainfall to be easily shed. This discourages growth of bacteria and fungi on the leaf surface.
- The trees grow rapidly upward towards the light. This results in tall slender trunks with few branches. All the leaves are concentrated in the umbrella-like crowns where they can absorb sunlight and photosynthesise. Because some of the trees are so tall, the tops can be subject to strong winds and so flexible trunks and buttresses help hold them up because the roots themselves are so shallow. Bark is thin because there is no need for protection against low temperatures.
- Most of the minerals needed by the trees are close to the surface of the soil and so the roots do not penetrate deep into the soil but spread outward, just below the forest floor.
- Climbing vines such as lianas are found throughout the rainforest. They have thick, woody stems and come in various lengths (up to 1,000 m) and varying shapes. They begin life on the forest floor but depend on trees for support as they climb upwards towards the sunlight. They attach themselves to trees with sucker roots, or tendrils wind themselves round a tree's trunk. When they reach the top of the canopy they often spread to other trees or wrap themselves around other lianas.
- Because the forest floor is so dark there are many epiphytic plants that grow on branches, trunks and even the leaves of the trees. Bromeliads, for example, have thick, waxy leaves that form a bowl shape in the centre for catching rainwater that is absorbed through hairs on their leaves. They can hold several litres of water and are miniature ecosystems in themselves, providing homes for several creatures including frogs and their tadpoles, salamanders, snails, beetles and mosquito larvae.
- Epiphytic orchids have aerial roots that cling to the host plant, absorb minerals and absorb water from the atmosphere.

Rainforest fauna make up more than half of the world's animals. They too have developed and adapted to the environment. Each layer of the rainforest has its own assemblage of animals adapted to each particular environment. Large predators such as the jaguar and omnivores like wild pigs live on the forest floor while monkeys, snakes, bats, birds and some amphibians live in the canopy. Some methods of adaptation include:

- camouflage, for example, the jaguar
- bright colours, for example, the poison arrow frog
- excellent climbing ability, for example, the gibbon.

Human activity and its impact on the tropical evergreen forests

The main impact of human activity is forest clearance. Although deforestation in Brazil is much lower than at its peak in 2004, it has increased since 2012. Between August 2017 and July 2018, 7,900 square kilometres were deforested in Brazil – a rise of 13.7 per cent – and the largest area cleared since 2008. Up to 95 per cent of this deforestation is illegal and about 80 per cent of it becomes pasture for cattle. Figure 6.48 shows how patches of forest clearance are slowly merging together.

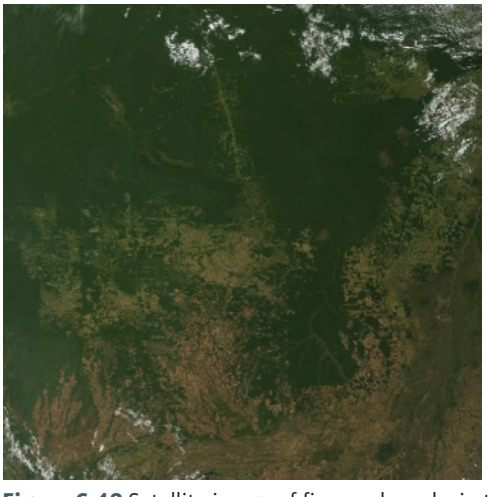


Figure 6.48 Satellite image of fires and smoke in the Amazon Basin rainforest in 2019

In Brazil, the INPE, or National Institute of Space Research produces deforestation figures annually. The figures for 2003–2019 are shown in Table 6.5.

Table 6.5 Estimated rainforest loss in Amazonia

Year	Estimated remaining forest cover in Amazonia (km^2)	Annual forest loss (km^2)	Percent of 1970 cover remaining	Total forest loss since 1970 (km^2)
2003	3,458,885	25,396	84.4	641,115
2004	3,431,113	27,772	83.7	668,887
2005	3,412,099	19,014	83.2	687,901
2006	3,397,814	14,285	82.9	702,186
2007	3,386,163	11,651	82.6	713,837
2008	3,373,252	12,911	82.3	726,748
2009	3,365,788	7,464	82.1	734,212
2010	3,358,788	7,000	81.9	741,212
2011	3,352,370	6,418	81.8	747,630
2012	3,347,799	4,571	81.7	752,201
2013	3,341,908	5,891	81.5	758,092
2014	3,336,896	5,012	81.4	763,104
2015	3,330,689	6,207	81.2	769,311
2016	3,322,796	7,893	81.0	777,204
2017	3,315,849	6,947	80.9	784,151
2018	3,308,313	7,536	80.7	791,687
2019	3,298,551	9,762	80.5	801,449

Source: based on data from the National Institute for Space Research and the Food and Agriculture Organization (FAO)

Skills focus

a What type of graph would be most suitable to display the information in Table 6.5?

b Construct your chosen graph using the figures in Table 6.5. Describe the trends shown by the graph.

Brazilian deforestation is strongly correlated to the economic health of the country: the decline in deforestation from 1988 to 1991 matched the economic slowdown during the same period, while the fast rate of deforestation from 1993 to 1998 paralleled Brazil's period of rapid economic growth.

Reasons for forest clearance

- **Poor subsistence cultivators**: Shifted cultivators (a term used to describe small-scale farmers and their families) migrate from other areas to the last unoccupied lands available in the forest. They generally have little understanding of the ecology of tropical forests and so they practise unsustainable forms of exploitation. The numbers are somewhere between 200 million and 600 million, accounting for roughly two-thirds of deforestation. They represent one of the biggest human migrations in such a short space of time.
- **Fire**: Though these controlled fires are intended to burn only limited agricultural plots, they frequently char pristine rainforest. There is often a surge in wildfires during dry years. In 2019, Brazil's space research centre reported that there were more than 53,000 fires in the Amazon biome – the highest number since 2010.
- **Commercial resource exploitation**: Timber, such as mahogany from Brazil and teak from Myanmar, has been removed in an unsustainable and illegal manner.
- **Cattle pasture**: A small number of large landowners clear vast sections of the Amazon rainforest for cattle pasture to feed the world's increasing demand for cheap beef. Studies from Brazil's INPE institute calculate this amount to equal 45 million hectares or 62 per cent of the total cleared area. Permanent agriculture comprises a smaller percentage of the cleared land at 3.5 million hectares.
- **Soya production**: Soya has become one of the main contributors to deforestation in the Brazilian Amazon. Brazilian scientists have developed a new strain of soya bean that flourishes in the rainforest climate which, combined with high prices, has served as an impetus to expanding soya cultivation.

  Soya farms cause some forest clearing directly, but they have a much greater impact by occupying already cleared land, thereby pushing ranchers and subsistence farmers deeper into the forest. Soya farming also provides a key economic and political impetus for new highways and infrastructure projects, which accelerate deforestation by others.

  The greatest increase in area planted to soya bean was in Mato Grosso, the Brazilian state with the highest deforestation rate (40 per cent of new deforestation).
- **Government policy**: Favourable taxation policies and government subsidised agriculture and colonisation programs have encouraged destruction of the Amazon rainforest. Low taxes overvalue agriculture and pasture and make it profitable to convert natural forest for these purposes.

Impacts of deforestation

Impacts can be physical, economic, social and cultural. They mainly occur on a local scale, but can be global.

- As habitats shrink, plant species become endangered and the food chain within the forest is disrupted. Some animal species, for example, tigers and orang-utans, are threatened by extinction.
- The vegetation protects the latosol soils from the regular heavy tropical downpours. Once the trees are removed the topsoil is open to erosion and to leaching of nutrients and minerals. Run-off causes sediment to block river channels and increases flooding.
- During the dry season, areas downstream of deforestation can be prone to long droughts which interrupt river navigation, wreak havoc on crops and disrupt industry.
- The microclimate of the forest is disturbed by deforestation – the daily water cycle of rapid evapotranspiration followed by afternoon precipitation cannot occur and there is less cloud cover, so a greater temperature range occurs. Rainforest is also the source of some of the rainwater that falls on agricultural areas outside the forest. Forest removal could lead to drought and agricultural failure.
- Burning associated with forest clearance leads to local air pollution. In 2019, fires in the Amazon rainforest produced such heavy smoke that it darkened the skies over the city of São Paulo. There was a rise in deaths on days when air pollution increased, caused by people inhaling into their lungs a particulate matter (very fine ash).
- The forests' ability to absorb the carbon dioxide is reduced. At the same time, burning trees release carbon dioxide, leading to increased atmospheric carbon dioxide and climate change.
- Deforestation can have economic benefits – income from mining, farming and exports of hardwood. However, the culture of indigenous people is destroyed and they may be forced to move from their land.

Development issues

Population change

There are about 50,000,000 **indigenous** people living in and depending on the world's rainforests. The 'slash and burn' or 'shifting' agriculture practised by these people is sometimes blamed for the destruction of large areas of tropical forest. In fact, they have developed the only demonstrably sustainable agriculture for tropical forests.

Many of the rainforests' native societies have already been destroyed. In Brazil alone, 87 tribes were wiped out by Western diseases or genocide in the first half of the twentieth century. In the Indonesian province of Irian Jaya, the military is estimated to have killed up to 150,000 tribal people under the guise of counter-insurgency measures. In the Congo Basin the Baka, BaAka and Bakola pygmies were formerly hunter-gatherers but are now becoming increasingly settled, both through their own choice and because of government policies.

Whilst the indigenous population is in decline, the overall population of the rainforests is increasing. In Amazonia, according to the Brazilian Institute of Geography and Statistics (IBGE), Manaus has grown to an estimated population of 2.2 million in 2020. It is the most populous city of the Amazon rainforest. In fact, the bulk of population in the Amazon Basin is found in cities which have emerged from the surrounding rainforest and become significant population centres. Outside the cities and towns, the region is sparsely populated. Also, whereas Brazil has mostly experienced steadily declining rates of population increase, the Amazon region continues to see recurrent spikes and contractions in growth rates, reflecting waves of migration to different parts of the region. This growth is mirrored in the Peruvian Amazon area.

Economic development

The national economies of the nine countries that have territory within the Amazon basin have all been boosted by the activities that have led to forest clearance. Mining, forestry, oil exploration and agriculture have brought economic benefits in the form of local employment and tax revenue. Some of this money has been re-invested in infrastructure (mainly roads) that has accelerated development and forest clearance.

In Brazil, in and around the city of Manaus there is an international airport, a free economic zone (Industrial Park of Manaus) and a river port (a free port) that can cater for sea-going vessels despite being 1600 km inland. Through the port, it is able to export Brazil nuts, timber, rubber, jute, rosewood oil and

other forest products. Industries include electronics, chemical and petrochemical plants, with multinational manufacturers (LG, Samsung and Philips) present. The city also attracts crowds of tourists for land and boat trips into the jungle. The increasing workforce led to the decision to build the Rio Negro bridge, the first bridge over the Amazon, to open up the south bank of the river to development (Figure 6.49).

Figure 6.49 The Rio Negro Bridge over the Amazon

Opened in late 2011, the bridge gives more commuters access to dormitory towns where developers are building more housing. Environmentalists fear that the bridge, combined with new gas pipelines, roads and rising populations, could open up the rainforest to further destruction.

Research has shown that forest clearance may eventually lead to a collapse in agriculture because of regional climate changes (less rainfall, higher temperatures). In Brazil, based on existing trends of deforestation, loss of carbon sequestration and related feedbacks on rainfall, temperature and biomass, the researchers project a 34 per cent fall in pasture productivity and a 28 per cent decline in soya bean yields by 2050.

Agricultural change

The traditional rainforest economy is a mix of **hunting/fishing/gathering** and **shifting cultivation**. Numbers of people were so small and the area so large that these activities remained sustainable. Small areas of forest were cleared and farmed for two to four years. As the soil became less fertile new areas were cleared and the old ones abandoned. Soil fertility had time to recover before the area was again cleared many years later.

Population pressure from areas outside the rainforest has increased the demand for land and food. This has resulted in two categories of farmer in the rainforest: the **shifted cultivators** and **large corporations**.

On poor land that is only marginally useful for agriculture, farmers use slash-and-burn techniques to clear it, but on a much larger scale than traditional practices. Instead of burning only a few hectares, they burn hundreds to thousands of hectares after felling a tract of forest and leaving it to dry. This releases the minerals locked up in vegetation and produces a layer of nutrient-rich material above the otherwise poor soil. The cleared area is quickly planted and supports vigorous growth for a few years, after which the nutrient stock is depleted and large amounts of fertiliser are required to maintain production. This fertiliser can then be washed into rivers, affecting fish and other aquatic life. The land is abandoned when the soil is so deprived of nutrients that even fertiliser cannot restore it. It reverts to scrub. Drought-resistant grasses may move in or cattle ranchers may plant imported African grasses for cattle grazing. The land is now only marginally productive and only a limited number of cattle can subsist in the area.

Where the land is suitable for agriculture, large, single, cash crops like rice, citrus fruits, oil palms, coffee, coca, opium, tea, soya, cacao, rubber and bananas are cultivated. Some of these crops are better adapted to such conditions and last longer on cleared forest lands. Problems with this type of monoculture (single crop plantations) in the tropics, besides the loss of forest are:

- planting of a single crop makes the crop highly vulnerable to disease and pests. In natural rainforest, widespread infestations are rare because individuals of a given species are widely dispersed
- the planting of monocultures can be economically risky because of price fluctuations common in international commodities markets
- a single cold spell or drought can devastate a substantial segment of the agricultural economy.

Agricultural companies are clearing more rainforest than ever before, especially in the Amazon, where large tracts of rainforest are being converted into soya bean plantations. In Asia, especially Malaysia and Indonesia, large areas of rainforest are being cleared for oil palm plantations to produce palm oil, which

is used widely in processed food, cosmetics and soap (Figure 6.50). Today palm oil is found in some 50 per cent of packaged snack foods, a proportion that is growing because palm oil is the cheapest type of vegetable oil.

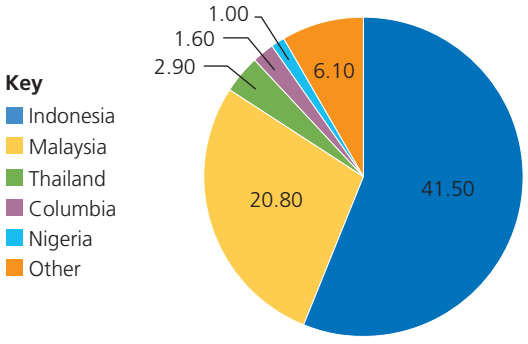


Figure 6.50 Global palm oil in million metric tonnes (MMT) and global market share

Source: U.S. Department of Agriculture

Extensive cattle ranching has caused deforestation in virtually every Amazon country; it accounts for up to 80 per cent of current deforestation. Brazil has 88 per cent of the Amazon herd, followed by Peru and Bolivia. Grazing densities are low as is productivity. There have been various attempts to slow down the rapid growth in ranching but poor communications, government inefficiencies and corruption have reduced their effect.

Implications for biodiversity

The destruction of natural habitats deprives the animals of the rainforest of the basis for their existence, causing an irreversible loss of biological diversity. One example, the orangutan, is particularly vulnerable because they are dependent on large contiguous forest areas. Fragmentation occurs when a once-continuous habitat becomes divided into separate fragments, resulting in limited breeding opportunities, reduced food resources and increased human conflicts for wildlife. In their search for food, orangutans often get lost in the palm oil plantations, where they are regarded as pests. They have been shot or clubbed to death by the palm oil plantation workers. According to the UN, there was a risk that no wild orangutans would remain outside of protected areas by 2020.

Sustainability

One example of how rainforests can be managed sustainably is taking place in Brazil, in parts of Amazonia such as Paragominas, where much local forest is either destroyed or damaged. This is because the usable timber would be ripped out of a stretch of forest and the rest would then be burned. When farming was tried on this cleared land, it was frequently found to be unprofitable. As a result, there are now about 165,000 km² of abandoned land in Brazilian Amazonia.

Changing attitudes has meant that timber firms are coming to see unharmed woodland as an asset that, properly managed, can yield a good income forever. Studies have shown that sustainable management of forests, also known as **reduced-impact logging (RIL)**, can be more profitable than the conventional methods of timber extraction. One such study, conducted near Paragominas, found that RIL was 12 per cent cheaper than conventional logging.

In RIL schemes, the area to be exploited is divided into up to 30 blocks, one of which has timber extracted each year, before being left alone for 29 years. This is enough for the forest to regenerate successfully, because in addition to rotation, the schemes take care to leave the oldest specimens of the exploited species standing. As well as providing cover from the tropical sun, the spreading branches of these tall trees re-seed the block with new specimens. RIL reduces the damage further by plotting the position of each block's valuable trees on a computer, which then works out the shortest set of access roads that needs to be carved out to remove the felled trees. Lumberjacks are also taught ways of felling trees that avoid damaging those around them.

Biome 2

The savanna grassland biome

It is to be found in:

- Northern-central Australia, straddling the tropic of Capricorn
- The Cerrado grassland of southern Brazil and the Llanos of Colombia and Venezuela
- Parts of India and Pakistan straddling the tropic of Cancer
- Sub-Saharan Africa, both north and south of the Equator. In East Africa, in the upland areas, these grasslands are found right on the Equator. It is the African Savanna that most people associate with this biome and which will be used as an example for the rest of the chapter.

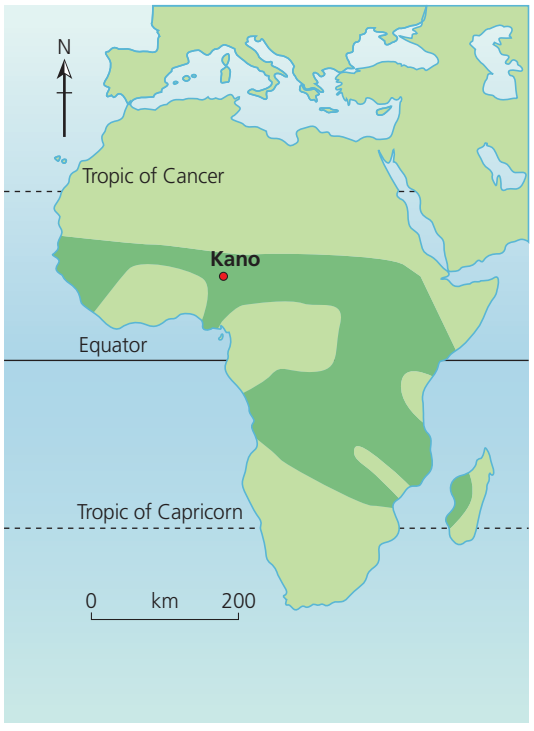


Figure 6.51 The savanna regions of Africa

Main characteristics of the savanna grassland biome

The climate

The typical climate is a tropical wet and dry climate that shows seasonal variations in wind direction, precipitation and temperature. It is transitional between the equatorial rainforests, where rain can be expected all year, and the hot deserts, which have minimal precipitation. Variations occur with increasing distance from the Equator. Ninety per cent of the annual precipitation falls within the wet season (Figure 6.52) and although the sun is higher in the sky, the rain clouds reflect insolation back out of the atmosphere and reduce the temperatures. During the dry season, with lack of cloud, heating can be intense and temperatures high.

The savanna areas of Africa (Figure 6.51) lie between:

- the **subtropical high pressure zone** where the falling air is very dry. (This has a desert climate)

and

- the **low pressure** equatorial region where there is rising air. This is called the **inter-tropical convergence zone (ITCZ)** because it is located where the **NE trade winds** meet (converge with) the **SE trade winds.**

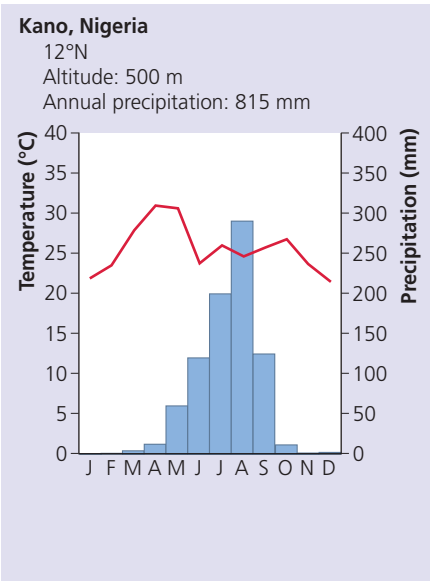


Figure 6.52 Climate statistics for Kano, Nigeria

The low pressure ITCZ migrates throughout the year. This migration is, in part, due to the migration of the overhead sun. The higher the sun is in the sky, the greater the heating effect.

In June, the sun is overhead at the Tropic of Cancer (23½ °N). The ITCZ then migrates northwards. It can reach as far north as 20 °N but often does not. By December, the overhead sun has moved to the Tropic of Capricorn (23½ °S). The ITCZ follows the overhead sun in East Africa, but because of the shape of the continent and the fact that the land heats up more than the sea, the ITCZ remains at between 5 °N and 10 °N in West Africa.

The ITCZ draws in both hot dry air from the sub-tropical desert and hot moist air from the warm Atlantic Ocean.

In January, this moist air is drawn on to the coastal area of West Africa. As it reaches the ITCZ, it rises and heavy **convectional rainfall** occurs. To the north of the area, air is drawn south-eastwards from the sub-tropical high pressure area of the Sahara desert. This is dry and dusty. It is a regular occurrence and is called the **harmattan**. This evaporates any moisture and is responsible for the dry season and drought of the inland part of West Africa.

As the ITCZ migrates northwards, this moist air is drawn further north, causing the belt of rain to move with the ITCZ. By July, the ITCZ reaches its northern limit.

The result of this is that there are variations in precipitation and temperature such as:

- on the equatorial rainforest margins more than 1,000 mm of rain falls per year; this is the rainy season and can last up to 11 months
- on the desert/semi-arid margins there is less than 500 mm of rainfall per year with only 1 or 2 months of rainy season. As one moves further away from the Equator the reliability of the rainfall decreases
- on the equatorial rainforest margin, temperatures range from 22 °C in the wet season to 28 °C in the dry season
- on the desert margins the temperatures range from 18 °C in the wet season to 34 °C in the dry season.

Skills focus

Contrast the savanna climate as illustrated by Figure 6.52 (page 274) with that of Manaus (Figure 6.44, page 267.

The soil moisture budget

The vegetation of the savanna depends on the amount of water available (the moisture budget) for plant growth. This is illustrated by the use of a soil moisture diagram.

Figure 6.53 shows that precipitation is greater than potential evapotranspiration between July and September, whereas the reverse is true between October and June. There are four distinct periods that can be seen in the soil moisture budget for this area:

- **Soil moisture** recharge occurs through July and early August and this is when precipitation first becomes greater than evapotranspiration. Rainwater begins to fill the empty pores in the soil. When they are full the soil is said to have reached its field capacity.
- **Soil moisture surplus** occurs in late August and September. When the soil becomes saturated rainwater has difficulty infiltrating the ground. This causes surface runoff and explains the high river levels and flooding.
- **Soil moisture utilisation** occurs from October, as evapotranspiration begins to exceed precipitation. There is more water evaporating from the ground surface and being transpired by plants than is falling as rain. Water is also drawn up the soil by capillary action and this leads to further evapotranspiration.
- **Soil moisture deficit** occurs by December when the soil moisture is used up and there is a water deficit. Plants can only survive by being drought resistant or through irrigation. This period lasts until precipitation again becomes greater than evapotranspiration in early July and soil moisture recharge can begin.

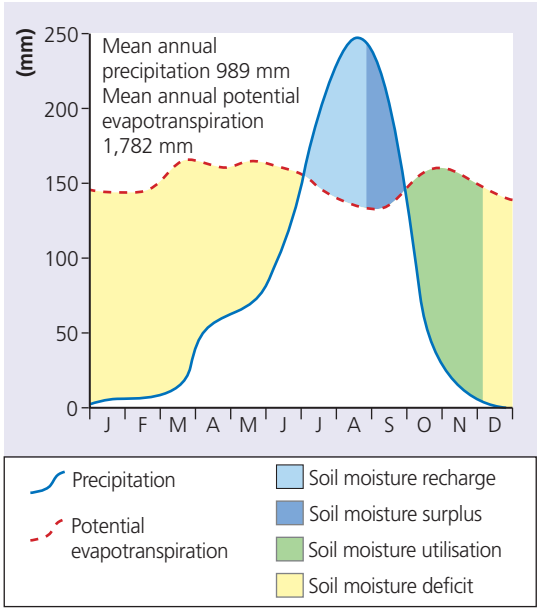


Figure 6.53 Soil moisture for Navrongo, northern Ghana (11°N)

This area in northern Ghana therefore shows the following characteristics:

- a lengthy period of moisture deficit
- a short period of moisture surplus
- total annual potential evapotranspiration greater than total annual precipitation.

The soil

The rapid downward movement of water in the soils during the rainy season washes away (leaches) the silica from the soil, leaving a layer rich in iron and aluminium oxides that is characterised by a red colour. When the thin topsoil is removed, the exposed layer of iron and aluminium can harden to form a cement-like crust called a **laterite** (Figure 6.54).

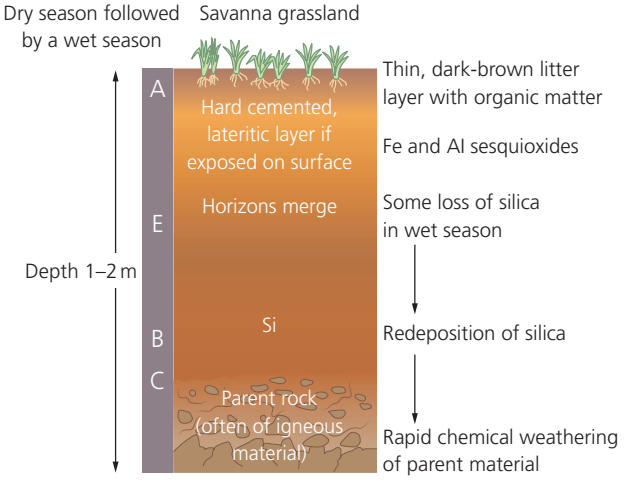


Figure 6.54 Profile of a ferruginous soil (laterite)

Nutrient cycling

Ferruginous soils are characterised by their infertility, although there are exceptions. The very old and already well-weathered parent material has, in many areas, been leached of nutrients by the very high rainfall of the wet season. Those areas of greatest wet season rain tend also to be the most infertile.

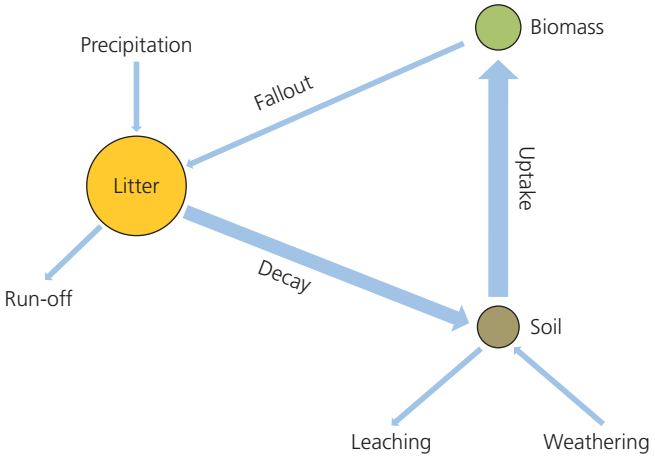


Figure 6.55 The mineral nutrient cycle for savanna grassland

During the dry season, plants die back and litter builds up which, if not destroyed by fire, forms a thin layer of humus. This low level of nutrients is further reduced by aerobic bacteria which are particularly active in areas with a long wet season and where the temperature remains above 25°C for substantial periods of time. The activity of these bacteria under such conditions occurs at a rate above that of plant growth, thus breaking down plant matter faster than it is produced (Figure 6.55).

During the wet season leaching removes nutrients from the soil and that can result in a hard crust forming that hinders plant growth. Biomass varies seasonally, with elephant grasses growing up to 2 metres high in the wet season.

Ecological responses

Adaptations by vegetation

Figure 6.56 shows that the vegetation of the wetter areas consists of tall coarse grasses (elephant grass) with many deciduous trees. It is known as the tree savanna. In drier areas towards the desert margins, shorter tussock grass becomes dominant, with bare soil between the tufts of grass. These areas are known as the grassland and shrub savannas.

Trees in these areas tend to be deciduous, losing their leaves in the dry season, although some evergreens are also present. These have hard leathery leaves to reduce transpiration losses. Other plants are microphyllous (small-leaved), for the same reason.

In the **tree savanna**, a parkland exists. Here isolated trees have low umbrella-shaped crowns that shade root areas and reduce soil moisture evaporation. The trees show xerophytic characteristics (adaptations to dry surroundings), with dense cell fluids, hard waxy leaves, thorns and protected stomata, which all reduce water loss. The two main trees of the savanna are:

- Acacia (Figure 6.57), which has a crown structure often flattened by the trade winds. It loses its leaves in the dry season. It has long tap roots that are able to get water from deep underground. There are also roots close to the surface that are able to make use of dew from the area shaded by the tree's crown.

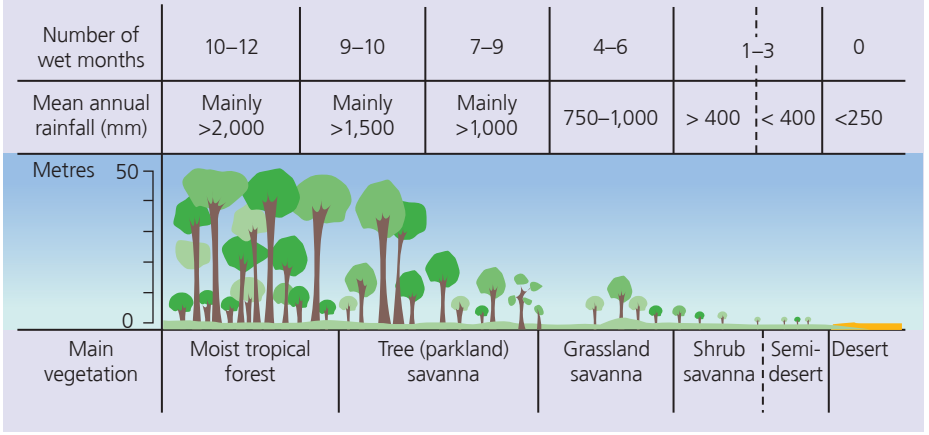


Figure 6.56 Section south to north through West Africa showing variations in vegetation

Figure 6.57 Acacia in the savanna of Amboseli National Park, Kenya

- Baobab ('upside-down tree'), which has a thick spongy trunk, long tap roots and bears leaves for only a few weeks. Like the acacia, the baobab is pyrophytic – it can withstand fire, mainly because of its insulating bark.

In the **grassland savanna** the grasses between the trees are shorter and sparser. They are perennial, dying back during the dry season and regrowing from root nodules when it rains. The grasses are tussocky, enabling them to retain some moisture. The naturally created straw dies down and protects the roots. They also have adaptations that discourage animals from grazing on them; some grasses are too sharp or bitter tasting for some animals, but not for others. This means that all animals have something to eat. Many grasses grow from the bottom up so that growth tissue does not get damaged by grazers.

In the **shrub** or **scrub savanna** there are many acacia trees, thorn bushes and short tufted grasses. Many species generate short stems from a single stock, with deep, branched roots and dormant seeds that compete for water. In some plants, even the stems may be capable of photosynthesis, so plants can have fewer leaves and retain more water. Some grasses are feathery and wiry and turn their blades away from the strong sun to reduce water loss.

The animal life (fauna)

Most of the animals on the savanna have long legs (for example, wildebeest) or wings (for example, swifts) that enable them to go on long migrations. Many burrow underground (for example, the meerkat) to avoid the heat or raise their young. The savanna is a perfect place for birds of prey like hawks and buzzards. The wide, open plain provides them with a clear view of their prey, hot air updrafts keep them soaring, and there is the occasional tree to rest on or nest in. Animals do not sweat to lose body heat, so they lose it through panting or through large areas of exposed skin, or ears like those of the elephant.

The savanna has a large range of highly specialised plants and animals. They all depend on each other to keep the environment in balance. There are over 40 different species of hoofed mammals that live on the savannas of Africa. Up to 16 different species of browsers (those who eat leaves off trees) and grazers can coexist in one area. They do this by having their own food preferences, browsing/grazing at different heights, having different times of day or year to use a given area and having different places to go during the dry season.

These different herbivores provide a wide range of food for carnivores such as lions, leopards, cheetahs, jackals and hyenas. Each species has its own preference, making it possible to live side by side and not be in competition for food.

Other fauna include locusts, which can destroy large areas of grassland with devastating speed, and termites, which aerate the soil and break down up to 30 kg of cellulose per hectare each year. In some areas up to 600 termite hills per hectare can be found; this can have a significant effect on the upper layers of the soil.

Human activity and its impact on the savanna in Africa

Most of the African savanna vegetation types are used for agriculture, mainly grazing by cattle, goats, sheep or game. In some areas of southern Africa, crops and subtropical fruit are cultivated. Urban growth is limited because the hot, moist climate and diseases (such as sleeping sickness and malaria) hinder dense human settlement.

Human activity has two main effects on the vegetation:

- Grass is burnt off to ensure better growth of young grass next season for grazing. This regular burning makes it difficult for young trees and bushes to become established. Their place is taken by herbaceous plants and by the few indigenous woody plants that can survive fire, like acacia and baobab, which are therefore common in the savanna.
- Woody plants are killed by cattle eating their foliage. Thorny, animal-repellent trees and shrubs, such as acacia, therefore become numerous.

There is a belief among some biogeographers that humans have had a much greater influence than climate on the development of savanna vegetation. Some even suggest that grassland may therefore not be the climatic climax community but a plagioclimax.

For thousands of years, herders of the East African savannas have penned their cattle overnight in brush-walled corrals, called bomas. These remain in use for about a year, resulting in tons of manure that fertilises these small areas. After abandonment, a lush carpet of grass springs up. These fertile 'glades', some as large as a football field, remain distinct from the surrounding area for over 100 years. Researchers in Kenya found that trees close to the edges of glades grew faster and were generally larger than trees elsewhere in the savanna. They also found more animal life there. The conclusion was that for many years after the pastoralists move on they leave fertile footprints across the landscape that significantly alter the dynamics of the entire ecosystem.

Conservation of the savanna is variable. There are huge areas of National Parks (e.g. the Kruger National Park in South Africa) as well as game reserves such as the Central Kalahari Game Reserve. Much of the area is used for game-farming and can thus be considered effectively preserved, provided that sustainable stocking levels are maintained. The importance of tourism and big game hunting in the conservation of the area must not be underestimated.

Development issues

Population change

Population densities for the African savanna areas have traditionally been low. The nomadic way of life does not support large numbers of people and permanent settlers are presented with a variety of challenges from endemic disease, unreliable water supply and poor soils.

Recently, as human population carries on growing, the risk of human–wildlife conflict increases. This is particularly true of parts of the African savanna where new settlements have expanded into traditional wildlife habitats. One conflict is between settlers and the growing number of elephants. Many poor farmers have taken drastic action against elephants to protect their crops. This includes shooting, poisoning and other physical attacks. They also turn a blind eye to poaching. This has led to the slaughter of up to 40,000 elephants a year.

A more humane form of management has recently been developed in Kenya. Wire fencing is erected around agricultural plots. Honey bee nests are built on the fences. As the elephant brushes against the wire it disturbs the bees and they become angry. The sound of the bees frightens the elephants away.

This method also allows the farmer to harvest the honey and make an income from what had formerly been a problem.

Agricultural change

Population growth outside the savanna areas of Africa is putting pressure on the continent's food resources. By 2050, the continent's population is expected to double and per capita GDP income will triple. This combination of people and prosperity means that Africa's food needs will increase by 400 per cent. It is estimated that satisfying this increased demand will require at least 140 million hectares of new farmland. This is made worse by growing efforts to turn Africa into a source of food and fuel for the rest of the world.

Scientists believe that growing crops like maize and soya in the savanna region would not be carbon efficient. Indeed, converting Africa's wetter savannas will have high environmental costs, which is unsurprising if we look at savannas elsewhere, such as Cerrado and Gran Chaco; these regions have recently undergone rapid agricultural transformation, resulting in large carbon and biodiversity losses.

Economic development

Although there are nodes of development based on natural resources such as metal ores and diamonds, Africa's savanna areas are dominated by agriculture and game conservation. Throughout the twentieth century game parks and reserves have been created. These in turn have attracted tourist money and created prosperity and jobs for local people. Examples of these reserves include the Kruger National Park of South Africa, the Hwange National Park in Zimbabwe and the Serengeti National Park in Tanzania.

Implications for biodiversity and sustainability

Two opposing forces are at work in much of the African savanna. The need to grow food puts pressure on the land and reduces biodiversity. The attraction of tourist income creates the need for reserves and protection programmes to maintain biodiversity.

Review question

17 Removal of topsoil often happens.

- Under what circumstances might this happen?
- What effects could it have on the landscape?
- What effects could it have on agriculture?

Case study of a specified region experiencing ecological change: The Sundarbans

Links

See Chapter 3, Coastal systems and landscapes, pages 131–5, for further background information.

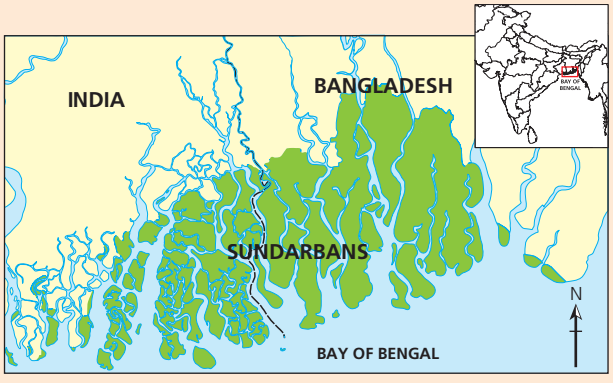


Figure 6.58 Location of the Sundarbans

Geographical setting

The Sundarbans is a region in the Bay of Bengal composed of vegetated low lying islands with elevations ranging from 0.9–2.1 m above mean sea level, interspersed with a maze of tidal waterways (Figure 6.58). The region is spread across West Bengal in India and southern Bangladesh and is famous for its unique mangrove forests which, at over 10,000 km², are the most extensive in the world. The Sundarbans are part of the delta formed by the confluence of the Ganges, Padma, Brahmaputra and Meghna rivers. They have been designated by the Worldwide Fund for Nature (WWF) as one of 867 **ecoregions**. An ecoregion is a contiguous area characterised by well-defined similarity in flora and fauna as well as geomorphology, climate and soils. Ecoregions do not follow any political boundaries or landscape alterations by humans. The Sundarbans have also been recognised by UNESCO as a World Heritage Site.

The Sundarbans forest is about 10,000 km² across India and Bangladesh, of which 60 per cent lies in Bangladesh. It is home to many rare and globally threatened wildlife species such as the estuarine crocodile, the royal Bengal tiger, the water monitor lizard, the Gangetic dolphin and the olive ridley turtle.

In June to September the monsoon regularly brings heavy rains and frequent devastating cyclones that cause widespread destruction. Annual rainfall can exceed 3,500 mm and daytime temperatures can exceed 48°C during these monsoon months.

The Sundarbans have a wide range of flora (334 plant species) and fauna (693 species), including 210 species of fish and 38 species of crustaceans.

The Sundarbans delta is an ecologically fragile and climatically vulnerable region that is home to over 4.5 million people.

Threats

The Sundarbans are subject to threats from both human activity and natural occurrences. These are summarised in Figure 6.59 (page 280).

Threats to the Sundarbans have resulted in:

- a decline in standing volume of the two main commercial mangrove species
- depletion of its biodiversity. Five mammal species – the Javan rhinoceros, water buffalo, swamp deer, gaur and probably hog deer – have become locally extinct since the beginning of the twenty-first century.

Issues and/or impacts arising from the environment and/or changes to that environment

- Between 2001 and 2008, the area under agriculture in the Indian Sundarbans had reduced from 2,149 km² to 1,691 km² and it continues to fall.
- Because of rising sea levels and shrinking forest, humans and tigers are fighting for space. Farmers are forced into the forest to hunt for honey, fish or collect crabs, putting them at risk of tiger attack.
- As a result of over-exploitation of aquatic species in the last 15 years, coastal fishing has seen a decline in catch-per-unit effort – from 150–200 kg per haul to 58–65 kg per haul.

Responses to the issues outlined above

- The entire Sundarbans is reserved forest, established under the Indian Forest Act, 1878, when both India and Bangladesh were part of the British Empire. Since then, a number of wildlife refuge areas have been created on both sides of the border. These sanctuaries are supposed to prohibit the entry of people, cultivation of land, damage or destruction of vegetation, hunting or capturing of wild animals, introduction of exotic species, straying of domestic animals, causing of fires and pollution of water. Much of this is not working as population pressure grows.
- A large part of this ecoregion has been designated as a World Heritage Site both in India and Bangladesh. This is because of the wide variety of species present and the fact that the area is an example of delta formation and the subsequent colonisation of the newly formed deltaic islands and associated mangrove communities.
- Temporary fishing camps are established every season by local people in order to ensure a fresh supply of fish to the local markets. Men also go out into the forest to hunt, collect honey or cut firewood. Over the years, thousands of people have been killed by storms, eaten by tigers and crocodiles, or died of disease. The widows of those killed by tigers are known as 'Tiger widows', often being forced from their homes because they are unable to support their families.
- The Indian government has restricted the number of boats that have the right to fish in Sundarbans' waters by issuing Boat Licensing Certificates (BLCs) – a document that authorises fishing activities. In Bangladesh, to increase compliance with fishery rules, there are community outreach initiatives as well as harsh penalties and boat confiscations for those who break the rules.
- It is thought that roughly 60 per cent of the male workforce in the Indian Sundarbans is moving away from the region. They have moved to work in construction and mining jobs in Mauritius or in the Indian states of Kerala, Gujarat, Karnataka and Maharashtra.
- Many others, unable to make ends meet, are opting to work in the brick kilns, which have replaced the water bodies that earlier dotted the delta.
- Cyclones alone account for about 45 per cent of all deaths in the Sundarbans. Since some islands are far more vulnerable to extreme climate conditions, plans have been put forward to move people from such areas to more stable zones and urban areas outside the Sundarbans. This would significantly decrease the population at risk and help land left behind to be converted into protected areas.

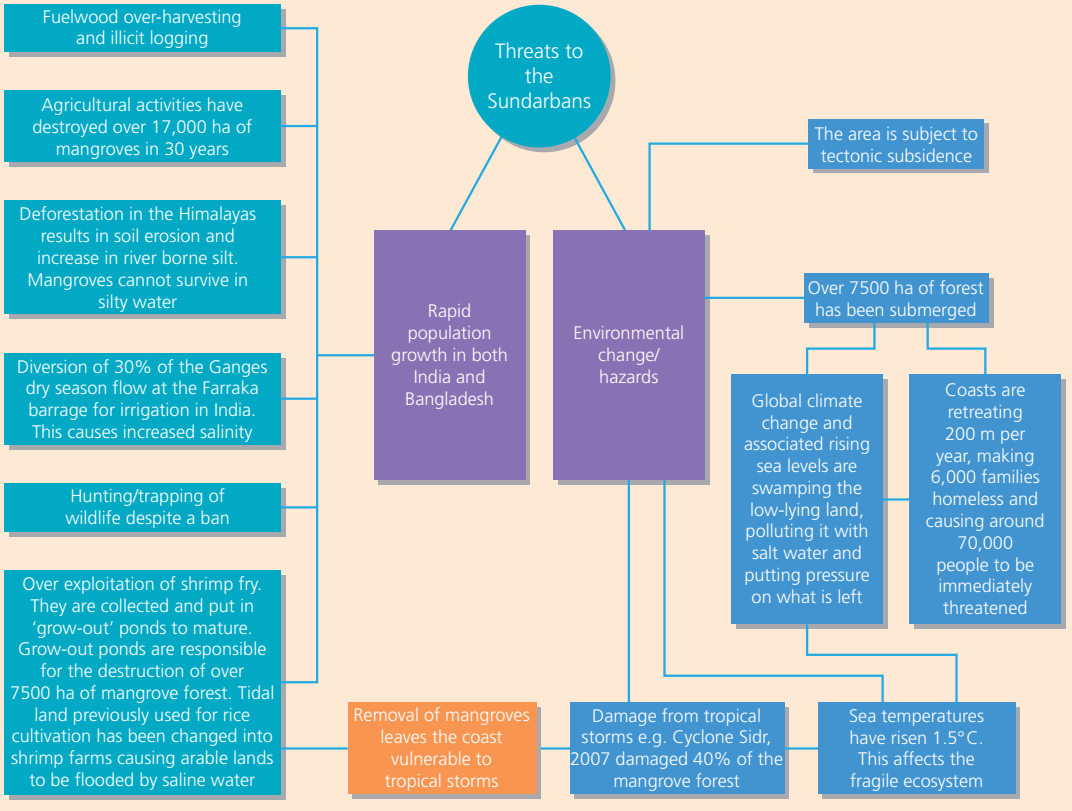


Figure 6.59 Threats to the Sundarbans

- Cyclones have badly damaged much of the 3500 km of river embankments put in place by the British in the 18th century. The Indian Government has initiated the Sundarbans Embankment Reconstruction Project, which is using modern machinery to support local efforts to rebuild new and much more sturdy embankments.

Fieldwork opportunities

1 Ponds offer a chance for fieldwork. Here are two questions that could be followed up:
- How does vegetation change from the edge of the water inland?
- How do soil characteristics/microclimate change with distance from the water?

2 Sand dunes offer a chance for fieldwork. Three questions that could be studied are:
- How does vegetation change from the seashore inland?
- How do soil characteristics/microclimate change with distance from the seashore?
- Compare two areas of the dune system, for example, ungrazed/grazed sections of the fixed dunes.

Further reading

Living Planet Report, WWF

The *Living Planet Report* is the world's leading, science-based analysis of the health of our planet and the impact of human activity.

The Social Dimension of Ecosystem-based Adaptation, No 12, 2013

The UK National Ecosystem Assessment

The UK National Ecosystem Assessment (UK NEA) was the first analysis of the UK's natural environment in terms of the benefits it provides to society and continuing economic prosperity. Only available from the UK National Ecosystem Assessment website.
http://uknea.unep-wcmc.org

Biodiversity and Climate Change – a Summary of Impacts in the UK, Joint Nature Conservation Committee.

This gives an overview of how biodiversity of different ecosystems in the UK is being affected by climate change.

The UK Biodiversity Action Plan (BAP)

This is also available from the Joint Nature Conservation Committee at https://jncc.gov.uk/

Moors for the Future Partnership is an organisation that aims to deliver a landscape scale programme of blanket bog restoration across the Peak District and South Pennine Moors.

www.moorsforthefuture.org.uk/

The reasons for the designation of The Belize Barrier Reef System as a World Heritage Site can be found at:

https://whc.unesco.org/en/list/764/

Question practice

A-level questions

1 Outline the causes and impacts of declining biodiversity in a named region/area. [4 marks]

2 Figure 6.6 (page 235) shows a simplified food web for an English woodland. With reference to the figure, analyse the relationships between different species within the food web. [6 marks]

3 With reference to a local ecosystem you have studied, evaluate the success of conservation strategies in managing that ecosystem. [9 marks]

4 Analyse how population change can affect the sustainability of warm water coral reefs. [9 marks]

5 With reference to a region that you have studied that is experiencing ecological change, to what extent has the region been affected by economic change? [20 marks]

Part 2

Human Geography

Chapter 7 Global systems and global governance

Chapter 8 Changing places

Chapter 9 Contemporary urban environments

Chapter 10 Population and the environment

Chapter 11 Resource security

Chapter

7 Global systems and global governance

Our world is more connected now than it has ever been. The global economy and society have altered significantly in recent years as a result of the process of globalisation. Arguably, this has been positive for overall economic development but has also produced some negative consequences in terms of socio-political conflicts and environmental sustainability. There are few subjects either as controversial or more in need of better awareness than attempts to manage and govern human affairs on a global scale.

As a geographer, being able to recognise and explain links between economic, political and social change, as well as engaging with contemporary world issues, is important for understanding both international trade and the workings of the global community.

In this chapter you will study:

- globalisation
- global systems
- international trade and access to markets
- global governance
- the 'global commons'
- Antarctica, the Southern Ocean and their protection
- a critique of globalisation.

Links

This chapter covers a core topic that can be linked to aspects of both human and physical geography. You may be asked examination questions requiring you to connect other topics you have studied and explain how they relate to or are affected by global systems and governance.

7.1 Globalisation

> It has been said that arguing against globalisation is like arguing against the laws of gravity.
>
> *Kofi Annan,* former Secretary-General of the United Nations

In the late 1960s, Canadian philosopher Marshall McLuhan predicted the advent of a 'global village' – a 'flat world' where free rein is given to economic and information flows. This would be reflected by the increasingly international manner in which organisations would operate:

- by thinking globally (not within national boundaries)
- by acting globally (being present in many countries)
- by making 'planet-wide' decisions.

The notion of McLuhan's global village may be exaggerated, even with regard to international trade. However, the interaction of prices, supply and demand increasingly operates at a global scale giving rise to a 'world economy'. This **world economy** is affected by global events that have similar impacts in most countries. For example, the 2008 global financial crisis began in the USA but developed into an international crisis that threatened the world financial system and caused a global economic downturn. Recovery from the crisis has not been smooth, with the world economy being hit by further setbacks including trade wars, falling commodity prices and international conflicts.

Globalisation, however, is not just about an increase in trade. In 2020, the world faced another global threat that would shine a light on the idea of a 'global village'. The outbreak and rapid international spread of a newly discovered coronavirus disease (COVID-19) caused the World Health Organization (WHO) to announce a global pandemic. To mitigate the impact of this highly infectious disease many countries went into an extended lockdown. This could have more devastating social and economic impacts on a global scale than those caused by the 2008 financial crisis.

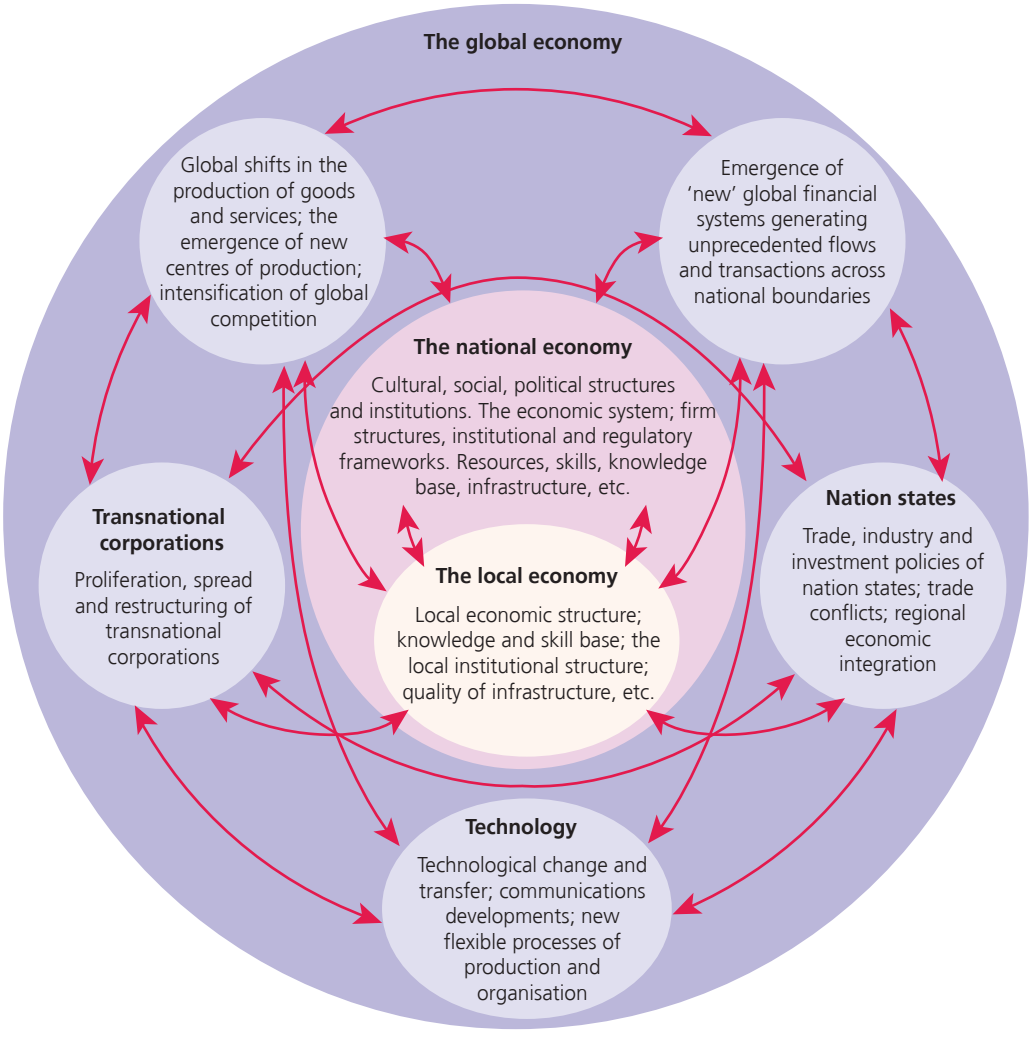


Figure 7.1 The global economy as a 'system' of interconnected elements and scales

Dimensions of globalisation

Key terms

Capital flows – The movement of money for the purpose of investment, trade or to produce goods/provide services. Usually regarded as investment into a production operation.

Globalisation – A process by which national economies, societies and cultures have become increasingly integrated through the global network of trade, communication, transportation and immigration.

International trade – The exchange of capital, goods and services across international borders. Inbound trade is defined as imports and outbound trade as exports.

The term 'globalisation' was used to describe a rapid integration of the world's economic systems in the 1990s under the combined influence of the information technology revolution and the opening up of former Communist Bloc countries to the market economy. The focus of globalisation has been primarily on economic relationships such as international trade, **foreign direct investment** and international capital flows.

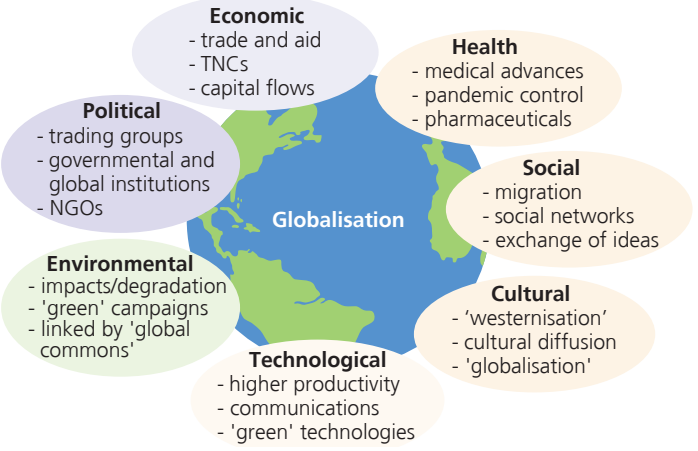


Figure 7.2 Dimensions of the globalisation process – aspects of these are transferred more easily through global networks

Globalisation has since been expanded to encompass a wider range of dimensions including cultural, social, technological, political, environmental and also health-related factors, for example, the control of pandemics or the impacts of climate change. The process involves the spread of ideas and information across the world, but in doing so it creates a growing uniformity shared by very different places (Figure 7.3).

Factors of production

The roots of globalisation lie in international trade and the increasing accessibility of markets, which open up to the wider global community. This means that goods and services have to be produced that can be traded. A number of productive resources have to be combined in order to provide these goods and services; economists call these the '**factors of production**'. They are as follows:

- **Land** – All natural resources provided by the Earth including minerals, soils, water, forests, etc.
- **Labour** – The human resource available in any economy. The quantity, quality and cost of the workforce are key considerations to any producer of goods and services.
- **Capital** – In economic terms this refers to any physical resource that can be regarded as a man-made aid for production, such as buildings, factories, machinery, etc. '**Capital flows**' can involve the transfer of these physical resources from one place to another. In reality it usually refers to the flow of investment finance used to provide this capital.
- **Enterprise** – This is a very particular form of human capital describing those who take the risk of establishing businesses and organising the production of goods or provision of services.

The increased international flows of capital, labour and enterprise have been both a cause and an ongoing consequence of the globalisation process.

Figure 7.3 Globalisation can lead to **uniformity** through recognition of international brands. The Body Shop and Starbucks are familiar to UK high streets but these are located in Beirut, Lebanon

Flows of capital

For the purpose of understanding '**international capital flows**', capital includes all money that moves between countries which is used for investment (for example, in land and physical capital), trade or production.

In the late twentieth century, **deregulation of world financial markets** meant that the activities of financial institutions such as banks, insurance companies and investment companies were no longer confined within national boundaries.

Figure 7.4 (page 286) shows the capital flows experienced in the global economy system. It distinguishes between a 'core' area and a 'periphery', both of which were included in Frank's 'Dependency Theory' and Wallerstein's 'World System Theory' models.

These models assume that global power is concentrated in blocks of highly developed economies (HDEs),

Table 7.1 The main dimensions of globalisation

Form of globalisation	Process caused by:	Characterised by:
Economic	increase in free trade growth of transnational corporations faster, cheaper transport global marketing	long distance flows of goods, capital and services as well as information and market exchanges
Cultural/social	migration global communication networks impact of western culture through media, sport, leisure and celebrity	spread of ideas, information and images
Political	growth of Western democracies and their influence on poorer countries decline of centralised (communist) economies (though communist political control is still strong in China and Russia)	the diffusion of government policy and development of market economies in former communist states

which they called the 'core'. Periphery countries are those with less developed economies (LDEs), which have been exploited and have suffered from a lack of investment, leakages and out-migration.

The flows of capital shown in Figure 7.4 still exist in the world economic system, but it is more difficult to distinguish between 'core' and 'periphery' regions. The rapid growth of large and medium **emerging economies** such as the BRIC and, more recently, MINT countries means there is now more of a continuum of development.

Key terms

BRIC – An acronym used to identify a group of four countries – Brazil, Russia, India and China – whose economies have advanced rapidly since the 1990s. Sometimes South Africa is added to this list so the acronym becomes BRICS.

Diaspora – A large group of people with a similar heritage or homeland who have moved and settled in places all over the world.

Leakage (economic) – Refers to a loss of income from an economic system. It most usually refers to the profits sent back to their base country by transnational corporations – also known as profit repatriation.

MINT – An acronym referring to the more recently emerging economies of Mexico, Indonesia, Nigeria and Turkey.

Figure 7.4 shows the salient flows of money between richer areas and emerging markets in the 'periphery'.

These flows assist emerging economies to develop and can reduce disparities between the core and the periphery.

- **Foreign direct investment (FDI):** This is investment made mainly by TNCs (sometimes by governments) based in one country, into the physical capital or assets of foreign enterprises. The investing company may make this overseas investment in a number of ways. For example, by setting up a subsidiary company, by acquiring shares or through merger or joint venture.
- **Repatriation of profits:** TNCs investing in overseas production will normally take any profit made from that investment back to their home-country. This is known as an economic leakage as the income is 'leaked' from the country that received the investment. The majority of these flows return to companies based in richer countries.
- **Aid:** This is an important source of financial support for poor countries. It can take many forms and can be provided through the UN (multilaterally) from contributions made by a number of richer countries (sometimes known as **Official Development Assistance – ODA**). It can also be provided bilaterally from one government to another, usually with mutual co-operation conditions applied.
- **Migration**: The majority of out-migration of labour takes place from poorer to richer countries. This will exacerbate disparities as the less developed

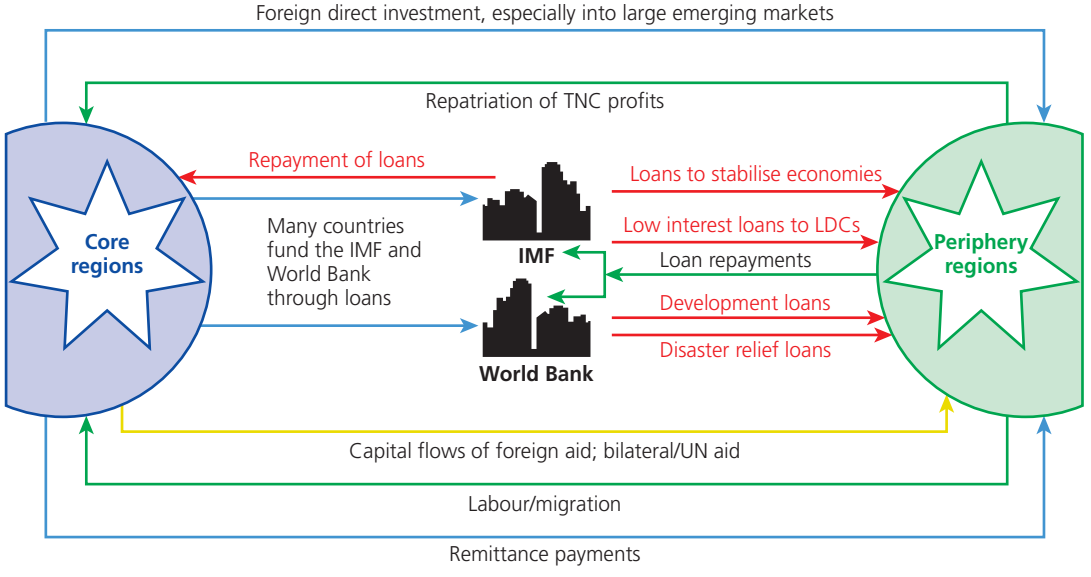


Figure 7.4 Flows of finance and capital between regions in the global economy

nations lose their most skilled and talented labour, who will pay taxes and spend much of their earnings in their destination country (although they often send **remittances** back to their country of origin).

- **Remittance payments:** These are transfers of money made by foreign workers to family in their home country. Figure 7.5 shows that by 2018 remittances have become as important as FDI as a source of income in developing countries. India is the world's top recipient of remittances with its **diaspora** sending US $79 billion back home in 2018.

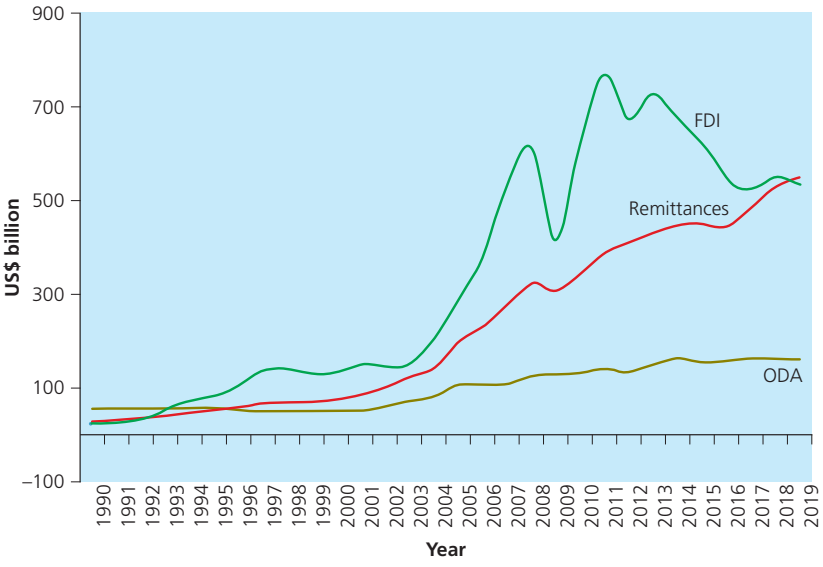


Figure 7.5 Sources of income received by low- and middle-income countries (including China) 1990–2018. Remittances are becoming an increasingly important source of income for developing countries.

Source: World Development Indicators and World Bank estimate

Remittance flows – 'rivers of gold'

The money that workers send home to their families from abroad has become a critical part of many developing economies around the world. Officially recorded remittances amounted to a record US $529 billion in 2018, which is more than three times the volume of ODA (Figure 7.5). They are expected to grow in importance as a source of external financing in LDEs.

The net flows of remittance payments are generally sent from wealthy and resource rich economies to developing economies. Among the main sources of remittances are the USA, Saudi Arabia, United Arab Emirates (UAE), UK, Canada, Russia and EU countries including Germany, France and Italy. Figure 7.6 shows the top ten remittance recipients and the amount they received. In smaller economies, remittances surpass 25 per cent of total GDP, for example, in Tonga, Kyrgyz Republic, Tajikistan and Nepal.

The benefits of remittances are that they:

- present a more stable and steadily rising source of income, unlike FDI which fluctuates and dips significantly during recessions
- are an effective way to alleviate poverty because they go directly to families.

However, the benefits are reduced by the high charges made by banks, post offices and other transfer agencies for sending the money. The global average cost for sending $200 in 2019 was around 7 per cent but it can be as high as between 15 to 18 per cent for some transfers. The UN recognises the importance of remittances for development and reducing remittance costs to 3 per cent by 2030 is a global target under the Sustainable Development Goals.

Another way to maximise the power of remittances is to encourage migrant workers to invest in their home countries more formally through **diaspora bonds**, which can be used to finance development.

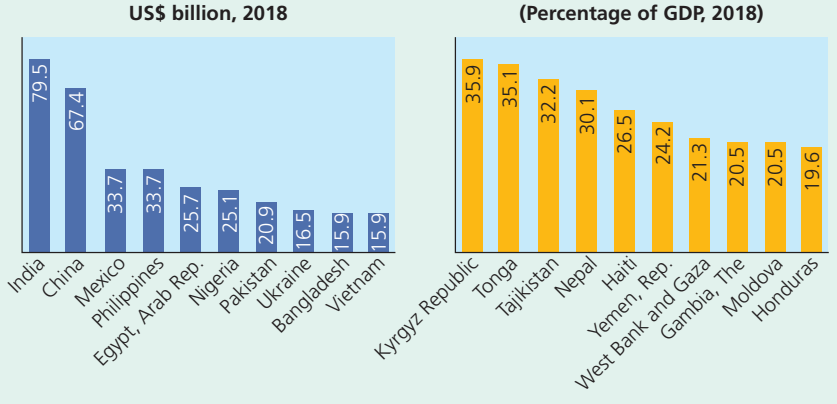


Figure 7.6 Top 10 remittance receivers 2018, by total amount and as a percentage of GDP

Source: The World Bank

Key to circular diagram

- The colour indicates the region each flow came from
- The width of a flow shows its size in millions
- The numbers indicate the total migration in and out of a region in millions
- The origin of the migrant flow is identified by being nearest to the region on the edge of the circle.

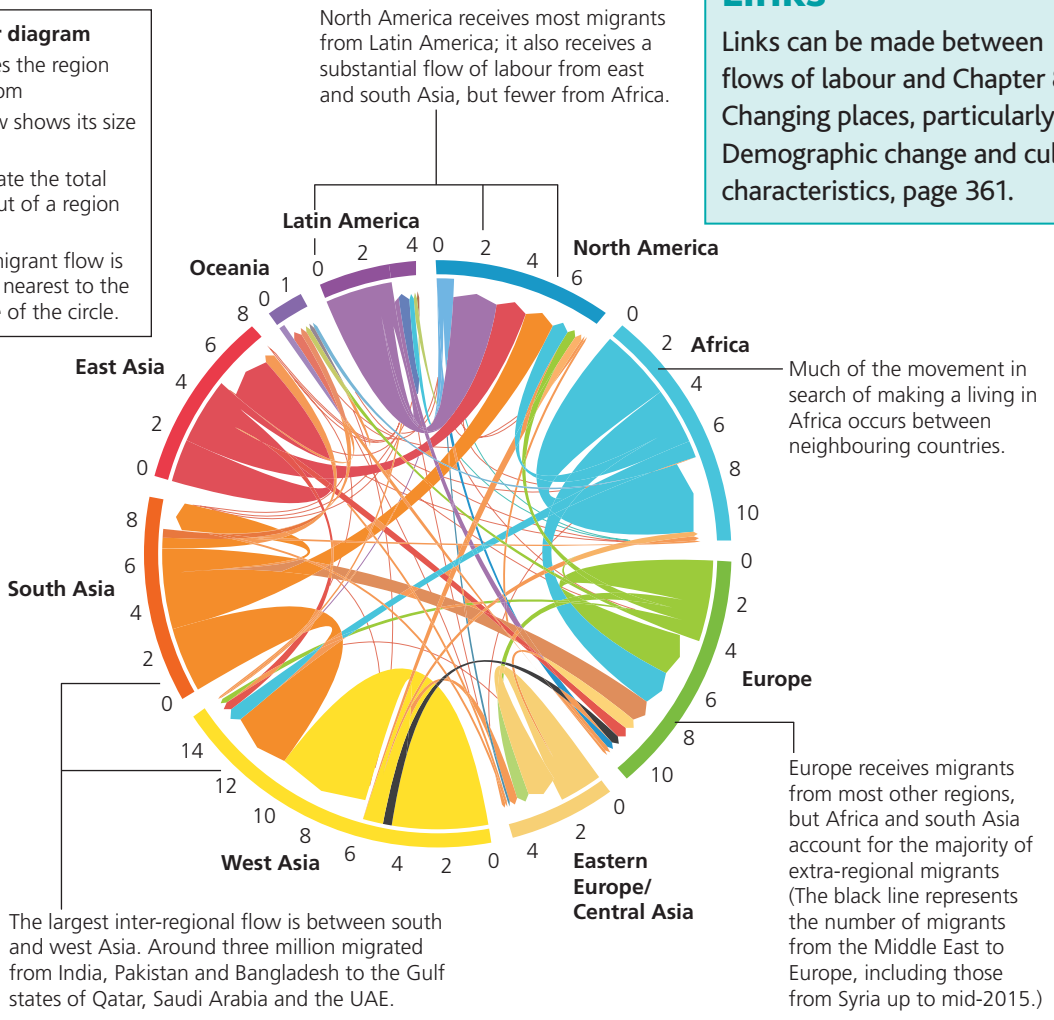


Figure 7.7 Circular diagram illustrating global migration flows from 2010–2015
Source: Abel, G. J., & Cohen, J. E. (2019), Spiegel

Links

Links can be made between flows of labour and Chapter 8, Changing places, particularly to Demographic change and cultural characteristics, page 361.

Flows of labour

Labour markets are not as free flowing as financial markets. People move less easily around the world than money because of restrictions on immigration. However, in recent years there has been a phenomenal rise in the numbers of migrants crossing international borders, mainly to seek better employment opportunities (Figure 7.7). Much of the movement has been from developing countries in South Asia, Africa and Latin America to the richer areas of North America and Europe. Another major destination for the movement of labour has been around the oil-rich Gulf States of Kuwait, Qatar, Bahrain, Saudi Arabia and the United Arab Emirates, where the construction boom has provided plentiful employment opportunities.

Key facts about the movement of labour:

- Despite increases in cross-border movements, most migrants move over short distances within the same region or between neighbouring regions (especially in sub-Saharan Africa).
- North America, Europe and the Gulf countries in western Asia attract migrants from furthest afield.
- The bulk of economic migrants moving between continents are not the poorest but are those with some education and financial means.

- The largest **inter-regional** flow of labour is in Asia. Between 2010 and 2015, around three million workers moved from south Asia to west Asia.
- The second largest inter-regional flow is from Latin America to North America.
- Over the past 25 years more people have migrated from Asia to North America and Europe but at the same time both the Gulf states and the Tiger economies of South East Asia have become attractive destinations for those seeking employment.

Flows of products

Flows of manufactured goods have increased significantly in recent years. This has been stimulated by demand from affluent populations in developed countries, combined with low production costs because of mass production and low wage economies in exporting regions in China and South East Asia.

The international movement of products is facilitated, especially for developing countries, by the reduction in costs and by the creation of regional trading blocs of nations.

- **Transaction costs** have been reduced by the improvements in flows of data and the ease with which capital can be transferred to pay for transactions.
- **Transport and time** costs have been reduced by the process of containerisation which has enabled more complex and long distance flows of products, and air transport, which can speed delivery and reduce costs of more valuable or perishable cargo.
- The most obvious regulatory barriers to trade are tariffs, which with the encouragement of the **World Trade Organization (WTO)** have generally been reduced in global trade.
- **Regional trading blocs** provide tariff-free trade or other favourable trading conditions to fellow member nations within the bloc.

Key terms

Containerisation – A system of standardised transport that uses large standard-size steel containers to transport goods. The containers can be transferred between ships, trains and lorries, enabling cheaper, efficient transport.

Protectionism – A deliberate policy by government to impose restrictions on trade in goods and services with other countries – usually done with the intention of protecting home-based industries from foreign competition.

Tariffs – A tax or duty placed on imported goods with the intention of making them more expensive to consumers so that they do not sell at a lower price than home-based goods – a strategy of protectionism.

Flows of services

Services are economic activities that are traded without the production of material goods, for example, financial or insurance services. They can also be sub-divided into:

- **high level services** – services to businesses such as finance, investment and advertising
- **low level services** – services to consumers such as banking, travel and tourism, customer call centres or communication services.

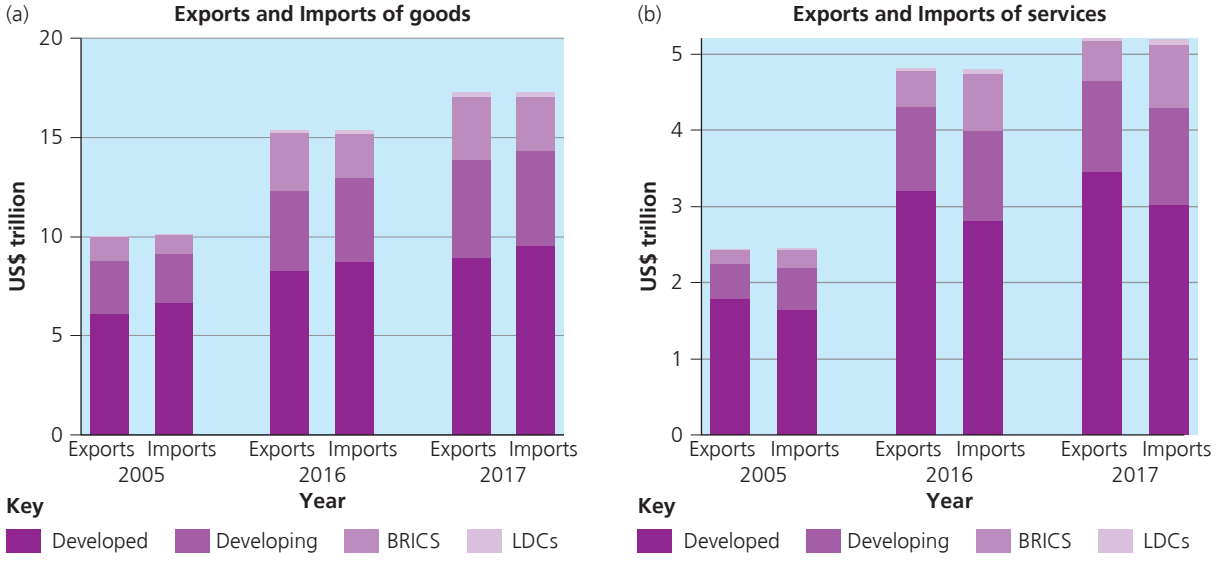


Figure 7.8 Values of trade in goods and services by region, 2005–2017

Source: United Nations

Services such as banking, insurance and advertising depend on communication and the transfer of information. They are **footloose** and can locate anywhere and advancing technology means they can still serve the needs of customers worldwide.

- High-level services are concentrated in cities in the more developed world, such as London, New York and Tokyo, which are the major centres of global finance.
- With the emergence of east Asian economies, Hong Kong, Singapore and Shanghai have also become major global financial centres.
- A growing number of transnational service **conglomerates** have emerged, seeking to extend their influence on a global scale, such as HSBC Holdings in banking, AXA (France) in insurance, WPP Group (UK) in advertising and TUI Group (Germany) in travel and tourism.

One trend has been the decentralisation of low level services from the developed to the developing world. Call centre operations, for example, have moved from the UK to India where labour costs are up to 20 per cent lower than in the UK. Much of India's economic success can be attributed to its growing service sector, especially information technology and online customer services, which provide support for companies in high-income countries (also known as 'outsourcing').

Flows of information

Information flows are governed by the movement of people through migration and by the speed of data and communication transfers. Both are responsible for the transfer of cultural ideas, language, industrial technology, design and business management support.

Digitisation and satellite technology have transformed these flows of information, which are now supported by:

- improvements to global telephone networks, making communication cheaper and easier
- mobile telecommunication technology
- email and the internet, which enable large amounts of information to be exchanged instantly across the globe
- live media coverage available on a global scale because of satellite technology.

The importance of information flows should not be underestimated because of their contribution to the expansion of **knowledge-intensive goods and services** (also known as the quaternary sector of an economy).

Such goods and services include those which have an intensive research and development (R&D) component and use highly skilled and educated labour. They include high-tech products such as pharmaceuticals, computer technology and business services including international law, accounting and engineering. These industries need the exchange of ideas and flow of expertise to flourish.

Key terms

Conglomerates – A collection of different companies or organisations which all report to one parent company – most transnational corporations are conglomerates.

Economies of scale – The cost advantages that result from the larger size, output or scale of an operation. Savings are made by spreading the costs or by rationalising operations.

Global marketing

Marketing is the process of promoting, advertising and selling products or services. When a company becomes a global marketeer, it views the world as one single market and creates products that fit the various regional marketplaces. It will usually develop a recognisable '**brand**' and employ one marketing strategy to advertise the product to customers all over the world. The ultimate goal is to sell the same 'or similar' product, the same way, everywhere. Having one marketing campaign on a global scale like this generates **economies of scale** for the organisation, which reduces their costs.

Coca-Cola is an example of a company with a single product; only minor elements are tweaked for different markets (Figure 7.9). The company uses the same formulas (one with sugar, the other with corn syrup) for all its markets. The bottle design is recognisable in every country but the size of bottles and cans conforms to each country's standard sizing.

Patterns of production, distribution and consumption

Globalisation created a new international division of labour with two main recognisable groups:

- the highly skilled, highly paid, decision-making, research and managerial occupations which, on a global scale, are largely concentrated in more developed countries

Figure 7.9 Coca-Cola in Africa: the unmistakeable Coca-Cola brand even reaches poorer parts of Senegal. Coca-Cola pay for paintwork and often for refrigeration units for vendors

- the unskilled, poorly paid assembly occupations, which have become increasingly located in newly industrialising countries, which have lower labour costs.

Many countries that were classified as less economically developed have become **newly industrialised countries** (NICs). These countries have developed their own industrial and commercial bases as well as markets for their own goods and services. TNCs have also emerged from these countries and have extended their global influence. This started with the four Asian '**tiger**' economies of Hong Kong, Singapore, South Korea and Taiwan, which were followed in succession by the **BRIC** economies, the most rapidly growing and prominent of these being China (see Figure 7.11, page 292). Development has come to regions surrounding these new 'cores', particularly to South East Asia.

Production

Over the past 40 years, manufacturing has become decentralised, moving away from the highly developed economies (HDEs) of Western Europe, North America and Japan. As a result of foreign direct investment by TNCs, many emerging economies have been able to develop competitive manufacturing industries.

Manufacturing operations transfer around the world relatively easily and the **global shift** in the pattern of production from HDEs to lower wage economies has been driven by a number of factors:

- lower land and labour costs
- incentives offered by governments, in the form of tax breaks or special economic zones, have encouraged TNCs to invest and relocate the **production** side of their business abroad
- the **transfer of technology** by TNCs has enabled countries in the developing world to increase their productivity, without raising their wages to the same levels as HDEs.

Other factors also influence the locational decisions made by entrepreneurs of large manufacturing companies, including:

- the availability of a skilled and educated workforce
- the opportunity to build new plants with the latest and most productive technology
- access to large markets without tariff barriers (enabled through trade agreements)
- availability of infrastructure including power supply, roads, ports etc. (though many TNCs are willing to invest in these assets if other conditions are favourable).

One of the main consequences of this global shift has been **deindustrialisation** in richer HDEs and a subsequent decline and loss of jobs in the manufacturing sector. Outdated production methods and lack of investment have also contributed to the decline of manufacturing in these regions. Governments in HDEs have reacted with varying strategies that attempt to reverse the decline. These include:

- encouraging foreign TNCs to invest in deindustrialised regions by offering incentives, such as tax breaks
- encouraging investment in skills and technology to upgrade manufacturing industry
- adopting more protectionist policies, such as import tariffs, to protect domestic production (this can prove counterproductive as it makes home industries uncompetitive globally).

Figure 7.10 (page 292) shows that employment in UK manufacturing has fallen by nearly 60 per cent since the late 1970s. This decline has steadied, with the numbers employed stabilising and even increasing slightly in recent years. Productivity in the UK has been rising, largely because of advances in manufacturing technology.

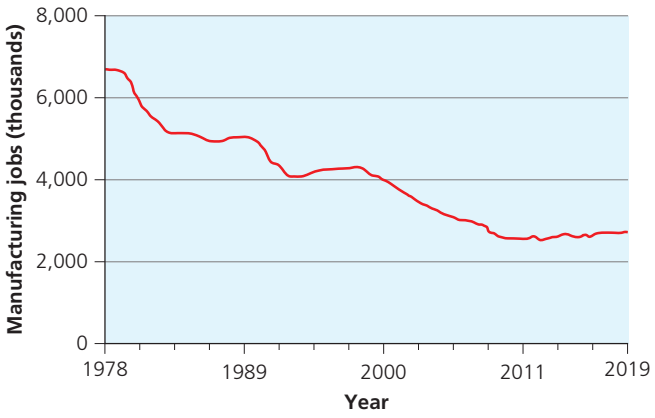


Figure 7.10 Loss of manufacturing jobs in the UK since 1978
Source: Statista

The global shift in production patterns has also caused political reaction against globalisation. The rise in **populist** and **nationalist** movements in developed economies is partially attributed to the decline in standards of living in areas suffering from deindustrialisation. In 2016, this reaction was arguably a key driver behind the Brexit vote in the UK and the US election in which Donald Trump was elected President (see Section 7.6, Globalisation critique).

Distribution and consumption

Product consumption still lies predominantly in the richer countries of the developed world. Products being manufactured in emerging NIC economies are largely exported and sold to countries in Europe, North America and Japan (Table 7.2).

However, the pattern is changing. As emerging NICs develop, their populations are becoming more affluent and are starting to demand similar consumer products to those being exported from their own countries. For example, Dyson, a UK-based manufacturer of household electronic goods, moved the manufacture of its products to Malaysia in 2002 and its headquarters to Singapore in 2019. It still sells the bulk of its products in the UK and other parts of Europe but 75 per cent of its significant growth in 2017 came from its Asian market and only one-fifth from Europe.

Table 7.2 Top 10 largest consumer markets by percentage share of goods and services consumed, 2015

Country	% share	Country	% share
1 USA	29%	6 UK	4.1%
2 Japan	8.5%	7 Italy	3.7%
3 = Germany	5.3%	8 = Brazil	2.4%
3 = China	5.3%	8 = Spain	2.4%
5 France	4.5%	10 Canada	2.3%

Source: World Economic Forum

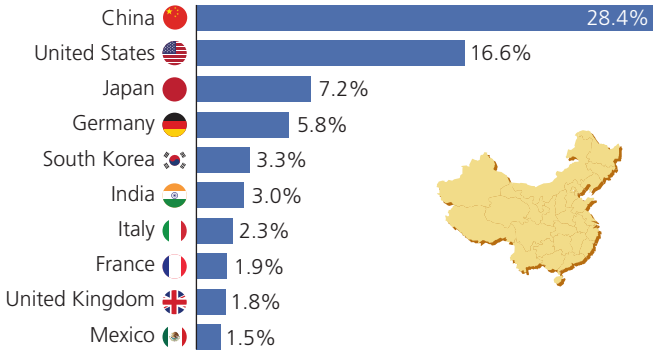


Figure 7.11 Top 10 countries by share of global manufacturing output, 2018
Source: ONS

Different patterns for distribution and consumption are likely to appear in the future with a marked transition from west to east as the centre of gravity of economic activity shifts. Forecasts suggest that:

- the USA, western Europe, Japan and China will continue to be the best destinations for exporters
- as Asia becomes more competitive a growing share of the region's exports will be to other countries in Asia
- China's 'Belt and Road Initiative' will open up access and diversify its exports to emerging markets
- finance corporations from HDEs have the potential to benefit from the expansion in financial services in the Asia-Pacific region but will face increasing competition from Chinese, Singaporean and Korean banks and insurance companies.

Factors in globalisation

A number of factors have combined to influence the breadth and depth of links between nations over the past 30 years (see Figure 7.12).

Communications technologies

In this digital age, there are few barriers to prevent the sharing of information and flow of data globally. Links between countries have grown significantly as a result of:

- the development of computer technology
- the advent of the internet, enabling speedy 24/7 global communication at the click of a button; there are 4.5 billion internet users worldwide (see Figure 7.13)
- the use of mobile phones (mobile phones are particularly important in LDEs as people and markets can be connected more easily); there are nearly **seven** billion mobile phone subscribers worldwide

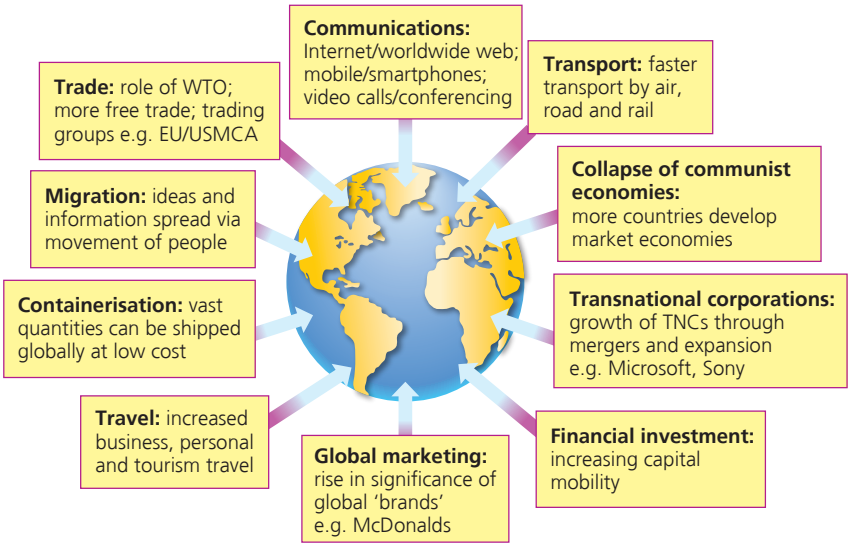


Figure 7.12 Factors that have influenced globalisation since the 1990s

- computer control and robotics technologies have integrated manufacturing operations
- computerised logistics systems have evolved to support supply chain distribution.

Financial systems

- The world became increasingly financially integrated in the 1980s and 1990s due to financial deregulation.
- Deregulation made it easier to move finance across borders, making it easier to trade with and invest in other countries.
- The **global financial system** (GFS) provides a framework that facilitates flows of capital for the purposes of financing investment and trade.

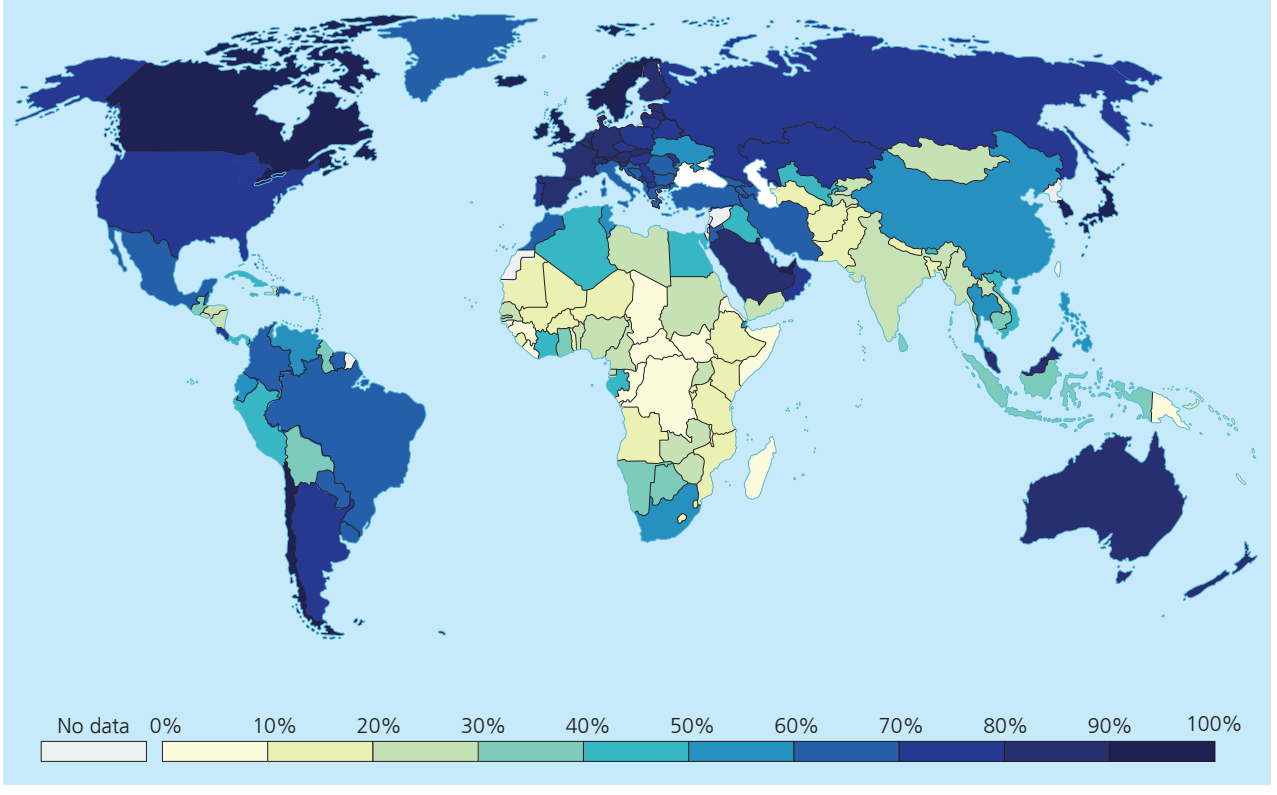


Figure 7.13 Share of the population using the internet, by country, 2017

Source: The World Bank

- The GFS is used by financial institutions, TNCs, consumers and investors and is monitored by regulatory bodies such as the International Monetary Fund (IMF).
- Integration of developing economies into the GFS has helped to reshape international trade.
- High-speed electronic transmission systems and global exchange connectivity means that transactions between importers and exporters can be completed securely with fewer concerns about exchange rates.
- One major disadvantage of deregulation and free movement of capital is that it leaves the system exposed to volatile capital flows; this triggered the global banking crisis in 2008–2009.

Transport systems

Products and commodities can be shipped more quickly and in larger quantities as a result of integrating the following technologies, which have all contributed to easing international trade:

- increased size of aircraft; integrated air traffic networks
- growth of low-cost airlines and air freight companies
- the use of standardised containers – by sea, rail, road and air (Figure 7.14)
- handling and distribution efficiencies
- high speed rail networks.

In some larger countries, where distances to sea terminals are long, especially in economies with less developed infrastructure, exporters are encouraged to use **dry ports**, located inland and near to their business. Dry ports save the exporter time and transport costs as all shipment arrangements and customs documentation are completed locally before the goods are shipped to a seaport. These ports provide local 'hubs' providing global connectivity. Pakistan, for example, has six inland dry ports where goods are deposited by producers before they are sent on to the port at Karachi for export.

Security systems

The global community faces a number of security issues as a result of globalisation. National boundaries are less of a barrier and easier travel means more fluid flows of products and people. This provides opportunities to reduce traditional security measures but also poses a number of threats such as:

- **terrorism** – screening and monitoring movements by security forces as part of counter-terrorism measures

Figure 7.14 Containers stacked at Southampton docks can be transferred from ships to rail or road transport

- **food imports** – ensuring that imported products meet required safety standards
- **biosecurity** – preventing the introduction and spread of harmful organisms or biochemical substances in order to minimise the risk of transmission of infectious diseases caused by viruses and bacteria to people, animals and plants
- **cybercrime** – reliance on the internet has led to breaches of secure information, which in turn can lead to online fraud and other online crimes
- **supply chains** – securing supply chains is critical for international manufacturing businesses that assemble components from all over the world. Supply chain security involves ensuring products are authentic, safe and can travel through borders relatively freely.

A number of initiatives are in place to mitigate these threats. Interpol enables police in most countries to work together to fight international crime. Counter-terrorism agencies in many countries co-operate to thwart terrorist attacks. The World Customs Organization has introduced the idea of 'Authorised Economic Operators' (AEOs) to tackle some of the issues relating to trade security. AEO status is awarded to export and import operators who meet a range of minimum standard criteria.

Management and information systems

Globalisation is influenced by how large companies operate to produce and distribute goods and deliver services worldwide. In manufacturing, high-volume production enables substantial economies of scale on

a global level. To benefit from these cost advantages, globally integrated companies have invested in:

- large production and assembly plants capable of exploiting the most advanced technology
- global marketing and distribution networks to ensure that sales keep pace with increased production.

Investments in international production, distribution and management are organised within **global value chains**, where different stages of the production process are located across different countries. Managing these chains requires remote management of production and distribution lines, which has been enabled by information systems giving businesses:

- virtually free telecommunications and video conferencing
- integrated ICT management systems, usually supplied by third-party service providers (for example, Infor Nexus) to facilitate greater supply chain organisation.

Similarly, '**just-in-time**' (**JIT**) systems give greater efficiency in the supply chain for manufacturers. JIT involves ensuring the correct quantities of materials, components and assembled goods are available on time, in the correct location at each stage of production. This reduces costs by having fewer goods and materials held in stock. The principle underlying JIT is that production is 'pulled through' by customer orders, rather than 'pushed through' to build up stock.

Management and information systems have led to:

- higher-order business activities (such as research and development, design and engineering, marketing and advertising) being based at corporation headquarters and strategic hubs around the world
- lower-order activities (such as production and assembly) being located at low production-cost locations or near to large markets for the finished goods
- global corporations focusing on key strategic activities and increasingly outsourcing non-strategic activities
- a rapid growth of the logistics and distribution 'solutions' industry.

Trade agreements

Links

Links can be made here to International trade and access to markets, which features later in this chapter (pages 306–14).

A trade agreement is a formal agreement between two or more countries that removes trade barriers between those in the agreement. A number of regional trade blocs have been formed around the world as a result of countries signing agreements to stimulate trade

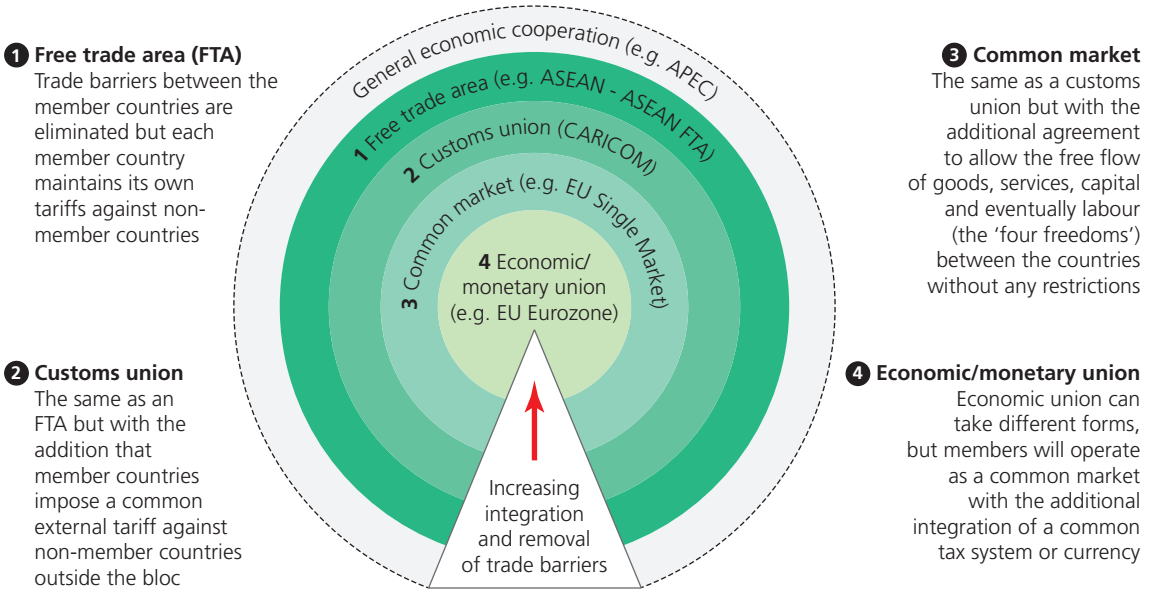


Figure 7.15 Types of trade groupings

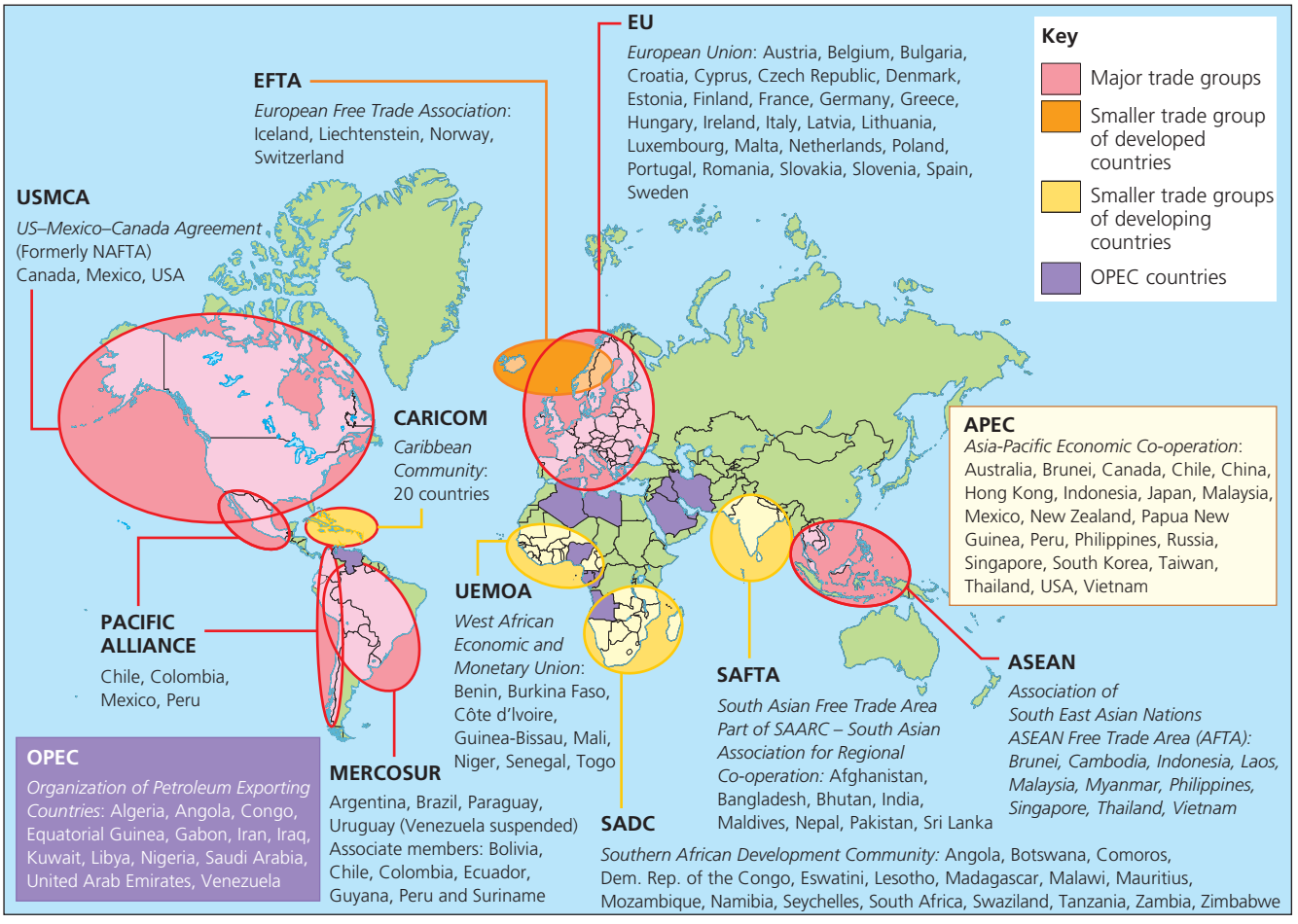


Figure 7.16 Some of the main global trading blocs

between themselves and to gain economic benefits from co-operation.

There are various forms that trade groupings might take, including free trade areas, customs unions, common markets and economic or monetary unions such as the European Union (EU) (Figure 7.15, page 295).

The distinctions between different types of trade groupings are those generally recognised by the World Trade Organization (WTO). In reality, distinguishing between them is difficult because there is some overlap between the different categories and it is more a continuum of greater integration. For example, Denmark is a member of the EU both economically and politically but not a member of its monetary union, the Eurozone, where the Euro (€) is the common currency. Some, such as UEMOA in West Africa, are subgroups of larger trading groups that exist in the same region (ECOWAS in West Africa). Figure 7.16 shows some of the major trading groups on a global scale. These are not exclusive and it is worth noting that many countries (for example, Mexico) are members of more than one trade agreement.

Not all trade agreements are regionally based. For example the **Organization of Petroleum Exporting Countries (OPEC)** is made up of members mainly from the Middle East but also from South America and Africa. Its focus is on the trade of oil globally, which is the single most important traded **commodity**.

Other groupings are more loose knit and contain members with an interest in co-operation and development of trade, but with no formal trade agreement, though these might be developed (for example, APEC).

On a global scale there are advantages to nations grouping together as trading entities, including:

- improving global peace and security and reducing conflict
- increasing global trade and economic co-operation
- encouraging social and economic development in middle- and lower-income countries.

On a national scale, there are both advantages and disadvantages to being part of a trading bloc (Table 7.3). For example, in 2016 the UK voted in a referendum to withdraw from the EU; a process that became known as 'Brexit'. In making this decision, voters were influenced by arguments both for and against remaining part of the EU's Single Market (the UK were not part of the Eurozone). The UK left the EU in January 2020 to enter a transition period during which a new trade agreement with the EU was negotiated and finalised for application from 2021.

Table 7.3 Advantages and disadvantages of being in a trading group agreement

Advantages	Disadvantages

• **Trade creation** – elimination of trade barriers gives member countries the opportunity to increase production and trade
• **Economies of scale** – mass production for a larger market reduces average costs of production
• **Investment** – attracts local and foreign direct investment because of a larger market size
• **Technology** – open borders lead to faster transfer of technology
• **International status**:
  • can compete on a global level with other trading groups
  • greater representation in world affairs
  • greater economic leverage as part of a group to negotiate trade deals with other groups
• **Employment** – in a 'common market' with the 'four freedoms', allows those seeking work to move between countries more easily
• **Exchange rates** – in monetary unions, developing a common currency prevents exchange rate fluctuations and simplifies transactions
• **Support for regions and sectors**:
  • for remote or declining industrial regions (for example, EU Regional Fund)
  • for sectors of the economy (for example, agriculture in the EU)
• **Social** – raising standards in healthcare and education across the region
• **Political** – promoting democracy and human rights among member states
• **Regional co-operation** – such as pooling resources in response to natural disasters, terrorism threats, pandemics etc.
	• **Trade diversion** – in a customs union, tariffs placed on external goods make them more expensive and member countries are forced to purchase from within the group as the price becomes artificially cheaper
	• **Loss of sovereignty** – decisions become centralised at a supra-national level (for example, in the EU loss of freedom to negotiate separate trade agreements)
	• **Legislation** – pressure to adopt central legislation, for example, European Court of Justice in the EU has ultimate legal power on many issues such as human rights (some argue this is beneficial as it provides an extra level to take cases to, others suggest it is less democratic and more bureaucratic)
	• **Employment** – freedom of movement of labour increases competition for work and can depress workers' wages in some sectors and may cause unemployment
	• **Finance/currency** – some loss of financial control to a central authority, especially in a monetary union (for example, the European Central Bank oversees monetary policy in the Eurozone; UEMOA has to adopt a common currency which is overvalued and damages exports)
	• **Dependence** – increased interdependence can mean that some members of the group become overly dependent on others
	• **Sharing/over-exploitation of resources** – certain sectors of the economy may be damaged by having to share resources (for example, the UK sharing its traditional fishing grounds with other EU nations, such as Spain and France)

USMCA – formerly the North American Free Trade Agreement (NAFTA)

The North American Free Trade Agreement (NAFTA) was signed by the USA, Canada and Mexico in 1994. NAFTA's main aims were to remove all trade barriers, increase investment opportunities and improve economic co-operation between the three countries. In 2017, the member economies generated $22 trillion in GDP, but NAFTA created 'winners' and 'losers' within each nation.

Supporters of NAFTA point out that:

- trade between member countries has quadrupled
- manufacturing grew in the USA, creating five million jobs and increasing economic output; this has helped North America compete with emerging economies in Asia
- foreign direct investment (FDI) more than tripled; Mexico, in particular, gained from this as TNCs locate production there in order to gain access to Mexico's trading partners' markets
- consumer prices in the US were lowered, especially for oil and food
- more competitive bidding for government contracts reduced costs
- the heads of state meet more frequently.

Opponents of NAFTA point out that:

- 'blue collar' jobs were lost in the US, particularly in the automotive sector, when manufacturers relocated to lower-waged Mexican **maquiladora** plants near the border
- job migration suppressed wages in US factories
- many Mexican farmers went out of business as they were neither able to compete with larger US agribusinesses nor with US and Canadian subsidised food surpluses being 'dumped' in Mexico

- unemployed Mexican farmers went to work in sub-standard non-unionised working conditions
- there has been an increase in negative environmental impacts as:
  - US and Canadian mining companies degraded the Mexican environment, taking advantage of its natural resources and less stringent pollution laws
  - Canada increased production of oil from the Alberta tar sands to supply US demand.

During the 2016 US presidential campaign, Donald Trump promised to renegotiate NAFTA. On becoming president, he negotiated a new trade deal with Canada and Mexico in 2018, which was signed and ratified by all three countries in early 2020. The new deal is known as the United States–Mexico–Canada Agreement (USMCA). Most commentators agreed that labour and environmental standards in NAFTA could be improved but Trump's main goal was to reduce the size of the US trade deficit that had built up with both Mexico and Canada. The key changes that USMCA has included are:

- **country of origin rules** – cars must have at least 75 per cent of their components manufactured in Mexico, Canada or the US to qualify for zero tariffs
- **labour laws** – 40 to 45 per cent of vehicle parts must be made by workers who earn at least $16 per hour
- **agriculture** – US farming conglomerates must get more access to Canada's lucrative dairy industry
- **intellectual property and digital trade** – copyright terms have been extended to 70 years beyond the life of the author and include new provisions to deal with the digital economy
- **sunset clause** – this means the terms expire after 16 years and are reviewed after that term.

Key terms

Maquiladora – a manufacturing operation (plant or factory) located in free trade zones in Mexico, usually near the US–Mexico border. Initially established in the 1960s to stem the flow of migrants, they have become an important part of Mexico's industrialisation, especially in the automotive industry. They import materials for assembly and then export the final product back to the USA.

Review questions

1 Explain why flows of capital are an important element of globalisation.

2 Outline changes in the global patterns of production, distribution and consumption.

3 How have developments in transport influenced globalisation?

4 Evaluate the impacts of trade agreements on globalisation.

5 Study the information on the new USMCA deal (pages 297–8) and suggest reasons for the key changes that have been introduced to the former NAFTA agreement.

7.2 Global systems

Globalisation means that countries and regions are more connected and dependent on each other than ever before. Systems have evolved to support the increased economic, political, social and environmental interdependence that exists in the contemporary world. Interdependency means that each country depends on others and so what happens in one place will increasingly have impacts on others. There is no better illustration of this than the COVID-19 pandemic of 2020–21, which started in Wuhan, China but spread, largely as a result of faster and more frequent travel, to have far-reaching social, economic and political impacts in nearly every country of the world. Such is the interdependence of global systems that any substantive 'event' might have global repercussions.

Form and nature of global interdependence

There are four main types of interdependence.

Economic interdependence

Countries' economies are linked with each other through:

- **trade** – more nations participate in international trade giving them access to products and services from many places. Countries rely on other countries to supply their needs and also to buy their exported products. Trading relationships are important as one nation's actions often have consequences for others. For example, Russia's policies on gas supply affect energy costs in Europe
- **advances in technology** – competing or co-operating with foreign businesses may bring innovations and new approaches, which can increase the quality of products and services and make them more affordable for consumers

- **employment** – jobs are generated in new areas but are sometimes lost elsewhere. For example, a UK manufacturing company might relocate to Malaysia, where they can increase production by hiring more workers at a lower cost. This would create jobs in Malaysia but cause job losses in the UK
- **international economic migration** – in 2019, international migrants accounted for more than 20 per cent of the total population in 48 countries. For example, around 20 per cent of the workforce in the United Arab Emirates are from the Philippines. Migrant workers are much needed in some sectors of the UAE economy, such as construction and engineering, where there are labour shortages. In turn, the Philippines relies on remittances sent back home from their workers in the UAE
- **TNCs and investment** – TNCs operate in several countries and outsource services and operations to foreign businesses. Companies from industrialised nations invest in developing countries, attracted by the lower cost of land and labour. They may form joint ventures with local companies, bringing technology, employment and infrastructure investment
- **supply chains** – components for any single manufactured product may be produced in different countries before being assembled. For example, in the manufacture of a particular vehicle model in the car industry, the different components, such as engines, wheels and body panels, will be produced in different countries. These components are brought together for assembly at another location before the vehicles are distributed globally
- **industrialisation** – globalisation has helped some countries, such as India, Brazil, Mexico, and Indonesia, to become industrialised. However, it has also led to deindustrialisation and structural unemployment in some regions of Europe and North America.

Political interdependence

- **Intergovernmental organisations** – global systems are supported by international organisations, which have been established to provide stability, dialogue and consensus between nations. For example:
  - the **International Monetary Fund (IMF)** and the **World Bank** facilitate international capital flows through the global financial system (see Figure 7.4, page 286)
  - the **World Trade Organization (WTO)** oversees the international trading system, seeking to establish barrier-free trade and fair access to markets
  - the **United Nations (UN)** is the leading organisation for global governance – (see Section 7.4 Global governance, page 327).
- **Security and stability** – one of the arguments for globalisation is that it will lead to greater political stability. Nations may co-operate more with each other as members of regional or international bodies, making it easier for governments to work together on common goals.
- **World peace** – some argue that wars are less likely to happen because of interdependence. In 1999, Thomas Friedman (a US writer on globalisation) outlined the 'Golden Arches' theory of conflict prevention. His theory suggests that 'No two countries that both had McDonald's had fought a war against each other since each got its McDonald's' because their economies and cultures are so interlinked. This theory has been disproved on a number of occasions since then.

Social interdependence

Technological advances in communications and falling costs of transportation have facilitated international exchanges and increased the interdependence of human societies.

- **Health** – another intergovernmental organisation, the **World Health Organization (WHO)**, has taken the lead in combating COVID-19 by co-ordinating advice on symptoms and development of vaccines; however, there have been distinctly different responses at a national level.
- **Education** – increases participation in student foreign exchange programmes (such as Erasmus in the EU). If more students study at foreign universities, it is beneficial both for the students and the institutions involved.
- **Culture** – social ties between countries can be strengthened through migration. For example, the large Indian diaspora settled in the UK has deepened the country's relationship with India, both economically and culturally. Extensive family networks now straddle the two countries.

Environmental interdependence

- '**Global commons**' – all countries are environmentally interdependent due to the impact of shared use of resources that are common, such as the oceans and the atmosphere. They are governed by the same international legislation (see Section 7.5 The 'global commons', page 333).
- **Global climate change** – concerns about the global environment, particularly climate change, have led to more international summits and the establishment of international agencies such as the **UN Framework Convention on Climate Change (UNFCCC)** and the **UN Environment Programme (UNEP)**. These encourage all nations and their citizens to work towards the shared goals of climate change mitigation and biodiversity protection.
- **Unsustainable practices** – some practices challenge environmental interdependency due to their unsustainable nature and because their causes or impacts affect more than one country. Examples include air pollution, acid rain deposition and deforestation. The underlying causes of rainforest deforestation in the Amazon or in South East Asia may be to supply hardwood or to clear land for commercial ranching or plantations to meet the demand from developed countries. Deforestation has impacts on a global scale and forest loss may become irreversible in some countries.

Issues associated with interdependence

Interdependence and globalisation have both positive and negative effects on different groups of people and places.

Links

There are links between Unequal flows of people and Chapter 8, Changing places – Demographic change and cultural characteristics, page 361.

Unequal flows of people

Globalisation enables workers to move more freely around the world. The main flows are, predictably, from developing countries towards highly developed economies (HDEs). This can have some positive effects but also causes inequalities, conflicts and injustices for people and places (Table 7.4).

Unequal flows of money

Unequal flows of money have other beneficial and negative impacts on people and places:

- **Remittances** – money sent home by migrants working overseas is an important source of income for developing economies (see page 287).
- **Loans** – developing countries borrow from the World Bank to fund projects, such as improving transport connectivity or for health and education programmes that increase standards, access and inclusion. Loans for engineering and construction projects such as transport or energy infrastructure can be a sound investment for countries. Most of the World Bank's lending to countries in 2020 was for emergency responses to the COVID-19 pandemic. Loans have to be paid back and this can only be done if the money has been invested effectively, which is more difficult if the loan is used for social rather than economic development projects.

Table 7.4 Positive and negative effects of flows of people

Positive effects (stability, growth and development)	Negative effects (inequalities, conflict and injustices)
• Reduced unemployment in places where there is a lack of work because of opportunities to work elsewhere • Addresses important skill and labour shortages in places • Reduces some inequality as foreign workers earn higher wages in HDEs • Remittances sent back to developing countries provide stability and opportunity for growth • Migrant workers increase workforce, pay taxes and spend money, which promotes growth and reduces dependency in HDEs with ageing populations • Some workers return to their country of origin equipped with new skills and ideas • Reduces population pressure on resources, such as food and water, and services, such as healthcare, in developing countries	• Developing countries lose younger, more talented workers attracted by higher wages – this 'brain drain' reinforces inequality and dependency • Loss of skilled workers adversely impacts on productivity, growth and development • Developing countries become over-dependent on remittances • Migrant workers and families may put pressure on health and education services in HDEs; they may be treated differently in these systems • If only workers are allowed to settle, families may be separated, which is unjust • Migrants may be segregated formally or informally into certain areas. • Resentment towards migrants may lead to ethnic and cultural conflict • Greater movement of labour may contribute to the risk of disease pandemics

- **Inflows of foreign direct investment (FDI)** – investment from TNCs or governments in Europe, Japan and the US to developing countries has raised average incomes and reduced poverty. It is thought that over one billion people have escaped poverty in Asia in the last 30 years. This would have been difficult to achieve without inflows of FDI for investment in key projects, but it can create dependency. Workers dependent on higher wages may be subject to poorer working conditions.
- **Growth of TNCs** – capital investment and taxes paid by successful global businesses to their host governments have stimulated economic growth. However, TNCs may pressure host governments to alleviate taxes or relax social and environmental laws.
- **Loss of profits ('leakages')** – the repatriation of profits by TNCs to their home country undermines the benefits gained from investment in developing countries. There may be limited 'trickle-down' from the FDI into the developing economy and this may exacerbate global inequalities in wealth.
- **Foreign aid** – aid will help low income countries in times of need, but it can reduce incentives for governments to help their own countries.

Unequal flows of ideas

Ideas generally flow from wealthier HDEs to developing countries that want to emulate the success enjoyed by richer nations. Many of these ideas are associated with global capitalism. Table 7.5 shows how these ideas may promote growth in some LDEs but how they can also cause other problems.

Unequal flows of technology

- **Information and data flows** – access to mobile and internet services is transforming people's lives in less developed economies. For example, the 'Village Phone' microfinance model pioneered in Bangladesh is now being used in Uganda. The project is supported by the World Bank and uses wireless technology. Loans allow the purchase of a smartphone, a car battery or solar charger to charge it and booster antennae that link to an improving network of mobile telephone masts. This enables access to the internet even in remote rural areas, improving people's quality of life. It enables small businesses, based on information provision, to be established, for example, providing information to local farmers on prices.

Figure 7.17 A man in Uganda charges a mobile phone using a solar charger. Access to mobile and internet services is transforming people's lives in less developed economies

- **Technology manufacture** – access to technology, such as computers and smartphones, in less developed countries is limited as technology is unaffordable to many. This may seem unjust given that the assembly of consumer technology products is often based in developing countries. Employees receive relatively low wages compared with the selling price of the goods.

Table 7.5 Advantages and disadvantages of flows of ideas

Benefits for stability	Ideas	Disadvantages, causing inequalities, conflicts or injustices
Dismantling state ownership of corporations can benefit consumers in LDEs by lowering prices	**Privatisation**	Profits are retained (rather than re-invested as is the case for nationalised industries), causing greater inequality and potentially inhibiting economic growth
Reducing government regulation and intervention can encourage enterprise	**Deregulation**	Deregulation can lead to more relaxed social and environmental laws in LDEs, causing social injustices and environmental degradation
Free trade allows global markets to develop and thrive and may help some LDEs to attract investment	**Free trade**	Free trade may not always be beneficial to some LDEs; they may be disadvantaged. LDE (infant) domestic industries may be outcompeted by free trade so some protection may be needed
Multi-culturalism enables developing countries to integrate into the global economy and to access markets	**Multi-culturalism**	Citizens may see it as a dilution of their culture and even a threat to their national sovereignty and identity

- **Technology transfer** – the transfer of labour-saving technologies from HDEs to developing countries can promote growth, but some technologies may put many people out of work, leading to high unemployment and poverty. This has happened with agricultural technologies and in the textile industry in some developing countries, such as India.

Inequality issues

Indicators suggest that globalisation is reducing global inequality through the transfer of capital and income from richer to poorer economies. Paradoxically, it may be increasing inequality within countries as richer members of societies cope better with the changes in jobs and technology.

Inequalities between countries

As communication and transport increase the integration of economies, developing countries are closing the gap with their rich-world counterparts. The development continuum, with the exception of some of the least developed countries, has become more condensed. The fastest growing economies continue to be in Asia and although countries in sub-Saharan Africa have a large gap to make up in living standards, some are now growing more quickly than most developed countries.

Table 7.6 shows the considerable differences in average income between a range of ten developed and developing economies in 2019. The average economic growth figures between 2014 and 2019, shown in the same table, support the idea that developing countries tend to have faster economic growth rates (Figure 7.18).

Skills focus

Consider methods of presenting the data in Table 7.6 graphically, in particular to show relationships between the size of GDP per capita and growth rates. This data could also provide the basis for a statistical analysis of relationships using a Spearman's rank correlation test.

Table 7.6 Inequalities in average income and trends in economic growth between 10 countries at different levels of development

Country	GDP per capita PPP* 2019 (US$)	Average GDP growth rate 2014–2019 (%)
Bangladesh	4,950	7.22
Ethiopia	2,312	9.27
India	7,034	6.83
Japan	43,236	1.06
Kenya	4,509	5.63
Mexico	20,411	2.18
Peru	13,380	3.08
Portugal	36,471	2.15
Turkey	27,875	4.28
UK	48,710	1.92

*PPP is Purchasing power parity = based on the equivalent of what you can purchase in each country

Source: World Bank (GDP) and IMF (Growth rates)

Inequalities within countries

The **Gini Index** is the statistical measure that is usually used to indicate levels of inequality of income distribution within a country. The Gini Index aggregates the inequalities in people's incomes into a single measure, giving a coefficient score of between zero and one. The higher the score within this range means greater income inequality.

The level of inequality varies widely around the world but there is usually more income inequality in LDEs than in wealthier countries – for example, South Africa has one of the highest scores with 0.63. The index indicates that, over the past 25 years, income inequality has increased in most developing countries, suggesting that globalisation has had a negative effect on income distribution. An increase in inequality also applies within highly developed economies. Many countries, including the UK, Canada and even traditionally egalitarian Sweden, have seen a rise in income inequality.

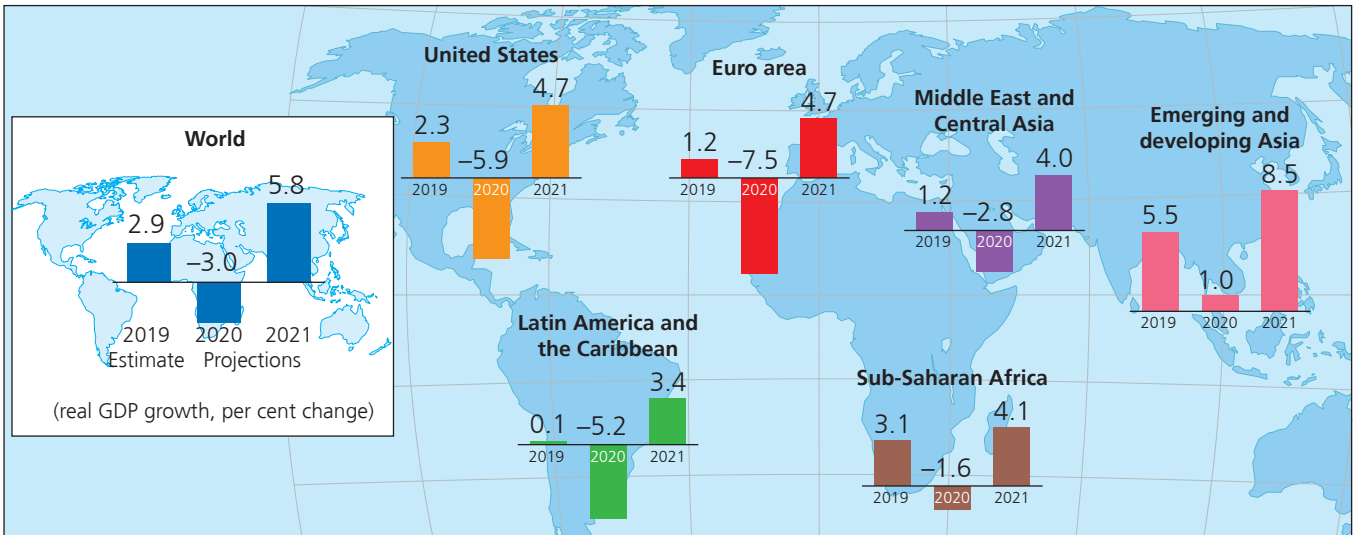


Figure 7.18 Global differential rates of projected GDP growth by region (including effects of COVID-19 pandemic in 2020) in per cent
Source: IMF

Unequal power relations in global systems

In general, wealthier developed countries wield more power, which enables them to steer global systems to their own advantage. There are a number of factors that reinforce this imbalance of power:

- HDEs have more wealth, advanced technology and, in many cases, more military power to use on a global scale. They provide aid, investment and transfers of technology and medical knowledge to assist developing countries but usually rely on geopolitical support in return.
- Wealthier countries have close relations with each other in groups such as the **G7**, **G20** and the **OECD**. By co-operating in these intergovernmental economic groups, they can synergise their efforts to be more influential in driving global economic and political systems.
- Developed countries have more influence in global governance through intergovernmental organisations such as the UN, the IMF, the World Bank and the WTO (see Section 7.4, Global governance). These have helped to improve stability and development, but they are led by the more powerful nations whose vested interests cause them to influence change to their own advantage.
- Developing countries have less influence and limited power to intervene so they have to depend on decisions made by wealthier countries. This means that less developed nations are constrained in how they are able to respond to geopolitical issues and may have to rely on geopolitical support from more powerful allies.

Key terms

G7 – The '**Group of Seven**' is an intergovernmental organisation made up of the world's seven largest advanced IMF economies: Canada, France, Germany, Italy, Japan, the UK and the US.

G20 – The G20 group is an international forum for the governments of 20 major economies. It includes the G7 countries and the EU as a single member. It was established in 1999 to give a voice to the major developing economies. This group includes China.

OECD – The Organisation of Economic Co-operation and Development is a group of 34 of the richest and most influential countries globally.

Geopolitical events

Geopolitical issues arise for a variety of reasons but are usually rooted in political or economic conflicts between two countries or two groups of countries. Conflicts may be based on resource shortages or on strategic territorial claims, exploration rights or supporting political allies or minority groups within countries. For example, the US permanent position on the UN Security Council may have shielded Israel from wider UN criticism for policies disadvantaging Palestinians. The examples below demonstrate how two states, Russia and China (both permanent members of the UN Security Council), are exerting their considerable political or economic power in global systems to influence geopolitical events to their advantage.

Russia's annexation of Crimea

When Russia annexed Crimea in March 2014, it was considered to be an exertion of power by a military superpower, invading Ukrainian sovereign territory. Russia's President Putin would have expected a reaction from the international community to this annexation, especially from Western powers such as the US and those in Europe. Crimea is strategically important to Russia and Putin would have gauged the likely response and calculated that it would not escalate into armed conflict with the West.

Background to the annexation:

- Crimea was once part of Russia but was ceded to Ukraine during the Soviet era when both states were part of the Union of Soviet Socialist Republics (USSR).
- After the collapse of the USSR in the 1990s, Russia continued to use Sevastopol (a much-needed warm water naval port, which is ice-free year round) on Crimea's Black Sea coast, by leasing it from Ukraine.
- Following independence, Ukraine remained largely pro-Russian. As a legacy of the Soviet era, Ukraine has a large ethnic Russian population and is also dependent on Russia for energy supplies.
- However, in 2014 a revolution took place in Ukraine when the 'Euromaidan' political movement demanded closer ties with the EU and eventually overthrew the pro-Russian government.
- This geopolitical shift to the 'West' by its former ally prompted Russia to take decisive action by taking military control of the Crimean Peninsula. Putin stated he was safeguarding the majority 60 per cent Russian population living there and the special status of Sevastopol.
- The Crimean Supreme Council had voted to accede to Russia, which gave some authority to Russia's claim, but 114 UN member states do not recognise the annexation; 17 pro-Russian states do.
- Since the annexation, Russia has strengthened its Black Sea fleet and supported further Russian separatist groups in eastern Ukraine where conflict still ensues.

Ukraine lacks the leverage to restore sovereignty over Crimea; they were not a member of western groups such as the EU or NATO, so no military power from the West came to assist. The response from developed nations in the West was largely political and economic:

- Russia's membership of the G8 was suspended so it became the G7.
- The EU and the USA imposed trade sanctions on Russia as a result of the annexation and because of its supposed backing of separatists in eastern Ukraine.
- The trade sanctions have had a negative impact on the Russian economy but those imposed by the EU were limited as the EU is heavily dependent on Russian energy supplies.
- To counterbalance the shift in power in the region, it is possible that the US will start to strengthen its Black Sea naval presence at its bases in Romania and Turkey (both NATO allies).

Analysts argue that this ongoing political battle between Russia and the West is a feature of increased '**nationalism**' that is actually slowing down globalisation.

China's expansive foreign policy

China's influence as an emerging economy has not been confined to development within its own borders. China has chosen to strategically expand its wealth and influence by investing in other parts of the world. This is reflected by two related foreign policies to develop infrastructure:

- investment in Africa
- the '**Belt and Road**' initiative.

Investment in Africa:

China's role in Africa is varied (Figure 7.19), but it has a clear objective to extract a range of raw materials, including energy resources, minerals (such as investment in copper mines in Zambia) and precious metals, to support industrial expansion in China. China also needs to ensure a secure flow of these materials to coastal ports by developing infrastructure and needs to gain political support and stability by investing in poorer African countries to develop housing, healthcare and education. Notable investments include:

- a joint project with Ethiopia to build the Grand Renaissance Dam to provide HEP
- development of the port of Mombasa in Kenya with a $14 billion road and bridge link to Nairobi and on to South Sudan's oil fields
- modernisation of the Benguela railway linking the Democratic Republic of Congo's resources (cobalt and copper) to the Angolan port of Lobito by the Chinese Railway Engineering Corporation (CREC)
- construction of a new international airport in Luanda, Angola.

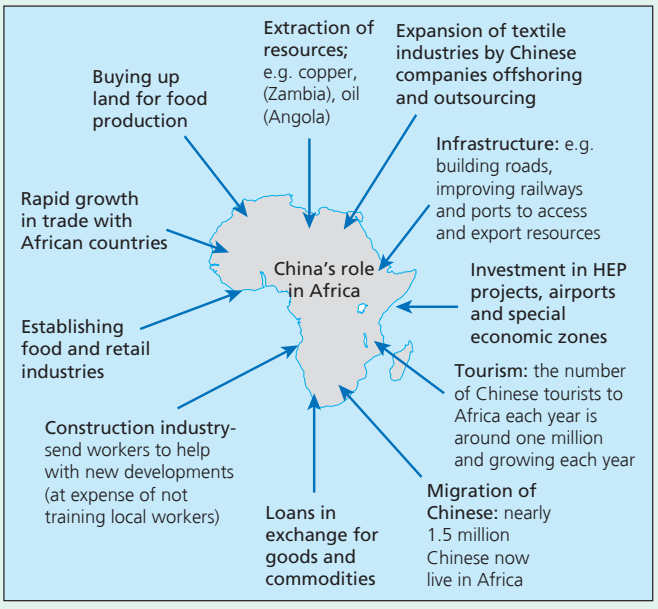


Figure 7.19 Chinese interests in Africa

The Belt and Road Initiative (BRI):

In 2013, President Xi Jinping launched an ambitious infrastructure development to boost trade and stimulate economic growth across Asia and beyond, into Europe and Africa. The BRI is a project that involves building a network of overland road, rail and pipeline corridors (the 'belts') and maritime routes through ports and shipping lanes (the 'roads') (Figure 7.20). The project is expected to cost more than US$1 trillion and involves construction in more than sixty countries. It has been described as a 'modern Silk Road'. Some notable features of the project include:

- the China–Pakistan Corridor – a shared $46 billion project between the two countries to build a highway of roads, railways and pipelines from China to the port of Gwadar on the Arabian Sea. It gives China more direct access to imported commodities from Africa and also oil from Iran, whilst also bypassing the shipping lane choke point of the Strait of Malacca
- a rail and road corridor into Europe that crosses central Asian countries and should help these landlocked economies to develop with improved access to markets for their products.

The Chinese state that they are taking responsibility for improving international co-operation and that the trade generated may kick start a new era of globalisation, though it has raised a number of geopolitical concerns:

- It is a form of economic imperialism that gives China too much leverage over other countries, often those that are smaller and poorer.
- China will be using its foreign exchange assets to underwrite many of the projects but will expect some input or repayment from the countries in which it is investing. Some observers suggest that poorer countries, especially those in Central Asia, will not be able to afford to repay these debts.
- Part of the BRI includes China's claim to sovereignty of the South China Sea and the energy supplies that are believed to lie beneath it. This is based on a map drawn with a 'nine-dash line' that encircles the numerous reefs and atolls, such as the Spratly Islands, which China lays claim to.
- The South China Sea is, by most definitions, in international waters and so many other countries including Malaysia, Philippines, Brunei and Vietnam also have territorial claims as well as fishing rights. These countries are all ASEAN members, but China prefers to negotiate on an individual basis, which restricts members' ability to respond.
- China's commercial presence is likely to lead to expanded military presence.

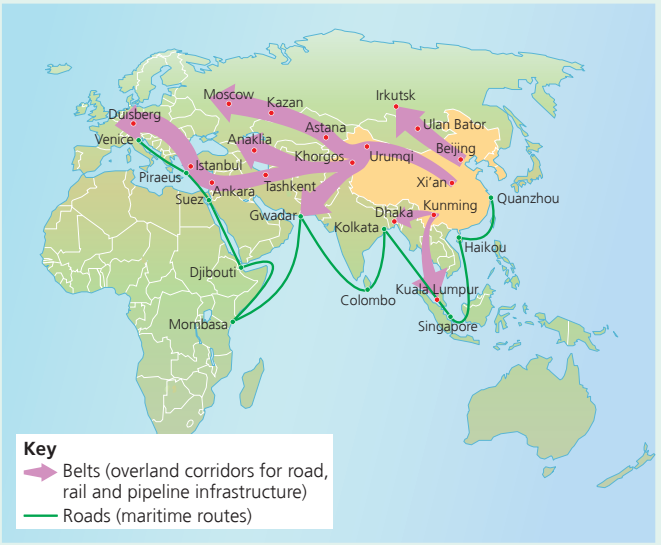


Figure 7.20 China's 'Belt and Road Initiative'

Source: Guardian News & Media Limited

Arguably, these infrastructure projects are beneficial for economic growth and the development of poorer nations and may help global integration. However, the state with the greatest power will usually derive most benefit from the system.

Review questions

6 Identify the forms that economic interdependence takes in the globalised world.

7 Explain how unequal flows of people within global systems can cause inequalities.

8 Suggest how unequal flows of money can promote stability.

9 Explain why some countries can only respond to geopolitical events in a constrained way.

10 Using an example of a country or group of countries with global power, assess how they have steered global systems to their own advantage.

7.3 International trade and access to markets

The notion of **comparative advantage** is why most nations trade. The theory behind comparative advantage suggests that countries should specialise in providing goods and services that they excel at producing. They trade these for the things they are not so good at producing. It is easier to make these trading exchanges if there are fewer barriers. Theoretically production should increase in each country, and globally, because each country is concentrating on what it does best.

Key term

Multilateral trade agreement – A trade agreement negotiated between more than two countries or groups of countries at the same time, usually facilitated by the WTO.

International trade operates in a 'rules-based' system which is governed by the World Trade Organization (WTO) (see Section 7.4, Global governance).

Barriers to trade and protectionism

A barrier to trade is a government-imposed restraint on the flow of international goods or services. The most common barrier to trade is a **tariff** – a tax on imports. Although free trade has eroded the practice of protectionism, it is still a strategy used by governments in both developed and developing countries to protect their established or emerging industries.

There are a range of other barriers to trade, which are used as either protectionist strategies or sometimes as a means to affect diplomatic relations.

- **Import licence:** A document issued by a national government authorising the importation of certain goods from a specific source.
- **Import quotas:** Quotas set a physical limit restricting the quantity of a particular commodity that can be imported into the country within a specific period of time.
- **Subsidies:** These are grants or allowances usually awarded to domestic producers to reduce their costs and make them more competitive against imported goods.
- **Sanctions:** These are restrictions on exports implemented for political reasons by countries and international organisations to maintain international peace and security.
- **Embargoes:** These involve the partial or complete prohibition of commerce and trade with a particular country. They are usually put into practice for political rather than commercial reasons.
- **Regulatory or technical restrictions:** These are restrictions placed on imports based on obstacles such as the quality standards of goods or how they are produced. For example, the EU attempts to put restrictions on the import of goods knowingly produced using child labour.

Trends in the volume and pattern of international trade and investment

The emergence of **free market** ideas at the end of the twentieth century saw a push to remove existing barriers to international trade. With an increase in the mobility of factors of production, especially capital, **regional trade agreements (RTAs)** (see pages 295–7) emerged and the global trade framework was strengthened by the World Trade Organization (WTO) which continued the work of gradually lowering the barriers to international trade, with **free trade** as its aim.

- Reaching agreements has not been easy but average trade tariffs on most goods have reduced.
- Since the global financial crisis of 2008–2009, multilateral trade agreements on a global scale have been difficult to achieve.
- International trade flows have fluctuated over recent years.

- During recessions, international trade inevitably stalls and the COVID-19 pandemic has caused a major fall in the volume of trade. It will be difficult to predict the recovery rate during the next ten years.

Volume and pattern of trade

The growth of international trade has stalled in recent years. In 2018, the value of world trade in goods was US $19.48 trillion, only US $0.50 trillion more than its value in 2013. The total value of services traded globally in 2018 reached US $5.8 trillion, a slightly larger increase from five years earlier.

The pattern of global trade has changed as a result of the greater integration of economies (Figure 7.21). This is partly explained by the increasing complexity of global supply chains for many manufactured products. For example, cars may have thousands of component parts, with production of different parts dispersed around the world and then imported to an assembly plant. Consequently, there are a number of factors driving current patterns of global trade:

- **comparative advantage** – countries specialise in producing and exporting goods that they can produce more efficiently at a lower cost
- **proximity** – countries are more likely to trade with their neighbours, partly because this reduces transport costs but also for cultural, historic or linguistic reasons. This is why trade is often organised on a regional basis and intra-regional trade still dominates the global pattern
- **agglomeration** – some industries tend to cluster in geographical areas as sharing of regional skills and specialist information saves costs
- **market size and strength** – exporters are drawn to larger, more affluent and growing markets where there is potential to increase volume and value of sales
- **geopolitical relationships** – political alliances are important in determining which countries co-operate and trade with each other; conversely, conflict may lead to sanctions or embargoes.

Geographically, trade is still dominated by a few large economic blocs, mainly North America, Europe and the Asia-Pacific region. The Asia-Pacific region is the fastest growing trading region and China is now the world's largest trader, surpassing the US in total trade value in 2018. Figure 7.21 shows the main pattern of global trade flows and can be summarised as follows:

- Intra-regional trade is particularly strong within Europe and within the Asia-Pacific region; it is also strong in North America but to a lesser extent.

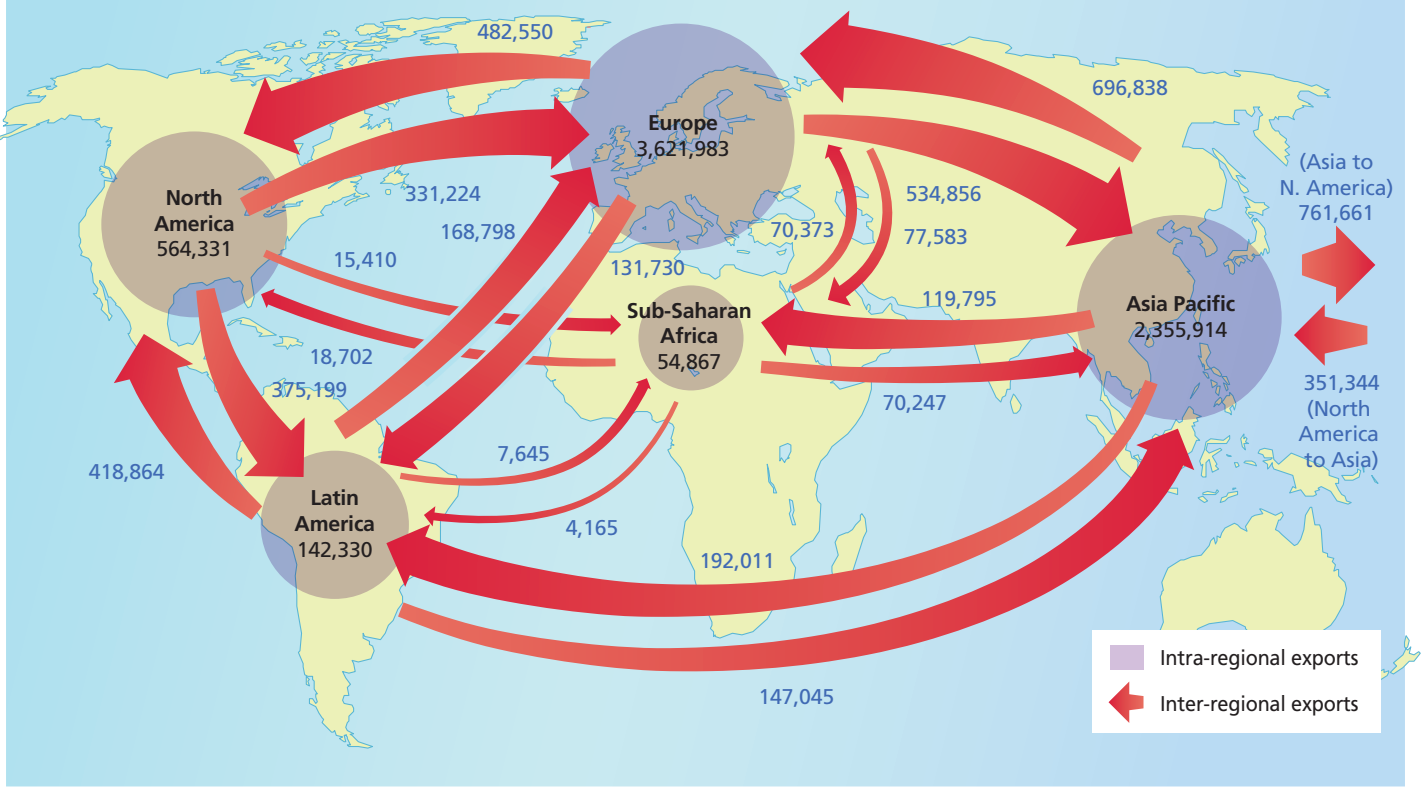


Figure 7.21 Patterns of intra-regional and inter-regional world trade, 2016 (US$ millions)

Source: United Nations

- In Latin America and Sub-Saharan Africa, there is less intra-regional trade as most trade involves imports and exports coming into or going out of these regions.
- Latin America has the strongest trade flows with North America, though trade with Europe and Asia-Pacific is also good; it has an overall trade surplus with North America and Europe but a deficit with the Asia-Pacific region.
- North America (mainly USA) has a trade deficit with all other regions whereas Asia-Pacific has a surplus flow of trade with all the other regions.
- Inter-regional trade is increasing between the Asia-Pacific region, Sub-Saharan Africa and Latin America.

The backlash against further liberalisation of global trade seen in the USA and parts of Europe (see below) may prompt them to pull back from further integration. As this happens, stronger links developing between the Asia-Pacific region, Africa and Latin America will become more important and may transform the pattern of global trade in the future.

Trends in investment

Foreign direct investment (FDI) is an important source of funding for development in all countries, especially for the least developed economies wanting to grow. Whether they are countries just beginning to modernise, or those of the rich **OECD** group, it is difficult to 'jumpstart' economic growth without an injection of capital.

In 2019, FDI inflows totalled US $1.39 trillion, which was the fourth consecutive annual decline. FDI to developed economies fell to US $640 billion, whereas flows to developing economies remained stable at around US $750 billion. A 2019 **UN Conference on Trade and Development** (UNCTAD) report found the following trends for FDI:

- Developing countries receive almost twice as much FDI as they initiate.
- FDI from North America has slumped.
- The Asia-Pacific region accounts for 40 per cent of FDI inflows.
- Eight of the top 20 recipients of FDI were developing economies.
- The largest FDI investors were Japan, China and France.

Table 7.7 shows the top ten recipients of FDI by total and also the top ten recipients of FDI as a proportion of their GDP (the global average is 3.4 per cent).

Table 7.7 Top ten recipients of FDI by total and as a percentage of their GDP

Recipient by total	FDI received, 2019 (US$ billion)	Recipient as a percentage of GDP	FDI as a percentage of GDP, 2018 (%)
1. USA	251	1. Congo	38.29
2. China	140	2. Malta	32.64
3. Singapore	110	3. Guyana	30.43
4. Brazil	75	4. Hong Kong	26.76
5. UK	61	5. Singapore	25.00
6. Hong Kong	55	6. Cyprus	20.86
7. France	52	7. Seychelles	19.35
8. India	49	8. Mozambique	18.20
9. Canada	47	9. Ireland	16.87
10. Germany	40	10. Mongolia	14.94

Source: UNCTAD and OECD

Skills focus

Find out the GDP for each country listed in Table 7.7. For each set of countries, rank them by GDP and using the data in Table 7.7, set up a Spearman's Rank relationship table. Calculate the correlation coefficient for each set of countries. Is there a significant positive or negative correlation? Which is the stronger?

There are a number of attractions that pull in FDI investment, but the main ones are:

- **manufacturing industries** – for example, foreign motor companies investing in the USA and UK or offshoring and outsourcing investment in China and India
- **natural resource development** – investment from mining corporations, for example, in Brazil, Congo, Guyana and Mongolia
- **financial business services** – attract investment in Singapore and Hong Kong
- **large and accessible consumer markets** – attract both manufacturers and service providers, for example, the EU is a single market of 500 million people
- **lower business taxes** – attract investment to Ireland and Cyprus, which have relatively low taxes and are also part of the large EU market; this also demonstrates that in some cases it is a combination of factors that attracts foreign investment.

Trading relationships and patterns

Highly developed economies (HDEs)

USA

The USA has traditionally been a 'protectionist' economy, lagging behind in entering any formal trade agreements. In the eight years of President Obama's administration, free trade agreements were being negotiated with Asia-Pacific countries in the Trans-Pacific Partnership (TTP) and with the EU in the Transatlantic Trade and Investment Partnership (TTIP). President Trump was elected in 2016 on a platform of protecting and reinvigorating US domestic industries. Consequently, the Obama strategy of 'opening up' to multilateral trade agreements has been reversed. As a result:

- Trump withdrew the US from TTP negotiations with countries around the Pacific. The other countries involved have since restarted the negotiations without US involvement
- the TTIP negotiations collapsed in 2016; attempts were made by the US to resurrect a similar deal in 2018 but US demands for access to EU markets and industries were too much for many EU countries to accept. Negotiations were finally declared closed by the EU in 2019
- the NAFTA free trade agreement with Canada and Mexico was renegotiated to form the United States-Mexico-Canada Agreement (USMCA) (see page 297).
- Trump aimed to pursue **bilateral trade deals** with individual countries. The US has twenty separate free trade agreements, including some with countries in the TPP group; others are mainly with central American or Middle East nations; South Korea is the only one in the Asia-Pacific region.

A change of administration, with the election of Joe Biden as President at the end of 2020, is likely to take the US in a different direction again on trade.

The European Union

As a customs union, the EU is protectionist in its own way. Though its 27 members trade freely with each other, there is an external tariff barrier for countries outside of the Union (with the exception of those in EFTA such as Norway and Iceland). This also means that the EU negotiates trade deals as a bloc representing all member states, so any agreements reached apply to all members.

- 65 per cent of the trade in the EU is intra-regional. This high level of trade between members fosters greater integration and interdependence.
- The EU has negotiated trade deals bilaterally with countries in different parts of the world. They are not all reciprocal free trade agreements but can offer reduced tariffs and easing of barriers. These include agreements with Australia, New Zealand, Mexico, Japan, Singapore and Vietnam.
- One of the most recent deals was the Canada–EU deal, approved by the EU Parliament in 2017. The comprehensive economic and trade agreement (CETA) is a broad-ranging agreement with virtually free trade and regulatory compatibility.
- In 2019, the EU agreed a trade deal with Mercosur, the South American trading bloc, which has still to be ratified by all member states.

Anti-free trade movements have gained popularity in the EU over recent years. They started with successful campaigns to stop trade negotiations with the US, and also put the Mercosur agreement in doubt:

- Farmers in countries such as Belgium, France and the Netherlands are concerned that imports of cheap beef and sugar will lead to unfair competition.
- Environmental groups suggest that Brazilian and Argentinian cattle are injected with chemicals and that Mercosur countries do not meet EU standards in working conditions, food production or environmental protection. They also object to Brazilian President Bolsonaro's attitude towards the commercial development of the Amazon rainforest.

With such opposition to cheap imports and differing food production and quality standards, it is difficult to foresee any trade agreement between the EU and the USA.

The loss of the UK from the EU is potentially detrimental as the UK was a major contributor to EU funds, it attracted investment in manufacturing and provided a sizeable market. During the 2020 Brexit transition period, the UK negotiated a new free trade deal with the EU. It was important for many UK firms that a good deal was agreed as decades of integration mean that they have built up supply networks across the EU.

Emerging economies

China

China's phenomenal economic growth rate in the last 30 years has been based on the expansion of manufacturing of consumer goods for richer countries. In the last few years:

- China has emerged as the world's major trading power, responsible for nearly 20 per cent of all global exports and around 15 per cent of all global imports
- there has been a clear pattern of importing mainly raw materials from developing countries, primarily from Latin America and Sub-Saharan Africa, and exporting processed metals such as steel and manufactured electronic consumer goods to developed countries, particularly those in Europe and North America
- the export of cheaply produced, state-subsidised steel onto the world market has triggered a trade war with the USA, which has escalated to include imposition of high tariffs on traded goods between the two countries; some analysts argue that this is the main cause of current static growth of world trade.

More information on China's intentions to boost trade through its Belt and Road Initiative is detailed in the case study on pages 304–5.

India

India has enjoyed sustained and rapid industrial growth over the past 25 years, though its emergence as a major economic and trading power has been slightly less meteoric than that of China. Arguably, India has a more diverse economy with a more globally integrated commercial and services sector. As in China, food security is a crucial issue for India; both countries experienced devastating famines in the 1960s. The success of the Green Revolution in India, in terms of staple food production, has alleviated these concerns. At the same time, the revolution caused agricultural labourers to be laid off and many migrated to towns and cities in search of work. Wholesale change in rural India has not been without problems but the legacy has given India a demographic dividend of a large, young workforce that has attracted investment for industrial development.

India has traditionally been more westward looking in terms of trading partners and this is still largely the case, possibly as a result of a high proficiency of English language amongst its population. However, trade with emerging markets in Asia and the Middle East has expanded significantly, giving India the following trade pattern in 2018:

- Top exports are refined petroleum, chemicals, gold and diamonds, rice, cars and textiles.
- Top imports are crude oil (for refining and exporting), coal briquettes and gas.
- Main export destinations are the EU (mainly Germany and the UK), the USA, the UAE, China, Hong Kong and Singapore.
- Imports come mainly from China, the USA, Saudi Arabia and the UAE.
- India's imports have expanded faster than its exports, so balance of payments problems make the economy vulnerable. In 2018, there was a deficit of US $166 billion.

Less developed economies

Latin America

Latin America's relations with other trading entities will be important in the future. Latin America has two distinct trading blocs – Mercosur and the Pacific Alliance – each with different approaches to developing their trading links (Figure 7.22).

Figure 7.22 Mercosur and the Pacific Alliance in Latin America

Mercosur: Formed in 1991, it is the larger of the two groups in terms of GDP (US $3.4 trillion) and comprises Brazil, Argentina, Uruguay and Paraguay. Venezuela was suspended indefinitely in 2016 as it failed to comply with many of the group's trade regulations.

- Mercosur is a single market and customs union and tries to emulate the EU. It is called the 'Common Market of the South' and allows the free movement of labour between member states.
- The most exported products are raw materials, energy, mineral and food resources, which enable Mercosur to trade globally, but mean it is vulnerable to fluctuating global trade prices.
- With the exception of mineral exports to China, Mercosur tends to view the EU and North America as its main markets. Its proposed trade deal with the EU will be a major development once it is ratified.

Pacific Alliance (PA): Formed more recently in 2011 and comprises Chile, Peru, Colombia and Mexico, with a combined GDP of $2.2 trillion. Ecuador is soon expected to join the Alliance. Its key differences to Mercosur are that:

- it has been more open to making bilateral trade agreements with other nations
- it tends to see the Asia-Pacific as its main market
- some PA countries, notably Peru and Colombia, are experiencing faster growth both in terms of individual economies and volume of trade. This may reflect the trend that trans-Pacific trade has been more dynamic than trans-Atlantic trade over recent years.

Since 2017 there has been more likelihood that the two groups could merge to form one large free trade area. This would be directed by the intergovernmental body, the Union of South American Nations (UNASUR), and would involve all Latin American countries being members. Any such agreement would increase intra-regional trade as well as giving all Latin American countries greater trading power on a global scale. The problems for greater integration in Latin America are the varying levels of economic development and the different political ideologies of the nations involved. This is typified by the political unrest seen in both Venezuela and Ecuador in recent years.

Sub-Saharan Africa

Trading relationships in sub-Saharan Africa are similar to Latin America but are more disadvantaged and less integrated.

- There is minimal intra-regional trade and Africa's main exports remain primary resources and commodities such as minerals, energy and food resources.
- Lack of skills, poor transport and energy infrastructure and widespread corruption have discouraged investment into industrial development that would enable African countries to add value to their exports. They rely on imports of most manufactured goods.
- Traditionally, Africa's main trading partner has been Europe, but this has changed with China's involvement in the continent.
- Africa is a larger, poorer and more divided continent than South America. There are many more language, ethnic and cultural divisions.

Two factors may change trading relationships for Africa now and in the future:

1. **China's investment** – Chinese investment in developing infrastructure has undoubtedly been for China's own benefit. Nonetheless, it has also increased the connectivity and integration of different African nations (see Figure 7.19, page 305).
2. **The African Continent Free Trade Area (AfCFTA)** – Africa has five main regional trading groups with some overlap but minimal integration between them. Even within these groupings there are linguistic and cultural divisions. In 2018, after years of talks led by the African Union, these five groups agreed to work together and created AfCFTA, the world's largest free trade area and, with a population of 1.3 billion people, the largest single market. It came into effect in 2019 without Nigeria on board but it joined shortly afterwards so that 54 of the 55 African nations have now signed the deal. AfCFTA has the following trade implications for African nations:
   - Its main purpose is to increase **intra-regional trade**, which is historically low at around 16 per cent. This is seen as critical for growth and job creation on the continent.

- It gives African nations more voice and leverage in global trade.
- As a result of the existing fragmentation and different stages of development of nations, trade liberalisation has the potential to disadvantage the poorest people within some countries.
- The African Union needs to ensure that prosperity is shared by creating supportive policies, eliminating monopolies and other uncompetitive behaviour, such as state subsidies.

Fair trade in LDEs

Fair trade is a social movement whose goal is to help producers in LDEs achieve better trading conditions. The movement focuses mainly on agricultural-based products including coffee, tea, cocoa, sugar, bananas, cotton and chocolate.

Fair trade supporters argue that those producing the commodities do not get an equitable deal from the organisations they supply their produce to. Buyers such as transnational corporations, food processing companies or supermarket chains from developed countries are able to force down the prices because the individual suppliers:

- have little market influence
- are extremely reliant on the income from their goods.

The fair trade movement advocates the payment of higher prices to producers, as well as helping them to achieve improved social and environmental standards. Fair trade organisations assist producers to form into co-operatives giving them more influence in governing market conditions and the power to negotiate better deals with buyers or to supply direct. The goods are labelled as 'fair trade' goods with the International Fairtrade Certification mark, so that 'ethical consumers' can recognise that they are buying goods for which the producer received a fair price. It is important not to confuse fair trade with free trade, especially as the two ideas can be conflicting. The terms are quite different – Table 7.8 summarises the key differences between the two.

Table 7.8 Fairtrade is NOT free trade

	Free trade	Fair trade
Main goal	To increase nations' economic growth	To empower marginalised people and improve the quality of their lives
Focuses on	Trade policies between countries	Commerce among individuals and businesses
Primarily benefits	Multinational corporations, powerful business interests	Vulnerable farmers, artisans and workers in less industrialised countries
Critics say	Unjust to marginalised people; sacrifices environmental sustainability	Interferes with free market, inefficient, too small scale for impact
Major actions	Countries lower tariffs, quotas, labour and environmental standards	Businesses offer producers favourable financing, long-term relationships, minimum prices and higher labour and environmental standards
Producer income determined by	Market and government policies	Living wage and community improvement costs
Supply chain	Includes many parties between producer and consumer	Includes fewer parties, more direct trade
Key advocate organisations	World Trade Organization, World Bank, International Monetary Fund	Fairtrade Labelling Organizations, International World Fair Trade Organization

Source: Adapted from GlobeIn

Differential access to markets

Each country's access to markets is determined by its ability to trade with other countries. Access to international markets is limited by any barriers that might be in place, such as tariffs, quotas or trade regulations, and this will depend on the degree of protectionism used by other countries or trading blocs.

Access associated with levels of economic development

Wealthier, more developed economies generally have better access to markets as:

- they can afford to pay higher tariffs on exported goods

- they can increase access by investing into foreign markets; for example, TNCs can relocate their production inside foreign markets to avoid tariffs
- HDEs often group together to form trading blocs and customs unions, such as the EU, to allow free trade with each other but impose barriers on countries outside the union.

In contrast, LDEs struggle to pay high tariffs or invest in foreign markets and have difficulties entering into trading agreements with richer countries. They often focus on exporting primary products that are traditionally subject to price volatility. This exacerbates the poor access to markets that LDEs already experience and reinforces inequality.

Trading agreements

Trade agreements have benefits and disadvantages for countries seeking to improve their access to foreign markets. Negotiating a trade deal or entering a trade bloc agreement improves access to the markets of other countries within that bloc. This is beneficial for lower income countries that may be able to trade freely at lower prices, especially if they have exchange rate advantages. However, if a country becomes part of a customs union, it will restrict its ability to negotiate new deals and expand trade beyond the members of that bloc. Free trade agreements may also remove the protection given to growing domestic industries before they are strong enough to compete. On balance, it is better to be part of a trade agreement than to be left out because:

- being part of a trading group gives more leverage when trying to gain access to other markets
- agreements between two or more trading groups can open up wider markets, especially for LDEs
- non-inclusion reinforces a position of isolation and having to overcome trade barriers to gain access.

Impacts on economic and societal well-being

Trade sanctions are used by nations as a political and economic 'weapon' against individual countries with whom there is a dispute. In 2006, the UN Security Council imposed sanctions against Iran because of its refusal to suspend its nuclear programme. This severely harmed Iran's economy until a deal was made to lift sanctions in 2016. Whether enforced by sanctions or whether due to unequal levels of development, poor access to international markets can have devastating economic and social impacts on nations and their populations:

- Limited revenue from exports and business taxes means lack of money for investment or development.
- Poor access to markets deters foreign investment.
- Balance of payment problems will increase debt and restrict economic growth.
- Limited investment, growth and development will cause unemployment and poverty.
- There is a limited range and a lack of affordability of imported goods for consumers.
- In the worst situations, a lack of ability to trade can even threaten food or energy security.

Measures to combat differential access

Special and differential treatment agreements (SDTs)

Special and differential treatment (SDT) agreements have been a feature of the WTO's (previously GATT's) multilateral trading system for the past sixty years. In 1971, the UN General Assembly created a category of 'Least Developed Countries' (LDCs) to provide developing countries with support measures to overcome their trading disadvantages. These measures gave LDCs **preferential access** through SDT to developed markets in trade agreements. SDT agreements meant that LDEs should:

- enjoy privileged access to the markets of their trading partners, particularly developed countries
- have the right to restrict imports to a greater degree than developed countries
- be allowed additional freedom to subsidise exports.

Trade plays an important role in promoting economic development and preferential access has allowed many LDCs to diversify their economies, achieve economic take-off and lift some of their population out of poverty. SDTs, however, are not without their problems:

- The WTO Doha Development round of trade talks made little progress in making SDT measures more effective for developing countries (see WTO on pages 329–30).
- Richer nations are concerned that non-reciprocal preferential trading agreements cause unfair trade as many emerging economies, including China, still have 'developing country' status at the WTO. China refuses to give up their SDT rights and is still able to subsidise exports. HDEs fear that this will result in

cheap imports flooding markets and undermining their own industrial base.
- The agreements are difficult to apply in a fair or standardised manner and the lack of reciprocity in the agreements has deterred some developed countries from participation.

In many cases, SDTs have been replaced by bilateral trade agreements between poorer countries and richer trading partners. For example, Mexico was previously a beneficiary of preferential access to Canadian and US markets but has arguably benefited more as a member of NAFTA (now USMCA).

Nature and role of transnational corporations (TNCs)

Links

There are links here with the impacts of external forces in Chapter 8, Changing Places, pages 361–2.

Transnational corporations (TNCs) are companies that operate in at least two countries, with a headquarters based in one country but with business operations usually in a number of others. TNCs take many different forms and are based in different economic sectors. They no longer only originate from more developed countries; emerging economies also have TNCs which are major global companies. Some TNCs are very powerful, with many having larger revenues than the total GDP of nations (see Table 7.9). They can have considerable political influence, for example, persuading governments of developing economies to reduce taxes or to create special economic zones to encourage the TNC to invest in their country.

TNCs may operate in more than one country for a number of reasons:
- **to escape trade tariffs** – for example, Nissan's decision to produce cars in Sunderland was largely to gain barrier-free access to the EU market
- **to find the lowest cost location for their production** – for example, Hewlett-Packard in Malaysia
- **to take advantage of foreign exchange rates that make exports cheaper** – for example, Dyson in Malaysia
- **to reach foreign markets more effectively** – for example, McDonald's
- **to exploit mineral or other resources available in foreign countries** – for example, BP in Azerbaijan.

There are certain characteristics thought to be common to TNC organisations, including:
- maximising global economies of scale by organising production to reduce costs
- sourcing raw materials or components at the lowest cost
- controlling key supply chains
- control of processing at each stage of production
- branding of products/services so they are easily recognisable.

Spatial organisation

Transnational corporations are globally integrated enterprises because they locate different functions to gain an effective blend of cost, skill and environment.

Traditionally, the company headquarters are based in a major city in the home country. Most TNCs have subsidiary headquarters in each continent, or in countries where their main operations are based.

In order to maintain their position competitively, TNCs engage in research and development (R&D) activities. These also tend to be based in the country of origin and will often locate near to centres of higher education, to take advantage of a graduate labour market, or to make use of university research facilities. Manufacturing operations are often located in LDEs, where land, labour and material costs are lower (**offshoring**).

Offshoring and outsourcing

The practices of offshoring and **outsourcing** are slightly different concepts but have similar causes and consequences. Both practices are used by TNCs to organise production or support services more cost effectively. They are both largely one directional and

usually involve exporting manufacturing or service jobs from costly economies in Europe and North America to lower wage economies.

- **Offshoring** involves relocating a part of the organisation, such as manufacturing operations, to an overseas location to take advantage of lower costs or alternatively to access foreign markets. This includes TNCs locating factories in higher wage economies, such as Toyota in the UK and US to access these wealthier markets.
- **Outsourcing** is a strategy that involves subcontracting part of the business operation to another company, usually in another country where costs are lower.

Both practices can contribute to growth, development and increasing skills and technology in LDEs but they also have negative effects, causing inequalities and injustices in both HDEs and LDEs (Table 7.10, page 317).

Table 7.9 List of largest 10 TNCs by revenue in 2016

The Fortune 2016 Global 500 companies	
1. Walmart (10)	United States
2. State Grid (14)	China
3. China National Petroleum (15)	China
4. Sinopec Group (16)	China
5. Royal Dutch Shell (18)	United Kingdom–The Netherlands
6. Exxon Mobil (21)	United States
7. Volkswagen (22)	Germany
8. Toyota Motor (23)	Japan
9. Apple (25)	United States
10. BP (27)	United Kingdom

Figures in brackets are the rank of the TNC if ranked against national GDPs in the world.

Source: Fortune, 2016

Production

Production operations of TNCs involved in the **primary sector** will be based wherever there are unexploited resources. This tends to be in developing economies as reserves in more developed countries have largely been depleted. However, a combination of rising world prices and new technologies can make access to new reserves of raw materials viable in the home country. For example, the recent development of hydraulic fracturing, known as 'fracking', has revitalised oil and gas industry bases in North America.

Figure 7.23 South Korean company Samsung has located LED TV production in Kaluga in western Russia. Russian industry gains from the transfer of Samsung's technology, skills and capital investment

For TNCs in the **secondary sector**, production operations have largely been located in the manufacturing regions of developing countries, especially in South East and South Asia. These areas attract TNCs because:

- labour costs are lower
- there is investment in education, which makes it easier to train workers
- the work ethic means workers are willing to work long hours, with relatively few holidays, in a non-unionised labour environment
- there may be government incentives such as tax-free breaks, special economic zones (SEZs) with low business rates or less restrictive environmental regulations.

The spatial organisation of production has allowed for TNCs from emerging economies to locate production where it gives access to large markets in developed regions. Kia Motors (South Korea) has factories in Slovakia and the USA to gain access to EU and USMCA markets.

Service-based TNCs in the tertiary sector are more footloose and will locate operations either where there are relatively low labour costs balanced with good education or in proximity to their markets. Language is another important consideration here: lower level services, such as call centres, are **outsourced** to India, for example, because of the high proportion of well-educated English-speaking workers, who offer a lower cost alternative.

Key terms

Agglomeration – When companies in similar industries locate near to each other because of the benefits gained by sharing ideas and resources – called 'agglomeration economies'.

Multiplier (effect) – A situation where an initial injection of investment or capital into an economy (at any scale) in turn creates additional income by, for example, increasing employment, wages, spending and tax revenues.

Linkages

One of the features of TNCs is the ability to expand and gain more control of their industry and markets. They do this by developing links with other companies and between countries. Links with other countries can be made by investing in them and by establishing **joint ventures** with national or state-controlled companies.

Expansion can also be achieved through takeovers or mergers with other companies to integrate different parts of the business.

There are two types of integration:

- **Vertical integration:** An arrangement in which the **supply chain** of a company is owned entirely by that company, from raw material through to the finished product. This gives the TNC control over its supplies and stocks and reduces costs because of economies of scale. A good example is BP in the oil and gas industry. BP has negotiated exploration and production rights in over 50 oil and gas fields worldwide, which it calls its upstream activities. It owns or jointly owns (with other energy companies) 11 oil pipelines and has its own shipping fleet. The company owns refineries, usually based in the countries where end products are sold, and globally has nearly 19,000 retail service stations. These are BP's 'downstream' activities.
- **Horizontal integration:** A strategy where a company **diversifies** its operations by expansion, merger or takeover to give a broader capability at the same stage of production. This can be either complementary or competitive to its existing business. For example, Kraft Foods' takeover of Cadbury in 2010 gave them a more diverse base in the grocery and confectionery market. Following a series of de-mergers, the newly formed Kraft Foods Group merged with Heinz in 2015 to form Kraft Heinz, for the same purpose.

Trading and marketing patterns

- The majority of TNCs trade in highly developed economies in North America, Europe and the Far East (Japan and Korea).
- Emerging economies in South and East Asia, the Middle East and Latin America are now creating a rapidly increasing demand for consumer goods.
- TNCs' activities in emerging economies have developed from a focus on production for export to becoming suppliers to internal markets to meet ever-growing consumption in those countries.
- TNCs use global marketing strategies to take advantage of marketing economies and to give clear identity through branding of their products.

Benefits and costs of TNCs

Table 7.10 Outline of how TNC offshore and outsourcing operations affect different participants

	For the host country	For the TNC	For country of origin (TNC base)
Benefits	Generates jobs and income Brings new technology Gives workers new skills Has a **multiplier effect** Improved energy and transport infrastructure	Lower costs because of cheaper land and lower wages Greater access to new resources and markets Fewer controls such as environmental legislation	Cheaper goods Specialise in management, financial services, R&D and other higher skill occupations
Problems	Poor working conditions Exploitation of resources Negative impacts on environment and local culture Economic leakages/repatriation of profits	Ethical issues such as the image of environmental damage or 'sweatshops' can be detrimental to their reputation Social and environmental conscience	Loss of manufacturing jobs Deindustrialisation Structural unemployment Demultiplier effect leads to spiral of decline in former manufacturing areas Can mean derelict factories, areas of deprivation and poverty

Apple Inc.: a transnational corporation

Apple Inc. – producer and retailer of computer technology and mobile electronic devices

Apple Inc. is a US transnational electronic technology corporation with its company headquarters based at Apple Park in Cupertino, northern California (Figure 7.24).

Apple produces a range of familiar high-tech products and services that are marketed globally either under its universal 'i' brand or as 'Apple'. These include iPhones, iPads and more recently Apple Watch, Apple TV and its most powerful device, the MacPro. In 2018 it launched a new smart speaker called HomePod.

Figure 7.24 Apple Park

The company started business in 1976, in the early days of personal computer manufacture. In 1982 it took over the smaller MacIntosh organisation and launched a new brand of desktop computer that later became AppleMac computers. Apple computers earned a growing reputation for quality and they attracted a growing niche market of brand-loyal customers. Since 2000, it has experienced phenomenal growth as an organisation because of its development of mobile and Wi-Fi devices. In 2019, Apple was the world's:

- largest IT company by revenue
- third-largest mobile phone manufacturer
- number one global brand by value (US$234 billion).

It has 137,000 full-time employees and over 510 retail stores in 22 countries.

Apple's success has been due to a number of factors:

- stylish and well-designed products
- slick marketing and branding – generating a growing number of customers with brand loyalty
- innovative products that are regularly updated
- focus on highly mobile devices, which fit their market's needs
- selling its products and services online together with ancillary products such as music and TV programmes
- diversification into new technologies and platforms as they evolve.

Spatial organisation and linkages

Apple is a truly global company with a distinctive geography. Their main products are designed in Silicon Valley, California, made in mainland China by Foxconn, a Taiwan-based company, and sold all over the world. The different components that are used in their products are

sourced from companies with bases in many countries all over the world, for example Wi-Fi chips can originate from any one of 20 countries in the Americas, Europe or Asia.

Most of Apple's employees are based in HDEs, mainly in the USA, reflecting the high-tech, knowledge-intensive nature of its business. Those manufacturing Apple products however are employed by other companies and number up to 2.5 million, nearly 20 times those employed by the company itself.

Apple has expanded its activities in the USA in recent years:

- Its headquarters and research and design centre are together at Apple Park in Cupertino, California.
- It has a further large campus and MacPro assembly facility employing 6,000 people at Austin, Texas, which it plans to expand to employ a further 10,000.
- It has five data centres in the USA – two in California at Newark and Santa Clara, a further two in western USA at Reno, Nevada and Prineville, Oregon and one on the eastern seaboard of the US at Maiden, North Carolina. A further large centre is planned for mid-2020s at Waukee, Iowa in central US to better serve North American users of iMessage and Siri.

Apple now employs 90,000 people in all 50 states of the USA.

The concentration of R&D locations in Silicon Valley, California is a good example of **agglomeration**, which is typical of high-tech industries where information exchange and access to well-qualified and expert staff are crucial.

- Apple has its European, Middle Eastern and African headquarters in Cork, Ireland, where it employs 6,000 workers on its iMac production line. Cork also acts as the global transport logistics hub for Apple, organising shipments of assembled iPhones and iPads around the world.
- Assembly of Apple's main products, iPhones and iPads is **outsourced** to partner companies Foxconn and Pegatron – both Taiwanese electronics corporations. Foxconn is Apple's longest-running partner; its main production base is in Foxconn City, Shenzen near Hong Kong. Foxconn also has large assembly complexes at Zhengzhou, Henan province and Chengdu, Sichuan province in central China. These three large sites alone employ nearly one million workers.

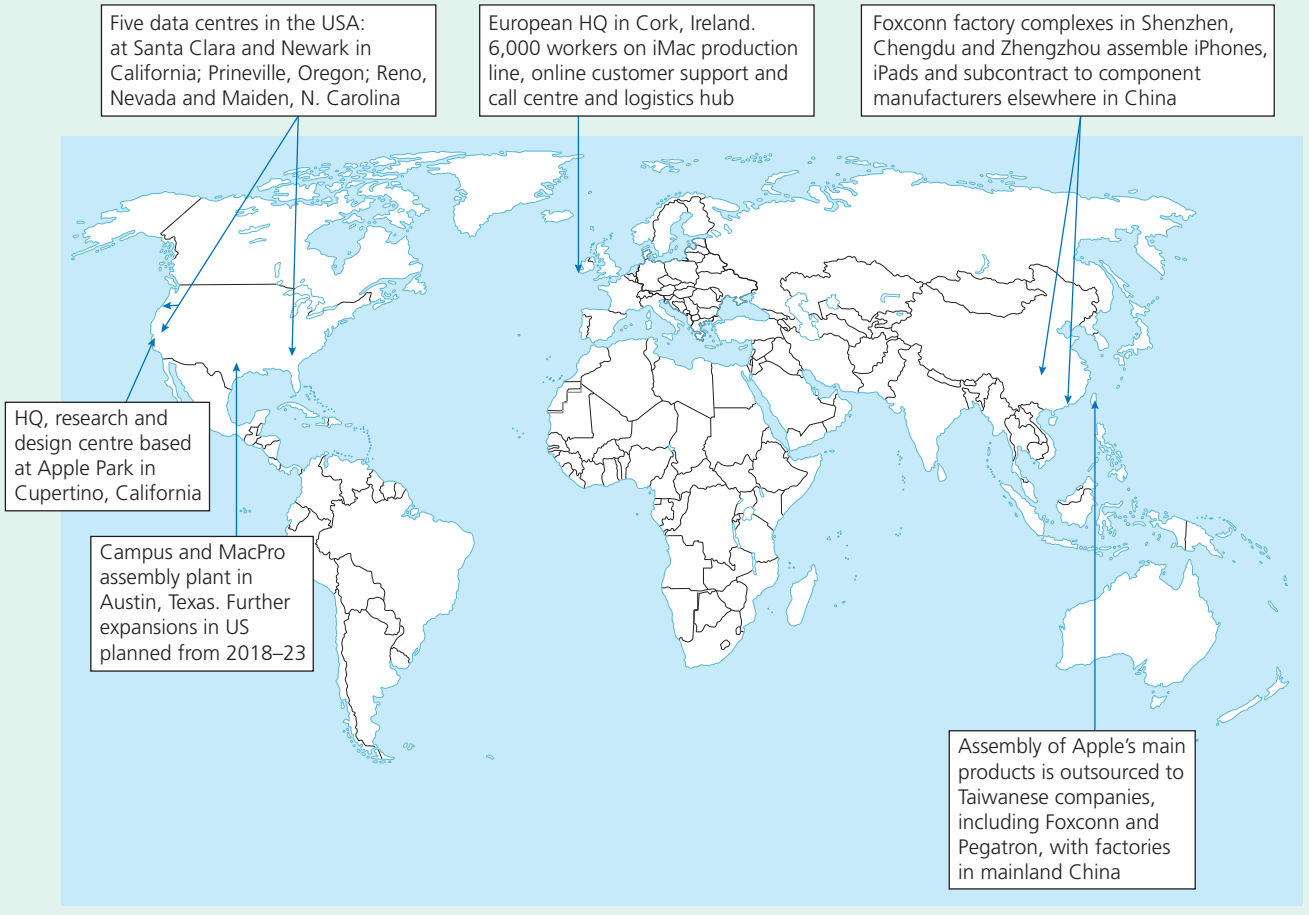


Figure 7.25 Location of Apple's main global operations

- Of its 511 retail stores, over half are in the US, 115 are in Europe and the Middle East and 42 are in China, where it has increased its market.

Production and linkages

Apple's mainstream products are produced in China for a number of reasons:

- they are able to outsource production to established Taiwanese electronics firms such as Foxconn, who competed for Apple's manufacturing contract on the basis of low production costs
- there is a large, young, skilled and non-unionised workforce who will work long hours for relatively low wages
- Shenzen was the location of China's first and most successful **Special Economic Zone** (SEZ) offering a number of incentives to attract foreign companies.

Figure 7.26 Workers on the production line at the Foxconn complex at Shenzen

As a global leader in supply chain management, Apple has a lean production system that makes the most of its spatial organisation. It arranges for different components to be shipped directly from their various original sources to the assembly complexes in China. This reduces stockpiling and the number of warehouses to one centralised location in California. This exceptional inventory management reduces costs and generates higher profitability compared to its competitors.

Since 2018, Apple has expanded production within the USA, notably in Austin, Texas where it manufactures the MacPro for the domestic market. Components are shipped from around the world, including from China, to the assembly plant there.

Impacts on countries in which it operates

Since its rapid growth, Apple has been the subject of a number of controversial claims about its business operations and the impacts they have in the countries in which they operate. This demonstrates the risks of outsourcing production to foreign companies and locating overseas, where employment regulations are less rigid and reflects general criticisms levelled at TNCs.

Ireland

Apple's European HQ is based at Hollyhill Campus, on the north side of Cork. It is the only fully Apple-owned manufacturing facility in the world. Along with a number of other foreign blue-chip firms, Apple was lured to locate in Ireland in 1980, at a time of high unemployment, by the government's 12.5 per cent corporation tax, the second lowest in the EU. Generally Apple's impact in Ireland, and in Cork in particular, has been positive:

- It employs 6,000 workers at Hollyhill, Cork's largest private employer. Most work on the iMac production line or in the online customer support and call centre.
- Cork's designation as Apple's logistics hub enhances Ireland's reputation and its global connectivity. Over 100 people work in the logistics team, remotely organising the shipping of products to over 50 countries.
- Apple's presence in Cork has generated around 3,000 jobs for workers employed in ancillary work.
- The company's presence in Cork has attracted other high-tech firms to the area.
- It has attracted a highly skilled workforce and provided an inspiration for local education, research and development.
- The company has expanded and contributed to infrastructural improvements in the city.
- Together with other companies locating in Dublin and in the south and west of Ireland, Apple has enhanced Ireland's reputation for hosting high-tech TNCs.
- Many of the more highly skilled workers at Hollyhill are foreign nationals (mainly from the EU), so Apple is criticised for not creating sufficient work for local people. There are 90 nationalities working there, including 700 French and 460 Germans – 53 different languages are spoken. The counter-arguments are that this has helped Cork become a more vibrant cosmopolitan city and that at least 60 per cent of the workers are Irish.

Tax practices

As a TNC operating in a number of countries, Apple uses subsidiary firms in other countries to declare profits and pay a lower rate of tax. In 2014, both the US and the EU claimed that Apple had received favourable tax treatment from the Irish government. It caused

tension between the USA and the EU, as each claimed that Apple owed them tax on their profits. In 2016, the EU Competition Commissioner concluded that Apple had received illegal state aid from Ireland and ordered them to pay €13 billion in unpaid taxes for 2004–2014. These tensions expose a flaw in taxation systems in an increasingly globalised world. Systems need to be updated to accommodate large TNCs such as Apple, which manufactures most of its products in China but derives the majority of its profit from 'intellectual property', marketing, patenting and branding, which can be funnelled through a maze of subsidiaries.

Apple is far from alone in claiming that profits from large chunks of its intellectual property are generated outside the USA.

China

Foxconn employs 1.3 million employees working on Apple products at its business parks in China and claims to support nearly five million jobs in the country, as they also supply other leading high-tech firms.

Labour practices

- **Working conditions**: Workers in Foxconn's high-security factory complexes mostly live on site, where there are dormitories, shops and cafes. Most are young migrant workers from other parts of China. They work an average of 62 hours each week for around $300 per month, half of which is taken up by living expenses. Media reports suggest that excessive hours and other problems may persist, despite Apple's attempts to reform factories following reports of alleged poor working conditions.
- **Health and safety**: Workers in factories producing Apple products have in the past been exposed to toxic chemicals used to clean the products. It has also been suggested that earplugs and goggles have not always been available to employees when working on certain machinery.
- **Student and child labour**: In 2017, student interns working for Foxconn were found to have been working excessive hours against its own policies. In 2018, Quanta Computer, another contractor which manufactures Apple Watches, was found to have been employing high school students as interns. Factories are allowed by provincial authorities in China to employ under 18s as interns or trainees, but Apple do not because of their commitment to social responsibility. Following these cases, Apple now demands that its contractors limit student interns to 10 per cent of the factories' workforces. (This is an example of how different 'norms' are accepted and interpreted by different societies and how this poses problems for many large 'western' companies seeking to outsource their production.)

In response to these cases, Apple now has a Supplier Code of Conduct and it audits supplier factories regularly.

Environmental issues

Following criticisms of its practices from Greenpeace in the 2000s, Apple responded with a 'Green my Apple' campaign. As a result, it:

- uses 100 per cent renewable energy in its facilities in the US and has established its own energy company, Apple Energy, generating electricity mainly from solar power

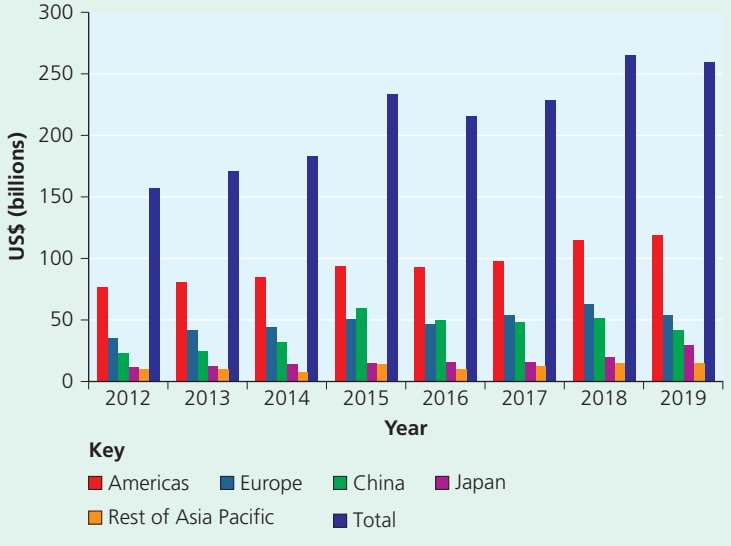


Figure 7.27 Growth of Apple's global revenue by country or region, 2012–2019

- has removed all PVC plastics and toxic chemicals from its products
- promotes recycling of its products, though lack of 'repairability' is still an issue.

In 2016, Apple issued a $1.5 billion green climate bond; proceeds from this are dedicated to the financing of environmental projects.

As a result of Apple's ongoing commitment to tackle green issues, it moved from eleventh in Greenpeace's 'Green Electronics' rankings in 2006 to second by 2017.

Trading and marketing

Apple's market is predominantly in richer, developed countries. In 2019, 45 per cent of product sales were in the USA. In 2015–16, China temporarily became the second largest market for Apple products, but Europe has since re-established itself as Apple's second largest market (Figure 7.27, page 320).

The services part of Apple's business is now playing a larger role than ever before in its success. Services revenue has increased more than threefold since 2013, from $13.5 billion to more than $50 billion in March 2020.

- Apple has been caught in the crossfire of the trade war between the US and China. This imposes heavy tariffs on Chinese imports to the US, meaning that Apple has to pay for importing its products back to its main market in the US.
- Since 2017, Apple has fallen in line with the US administration's aim to re-invigorate domestic manufacturing. Supported by incentives, they are investing more in US domestic suppliers and manufacturers. As part of a commitment to support the US economic resurgence, Apple pledged to create 20,000 more jobs and invest $5 billion to support innovation among US manufacturers from 2018 to 2023.
- Apple's decision to invest in the US is encouraged by federal tariff exclusion zones in Austin, so they do not have to pay tariffs on imported components there. Further expansions in the US are planned.
- Since its peak in 2016/2017, the market for Apple products in China has plateaued and may have reached saturation point. It remains a good potential market because of its strong iPhone user base, but Apple is facing competition from Chinese smartphone manufacturer, Huawei, who can undercut iPhone in China.
- Apple has attempted to enter the rapidly expanding Indian market. However, this has largely been unsuccessful as the smartphone and tablet market there is led by cheaper competitors.

Conclusion

As a TNC, Apple is typical in demonstrating how it can respond easily to incentives, growing markets, changes in terms of trade and geopolitical developments by redirecting investment in different parts of its operation to different locations. Apple has diversified to reduce risk and capture other markets. Apple also demonstrates that having a well-organised global supply chain reduces costs and increases profits. Its responses to criticisms about labour practices and environmental issues illustrate the concern shown by TNCs regarding corporate social responsibility and ethical purchasing.

World trade in bananas

The banana trade raises a variety of environmental, economic, social and political problems. Bananas are the fourth most important food product within least developed countries, being a staple food for around 500 million people. Of all fruits bananas are the most internationally traded, generating revenues of over $15 billion per year. They contain large quantities of energy and are a rich source of vitamins, fibre and protein. A single banana provides more than an adult's daily potassium requirement.

Globally, bananas are the fifth most traded agricultural commodity with global exports, primarily from Latin America and the Caribbean, estimated at 23.3 million tonnes. They are an important cash crop, although 80 per cent of bananas that are produced are for local or national consumption.

The banana industry

Bananas are grown predominantly in hot, rainy lowlands of tropical regions. Global production of bananas in 2019 was 120 million tonnes. The top four producers, India, China, Indonesia and Brazil, are also the four largest consumers as their output is mostly consumed domestically, as is most of the production in African countries.

With the exception of the Philippines, the main regions that produce bananas commercially for export are concentrated in Latin America and the Caribbean. Some of the countries in this region, notably Ecuador, Costa Rica and Guatemala, are highly dependent on banana exports. The Philippines have become an increasingly important producer and exporter. They export bananas to Japan and other parts of East Asia and the Middle East.

Banana varieties are susceptible to diseases and almost all bananas are treated with chemicals throughout the production cycle. Commercial plantations operated by large TNCs apply around 30 kg of active ingredients per hectare, per year. These include fungicides, insecticides and herbicides. In addition, fertilisers are applied regularly and after harvesting the fruit is washed with a disinfectant. With the exception of cotton, the banana industry has the largest agrochemical input into the environment.

Banana plantations also cost the environment in terms of deforestation, soil fertility (because of contaminants) and loss of biodiversity (especially aquatic life as pollutants run into water courses).

The global banana trade

World trade is dominated by two different groups of producers: the ACP group (Africa, Caribbean and Pacific), which includes mainly small- and medium-sized producers, and the so-called 'dollar producers' of Latin America (primarily Ecuador and Colombia) controlled by large US TNCs. Most bananas for export are grown on large **monoculture** plantations, especially in Latin America and increasingly in Africa. Banana production in the Caribbean is controlled less by TNCs and more by small- and medium-scale growers on family farms and co-operatives. Traditionally, production in both the Philippines and Africa has a similar structure of small and medium growers but there is now more TNC involvement in both regions.

Trade follows the traditional pattern of developing regions exporting a low-value primary product to more developed countries. Exports are dominated by Latin America and the Caribbean countries, which produced 17 million tonnes of bananas for the export market in 2018. The leading producers are Ecuador, Costa Rica, Colombia and Guatemala. Smaller countries in Central America, such as Panama, are now exporting at a faster rate. In Asia, which produces 17 per cent of the export market, the main commercial producer is the Philippines, while in Africa (where exports are smaller) the two main producers are Côte d'Ivoire and Cameroon.

The largest importers are the EU and the USA.

As with many commodities produced in developing regions but consumed in developed countries, around 85 per cent of the price paid by the end consumer stays in the richer country and never reaches the producer, who has most of the risks of producing a perishable fruit. The largest slice is taken by retailers and bananas are one of the biggest profit-makers in supermarkets. Workers, on average, receive between 5 and 9 per cent of the total value of bananas, while retailers manage to capture 36–43 per cent (Figure 7.29).

Table 7.11 The top six global producers and consumers of bananas, 2017–18

Top producers		Top consumers	
Country	Total volume (millions of tonnes)	Country	Total volume (millions of tonnes)
India	30.0	India	30.0
China	11.0	China	12.0
Indonesia	7.2	Indonesia	7.1
Brazil	6.7	Brazil	7.0
Ecuador	6.5	EU (as a group)	6.4
Philippines	5.8	USA	3.8

Table 7.12 The top six exporters and importers of bananas, 2018

Top exporters		Top importers	
Country	Total volume (millions of tonnes)	Country	Total volume (millions of tonnes)
Ecuador	6.6	USA	3.6
Philippines	2.9	Russia	1.5
Colombia	2.6	Belgium	1.4
Costa Rica	2.4	Germany	1.4
Guatemala	2.1	UK	1.2
Netherlands*	1.7	China	1.0

*Netherlands (also Belgium and Germany) often appears in lists of top banana exporters. Although they have started growing glasshouse bananas, the figures are derived from imports which are then re-exported.

Banana exporters – main banana exporting regions in the world

Banana importers – main areas of the world importing bananas

Banana consumers – these regions eat most bananas on a total consumption basis (tonnes). (Most are major growing regions eating their own produce and not exporting, for example, India.)

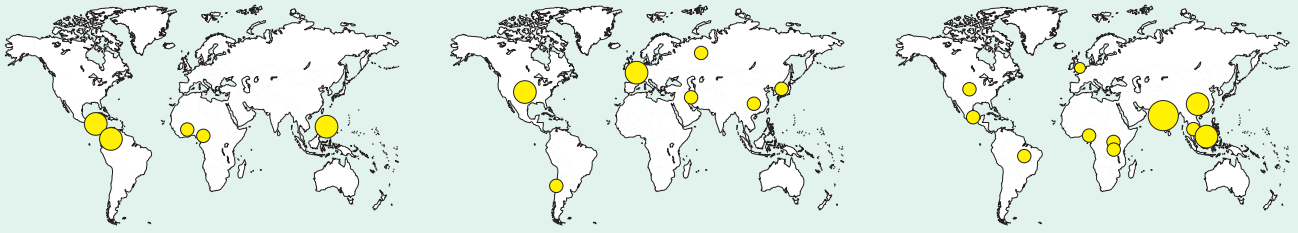


Figure 7.28 Pattern of main exporters, importers and consumers of bananas

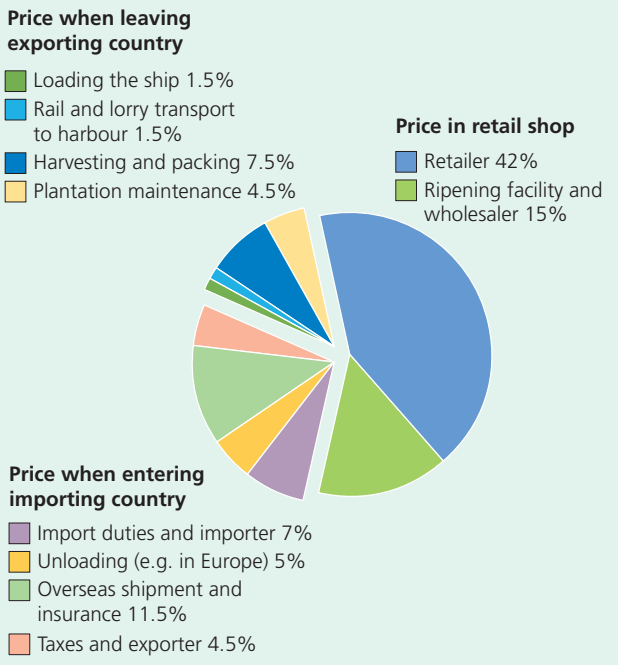


Figure 7.29 Contribution to prices at each stage of banana production, export and sale

In the past, the banana trade was dominated by just four large transnational companies **Chiquita, Dole, Del Monte** (all US-based TNCs) and **Fyffes** (based in Ireland). The other important producer is **Noboa**, which is a national corporation based in Ecuador. These companies are vertically integrated in the supply chain. They own or contract out plantations to other producers; they have their own sea transport and ripening facilities, and their own distribution networks in consuming countries.

Organisation of the banana trade has changed in recent years. In 2002, the big five companies controlled 70 per cent of the market but their share fell to less than 45 per cent by 2017 (Figure 7.30, page 324).

The big companies have freed themselves of direct ownership of plantations, in favour of guaranteed supply contracts with medium- and large-scale producers. An increasing number of national growing companies based in Ecuador, Costa Rica and Colombia sell their produce either to the TNCs (as distributors) or directly to retailers in the developed countries, for example, Wal-Mart and Tesco. A shift in power has given retailers an increasingly dominant role in the supply chain. As market share becomes concentrated in the hands of fewer retailers, suppliers have little option but to accept conditions such as low prices, discounts and delayed payments or otherwise risk being taken from the supplier list.

Issues

Trade wars

Bananas were the subject of one of the longest trade disputes in history, lasting 20 years from 1992 until the 2009 Geneva Banana Agreement was reached, coming into effect in 2012.

- The dispute started in 1975 when EU countries negotiated a trade agreement with former colonies. The agreement was the **Lomé Convention** and was made with 71 African, Caribbean and Pacific countries (ACP countries), many of whom were banana producers.
- These countries were given special and differential treatment (SDT) with preferential tariff-free import quotas to supply EU markets. The intention was to enable these former European colonies to develop independently without having to use overseas aid.
- The agreement was extended to a list of suppliers including Cameroon, Dominican Republic, Belize, Ivory Coast, Jamaica, Ghana, Surinam and the Windward Isles.
- This protected mainly smaller, family-run farms in the Caribbean and Africa from competition with the large Latin American producers, whose bananas were produced more cheaply on mechanised plantations.
- At the time, the US TNCs that controlled the Latin American crop were supplying around 75 per cent of the EU market, while only 7 per cent came from Caribbean suppliers. Despite this, in 1992, the TNCs filed a complaint to GATT (now WTO) that the practice constituted unfair trade. In 1997 the WTO ruled against the EU and ordered the EU to cease the discrimination.
- This led to a trade war between the USA and the EU. As a result, the US government imposed WTO-approved sanctions on a range of EU products.
- A compromise was eventually reached in Geneva in 2009 with the EU agreeing to gradually reduce tariffs on Latin American bananas from 2012 onwards.
- Tariffs came down from €176 to €75 per tonne between 2012 and 2018. Following EU trade agreements, the deal now also applies to other Latin American countries. This increases the chance of over-supply in the EU market so, to safeguard other producers, the EU will not reduce the tariff further.

There are concerns from the ACP producers that they are not able to compete. Of the Caribbean countries, only the Dominican Republic, Belize and the Windward Isles are competing successfully with the larger producers. Their focus on producing organic and fair trade bananas is meeting the needs of a growing market for ethical, sustainably-produced food in richer EU countries.

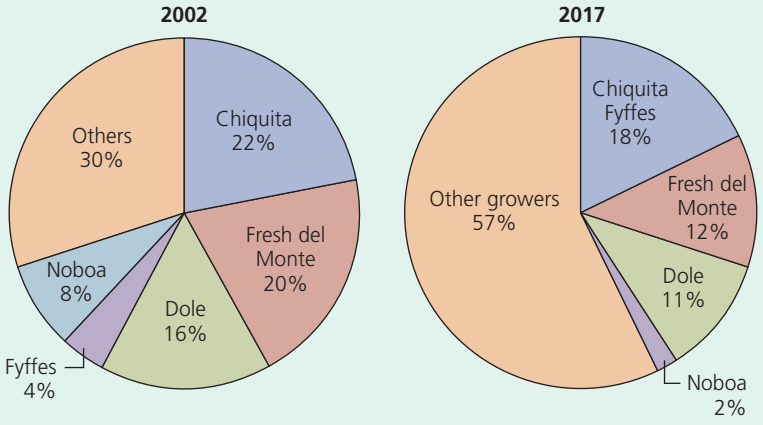


Figure 7.30 Changing market share of the largest banana traders

Source: Banana Link

Race to the bottom

Because of the low prices paid to suppliers by supermarkets, many of the larger companies are relocating their plantations, increasingly to West Africa, as companies search for lower labour costs and weaker legislation. This is called pursuing a 'race to the bottom' in terms of social and environmental standards. Plantation work is based on sub-contracted casual labour. The work involves long shifts in unbearable heat and many workers fail to earn enough to cover their basic needs.

Fair trade and organic bananas

There has been a steady growth in the sales of so-called 'sustainable' bananas, which includes both fair trade and organic produce. This trend helps smaller-scale producers in the Caribbean and in parts of Africa and will partially counter the deterioration of conditions in banana production. There is undoubtedly a growing market segment of ethical consumers in richer nations who are becoming aware of the shortcomings in the supply chain and are willing to pay a higher price for a certified product.

Disease

There are many different banana varieties around the world but the Cavendish cultivar, which accounts for 47 per cent of bananas grown, makes up 99 per cent of global trade. It has been grown commercially since it was discovered to be resistant to tropical race 1 (TR1), a fungus which took the previous most popular variety, the Gros Michel, to the brink of extinction in the mid-twentieth century.

A new disease, TR4, is now ravaging Cavendish plantations throughout South East Asia and Australia and is spreading to Africa. TR4 can lie dormant in soil for decades and has proven resistant to fungicides so there is concern that it will take hold in Latin America. The Cavendish variety is ripe for attack and as bananas have no seeds and reproduce asexually, they cannot be hybridised. The only options will be to genetically modify the variety or to encourage consumers to start eating other varieties.

Summary

The world trade in bananas demonstrates a number of relevant points about trade, especially in primary commodities:

- Mass production in developing countries will have negative environmental consequences.
- TNCs have a large degree of control of markets and can influence political decisions.
- WTO will support free trade against protectionist activities or agreements at all costs, even if the protection may be to help development (the lobbying power of US TNCs took precedence over the SDT agreements for least developed countries, which were not very effective in this situation).
- Geopolitical processes mean that trade disputes can spread and escalate to trade wars between regional trading blocs.
- Power and control of food production has shifted away from growers and towards retailers in high-income countries.
- More ethical, sustainable consumer markets are growing but relatively slowly and only in places that can afford products bought at a higher price.

Geographical consequences of international trade in global systems

The following have been the main beneficiaries from the process of globalisation of trade.

- **Emerging economies:** Some medium-income nations have developed rapidly as a result of inward investment and have emerged to become major economic powers competing with the richer developed regions of the world.
- **Transnational corporations:** Large companies have grown in a number of different industrial and service sectors and although they are mainly based in the developed countries, TNCs from emerging economies have also become global powers.
- **International organisations:** Organisations such as the IMF, the World Bank and the WTO have all contributed to the integration of economies and have consolidated their position and control of global systems.
- **Regional trading blocs:** More trade agreements have evolved which benefit their members, though some would argue that it is at the expense of the nation state. Free trade areas and customs unions are expanding among developing countries in Asia, Latin America and Africa. This improves the negotiating leverage of their members in trade negotiations and will offer more access to markets for developing countries in the long term.

Economic consequences

Free trade within customs unions, such as the EU, can be beneficial for consumers but has disadvantages for some sectors of the workforce.

For example, Japanese car manufacturers Honda, Nissan and Toyota saw the UK as a base for car production in the 1990s and invested there as it gave them tariff-free access to the whole EU market of half a billion people, most of them car-owning households. As a result:

- people in areas undergoing deindustrialisation, such as Sunderland, gained employment at Nissan's factory.
- imports of cars from Europe were cheaper as there were no tariffs.
- the combination of no tariffs and economies of scale gained by manufacturers has in real terms meant reduced car prices.

The price of many electronic goods, such as computers, has fallen as they are produced in lower-cost locations.

Major retail TNCs have more control of food supply chains and supermarkets pay less to import food from other countries. As a result, consumers are able to buy food relatively cheaply all year round from a wide range of countries.

Fair trade has become an option for some consumers.

However, as a result of offshoring and outsourcing, traditional manufacturing jobs in developed countries have been lost and wages, in real terms, have remained stagnant.

Political consequences

Not everyone has benefited from the increased globalisation of trade and this is reflected in the greater inequality that has grown in most countries. In recent years, there has been a backlash against further integration, especially in developed countries. This rise in 'national populism' has seen the UK electorate vote to leave the EU and the election of Donald Trump in the US on a very nationalistic and protectionist manifesto.

Links

These political trends are discussed in more detail in Section 7.6, Globalisation critique, on pages 346–7.

Table 7.13 Benefits and costs of free trade access to markets

Benefits	Costs
Lower prices for consumers Greater choice Access to larger, wealthy markets for TNCs Greater economies of scale through increased specialisation Greater foreign competition may weaken domestic monopolies (for example, UK supermarkets) Competition leads to greater innovation Access to cheaper raw materials for TNCs	The **injustice** of free trade not giving sufficient 'protection' to emerging industries in developing economies so they cannot compete with developed countries More developed economies are still protected by tariffs on agricultural imports Diseconomies of scale as a result of difficulties co-ordinating subsidiary companies

Social and cultural impacts

- Globalisation has allowed for greater sharing of ideas, lifestyles and traditions. It means that people have greater access to foreign culture such as film, music, food, clothing and other goods and services that were not previously available domestically.
- The over-standardisation of many goods and services has led to increased cultural homogeny.
- In a bid to offset criticism of the standardisation of products on a global scale, a strategy called **glocalisation** has been adopted by some TNCs. This involves thinking globally but acting locally to reduce threats to cultural dilution. Products and services are likely to be more successful when they are customised for the local market. McDonald's is a commonly cited example of this strategy as the restaurant's menu is often customised to suit local tastes (Figure 7.32).

Key term

Glocalisation – A term used to describe products or services that are distributed globally but which are fashioned to appeal to the consumers in a local market.

- Increased awareness of global news and events – which can be positive in informing people of environmental issues such as climate change, but can be used negatively, for example, for propaganda purposes by terrorist organisations.

Figure 7.32 Glocalisation – global brands going local. McDonald's Mac Burrito in Mexico City

Environmental consequences

Environmentalists argue that globalisation is concerned primarily with economic costs and largely ignores environmental costs. Problems include:

- more transportation, which increases greenhouse gas emissions
- depletion of non-renewable resources
- TNCs outsourcing production to countries where environmental standards are less strict
- weak or non-existent controls allowing pollution of the air, land, rivers and seas
- more waste from packaging
- IMF-enforced spending cuts reducing many nations' spending on the environment
- greater movement giving a higher risk of diseases and invasive species being introduced.

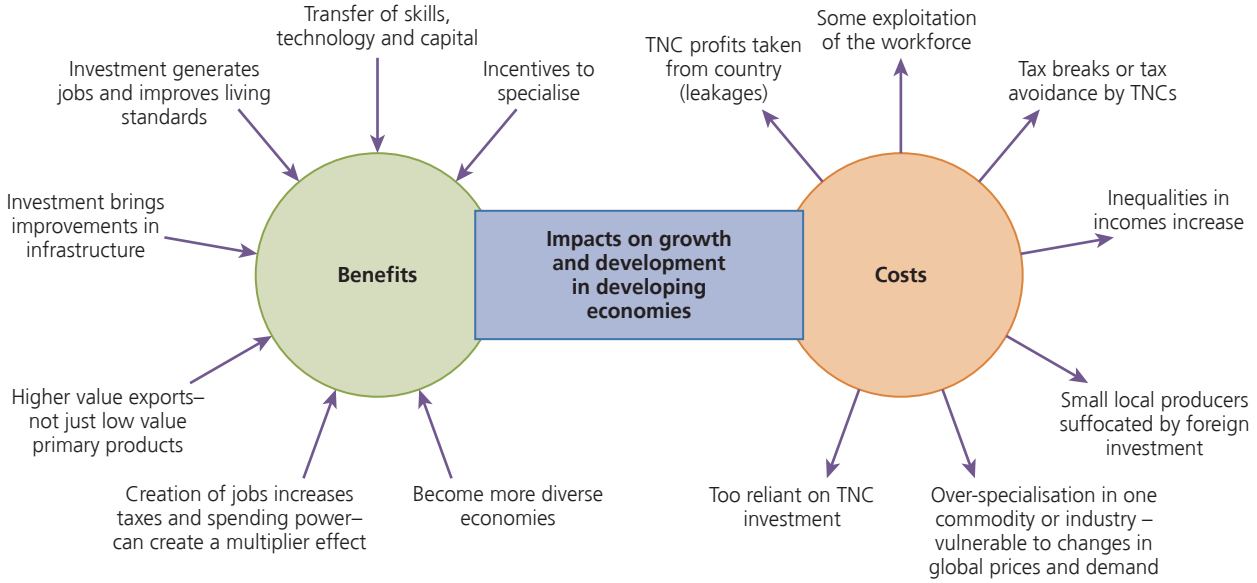


Figure 7.31 Benefits and costs of globalisation affecting growth and development in developing economies

Review questions

11 Outline the key patterns of international trade and investment.

12 Explain the causes and consequences of differential access to markets.

13 Describe the main characteristics of TNCs.

14 Assess how international trade impacts on people's lives around the world.

7.4 Global governance

Global governance refers to the way in which global affairs that affect the whole world are managed. In international relations generally, consensus is usually reached between countries. They then agree to abide by similar rules given in signed treaties or international laws.

As the world becomes more interdependent, global governance is increasingly important for achieving sustainable development across all nations. This is especially true for governance of the 'global commons' where environmental sustainability is at the forefront of our concerns.

Regulating and reproducing global systems

Key terms

Environmental sustainability – A state in which the demands placed on the environment can be met without reducing the quality of the environment for the future.

Global commons – Resource domains or areas that lie outside the political reach of any one nation state.

Global governance – A movement of political integration aimed at negotiating responses to problems that affect more than one state or region.

Norms – The shared values, traditions and customs that govern individual and group behaviour in a society. They are the perception of what is acceptable or unacceptable in societies at all scales.

Reproduction – The way that existing global systems remain largely unchanged and are even extended and reinforced by regulation.

Global governance regulates interconnected global economic, social, political and environmental systems to keep them stable and in balance. Regulation means enforcing sets of rules that have been agreed by the nations involved. They will have committed to an agreement, such as the Antarctic Treaty System (ATS).

In response to tackling issues raised by increased integration, international agreements are established in the form of treaties, directives or protocols. These agreements are formulated, maintained and enforced by intergovernmental **agencies** or **institutions**, such as the UN, the World Bank, World Trade Organization (WTO) or similar organisations. The international **laws** (mandatory rules) or regulations contained within international agreements are based on the values and customs that exist and guide behaviour within society. These are also known as social norms. It is very difficult to establish laws or treaties that encompass social norms on a global scale as national or regional outlooks and societies differ so much. This may have consequences for how nations committing to international agreements interpret and enforce them.

Observers argue that the norms expressed and contained within international regulations are largely determined by the wealthier, more powerful and developed countries, mainly based in the western world. These norms are therefore based on free-market, capitalist economic systems (with varying levels of social intervention) and democracy. This suggests that for countries to access the benefits of being integrated into the globalised economy, they may have to adapt their economic and political systems. In this way, global systems are being **reproduced** to extend and strengthen their reach, which is in the interests of the wealthier countries that dominate them. This reproduction of global systems may help to keep them stable but arguably it may reinforce some of the inequalities and injustices that exist within the systems. In summary, regulation and reproduction of global systems take place through:

- the establishment of international legal agreements
- governance undertaken by intergovernmental institutions and agencies
- the extension and reinforcement of social norms on a global scale.

International agencies

A number of intergovernmental institutions and agencies take responsibility for global governance. Many of the organisations operating internationally are connected to the United Nations (UN).

United Nations (UN)

The UN is an international organisation founded in 1945, made up of 193 member states whose aim is to promote international peace, security and co-operation. Due to its unique international status and the powers vested in its founding charter, the UN is the intergovernmental institution with the greatest political authority at a global level.

The UN's work is guided by the principles contained in its charter, which include:

- taking action on issues that face humanity in the twenty-first century, such as peace and security, climate change, sustainable development, human rights, poverty and many more
- fostering co-operation by facilitating dialogue and negotiations between nations; the UN is a mechanism for solving problems and finding areas of agreement.

The UN comprises a number of other specialist bodies within its family. These are governed by the UN and take responsibility for managing specific global issues. Table 7.14 gives some examples.

Following the largely successful achievement of the Millennium Development Goals (MDG), UN members met at the Sustainable Development Summit in September 2015 and agreed a new **Agenda 2030** for global **sustainable development**, which defined a set of 17 **Sustainable Development Goals (SDGs)** from 2015 to 2030 (see details at https://sdgs.un.org/goals). An annual report and statistics on the progress of the SDGs is available at https://unstats.un.org/sdgs/report/2019/.

Key term

Sustainable development – Development which meets the needs of the present without compromising the ability of future generations to meet their own needs.

Table 7.14 Examples of UN special bodies and responsibilities for specific global issues

Organisation or body	Main area of responsibility	Evaluation of work
UN General Assembly	• To develop good relations and co-operation between nations • Settling disputes within international law through the International Court of Justice	• Works as an independent arbitrator to uphold international agreements • Each country has a representative, but the UN relies on funding from more developed nations, which leaves it open to criticism about impartiality
UN Security Council	• To maintain world peace and security • Has five permanent members (China, France, Russia, UK and the USA) plus ten temporary members who are elected every two years	• Has facilitated nuclear disarmament since the Cold War • Promotes geopolitical stability, for example, by sending 'neutral' peacekeeping troops to intercede and de-escalate in regions where there is conflict • Does have a history of not agreeing on resolutions in some major conflicts (for example, in Syria and on Russia's annexation of Crimea)
World Health Organization (WHO)	• To manage international public health issues by combating and eradicating spread of global diseases • To monitor outbreaks of disease in countries and declare global pandemics	• Achieved eradication of smallpox and near eradication of polio • Praised for its co-ordination of the global response to the 2020 COVID-19 pandemic • Equally, criticised for lack of leadership and efficacy on COVID-19, as well as on the Ebola virus in the past • Critics have called for improved vaccine stockpiling and responses to epidemics
UN Development Programme (UNDP)	• The eradication of poverty and the reduction of inequalities and exclusion	• Led the efforts to successfully meet most of the **Millennium Development Goals (MDGs)** • These eight anti-poverty targets were set in September 2000 and the world committed to achieve them by 2015

Table 7.15 Differences in the role and functions between the IMF and the World Bank

IMF	World Bank
Regulates financial flows and stabilises the global monetary system	Promotes the reduction of poverty in developing countries by providing assistance for development
Offers financial and technical assistance to its members, employing economists who monitor the economic health of member countries	Provides long-term investment loans for development projects in developing countries – through the **IBRD – International Bank of Reconstruction and Development**
Provides 'bail-out' loans to countries to resolve problems that might lead to a global economic crisis – the international 'lender of last resort'	Provides special interest-free loans, via the **International Development Association**, to countries with very low per capita incomes
Provides loans to help members tackle balance of payments problems and to stabilise their economies	Encourages and advises start-up enterprises in developing countries
Financial reserves are drawn from quota subscriptions of its 189 member countries	Assistance is long term and financed, mainly by the wealthiest nations, through the issue of government bonds

In its 75-year history, the UN has achieved a great deal in maintaining stability, reducing poverty and organising global responses to inequality, injustice, climate change and sustainable development. It is, however, often criticised for:

- a lack of agreement and subsequent inaction on many security issues
- limited power to enforce compliance to regulations by national governments
- its organisation and funding – in a way that supports the status quo, with more powerful nations setting the agenda for their own self-interest.

The International Monetary Fund (IMF) and the World Bank

The intergovernmental finance institutions, the IMF and the World Bank, were also established in 1945 at the end of the Second World War to stabilise the global economy and to provide financial stability. Their functions are pivotal in regulating and acting as intermediaries in the flow of international capital (see Figure 7.4, page 286). Their roles are quite different but they both underpin the interdependency that exists in the global flow of finance.

Table 7.15 outlines how both organisations have contributed to stability, development and growth in many parts of the world, but both have also been criticised for their approach to their respective roles:

- IMF funding is paid by members and influence is determined by their wealth. Developing countries criticise this as the larger contributions give wealthier nations a more powerful vote on decisions.
- As part of their conditions for loans, the IMF has been known to impose severe cuts on spending by governments in developing countries, adversely affecting education and welfare.
- Rescue loans are provided by the IMF only to stabilise international trade; eventually the borrowing country has to pay the loan back at high interest rates.
- Conditions attached to World Bank loans often insist on the reproduction of capitalist and free-trade market models in the loanee's economy, which do not always have the effect of reducing poverty.
- The World Bank has also been criticised for funding major '**top down**' projects, such as large multi-purpose dams to provide hydro-electric power, which have not helped to reduce poverty. The Bank now claims to support more sustainable '**bottom-up**' development projects.

Key terms

Bottom up – When local people are consulted and supported in making decisions to undertake projects or developments that meet one or more of their specific needs.

Top down – When the decision to undertake projects or developments is made by a central authority such as government with little or no consultation with the local people whom it will affect.

The World Trade Organization (WTO)

The WTO came into existence in 1995 as the successor to the General Agreement on Tariffs and Trade (GATT), which was established in the wake of the Second World War. The WTO currently has 164 members and covers 98 per cent of world trade.

Table 7.16 Evaluation of the WTO Doha Development Round (2001–2015)

Problems	Progress

• Disagreement mainly between HDEs (USA, EU and Japan) and developing economies (especially Brazil, India and China)
  • HDEs insisted that, in return for free access, developing economies open their markets to 'western' manufactured goods
  • Emerging economies insisted on large cuts in subsidies paid to US and EU farmers; US farmers demanded more market access abroad in return for subsidy cuts
  • USA, China and India would not compromise on the size of their tariffs
  • USA would not agree to allow India and China to use 'safeguard clauses', which allowed developing nations to impose emergency quotas on imports
• Divisions among LDCs, between net importers and net exporters of agricultural produce
• HDEs disagreed among themselves; EU more willing than the US for agricultural trade to be opened up
• China transformed into a global export powerhouse but continued to claim developing country status
• HDEs, especially the US, accused China of unfair trade as much of their production is state-aided
	• A **bilateral agreement** was reached between the EU and Latin American countries to end a long-standing dispute about trade in bananas
	• WTO achieved its first trade **multilateral agreement** in nearly twenty years with the Bali Package – an agreement between all members to speed up the movement of traded goods and reduce costs by removing red tape in customs procedures
	• Nairobi Package: some decisions were made to benefit WTO's poorest members:
	  • export subsidies for farm exports were abolished
	  • preferential treatment given for LDCs in services trade
	  • safeguarding measures introduced to ensure food security in LDCs
	  • improved measures for LDC cotton traders
	  • some trade barrier removal on goods of LDC origin

The WTO is responsible for facilitating international trade. It provides stability and gives members confidence in the international trading system by:

- undertaking the liberalisation of trade by encouraging the removal of barriers and protectionist policies adopted by some governments and trading blocs
- resolving trade disputes between member governments
- providing a forum for trade negotiation
- administering trade agreements that become the ground rules for international commerce.

The WTO holds 'rounds' of talks to reach **multi-lateral agreements** to advance global free trade. The most recent of these was the **Doha Development Round**, which started in 2001 in Qatar and effectively came to an end in Nairobi in 2015. The Doha round focused on reforming trade in agricultural produce, especially between advanced and developing economies, with the aim of helping less developed nations out of poverty. The annual talks repeatedly stalled with meetings in Geneva (2008), Bali (2013) and finally Nairobi (2015). Some progress was achieved to redress inequalities in access for LDEs to HDE markets (Table 7.16) but overall, the difficulty of reaching a multi-lateral trade agreement with 164 countries, each with economic and political self-interest, proved insurmountable.

While the Doha round was ongoing, many countries pushed ahead with bilateral or regional deals. It seems that bilateral deals between governments or trading blocs is a more effective way to establish free trade. Such agreements can be expanded and eventually made multi-lateral under WTO rules.

Interactions of governance between different scales

Success in global governance and in regulating the extremes resulting from unchecked globalisation can be achieved if there is clear communication and understanding of strategies at all scales, from global to local. Governance is not just about control but is also concerned with mobilising organisations or individuals to follow guidelines, which may be set out in regulations or agreements. This involves informing people about issues and persuading them to follow an appropriate course of action.

Multiscalar power is the ability to influence behaviour and attitudes, through interactions at different scales, to encourage or prevent change. Multiscalar relationships operate in a complex way. They involve different tiers of government but other 'players' are involved in decision-making at each level, including **non-governmental organisations (NGOs)** and **civil society** groups as well as individual citizens, all of whom may influence, organise or mobilise each other to take action (see Figure 7.33).

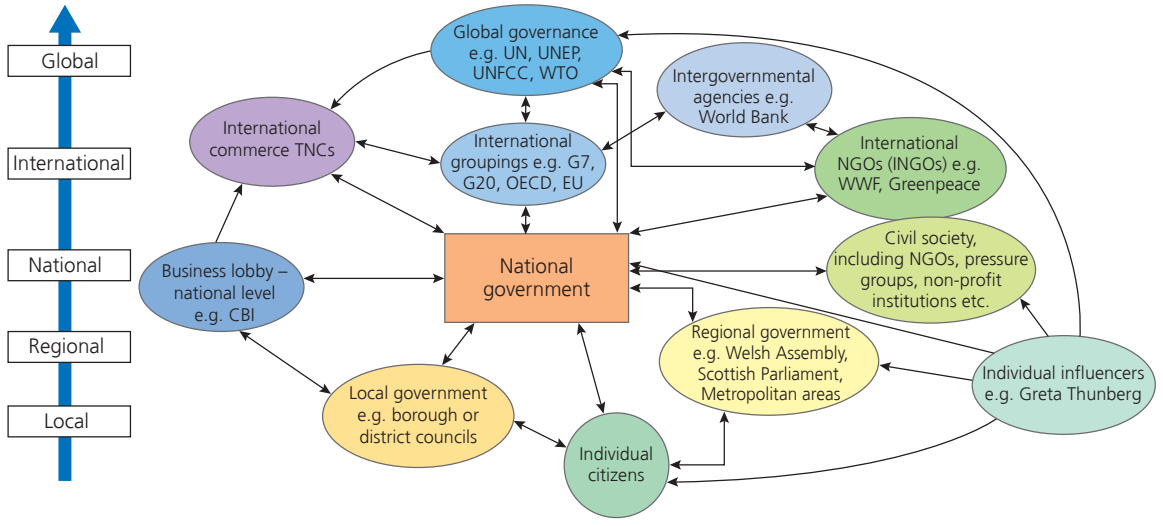


Figure 7.33 Multiscalar web of interactions that influence global governance

Key terms

Bilateral agreement – An agreement on trade (or aid) that is negotiated between two countries or two groups of countries.

Multiscalar power – When the ability to influence (or prevent) change is vested in citizens, governments, institutions and other interested groups (players) located and acting at different geographical scales.

Multilateral agreement – An agreement negotiated between more than two countries or groups of countries at the same time.

Non-governmental organisations (NGOs) – Any non-profit, non-commercial organisation with a common interest working independently of government influence. NGOs can be organised on a local, national or international scale.

Civil society – The 'third sector' of societal decision-making, distinct from government and business. It can include a wide array of non-profit organisations including NGOs, pressure groups, community groups, indigenous groups, charitable organisations, professional bodies and unions, etc.

One of the responsibilities of governance for the UN is to organise world summits, which bring together world leaders to agree on strategies to tackle global issues, such as sustainable development and climate change. These involve interactions at all scales and can be used to exemplify the multiscalar approach.

Sustainable development

At the **Earth Summit** in Rio de Janeiro in 1992, leaders of the international community adopted **Agenda 21**, a UN non-binding resolution that was seen as a blueprint for sustainable development for the early part of the twenty-first century. In total, 178 national governments signed up to follow the resolution. The plan focused on social and economic issues, such as combating poverty and the conservation and management of resources.

- Agenda 21 action plans were cascaded down from international agencies, such as UNEP, to national governments and then through to regional and local government decision-making, in order to influence the actions of individuals.
- Accordingly, each local authority in the UK has a Local Agenda 21 (LA21) strategy in line with the UN objectives and actions. The fact that most parts of the UK have various recycling options as part of their waste collection is a tangible outcome stemming from Agenda 21.
- It was seen as a '**top-down**' approach, which encouraged a '**bottom-up**' response, as the ideas are designed to filter down from the UN to encourage more sustainable lifestyles and behaviour.
- A feature of Agenda 21 was to strengthen the role of NGOs and civil societies in the decision-making process

The interactions to influence behaviour seem to be in a '**chain**' from a global resolution down to individual response and can be traced linearly through the diagram in Figure 7.33. However, as the '**web**' diagram suggests, the interactions are more complex.

Implementation of Agenda 21 attempted to engage governments and individuals at all levels, on the basis that good governance is essential for sustainable development. It has now been superseded by Agenda 2030, which is supported by the 17 SDGs.

Climate change

At the UNFCCC Paris Climate Summit in December 2015, a landmark agreement to combat climate change was struck by world leaders. The Paris Agreement came into force in November 2016 and has been ratified by nearly all national governments. (Significantly, the USA withdrew from the Climate Agreement in November 2020, but this decision has been reversed by the new Biden administration.) The Agreement's central aim is to keep global temperature rise this century to **well below 2°C above pre-industrial levels** and to pursue efforts to **limit the temperature increase to 1.5°C** in order to avert the worst effects of global warming.

The perceived lack of action by many governments to impose measures that would address the climate 'emergency' gave rise to a series of citizen-led protests. Events were organised by environmental pressure (civil society) groups such as the newly formed 'Extinction Rebellion' and NGOs such as Greenpeace. From 2019, 'School Strike for Climate' protests, inspired by Sweden's then 16-year-old Greta Thunberg, were supported by 1.5 million schoolchildren and students. A new civil society group of mainly under 18s, the UK Student Climate Network (UKSCN), was formed to organise these strikes.

In 2019, the UK government made a declaration of an 'Environment and Climate Emergency' as a result of pressure from citizens, civil society groups, NGOs and from local and regional governments. The declaration itself was largely symbolic with no legislation attached but it did acknowledge:

- the term climate 'emergency'; this has been used by activists in preference to climate change to signal that a more urgent response is required
- that politicians, at all levels, should increasingly engage with climate and environment issues, which should be a central and integral part of government policies, rather than an afterthought.

Non-governmental organisations (NGOs)

Non-governmental organisations (NGOs) operate across boundaries and at all scales and are important participants in the interactions that influence global governance. They can mobilise strong support from individual citizens, giving them more authority to challenge the excesses of TNCs and national governments that threaten equality, justice or the environment and to influence governance.

With greater communication, including social media, to raise awareness of issues, NGOs have expanded their scope and have increasingly become 'international' organisations. In the twenty-first century NGOs have emerged as a global force to:

- democratise decision-making (encouraging governments to consider 'bottom-up' approaches)
- protect human rights
- provide essential services such as food, water or medical help to the most needy.

There is a distinction to be made between **operational** and **advocacy** (campaigning) NGOs. Both are funded by donations and subscriptions, but some also receive money from governments and others from businesses (though this may compromise their independence).

- **Operational NGOs:** those providing frontline support services to the needy (for example, Oxfam, Médecins sans Frontières); they tend to raise money for each project they undertake.
- **Advocacy NGOs:** those who focus on campaigns to raise awareness to gain support for a cause (for example, Greenpeace); derive money from donations and, in some cases, from subscriptions.

NGOs work increasingly in partnerships with other stakeholders, such as governments, international organisations and other civil society groups. They are becoming increasingly important in supporting development and are often the only co-ordinated organisations to provide a voice for the poor.

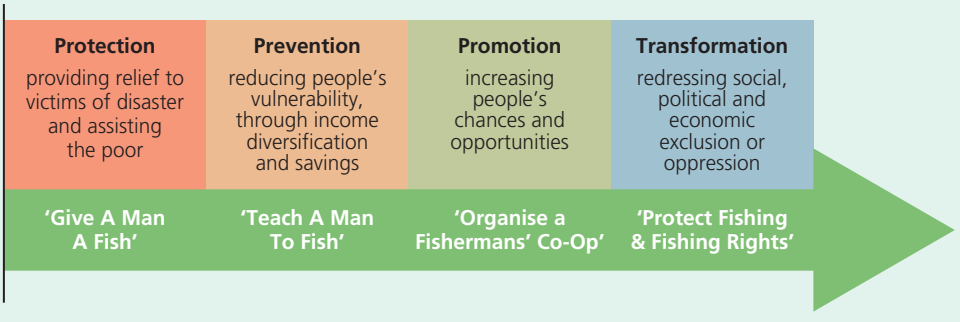


Figure 7.34 Different roles of NGOs

These examples demonstrate the interactions between governance and decision-making that take place at different scales and between different 'players' at each scale. Similar ideas can be applied to other global issues, such as the transmission of COVID-19 in the global pandemic. Not all issues affect all scales and although we live in a globalised world, most governance is retained at the national level. In the UK, for example, decisions about the building of a high-speed rail (HS2) link and stopping fracking for gas are made at national government level but are influenced from above and below.

Review questions

15 Give an example of how global systems are reproduced.

16 Evaluate the role of the IMF in promoting growth and stability.

17 Outline the key problems in achieving multilateral trade agreements on a global scale.

18 Explain the interactions that occur between different scales in global governance.

7.5 The 'global commons'

The 'global commons' refers to resource domains or areas that lie outside of the political reach of any one nation state. It is a term used to describe supra-national 'spaces' in which common shared resources can be found. International law recognises four global commons:

- the high seas
- the atmosphere
- Antarctica
- outer space.

Some commentators would argue that **cyberspace** has also emerged as a new domain which meets the definition of a 'global common' by being a resource that is shared by all but is not controlled by any single nation. This includes the development of the worldwide web, one of the main drivers of globalisation.

Rights and benefits of the global commons

International law relating to the global commons is guided by **the common heritage of mankind principle**. This principle affirms that the natural resources in defined territories or spaces are held in common by all nations. These resources are to be distributed equitably for the benefit of everyone, including for future generations; individual nations have no sovereign rights to exploit them.

Historically, access to these resources (with the exception of fishing and whaling on the high seas) has been difficult.

More recently however:

- advances in science and technology have given easier access to a range of resources
- greater scarcity of resources, especially minerals, fuels and food has put increasing pressure on the global commons to provide resources for a needy and developing world.

These changes mean that the concept of common heritage is being put under increasing pressure.

'The tragedy of the commons'

This concept explains why shared 'common access resources' of any type are likely to be overexploited. If individuals act independently and according to their self-interest, this will be contrary to the interest of the whole group because the shared resource will become depleted.

This is exactly what has happened on a global scale to fish stocks, which have been overfished, leading to depletion, marine pollution and loss of sustainability.

The rights of all to the benefits of global commons can only be secured through sustainable development, which requires effective governance at a global level. Clear rules given by the UN will ensure that the commons are protected from over-exploitation and preserved for future generations.

Protecting the global commons

Each of the global commons is covered by a number of international laws or treaties of one kind or another:

- the high seas by the UN Convention on the Law of the Sea (UNCLOS)
- the atmosphere by the United Nations Framework Convention on Climate Change (UNFCCC), plus other international environmental agreements, including the Paris Climate Accord
- Antarctica by the **Antarctic Treaty Systems** (ATS)
- outer space by the Outer Space Treaty (1967).

Antarctica as a global common

Location and geography

Most of Antarctica is south of the Antarctic Circle, with the exception of parts of the East Antarctic coastline and the Antarctic Peninsula which extends northwards from West Antarctica to about 63°S. Much of the continent's coastline is fringed by ice shelves. The largest of these are the Ross Ice Shelf in the Ross Sea and the Ronne Ice Shelf in the Weddell Sea. Each of these ice shelves covers an area greater than the British Isles.

In the spring of 2000, the International Hydrographic Organization delimited a fifth world ocean – the **Southern Ocean**, also known as the Antarctic Ocean, from the southern portions of the Atlantic, Indian and Pacific Oceans. It comprises the southernmost waters of the world, taken as being south of 60°S latitude.

The Southern Ocean extends from the coast of Antarctica north to 60 degrees south latitude, which coincides with the Antarctic Treaty Limit. At approximately this latitude, cold, northward flowing waters from the Antarctic meet with warmer sub-Antarctic waters. This is known as the Antarctic Convergence zone.

Key term

Antarctic Convergence – A zone continuously encircling Antarctica where cold northward flowing Antarctic waters meet the relatively warmer waters of the sub-Antarctic

The Antarctic Convergence is a zone approximately 32 to 48 km wide, varying in latitude seasonally and in different longitudes, extending across the Atlantic, Pacific, and Indian Oceans between the 48th and 61st parallels of south latitude. Here the cold, northward flowing Antarctic waters sink beneath sub-Antarctic waters. This forms associated areas of mixing and upwelling currents which are very high in marine productivity, especially Antarctic **krill.**

The Antarctic Convergence Zone (Figure 7.36) is a natural boundary that separates:

- two distinct hydrological regions
- areas of distinct climate
- areas of distinctive wildlife.

Figure 7.36 shows the Convergence Zone at around 60 degrees latitude south; this gives rise to the largest surface current in the world called the **Antarctic Circumpolar Current** which flows around Antarctica. This current effectively blocks warmer waters travelling southwards. The Circumpolar Current flows in an eastward direction being driven by westerly winds, and is otherwise known as the West Wind Drift. A lesser current, closer to Antarctica, flows in the opposite direction westwards; it is known as the East Wind Drift and is driven by easterly polar winds. The East Wind Drift features prominently in the Ross and Weddell Seas. The area where these opposite flowing currents meet is known as the zone of **Antarctic Divergence**.

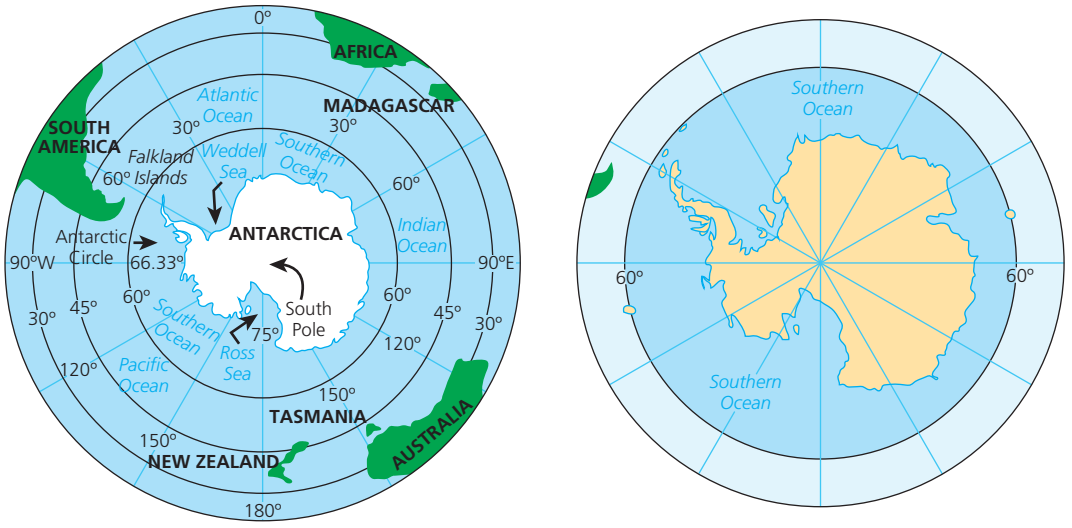


Figure 7.35 Map of Antarctica and the Southern Ocean

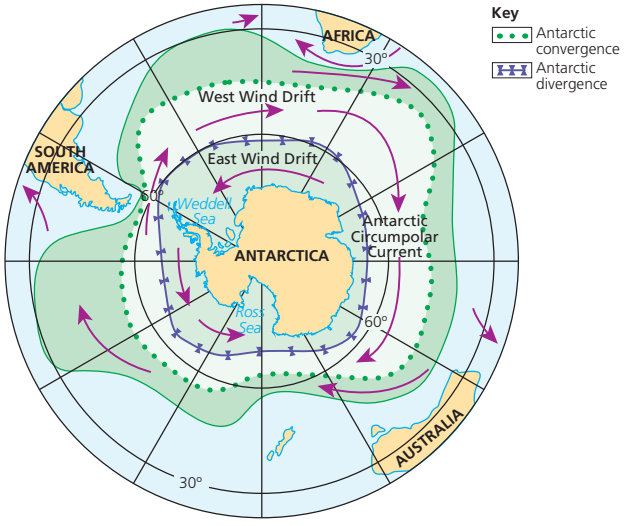


Figure 7.36 The Antarctic Convergence Zone – Circumpolar Current and Divergence Zone of easterly and westerly flowing currents

Physical geography and climate of Antarctica

Antarctica is a unique continent as its mountainous and rocky terrain is almost entirely (98 per cent) covered by glacial ice (Figure 7.37, page 336). The continent is unevenly divided into East (Greater) Antarctica and West (Lesser) Antarctica by the Transantarctic Mountains. The Transantarctic range contains many peaks above 4,000 m and it also divides the West and East Antarctic ice sheets. The East Antarctic ice sheet is larger, thicker and much older than the West Antarctic ice sheet. The Ellsworth Mountains are found in West Antarctica and contain the continent's highest peak – Mount Vinson (4,892 m).

With an average height of 2,500 m above sea level, Antarctica has the distinction of being the highest continent. While it does have high mountains, it is the thickness of the ice sheets which give it the highest average elevation of all the continents.

Climate

In addition to being the highest continent, Antarctica is also distinctive for being the coldest, windiest and driest. Climatic features include:

- average temperature is –50°C, though it can be as low as –89°C
- mean annual wind speed of 50 miles per hour (mph) resulting from the convergent katabatic winds; gales can reach 200 mph
- mean annual precipitation, falling predominantly as snow, is low (especially in the interior) at less than 50 mm per annum
- it can best be described as a **polar desert.**

The ice sheets, which are 2.5 km thick, are the result of an accumulation of small inputs of snow and frost, which far exceed ablation rates, over a very long period of time.

Very little of Antarctica is free from ice and any ice-free areas owe their existence to specific local-scale factors. For example, high mountain peaks called **nunataks** are pinnacles of rock emerging above the ice sheets, like islands in a sea of ice (Figure 7.38a, page 337). High winds and steep slopes prevent snow and ice from accumulating on parts of these mountains.

Dry valleys are another intriguing type of landscape found in Antarctica (Figure 7.38b, page 337). These are found in high altitude areas of extreme aridity.

Surrounding sea temperatures are warmer than the land and marine conditions support a diverse ecosystem. Upwelling of cooler water from the ocean depths brings nutrients, which in turn support **phytoplankton** (microscopic algae that are the base of the marine food web). Blooms of phytoplankton provide food for krill, which many species depend on as a food source.

Some coastal areas of Antarctica, particularly along the Peninsula, have micro-climatic and topographic conditions which cause enough melting during summer months to allow some land to remain free of ice. The areas on the western Peninsula are likely to see even more climatic differences because of climate change.

Threats to Antarctica

Fishing and whaling

The ocean's nutrient rich waters support an abundance of marine life and are the feeding ground for 80 per cent of the world's great whales. Whaling began in the nineteenth century. As the whale population of the North Atlantic became reduced due to massive exploitation, whalers turned their attention to the Southern Ocean.

Whaling was a highly profitable business and whaling stations were established on South Georgia and the

Figure 7.37 Physical map of Antarctica

South Shetlands. The range of products included oil, whalebone, meat extract and, in later years, frozen whale meat. By the 1960s, whaling was no longer commercially viable as whale populations had declined by over 90 per cent due to overkilling and many species had become endangered.

The **International Whaling Commission (IWC)** placed a moratorium on commercial whaling in 1986. However, Japan's position on whaling in the Southern Ocean caused international tensions over the issue (see page 343).

Whales in the Southern Ocean are currently protected, but there is still an underlying threat from Japan and other pro-whaling nations, notably Norway and Iceland (both IWC members), who continue to lobby for more commercial whaling.

Fishing has now replaced whaling in the area. In the 1960s, Russian ships began to exploit the Southern Ocean for a number of fish species, including the Antarctic toothfish. Concerns have been expressed over the quantity of fish being taken, particularly by Russian and Japanese fleets fishing for krill.

Figure 7.38a Nunataks – Una Peaks at Cape Renard in Graham Land on the Antarctic Peninsula

Figure 7.38b Dry valley – Wright Valley, one of the McMurdo Dry Valleys near the Ross Sea, with wind carved rock in the foreground

Krill

Krill are tiny shrimp-like crustaceans which swim in massive numbers and are a key part of the entire food chain of the Southern Ocean. Virtually everything in the Antarctic region is dependent on krill – seabirds, penguins, seals, whales and other fish, such as the toothfish, which are another major fisheries target. A loss in the krill biomass will affect most other species.

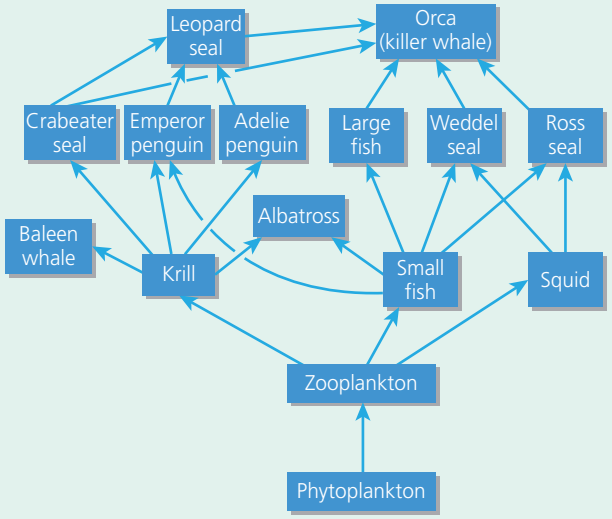


Figure 7.39 Southern Ocean food web reliant on the shrimp-like krill

Krill are being caught for processing into feed for farmed fish and as a nutritional supplement in the form of **krill oil** all over the world. They are an excellent source of protein. Krill oil is extremely rich in omega-3 fatty acids and is used to prevent hypertension, strokes and depression. Cartons of krill oil capsules can command very high prices (as much as £40 for 100 capsules for red krill oil) in health food retail outlets.

The sudden recent demand for krill products has caused a dramatic increase in fishing that coincides with a long-term decline in krill biomass. Unsustainable fishing practices may cause a krill population crash, which would have a catastrophic impact on the Antarctic ecosystem.

Commercial fishing is becoming a significant threat to the Southern Ocean and Antarctica. There is a high risk of overfishing of target species, by-catch (unintentionally catching non-target species while fishing for target species) and direct destruction of marine habitats by ships and fishing gear. Ships also dump waste into the ocean, such as discarded plastic, which contaminates the water. Marine life is also killed by ingesting or becoming ensnared on lost fishing nets and hooks.

Climate change

Since Antarctica is a large area, climate change is not having a uniform impact. Some parts of east Antarctica are accumulating small amounts of land ice but there are dramatic increases in sea ice extent on the eastern side. The west side, however, is losing significant quantities of both land ice and sea ice, causing substantial impacts on wildlife.

Western Antarctic Ice Sheet and Antarctic Peninsula

The Antarctic Peninsula is particularly sensitive to small rises in the annual average temperature; this area has experienced air temperature increases of nearly 3°C in the region in the last 50 years, which is a much faster rate than the average for global warming, according to the **Intergovernmental Panel on Climate Change (IPCC)**. The temperature of Western Antarctica has risen and, as a result, the West Antarctic Ice Sheet has thinned significantly.

This warming is not restricted to the land and Southern Ocean temperatures to the west of the Antarctic Peninsula have also increased by over 1°C since 1955. It has also been established that the Antarctic Circumpolar Current is warming more

rapidly than the global ocean as a whole. This warming has had the following effects on the physical and living environment of western Antarctica:

- distribution of penguin colonies has changed
- melting of snow and ice cover has increased colonisation by plants
- decline in the abundance of Antarctic krill
- glaciers and ice shelves fringing the Peninsula have retreated; some have collapsed entirely
- there is particular concern about the **Thwaites Glacier** on the west coast as it is melting more quickly than previously thought. It is an unusually broad glacier and is one of the largest contributors to global sea level rise. It is thought that complete collapse could cause a sea level rise of up to 63 cm
- the Ronne and Ross sea ice shelves will not add to sea level change but as they break up it increases the flow rate of glaciers behind them – it is the melting of these glaciers that will cause sea levels to rise.

The temperature of Antarctica as a whole is predicted to rise by a small amount over the next 50 years. Any increase in the rate of ice melting is expected to be partly offset by increased snowfall as a result of the warming.

East Antarctica and sea ice expansion

The temperature of Eastern Antarctica has risen by a much smaller amount and unlike the Peninsula, there is no significant loss of ice of any kind.

The paradox is that while the world is warming, sea ice is expanding on the eastern side of Antarctica towards the Indian and Pacific Oceans. In 2014, the extent of Antarctic sea ice was at an all-time record extent of 20.11 million square kilometres, larger than the continental land mass itself. Sea ice is a thin layer between one to two metres thick. It is formed when water is cooled sufficiently by the surrounding atmosphere. Once it is formed it can be blown by the wind or taken by currents, ultimately expanding and bonding with other floating sheets of ice.

There are four reasons why climate change may contribute to Antarctic sea ice expansion:

- Increasing westerly winds around the Southern Ocean, caused by climate change and ozone depletion are driving the seas northwards.
- More rain and snow resulting from climate change are layering the Southern Ocean with a cooler, denser layer on top.
- These storms are also freshening the local water (it becomes less salty) thus raising the temperature needed for sea ice to form.
- Increased melting of continental land ice creates more floating icebergs which contribute to sea ice formation.

Ocean acidification

The decline in pH of ocean water results from an increase in atmospheric carbon dioxide (CO_2) concentrations creating carbonic acid, which makes the ocean slightly less alkaline.

This poses potential threats to the marine environments of the Southern Ocean. The rapid rate at which acidification is progressing will make any adaptation difficult. Ocean water may actually become corrosive to unprotected shells and exoskeletons of organisms; losing these organisms will disrupt food webs.

The search for mineral resources

There has never been commercial mining in Antarctica as it is completely banned by the Antarctic Treaty. There are no current or known future plans by any of the Antarctic Treaty nations to reverse this decision. Nevertheless, the future demand for resources will undoubtedly put pressure on the vast mineral reserves that are to be found on the continent (Figure 7.40).

The difficult and dangerous Antarctic conditions make exploitation unlikely as it is unviable. The expense of extracting resources through very thick ice and then establishing transport systems across moving glaciers and melting streams would prevent any commercial operation.

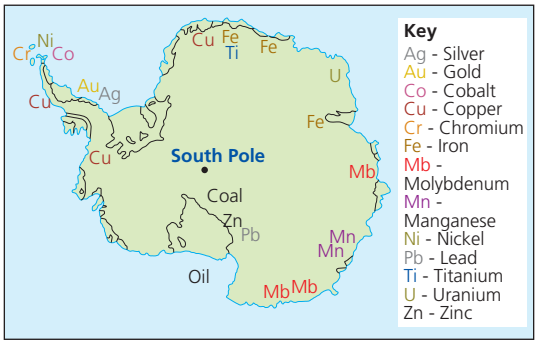


Figure 7.40 Mineral map of Antarctica showing known deposits of minerals

The initial moratorium on mining was voluntary but the issue was raised by some members of the Treaty in the 1980s who attempted to formulate a new convention allowing exploration for mineral and gas

reserves. Their proposed Convention on the Regulation of Antarctic Mineral Resource Activities was never ratified and later countered by the **Madrid Protocol**, which protects Antarctica from exploitation until 2048.

Tourism and scientific research

Antarctica is not populated except by scientists at a small number of permanent research stations. Scientists and tourists are the main human threats to Antarctica.

Scientific researchers working either in the interior near the South Pole or in coastal areas are well briefed about the need for care of the Antarctic environment, but any activity is bound to create some degree of disturbance to the environment and wildlife. Impacts are caused by vehicle exhausts, construction of buildings and related facilities such as fuel storage, runways and the disposal of waste such as rope, fuel drums and plastics, all of which endanger living organisms.

Figure 7.41 Penguin 'posing' for tourist, Antarctica

The summer tourism season in the Southern Hemisphere goes from mid-November to March. Antarctic tourism takes a variety of forms:

- **Seaborne tourism** – usually a cruise departing from Ushuaia in Argentina or Punta Arenas in Chile, with the opportunity to transfer to smaller boats for landing or cruising close to shore.
- **Airborne tourism** – flying from Australia, New Zealand or South America; can involve scenic flights over the continent or in some cases landing.
- **Fly-cruises** – for those with less time or wanting to avoid a long sea journey, flights land on King George Island from where the visitors take a cruise boat.

Tourists are attracted by the glacial landscapes and the wildlife, for example, penguins (Figure 7.41). They are also interested in visiting historic sites, such as McMurdo Sound with its huts dating from the Scott and Shackleton expeditions. Many go to experience the remoteness and isolation of a true wilderness. The tourist experience varies depending on whether they set foot on land or not.

Tourist numbers visiting Antarctica have grown steadily from just a few hundred specialist visitors in the late 1950s, to nearly 56,000 (of whom 45,000 landed) in 2018–19. For those who land, there are different itineraries with optional activities such as visiting penguin or seal colonies, kayaking, etc. With prices becoming relatively cheaper and more tour operators joining the **International Association of Antarctic Tour Operators (IAATO)** each year, there are concerns that what used to be a niche market is now edging towards mainstream tourism. However, the IAATO try to ensure there are measures in place to

Table 7.17 Negative impacts and threats of tourism, and mitigation, resilience and adaptation arguments

Negative impacts	Mitigation, resilience and adaptation arguments
• Disturbance of an extremely fragile ecosystem can leave an imprint for a long time (walking the same routes can create scars or paths) • Tourist season coincides with peak breeding season for most Antarctic wildlife • Terns are particularly disturbed by visitors and may abandon nests or vacate an area completely • Over-flying by light planes or helicopters causes stress to breeding colonies of penguins and birds • Accidentally importing invasive alien species, such as insects or seeds on boots, clothes etc. • Oil spills from ships and boats • The possibility of more permanent land-based tourism being developed	• Antarctic tourism is a well governed industry with strict rules set by the IAATO and a sound record for environmental concern • Regulations are widely accepted by tour operators and tourists • Visitors become ambassadors for Antarctica, voicing the need for its preservation from development • Visitors tend to be well educated and are receptive to the conservation message, which is reinforced by the tours • Seals and penguins seem to be resilient and have adapted to the presence of humans on land • Scott Polar Research Institute findings show that to date tourism has had minimal overall impact

minimise the impacts and maintain a sustainable tourism operation. They work with members of the Antarctic Treaty to formulate mandatory rules, including:

- no large ships of more than 500 passengers are allowed to visit; most of the ships are comparatively small with a capacity of between 50 to 200 people
- when landing, groups must be divided into boatloads of around 20 people, each group being led by a qualified guide
- briefing tourists on a code of behaviour when ashore, including adherence to health and safety requirements and rules about wildlife observation
- limits on visiting sites – each site may be visited only every three to four days to minimise impacts and allow recovery.

Although landings are only for a few hours, they are concentrated in a few accessible areas and are relatively high impact. The growth of tourism poses a long-term threat to the environment of Antarctica and the Southern Ocean. These threats are evaluated in Table 7.17 (page 339).

Vulnerability to environmental change

The threats to Antarctica may disturb its fragile environment and will impact on indigenous species and biodiversity. The **Scientific Committee on Antarctic Research (SCAR)** monitors the resilience and adaptation to these disturbances and also the effect of any mitigation measures put in place.

- **Resilience** gives an ecosystem the ability to recover from 'shock' events, disturbances or ongoing change, whether natural or caused by human activity. Endemic species are usually more resilient to shock events as they evolve strategies to allow their populations to rebuild. For example, sealing, whaling and hunting penguins pushed species to the brink of extinction in the past. Once these activities ceased, many species recovered and, in some cases, such as the king penguin, thrived.
- The ability of an ecosystem to return to an equilibrium status is also linked to the **adaptation** of species to changing conditions. For example, some organisms may benefit, at least in the short term, from climate change. Many others will be vulnerable as their ability to adapt is slower and environmental change is taking place too quickly for them.
- **Mitigation** is intervention by humans to eliminate or reduce the impacts presented by these threats. In the case of Antarctica, this requires governance of a continental land mass with agreements at a global level, so that mitigation efforts are effective.

Key terms

Adaptation – Any alteration or adjustment in the structure or function of an organism or system which enables it to survive better in changing environmental conditions.

Mitigation – Includes any actions, strategies, measures or projects undertaken (by mankind) to offset the known detrimental impacts of a process.

Resilience (ecological) – The amount of disturbance that an ecosystem can withstand without changing existing structures and processes.

Governance of Antarctica

Since its discovery and further exploration, during an era when colonial and territorial claims were still being made globally, Antarctica has been a politically sensitive area. There are still many geopolitical issues that need to be addressed to ensure its ongoing protection as a global common. Despite this, it is often cited as an example of good global governance and co-operation between nation states.

The Antarctic Treaty (AT)

Against a backdrop of underlying Cold War geopolitics, the Antarctic Treaty was drafted in 1959 and ratified by the 12 original member nations (see Figure 7.42) to come into force in June 1961. The treaty formalised the notion that Antarctica does not 'belong' to any single nation. Instead, it designated the whole area south of 60°S latitude as an area where countries would work together for the common causes of scientific research. The treaty:

- stipulates that Antarctica be used exclusively for peaceful purposes; military activities are banned
- guarantees continued freedom to conduct scientific research
- promotes international scientific co-operation including the exchange of research plans and personnel and sharing of research results
- freezes any claims to sovereignty for the duration of the treaty and prohibits any new or enlarged claims
- prohibits nuclear explosions and the disposal of radioactive waste

- allows ships, stations and equipment to be inspected by observers to ensure compliance
- requires parties to give advance notice of their expeditions
- includes a dispute settlement procedure and a review mechanism by which the treaty can be modified.

Members of the AT committed to operate on a consensus basis, regardless of size or status; thus, the superpowers of the time, the Soviet Union (Russia) and USA, had to work on an equal footing with others, such as Belgium, New Zealand and South Africa. Other nations were allowed to become signatories to the AT but to be able to vote had to demonstrate 'substantial scientific activity', such as establishing a research station in Antarctica.

International government organisations

Other than facilitating and supporting the concept of the treaty, the **United Nations (UN)** had little involvement in the early governance of Antarctica and responsibility was devolved to AT member states. By the 1980s, however, there was increasing international pressure for access to Antarctic resources.

- Other countries, including China and India, had become AT signatories with interests in gaining some influence in the control of resources in Antarctica.
- Some member states wanted to negotiate management of the fisheries.
- A Convention on the Regulation of Antarctic Mineral Resource Activities was proposed in 1988 that would have allowed exploration for minerals.

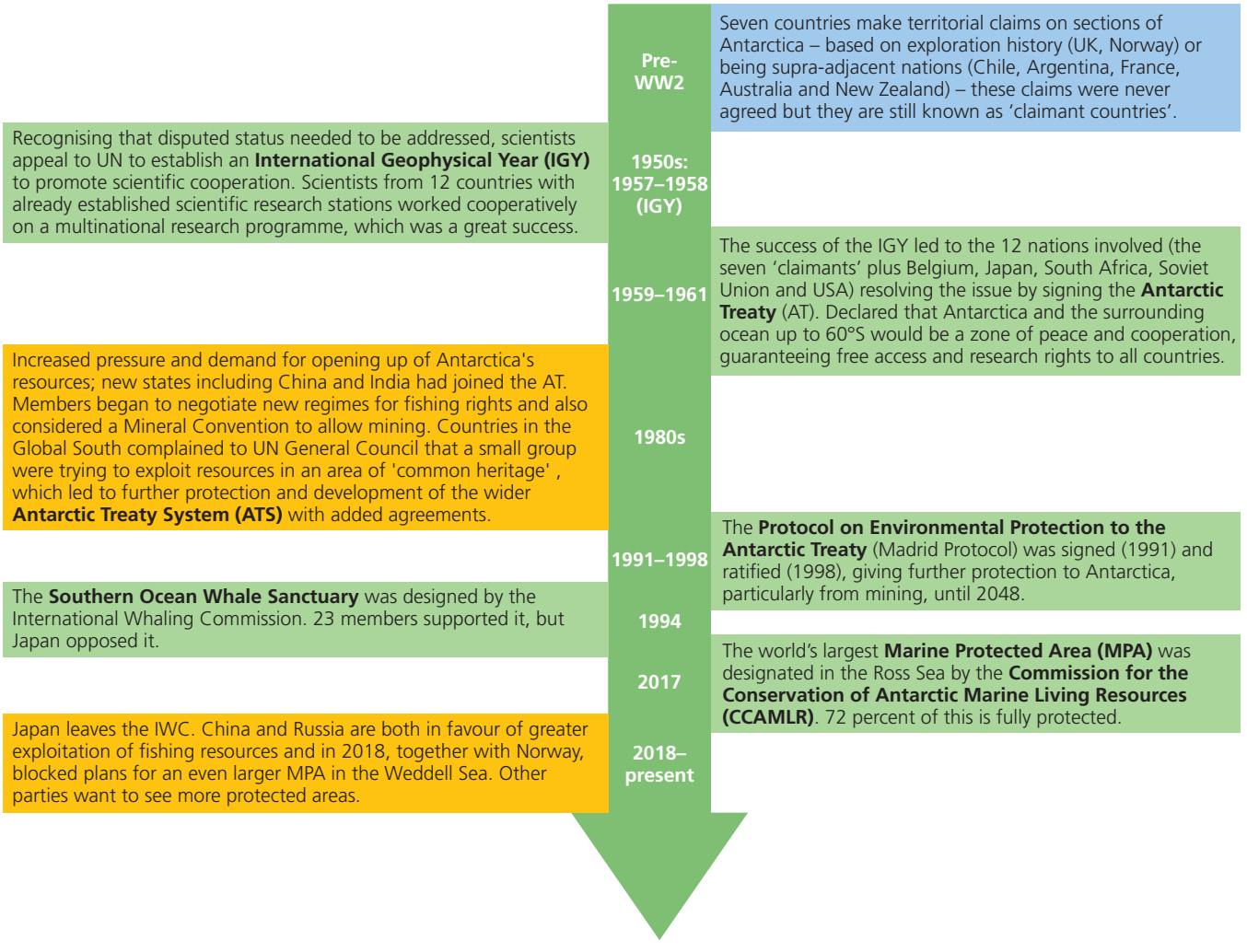


Figure 7.42 Timeline of key developments in the governance of Antarctica

As a result of these pressures, the governance of Antarctica was contested by some nations (AT members and non-members) who raised concerns in the UN General Assembly about its accountability, suggesting that it was in danger of succumbing to pressures of resource development. Under the legal provision of the 'common heritage of mankind', AT members accepted proposals from the UN that the **United Nations Environment Programme (UNEP)** would extend the Treaty by adding agreements that would strengthen and secure protection for Antarctica. This led to the establishment of the broader **Antarctic Treaty System (ATS)**.

The Antarctic Treaty System (ATS)

International governance of Antarctica is under the control of the ATS, which is managed by annual **Antarctic Treaty Consultative Meetings (ATCM)**, comprising the original 12 member states and a further 16 'consultative parties'. The ATS includes a number of other organisations, such as NGOs and scientific institutions that influence the decision-making process regarding activities taking place on the continent.

The ATS includes the AT at its core, together with a number of additional agreements:

- the Protocol on Environmental Protection to the Antarctic Treaty (Madrid, 1991) (see below)
- two separate conventions dealing with the conservation of Antarctic seals
- the Commission for the Conservation of Antarctic Marine Living Resources (CCAMLR).

These supplementary agreements have added to the scope of governance which now includes:

- protection of the Antarctic environment
- conservation of plants and animals
- preservation of historic sites
- designation and management of protected areas
- management of tourism
- collection of meteorological data and hydrographic charting.

The Protocol on Environmental Protection to the Antarctic Treaty 1991 (the Madrid Protocol)

The 'Environmental Protocol' was negotiated by **UNEP** and treaty members at an Antarctic Treaty conference in Madrid in 1991. The negotiation followed years of talks on controlling mineral resources in Antarctica. The 1988 Convention on the Regulation of Antarctic Mineral Resource Activities proposed for mineral exploration to go ahead in Antarctica. However, AT members decided not to adopt the Convention because of international pressure from France, Australia and many non-AT countries expressing the need for environmental protection. Instead, they followed a separate proposal that eventually led to the Environmental Protocol, which was ratified and came into force in 1998. The Protocol:

- designates Antarctica as a 'natural reserve, devoted to peace and science'
- establishes environmental principles for the conduct of all activities
- prohibits mining or mineral resource exploration including exploration of the continental shelf
- established a Committee for Environmental Protection
- requires operators to develop contingency plans to respond to environmental emergencies
- requires that waste of all kinds be returned to the country of origin wherever possible.

The ban on mining can only be modified if all parties agree. Otherwise, the next review of the mining prohibition is due in 2048. This means any modifications to the prohibition are unlikely beforehand and only then if 75 per cent of the consultative parties agree.

Systems for inspection and enforcement

Inspection and observation of all operations on Antarctica are compulsory under the terms of the Treaty.

Antarctica has no permanent population and so no citizenship or government. All personnel present on Antarctica are citizens of a nation outside of Antarctica and under that nation's jurisdiction, so there is no single unifying 'legal system'. Prosecutions against treaty rules are the responsibility of individual nation states through the **ATCM**.

The Madrid Protocol introduced a more rigorous regime. Observers are now designated by the ATCM. **Environmental audits** are carried out around scientific bases, on land and in the sea to assess the impact that bases and their activities are having on the surrounding area.

The scope of inspections under the Protocol includes inspection of stations, ships, aircraft and loading areas, unfortunately with a lesser focus on maritime areas, which are equally important. Any new activities by operators are subject to **environmental impact assessments**.

International Whaling Commission (IWC)

The IWC is the global body responsible for the conservation of whales and the management of whaling activities. It was established in 1946 to provide for the proper conservation of whale stocks and the sustainable development of the whaling industry.

Figure 7.43 Part of the Southern Ocean Whale Sanctuary

The main duty of the IWC is to oversee the measures laid down by the Whaling Convention which governs the conduct of whaling globally. These measures:

- provide complete protection of certain whale species
- designate specified areas as whale sanctuaries
- establish limits on the numbers and size of whales which may be taken
- prescribe open and closed seasons and areas for whaling
- prohibit the capture of suckling calves and female whales accompanied by calves
- require catch reports and other statistical and biological records.

In 1994, the IWC established the Southern Ocean Whale Sanctuary (Figure 7.43), an area of 50 million km^2, where all commercial whaling is banned, though Japan continued its 'research whaling' there.

The International Whaling Moratorium

The IWC recognises three different types of whaling:

- aboriginal subsistence whaling
- commercial whaling
- special permit whaling for scientific research.

In 1982, the IWC decided that there should be a suspension of commercial whaling on all stocks from 1986 onwards. This is referred to as the **whaling moratorium**, and it has remained in effect, despite efforts by Japan and other pro-whaling nations to overturn it.

- Japan continued to evade the moratorium by 'special permit' whaling which enabled them to exploit a loophole in the IWC regulations that permitted whaling for 'scientific research' in the Southern Ocean. (Japan stated that the purpose of its 'research whaling' was to monitor stocks to pave the way for the resumption of commercial whaling.)
- Japan's whaling activities in the ocean were continually monitored by the Sea Shepherd Conservation Society, a marine conservation NGO, and by the Australian government.
- In 2014, the International Court of Justice ruled that Japan's whaling fleet had, in effect, used 'research whaling' as a cover for commercial whaling.
- Norway and Iceland continue to hunt common minke whales commercially within their own exclusive economic zones (EEZ); neither takes whales from the Southern Ocean.

In 2018, IWC members met in Florianopolis, Brazil and rejected a proposal by Japan to remove the moratorium on commercial whaling. As a result of the discussions, the '**Florianopolis Declaration**' concluded that the moratorium would continue in perpetuity to allow the recovery of whale populations to pre-industrial whaling levels. This led to Japan withdrawing its membership of the IWC to resume commercial whaling within its own territorial waters and EEZ. Japan had to cease 'special permit' whaling activities in the Southern Ocean.

The moratorium does not apply to 'aboriginal subsistence whaling' carried out by indigenous cultures in Alaska, Greenland and parts of Canada. This is not viewed as commercial whaling, though it is regulated by the IWC, which sets catch limits every six years.

Non-governmental organisations in the Antarctic

Environmental NGOs are valuable and influential participants in the governance of Antarctica, especially as they represent those who consider environmental protection to be a top priority. There are a number of NGOs and international conservation organisations operating in the Antarctic. Table 7.18 outlines the role and background of two of the most important of these.

The role of NGOs and international conservation organisations in Antarctica

NGOs play a role in influencing decisions taken about development. Their ability to influence proceedings is enhanced because they have no sovereign national agenda and they can act as:

- **guardians** – to monitor and challenge the decision making of ATS members and promote alternatives. For example, ASOC, with strong involvement from Greenpeace, was successful in campaigning to block the Minerals Convention and instrumental in the development of the Madrid Protocol
- **technical advisers** – to make use of their expertise to gain acceptance and influence decisions made by ATS members. For example, ASOC's advice to adopt an 'ecosystem approach' to management of fisheries was embedded into ATS conservation strategies
- **partners** – to work with ATS members and other bodies to designate areas of protection. NGO groups worked with CCAMLR to establish the Ross Sea Marine Protected Area, from 2012 until it came into being in 2017. The Ross Sea MPA is the largest in the world, covering 1.55 million km^2 (1.12 million km^2 is fully protected) and will remain protected for 35 years
- **visionaries** – to formulate long-term and large-scale strategies. For example, ASOC's goal is to establish a network of MPAs, including the adoption of the proposed Weddell Sea, Antarctic Peninsula and East Antarctica MPAs. During negotiations over the Minerals Convention, Greenpeace's vision was the creation of a 'World Park Antarctica' that would ensure the entire continent is 'off limits' for any type of resource exploitation or development.

Figure 7.44 Antarctic and Southern Ocean Coalition logo

Table 7.18 International conservation organisations in Antarctica

Organisation	Background	Role and influence
International Union for Conservation of Nature and Natural Resources (IUCN)	An umbrella organisation that includes both government and NGO bodies	• A respected knowledge-based coalition of expert groups • Shares advocacy on some issues but primary role is to provide advice on conservation issues • Has had 'invited expert' status since 1987 and was first NGO to gain formal observer status on the full ATCM • Offers expertise but not a decision maker
Antarctic and Southern Ocean Coalition (ASOC)	A coalition group of over 20 NGOs interested in the protection of the Antarctic environment, including Friends of the Earth, the Pew Charitable Trusts, Greenpeace and WWF. ASOC members pursue their own individual organisational Antarctic work, in addition to working collectively through the coalition The only NGO group that dedicates all its time to Antarctica and the Southern Ocean and the only environmental NGO with official status at Antarctic governance meetings	• Provides co-ordination between national NGOs to develop shared positions on Antarctic conservation issues • Gradually earned respect from ATS members and gained observer status in the CCAMLR meetings • ASOC and its members participate in a wide range of activities and campaigns related to Antarctic environmental protection. E.g., from 1987–1991, Greenpeace established an NGO base where it drew attention to the need to protect Antarctica from oil and gas exploitation • Helped to establish the partner **Antarctic Ocean Alliance (AOA)** that was instrumental in supporting CCAMLR to establish the **Ross Sea Marine Protected Area (MPA)** • ASOC wants to see ATS members doing more to manage human activities and to protect biodiversity, including by establishing more terrestrial and marine protected areas

ASOC in particular campaigns to bring key issues to the attention of ATS members, to ensure that they are negotiated and, if agreed, implemented and then monitored, including:

- the creation of systems of marine and terrestrial protected areas to preserve biodiversity and foster ecosystem resilience in a changing climate
- the incorporation of climate change into all relevant aspects of decision making
- the precautionary, ecosystem-based management of Southern Ocean fisheries, including krill
- the regulation of tourism.

For more information on ASOC, visit the website at www.asoc.org.

Assessment of global governance of Antarctica and future consequences

The ATS is recognised as being one of the most successful international agreements. Tensions concerning territorial claims have been suspended and as a disarmament measure it has been effective. The involvement of the UN and ATS member countries, with the underlying influence of NGOs, demonstrates how governance of a 'global common' area has been integrated effectively. However, protecting Antarctica is not easy and there are still issues to be resolved concerning its future governance.

- **Duration of agreements** – The Environment Protocol underpinning the ATS system could be reviewed in 2048 and could be modified, as could the Ross Sea MPA in 2052.
- **Territorial claims** – The ATS has not rejected any existing claim, but it prohibits their assertion as well as the establishment of new claims. None of the seven 'claimant nations' have renounced their claim; the USA and Russia also maintain the 'right' to lay claims.
- **Impacts of human activity** – The increasing footprint from scientific activity, tourism and fishing does not help environmental protection.
- **Marine Protection Areas (MPAs)** – In the next 10 years, to meet UNEP global biodiversity goals, attempts to establish additional MPAs are likely to lead to disagreement between those ATS parties wanting to extend marine protection and those, such as China and Russia, who want greater access to fisheries. In 2018, these nations, together with Norway, blocked a proposal for the Weddell Sea MPA plan.
- **Climate change** – Antarctica is an essential area for research that will help us to understand and address climate change. In the short term, scientists can do little to prevent land glaciers melting and oceans becoming warmer, which makes confronting some of the other issues, such as MPAs, more complex.
- **Geopolitics of Antarctica** – Pro-conservation 'claimants', such as Australia, are concerned that China and Russia are trying to increase their claims to territory. In 2009, China established its third research station at Dome Argus on the Antarctic plateau. They proposed a code of conduct there, which is still under discussion, with other parties concerned about potential limits on their own scientific activities. However, similar efforts by Western powers to manage scientific research in other sensitive areas have been accepted. China has an astronomical observatory at Dome A, the best place globally for space observation. The area is also thought to contain uranium deposits.

The governance of Antarctica continues to prevail but relies on political compromises. It cannot be assumed that Antarctica will be protected forever, especially if some of the more powerful members, such as China or the USA, adopt a more nationalist approach and start to question whether Antarctica needs to be protected.

Review questions

19 Outline key developments in the establishment of the Antarctic Treaty as the accepted method for the governance of Antarctica.

20 Explain how and why environmental protection in Antarctica was extended.

21 Evaluate the effectiveness of the International Whaling Commission in reducing the threat to whales in the Southern Ocean.

22 Describe the role of NGOs in enhancing the protection of Antarctica.

23 Assess the possible geographical consequences of a disintegration of the current governance system in Antarctica and the Southern Ocean.

7.6 Globalisation critique

The quote at the start of this chapter suggests that globalisation is an unstoppable phenomenon with dimensions that will continue to expand. For the past three decades, there has been an almost continuous growth of world trade, global tourism and international co-operation.

Challenges of globalisation

Economic challenges

- The notion that globalisation would advance economic growth and development for all was challenged by the global financial crisis in 2008–2009, from which many economies are still recovering.
- Other indicators suggest that globalisation has either 'peaked' or at least stalled. For example, global trade **growth** has slowed significantly in recent years (Figure 7.45).
- Ongoing geopolitical tensions between the USA and China, with escalating tariffs being used by both nations, have limited trade growth.

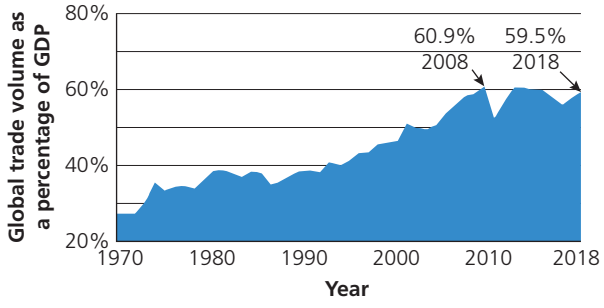


Figure 7.45 Global trade volume as a percentage of GDP 1970–2018 – has globalisation passed its peak?

Source: The World Bank

Social challenges

- Growth in international economic migration offers bilateral benefits for some nations, for example, migrants from the Philippines working in the UAE. Some destination countries, however, experience racial tensions and conflicts of culture.
- The transfer of manufacturing jobs from HDEs to lower cost economies causes unemployment, wage stagnation and related social problems of deprivation in some areas in HDEs. In communities experiencing this decline, there may be a feeling of neglect.

Political challenges

Many communities feel that they have not benefitted from global integration and that their lives have been made worse. The recognition or perception that globalisation has failed to be inclusive has led to the rise of 'populist' movements, which reject existing political establishments, especially in some more developed countries. Populist 'nation-first' campaigns contributed to two significant political events in 2016:

- the UK electorate deciding to leave the European Union, a process later entitled 'Brexit'
- the election of Donald Trump as US President on a 'Make America Great Again' platform.

Brexit and globalisation

When voters in the UK decided to reject membership of the EU, it was seen by many as a rejection of greater integration and the globalisation process. There may have been a number of reasons for this:

- **Sovereignty** – The 'Leave' campaign used the 'Take back control' slogan to persuade voters that the UK had lost the power to make its own decisions on legal matters or to negotiate its own trade deals.
- **Immigration** – With freedom of movement in the EU, nearly one million East European migrants had entered the UK since 2004; this caused tensions because of pressure on jobs, services and so on, as well as discrimination in some areas.
- **Inequality** – A widening gap between the rich and the poor; the Gini coefficient had increased from 0.27 in 1973, when Britain joined Europe, to 0.35 in 2016. Many communities felt 'left behind' because of decline as a result of deindustrialisation, and blamed EU membership for this.
- **Pro-nationalistic feelings** – Some of the UK electorate prefer the idea of an independent Britain making its own future in the world.

The UK left the EU on 31 January 2020 and entered a transition period until the end of 2020. Many Leave supporters argue that Brexit is not an 'antiglobalisation' vote and that it gives the UK the opportunity to trade more globally without the constraints of being a member of a customs union.

'Make America great again' and globalisation

In November 2016, a similar populist movement was mobilised by Donald Trump in the US, when he was elected US President with a promise to return jobs to American workers, with more protectionist policies. The previous Obama administration had started to open up the US to more global trade, but Trump reversed this by promising to renegotiate free trade deals and restrict movement of workers.

Examples of some of the US's 'America-First' measures since the election of Trump include:

- the removal of the US from the Trans-Pacific Partnership (TPP) trade agreement
- rhetoric about building a wall between the USA and Mexico to prevent illegal immigration from Central America
- the encouragement of US car manufacturers, such as Ford, to build new factories in the USA, rather than investing in production in Mexico
- instigation of US withdrawal from the Paris Climate Agreement
- the increase of tariffs on imported manufactured goods from China and other parts of Asia, in an attempt to stimulate production in the traditional manufacturing regions of the US.

Despite losing the 2020 Presidential election, support for Trump remained strong. However, since taking office in 2021, newly-elected President Biden has signalled an intent to reverse many of the Trump policies as his administration adopts a more 'internationalist' approach.

Summary of arguments

Table 7.19 presents some of the arguments both for and against globalisation that have been covered in this chapter.

Table 7.19 Benefits and costs of globalisation

Benefits of globalisation	Costs of globalisation
Growth and development • Investment in LDCs reduces poverty • TNCs create jobs and stimulate growth and help LDCs to develop • Internet provides more opportunities for people to participate and has supported enterprise in many LDCs • LDC economies are growing more quickly and there is less inequality between countries	**Inequality and injustice** • Inequality of income and opportunities within countries has increased • LDCs are still disadvantaged with access to trade, as barriers still exist that protect richer nations • Governments of HDEs have more influence in international organisations • TNCs have more political influence and can exploit weaker governance in some LDCs to exploit resources, workforce etc
Integration • Free trade removes barriers and gives more access to LDCs • Nations work together more co-operatively to tackle global environmental issues such as climate change • Greater cultural understanding and tolerance as different cultures are shared	**Environmental impacts** • TNC investment in LDCs to develop resources or manufacturing is not always environmentally sound or sustainable • Trying to achieve global action on climate change is difficult, especially as major industrial economies such as the USA and China continue to rely on fossil fuels
Stability • Global governance has meant greater sharing of information and more political integration, so arguably the world is a safer place • More co-operation demonstrated in response to global crises and hazard events	**Conflict** • Increasing international migration can lead to social and cultural conflict and the rise of extremist political groups • Reaction to inequality and injustice and loss of cultural identity can lead to extremist ideologies, which are more easily spread by internet, etc.

Fieldwork opportunities

Assess the level of globalisation affecting your local area:

1 Visit a local car park when full and take a systematic sample of 50 cars, making a note of the manufacturer.
   a) Identify the country of origin of the named manufacturer.
   b) Identify the place of final assembly (you will have to check this online).
   c) Plot the data graphically.

2 Visit a local supermarket. Pre-select 20 grocery items, to include at least five items of fresh fruit or vegetables.
   a) Identify the countries of origin of each item.
   b) Identify the percentage of items certified as either i) fair trade or ii) organic.
   c) Plot the data as a pie chart.

Further reading

General:

www.economist.com – *The Economist*, a magazine that follows aspects of global economics and politics

www.oecd.org/ – Organisation for Economic Co-operation and Development. Covers a range of issues in global systems and governance, such as trade, growth and development

www.globalissues.org – an independent website with information on contemporary global issues and a back catalogue of useful files

Globalisation and global systems:

www.worldbank.org – World Bank. Contains useful information and data on international development

www.knomad.org/sites/default/files/2019-04/Migrationanddevelopmentbrief31.pdf – World Bank, Migration and Remittances – Recent Developments and Outlook, April 2019

Phillips, N. 'Power and inequality in the global political economy', *International Affairs*, Volume 93, Issue 2, March 2017, pp429–444

International trade:

www.wto.org – World Trade Organization

http://comtrade.un.org/pb/ – UN International Trade Statistics Yearbook (Volume 1 by country/region and Volume 2 by product)

www.unctad.org – UN Conference on Trade and Development

Poon, Prof. JPH (2019), 'Global trade patterns: how are they changing?' *Geography Review*, Volume 33, Number 2, November 2019, pp26–29

Global governance:

www.un.org – United Nations

www.unenvironment.org/resources/report/environmental-governance-update-august-2019 – UNEP report: Environmental governance update 2019

Oakes, Dr. S. 'Scales of governance: Climate policy in the USA', *Geography Review*, Volume 33, Number 2, November 2019, pp30–33

Antarctica:

www.discoveringantarctica.org.uk – educational website for The British Antarctic Survey

www.asoc.org – Antarctic and Southern Ocean Alliance

www.coolantarctica.com/ – Cool Antarctica

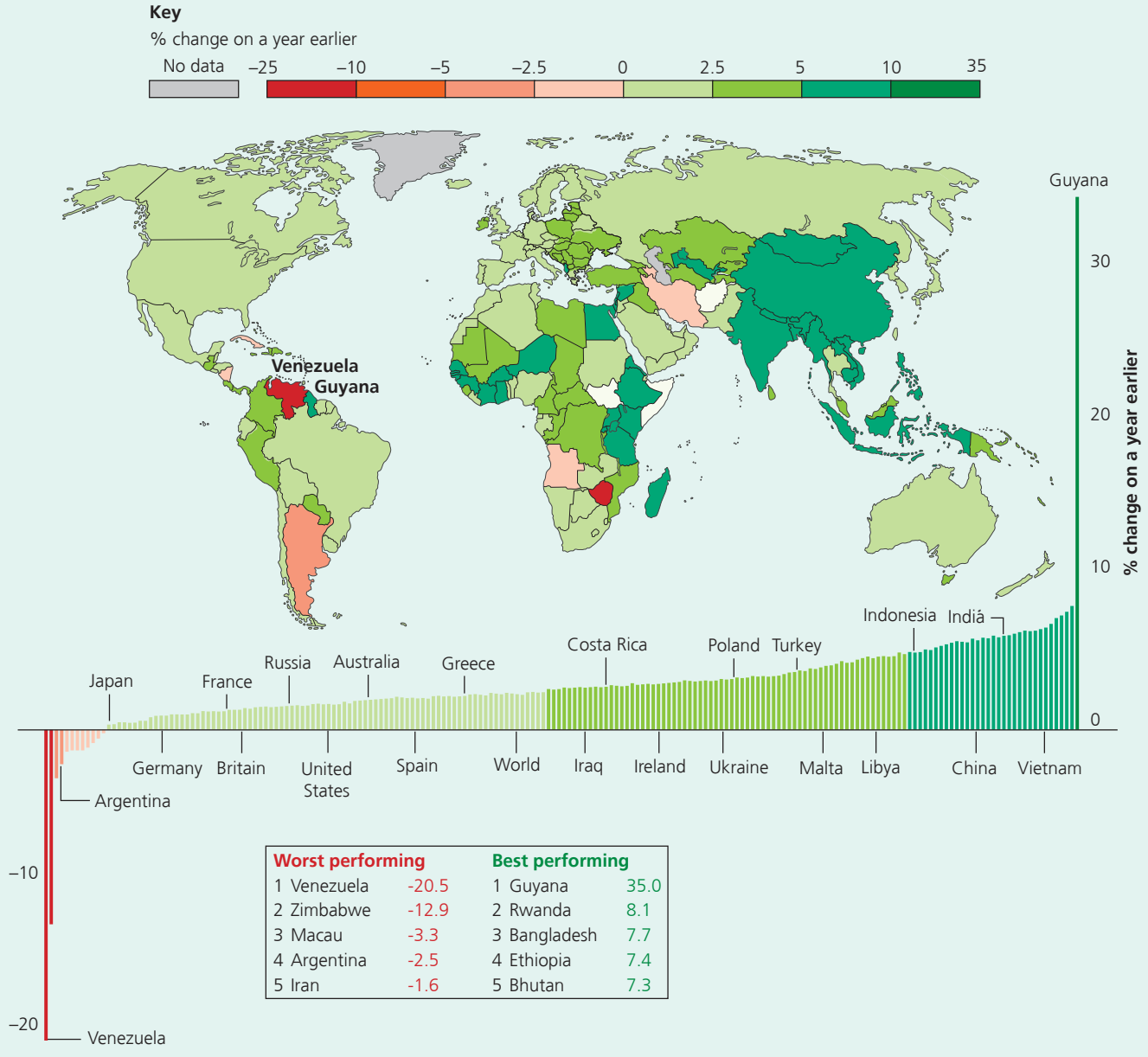
Dodds, Prof. K. 'Global governance update: The Antarctic Treaty', *Geography Review*, Volume 33, Number 4, April 2020, pp 15–17

Question practice

A-level questions

1 Explain how flows of labour have been affected by globalisation. [4 marks]

2 Analyse the data shown in Figure 7.46. [6 marks]



Worst performing		Best performing	
1 Venezuela	-20.5	1 Guyana	35.0
2 Zimbabwe	-12.9	2 Rwanda	8.1
3 Macau	-3.3	3 Bangladesh	7.7
4 Argentina	-2.5	4 Ethiopia	7.4
5 Iran	-1.6	5 Bhutan	7.3

Figure 7.46 GDP forecasts by country for 2020

Source: The Economist

3 Using Figures 7.47 and 7.48 and your own knowledge, assess the extent to which tourism poses a greater threat than whaling to Antarctica and the Southern Ocean. [6 marks]

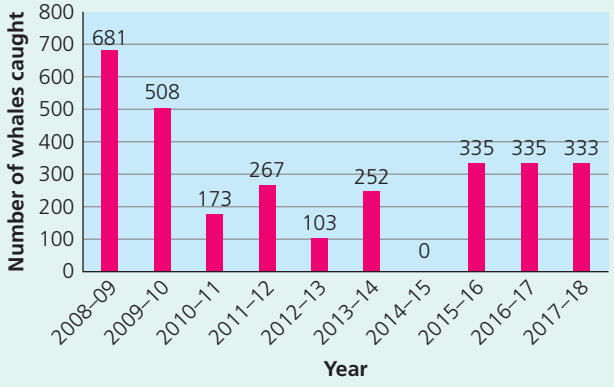


Figure 7.47 The number of whales caught in the Southern Ocean by Japan under a 'special permit for scientific research' from 2008–2009 until 2017–2018

Source: International Whaling Commission

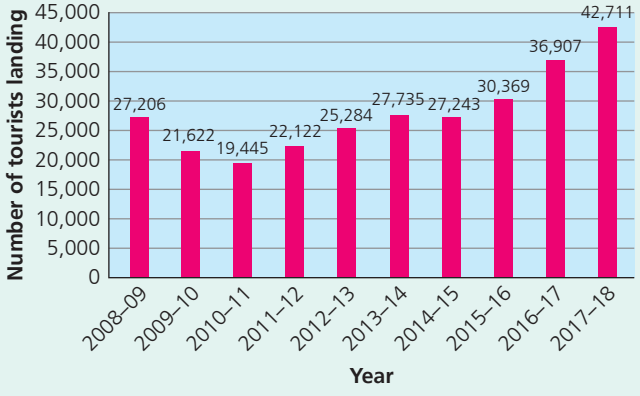


Figure 7.48 The number of tourists who landed on Antarctica from 2008–2009 until 2017–2018

Source: IAATO

4 'Removing barriers to trade and giving equal access to markets through trade agreements will help all countries to achieve sustainable economic growth and development.' Critically evaluate this statement with reference to a food commodity or manufactured product that you have studied. [20 marks]

This alternative 20-mark question requires a link between two core topics:

5 Evaluate the relative importance of global systems in contributing to the character of local places that you have studied. [20 marks]

Chapter

8 Changing places

Place is an everyday word and a key term within the study of Geography. Place can be seen as a definite location on a map or it may refer to the description of the human and physical characteristics of a particular location. Place differs to the abstract notion of space because places have **meaning** to people. Space becomes place as we get to know it better. For many, the most familiar example of place is their home as it is where they feel most attached and can be themselves. Place is more than just location, however, and academic geographers increasingly discuss the different aspects of place – the so-called multidimensionality of place. This may include **subjective** aspects such as emotional responses to place, or film, photography, art and **media** representations of place. The Nobel peace prize winner Malala Yousafzai conveys an emotional response to place when she compares life in Birmingham to that of her homeland of the Swat Valley in Pakistan:

> 'When I stand in front of my window and look out, I see tall buildings, long roads full of vehicles moving in orderly lines, neat green hedges and lawns and tidy pavements to walk on. I close my eyes and for a moment I am back in my valley – the high snow-topped mountains, green waving fields and fresh blue rivers – and my heart smiles when it looks at the people of Swat.'
>
> Prologue, *I am Malala*, 2013

Malala refers to two contrasting places above. As part of this topic you will be expected to investigate two places, one local and one distant.

As geographers, we traditionally focus on the description of place and its locality using maps and statistics but a more accurate reading and writing of a place must also include experience of place, whether through direct interaction or indirect representation.

In this chapter you will study:

- the different meanings and representations of place
- how humans perceive, engage with and form attachments to places
- the character of place and how this can change over time
- how external agencies seek to improve perceptions of place.

8.1 The nature and importance of places

The concept of place

The first place you may think of is your home, where you live and where your physical possessions and memories are gathered together. The place where you live is a particular place. However, the meaning of the term place has been hotly debated and varies according to discipline. To those in planning, place may refer to the built environment; artists and writers attempt to evoke place in their work; and to a philosopher, place may be a way of being-in-the-world. Broadly speaking, geographers refer to three aspects of place: **location**, **locale** and **sense of place** (Figure 8.1, page 352).

Location is clearly the starting point here but places take on a significance far greater than simply points on a map and this is where the other two dimensions of place begin to form the wider picture. We can recognise these three aspects in the example of the town of Glastonbury.

Location: Glastonbury is in the county of Somerset (Figure 8.2, page 352). It is located 23 miles south of Bristol. Latitude: 51.1456N and Longitude: 2.7144W. It is situated at a dry point on the low-lying Somerset Levels.

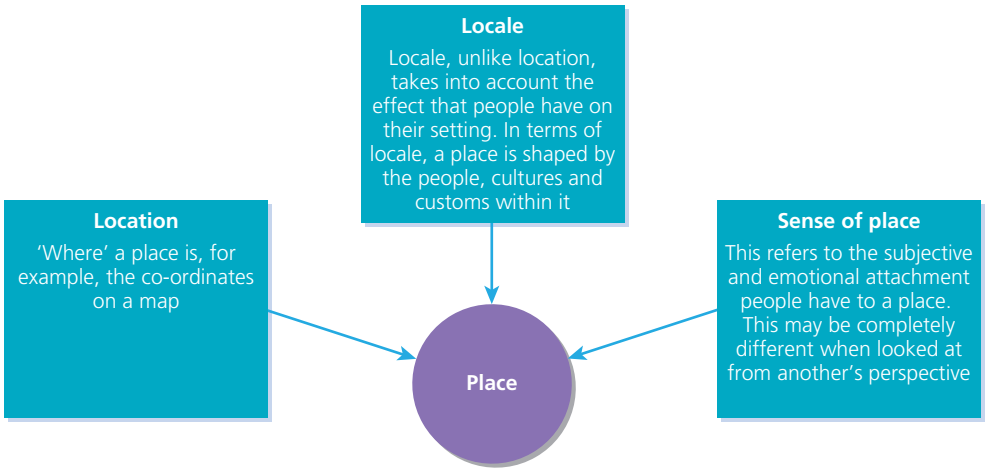


Figure 8.1 The different aspects of place

Locale: Glastonbury has its own unique character. It is home to a number of visitor attractions including Glastonbury Abbey and Glastonbury Tor (Figure 8.3). The *Visit Britain* website described it as:

> 'alive with a history that sits comfortably alongside colourful myth and legend. The Abbey was an early centre of pilgrimage and many believe King Arthur is buried amid the ruins. It is said that a young Christ also visited Glastonbury and that the Holy Grail is hidden nearby. Quirky independent shops abound selling everything from healing crystals to cakes.'

The National Trust describes Glastonbury Tor as being 'a spiritual magnet for centuries, for both Pagans and Christians. Tales have grown out of history, becoming blended and embellished leaving the truth, whatever that is, literally lost in the mists of Avalon'.

Figure 8.2 Location of Glastonbury

Sense of place: Glastonbury is a place of great spiritual importance for people interested in paganism, religious connections or the King Arthur affiliation. For many others, Glastonbury evokes emotions about the internationally famous music festival which takes place most years at Worthy Farm in Pilton on the edge of Glastonbury (Figure 8.4).

Figure 8.3 Glastonbury Tor is considered a spiritual site. Excavations at the top of the Tor have revealed the plans of two superimposed churches of St Michael, of which only a fifteenth-century tower remains

Figure 8.4 The first Glastonbury Festival took place in June 1970. It was attended by 1,500 people and cost £1 entry with free milk from the farm

Key terms

Locale – This is the place where something happens or is set, or that has particular events associated with it.

Location – 'Where' a place is, for example, the co-ordinates on a map.

Meaning – Meaning relates to individual or collective perceptions of place.

Media – Means of communication including television, film, photography, art, newspapers, books, songs, etc. These reach or influence people widely.

Place – Defined as a location with meaning. Places can be meaningful to individuals in ways that are personal or subjective. Places can also be meaningful at a social or cultural level and these meanings may be shared by different groups of people.

Placemaking – The deliberate shaping of an environment to facilitate social interaction and improve a community's quality of life.

Sense of place – This refers to the subjective and emotional attachment people have to a place. People develop a 'sense of place' through experience and knowledge of a particular area.

Subjective – Based on or influenced by personal feelings, tastes or opinions.

Theoretical approaches to place

The nature and meaning of place has long been debated in geography but there are three main approaches into which the study of place has been divided.

1. A descriptive approach

This is the idea that the world is a set of places and each place can be studied and is distinct.

2. A social constructionist approach

This approach sees place as a product of a particular set of social processes occurring at a particular time. For example, Trafalgar Square was built to commemorate a British naval victory in the 1800s and, using a social constructionist approach, could be understood as a place of empire and colonialism.

3. A phenomenological approach

This approach is not interested in the unique characteristics of a place or why it was constructed. Instead it is interested in how an individual person experiences place, recognising a highly personal relationship between place and person. The work of geographers Yi-Fu Tuan and Edward Relph has been particularly important here. Tuan developed the term 'topophilia' to describe the affective bond between people and place and argued that it is through human perception and experience that we get to know places. Relph argues that the degree of attachment, involvement and concern that a person or group has for a particular place is critical in our understanding of place.

There is clearly some overlap between these three approaches and it would be inaccurate to say that any one approach is more important than another. For A-level Geography it is important that you consider all three approaches. Places are not simply geographical locations but should be understood as dynamic areas where people, ideas and information come together. Furthermore, places hold meaning: different groups of people may see places differently and have different scope to change these places.

Trafalgar Square, London

The cultural geographer Jon Anderson argues that places can be defined and given meaning by the traces that exist within them:

- Material traces are physical additions to the environment and include things such as buildings, signs and statues.
- Non-material traces include events, performances or emotions which occur in that place.

In Trafalgar Square the immediate traces are the statue of Admiral Lord Nelson, the column on which this stands, a large public square and two fountains. However, you can look behind these traces to try and understand the meanings behind them and the aims of the 'trace-makers' who constructed them.

Figure 8.5 Trafalgar Square is a popular London tourist attraction

The statue of Admiral Lord Nelson is at the top of a column that bears his name in a square commemorating Nelson's naval victory over combined French and Spanish fleets in 1805. It has been suggested that the city planners of the 1800s wanted to commemorate British leadership and victory, and this place can be understood as a space of Empire. Furthermore, it sought to inspire pride and patriotism in the country. Trafalgar Square is still used to celebrate victory and is often the focus for national celebrations such as sporting victories and London's successful Olympics bid.

Traces change over time and this is evidenced by the Fourth Plinth Programme in Trafalgar Square. It was initially intended to hold an equestrian statue of William IV, but had remained vacant due to lack of funding. Since 1999 however, it has become a centrepiece for some of the world's most provocative contemporary public art in a bid to showcase London as the cultural and artistic capital of the world. Commissions have included a marble torso of a pregnant artist, a large blue cockerel and the controversial Antony Gormley plinth entitled 'One and other' which involved 2,400 selected members of the public spending one hour each on the plinth doing whatever they wanted. Art installations like this reflect the notion that although London is a place of tradition, it is also not afraid to take risks and embrace innovation and diversity.

Trafalgar Square has sometimes been used for protest, when people have rallied against the British state rather than supporting and celebrating it. Climate change activists have held many demonstrations in recent years and the traces left by the protestors made Trafalgar Square into a very different place at that time.

The importance of place in human life and experience

It is not hard to find examples of the ways in which place is important:

- People define themselves through a sense of place and by living in places and carrying out a range of everyday practices there. A person–place relationship is developed.
- The promotion of place is crucial in the marketing of holiday destinations.
- Food items are increasingly marketed in terms of the place from which they came and the popularity of particular events may be linked to the reputation of the place at which they happen. The Glastonbury Music Festival would be an example here.
- People may 'buy into' or 'consume' place. For example, those who like the countryside tend to holiday in rural locations, enjoy books and television programmes about these areas, spend money on walking gear and maps and even furnish their houses in a rustic country style. Numerous products are marketed so that people can buy into the notion of the rural idyll.

People's **lived experience** of place can be explored by looking at the impact of place on:

- identity
- belonging
- well-being.

The placemaking movement, which has expanded rapidly in recent years, places great emphasis on all three aspects.

Identity

Place can be critical to the construction of identity. Probably first and foremost is the sense of place developed in relation to our own home and local geographical area. Reading local newspapers, playing sport for a local team or attending a local fayre or event all foster a sense of local place.

Identity can be evident at a number of scales:

- Localism: An affection for or emotional ownership of a particular place. Localism rarely manifests itself in a political sense but can be demonstrated in 'nimbyism' (not in my backyard) which occurs when people are reluctant to have their local area affected by development.
- Regionalism: Consciousness of, and loyalty to, a distinct region with a population that shares similarities.
- Nationalism: Loyalty and devotion to a nation, which creates a sense of national consciousness. Patriotism could be considered as an example of a sense of place.

Historically, people have identified more with their local place or community because they have greater knowledge of this area and people. In some areas of the UK, this has led to calls for more regional government. In Cornwall, the Mebyon Kernow party has been leading the campaign for a National Assembly

for Cornwall. The party believes that the county of Cornwall, with its own distinct identity, language and heritage, has the same right to self-rule as other parts of the UK, such as Scotland and Wales which have already achieved a degree of devolution.

At a national level many people identify with place through:

- a common language
- a national anthem
- a flag
- cultural and sporting events.

A resurgence in the Welsh language and culture has highlighted a stronger national identity among the Welsh in recent years.

Religion

Religion, too, can be used to foster a sense of identity in place. At a local level, churches, mosques and synagogues are places where people from the same religious identity come together to worship. There may also be larger sacred places such as Bethlehem or Mecca where people go on pilgrimages.

Protests

The power of place in political protest is a feature of the early twenty-first century in reaction to unpopular political regimes and problems associated with capitalism, racial inequality and climate change. In 2011, Tahrir Square in Cairo was the focal point of the Egyptian Revolution against former president Hosni Mubarak. Despite the effective banning of public protests in Egypt, the square became a symbol for pursuit of democracy in Egypt (Figure 8.6).

Figure 8.6 Redevelopment of Tahrir Square, 2020

In London, from 2018 onwards, climate change activists, led by Extinction Rebellion, held peaceful protests centred around Parliament Square. Similarly recognisable sites were chosen around the world as the activists used the power of place to attract attention and lodge their message in people's memories.

A global sense of place

The economic and social geographer Doreen Massey wrote about a global sense of place, in which she questioned the idea that places are static. She argued instead that places are dynamic, they have multiple identities and they do not have to have boundaries.

She used her own local area to illustrate that place is influenced by constantly changing elements of a wider world.

> Take, for instance, a walk down Kilburn High Road, my local shopping centre. It is a pretty ordinary place, north-west of the centre of London ... Thread your way through the ... traffic diagonally across the road from the newsstand and there's a shop which as long as I can remember has displayed saris in the window. Four life-sized models of Indian women and reams of cloth. On the door a notice announces a forthcoming concert at Wembley Arena: Anand Miland presents Rekha, live, wih Aamir Khan, Jahi Chawla and Ravenna Tandon. On another ad, for the end of the month, is written, 'All Hindus are cordially invited.' In another newsagents I chat with the man who keeps it, a Muslim unutterably depressed by events in the Gulf.
>
> *Massey,* 1994

Massey argued that the character of a place can only be seen and understood by linking that place to places beyond. She concluded, 'What we need, it seems to me, is a global sense of the local, a global sense of place.'

Links

There are links to Chapter 7, Global systems and global governance.

Globalisation of place

Some argue that globalisation has made place less important as the forces of global capitalism have eroded local cultures and produced identical or **homogenised** places (Figure 8.7, page 356). This can be seen through

the increased presence of global chains such as Starbucks in high streets all over the world. The American novelist James Kunstler has talked of a geography of nowhere – where processes such as urban sprawl have led to community-less cities covering huge areas of countryside with identical shopping malls, car parks and roads. He argues that 'every place is like no place in particular'. In the UK, the term **'clone town'** has been used to describe settlements where the high street is dominated by chain stores. The term **placelessness** has also been used to describe such places.

Figure 8.7 Airports, hotels and high street shops can look the same wherever they are and can lead to the loss of place identity. This Hilton hotel is in Quito, Ecuador but how different does it look to the ones in New York, Beijing and London?

Some local places and cultures are resisting the power of globalisation, as shown by the anti-Costa campaign in Totnes, Devon in 2012 (see text box). Multinational companies are also increasingly having to adapt to the local marketplace. This is known as **glocalisation.** One example of this is the McDonald's franchise. There are currently more than 38,000 McDonald's restaurants in over 100 countries around the world, but the aim to increase profits has led to the company adapting its brand and product to the local market place. In Hindu countries, for example, beef has been removed from the menu while in Muslim countries, pork has been removed. The number of McCafes has increased in countries with a coffee culture.

'The independent coffee republic of Totnes' (so called in a BBC article)

Figure 8.8 Clonestopping poster displayed in Totnes

In 2012, the coffee chain Costa set about trying to open an outlet in the South Devon town of Totnes. Within weeks of the proposal, three-quarters of the town's population had signed a petition saying that they supported the independent high street and would boycott any coffee shop chain that came to Totnes. The opposition was not anti-capitalist, rather a community fiercely proud of its independently-owned outlets and one eager to prevent Totnes becoming a 'clone town'. After an eight-month battle, Costa dropped their plans, announcing: 'Costa has recognised the strength of feeling in Totnes against national brands and taken into account the specific circumstances of Totnes.' However, in 2019 a coffee shop from a different national brand did in fact open a branch in Totnes. Either the community's opposition to such brands has lessened or there is some inevitability in becoming a clone town.

Localisation of place

Place has become a political symbol for people fighting against global capitalism. One particular response has been a greater focus on 'local' place and the promotion of local goods and services.

Some places have introduced a local currency to encourage people to shop locally and keep their money in the local economy. Totnes in South Devon led the way in 2007, but the local currency ended in 2019. The Bristol Pound was launched in 2012. In 2020, Bristol was looking for ways of creating an e-wallet, a local version of GooglePay or ApplePay, and ways of adapting

to a cashless economy (Figure 8.9). Even before the impact of COVID-19 these schemes had experienced a number of issues and the online Bristol Pay was still being tested. COVID-19 restrictions will have further limited its development and success.

Figure 8.9 The Bristol Pound was launched in 2012 to encourage stronger community connections and keep money in the local economy

Belonging

In the context of place, to belong means to be part of the community. Belonging is increasingly seen as one of the key factors that makes a place sustainable and successful. Regeneration schemes now often focus as much on the social environment as on the built environment and the work of different agencies and individuals can have a positive impact on people's lived experience of the place in which they live. The Transition Town movement, detailed on page 358, is an initiative that places great emphasis on community involvement and has been successful in developing a clearer sense of belonging in places all over the world.

The extent to which one might feel a sense of belonging to a place can be influenced by a number of factors, such as age, gender, sexuality, socio-economic status, religion and level of education.

Race and ethnicity can also be linked to the feeling of belonging. Race is based on biological and physical characteristics such as skin colour. Ethnicity is a broader concept which can be defined as belonging to a social group that has a common national or cultural tradition such as language and ancestry.

As globalisation and migration have increased, many places, particularly cities, have become much more ethnically and culturally diverse. London is often referred to as one of the world's most multicultural cities and hosts an enormous number and variety of festivals which demonstrate this. It has the largest Chinese New Year festival outside China and other celebrations include Vaisakhi (Sikh New Year), Diwali (the Hindu Festival of Lights) and the Shubbak Festival of Arab culture. In spite of its multicultural status, there are still minority ethnic clusters in parts of London, including China Town in Soho and Banglatown in and around Brick Lane. These have tended to develop, with dedicated shops and services, for reasons of mutual support and cultural preservation.

Well-being

In 2020, the *Sunday Times* judges deemed Altrincham in Greater Manchester as the best place to live in Britain due to its revitalised town centre, schools, proximity to Dunham Massey (a large National Trust property and park) and the town's initial response to the COVID-19 pandemic. Regional winners included places with different attributes:

- Narberth in south Wales with its warm welcome to visitors and 'high-class high street'
- Ballycastle in Northern Ireland with its quality local produce and 'magical market'
- Leith in Edinburgh, Scotland with its 'repurposed industrial buildings' and 'trendy' restaurants
- Victoria Park Village in London described as 'wonderfully hip but sprinkled with history'.

Individuals will have different views about what makes a place great. Is it likely that all the residents of the places above agree with the *Sunday Times* survey? There are certain features which are generally accepted to be more important in promoting happiness and **well-being** within a place, as shown in Figure 8.10 (page 358), but different factors will be important to different groups of people, as the quality of lived experience and perception of place are influenced by a sense of belonging, age, gender, socio-economic status, religion, level of education and sexuality.

The Transition Town movement

Transition Network was founded in 2007, as a response to the twin threats of climate change and peak oil (the point in time when the maximum rate of extraction of oil is reached, after which the rate of production is expected to enter terminal decline). Since then, it can be said to have gone on to tackle many other issues including some associated with globalisation, such as dilution of place identity and loss of community and economy. One of the founders, Rob Hopkins, says:

'It's about what you can create with the help of the people who live in your street, reimagining and rebuilding your neighbourhood, your town. If enough people do it, it can lead to real impact, to real jobs and real transformation of the places we live, and beyond.'

There are now over 1,000 Transition initiatives worldwide and increasing numbers of places are embracing a life where communities come together to share skills, grow food, provide care for dependants and fight inequality.

Learn more about the Transition Network at www.transitionnetwork.org/

Insider and outsider perspectives on place

People have a stronger relationship with the places they are familiar with. 'To be inside a place is to belong to it and identify with it, and the more profoundly inside you are the stronger is the identity with the place' (Relph, 1976). This is the main reason why people living within a place are more likely to oppose developments within their local area than those from outside. The term Nimby is an acronym for the phrase 'not in my backyard' and can be applied to local opposition to developments such as new housing estates or wind farms and fracking proposals.

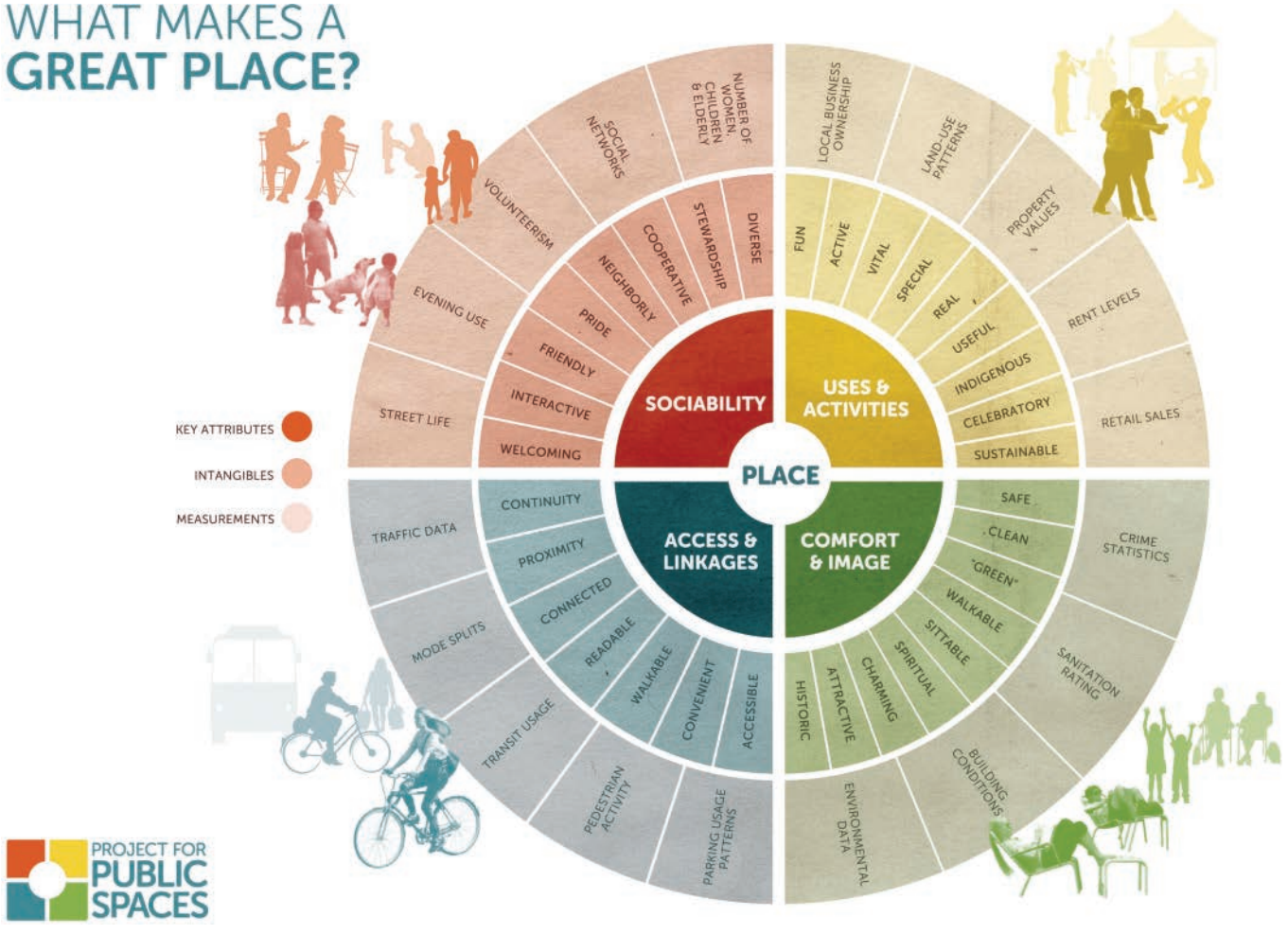


Figure 8.10 What makes a great place? The Place diagram is one of the tools PPS has developed to help communities evaluate places. The inner ring represents a place's key attributes, the middle ring its tangible qualities, and the outer ring its measurable data

People, activities and events can be seen as 'in place' or 'out of place'. For example, a business person may feel 'in place' in Canary Wharf in London, while a homeless person may not. If certain people feel 'out of place', why is this, and who has created this feeling? As immigration has become a significant focus of the politics of the UK and other countries, there is a greater need to understand the notion of place in the creation of insiders and outsiders.

'A place for everything and everything in its place' is a well-known phrase that suggests there is a particular ordering of things in the world. Cultural geographer Tim Cresswell argues that people, things and practices are often strongly linked to particular places and when these links are broken, or when people have acted out of place, they are deemed to have committed something of a crime. Graffiti on historic buildings or litter in an area of outstanding beauty are examples of this but increasingly we are seeing groups of people being treated as outsiders.

- In the past, and often still today, this has included travellers, protestors and lesbian, gay, bisexual and trans people.
- There were often stereotyped views of the kinds of places certain groups of people were associated with.
- Gender roles were often associated with specific places, for example, the view that 'a woman's place is in the home' was the stereotypical societal view held by many in the early to mid-twentieth century and this affected the type of places in which women felt comfortable.

A person's gender, race, ethnicity, religion, politics, socio-economic status and sexuality may affect their perception of place.

People may have a strong attachment to a place and feel like they belong there, because it is significant to their lives. For example, Mecca in Saudi Arabia is the most holy of religious places for Muslims, but for non-Muslims it is viewed more as an historical and cultural place.

People may also have a sense of belonging and positive perception of a place because of positive experiences there. Negative experiences of place create a negative perception and limited attachment. For example, the expected behaviours and norms associated with expensive or exclusive hotels or restaurants may make people from disadvantaged backgrounds feel alienated. These feelings may be reinforced with exclusionary rules such as dress-codes, so reinforcing negative perception. However, for those familiar with the routines, rituals and traditions of a place they will feel welcomed and thoroughly in-place.

Experiences of place change as we get older. Would you feel the same excitement going to the local playground now as you did when you were younger?

Migrants are often referred to as 'out of place'. Deaths of migrants in the Mediterranean have highlighted the plight of hundreds of thousands of people trying to escape conflict and reach a 'better life' in Europe, but as people-without-place, such refugees present a crisis to state power. In the United Kingdom, media reports and anti-migration groups have used metaphors associated with water, blood and disease to describe the influx of refugees. Such negative terminology has invariably meant that the presence of such migrants has met with great resistance and calls from some to protect 'our place' and 'our culture' against people who do not 'belong' here.

Categories of place

Near places and far places

The terms 'near' and 'far' have several potential meanings when it comes to place. They could refer to the geographical distance between places. Equally, they could describe the emotional connection with a particular place and how comfortable a person feels within that place. Some places feel more familiar than others partly due to personal experience, but also because of frequent representational exposure.

The key point is that geographically 'near' places do not automatically foster identities of familiarity and belonging and that in these days of globalised culture, travel and media, far-off places are not automatically strange, uncomfortable and different.

Experienced places and media places

Experienced places are those places that a person has spent time in, whereas **media places** are those that the person has only read about or seen on film. The 'reality' of a place can be far different to that put across by the media and this is most clearly seen through the portrayal of rural places.

Media portrayal of rural living

For a long time, the countryside in the UK has been 'idyll-ised' and countryside living has been stereotyped as involving a happy, healthy and close-knit community experiencing few of the problems of urban life.

Geographer Paul Cloke has looked extensively at rurality and argues that magazines such as *Country Life*, television series such as *Doc Martin* and children's books such as *Postman Pat* seek to reinforce these images by focusing on more nostalgic images of the countryside.

Stereotypical images of rural living permeate but the idyllic image of the countryside put forward by the media and advertising companies hides a host of problems, which have all sought to disadvantage low-income households in rural areas:

- unemployment and underemployment
- the scarce availability of affordable housing
- the reduction in public transport services
- rural homelessness.

Media portrayal of city living

In contrast, cities are often stereotyped in a negative way. Economic and social deprivation, homelessness, crime, vandalism and pollution are just some of the images routinely ascribed to British cities by the media. It is true that some of these problems are worse in urban areas, but it is wrong to assume that all cities are the same. Successful regeneration of urban areas has made city-living far more attractive in recent decades.

Factors contributing to the character of places

Endogenous factors

The **character** of a place refers to the physical and human features that help to distinguish it from another place. This character may be linked strongly to the natural environment but it is more often a combination of natural and cultural features in the landscape, and generally includes the people who occupy the place. These characteristics are known as **endogenous factors**.

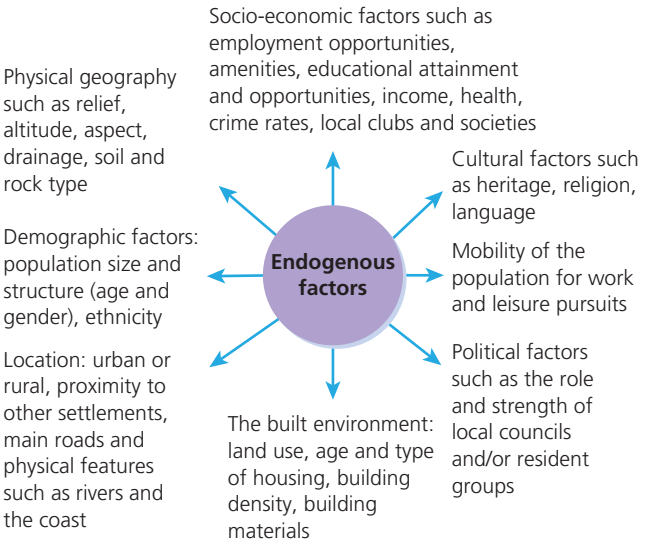


Figure 8.11 Endogenous factors affecting the character of places

Endogenous factors are those that originate internally and may include aspects of the site or land on which the place is built, such as height, relief, drainage, soil type, geology and the availability of resources. They also include the demographic and economic characteristics of the area as well as aspects of the built environment and **infrastructure**.

Key terms

Diaspora – A group of people with a similar heritage or homeland who have settled elsewhere in the world.

Endogenous factors – In the context of place, this refers to the characteristics of the place itself or factors which have originated internally. This would include aspects such as location, physical geography, land use and social and economic characteristics such as population size and employment rates.

Exogenous factors – This refers to the relationship of one place with other places and the external factors which affect this. The demographic, socio-economic and cultural characteristics of a place are shaped by shifting flows of people, resources, money and investment.

Infrastructure – Infrastructure relates to the services considered essential to enable or enhance living conditions. These primarily consist of transport (roads, railways, canals, and/or airports), communications infrastructure (broadband and phone networks) and services such as water supply, sewers and electrical grids. They may also include infrastructure such as the local education system, healthcare provision, local government and law enforcement, as well as emergency services.

As Figure 8.11 shows, there is clearly a wide range of factors which affect the character of places.

Some places may have an industrial story; others may have developed as agricultural places or tourist resorts. When places first started to grow, most had only one distinct **function**. This might be a good defensive position, a bridging point, availability of natural resources, natural routeways or trading centres.

Exogenous factors

As places developed, exogenous, external factors became more important and the importance of the initial functions diminished as technological advances occurred. **Exogenous factors** are those that have an external cause or origin.

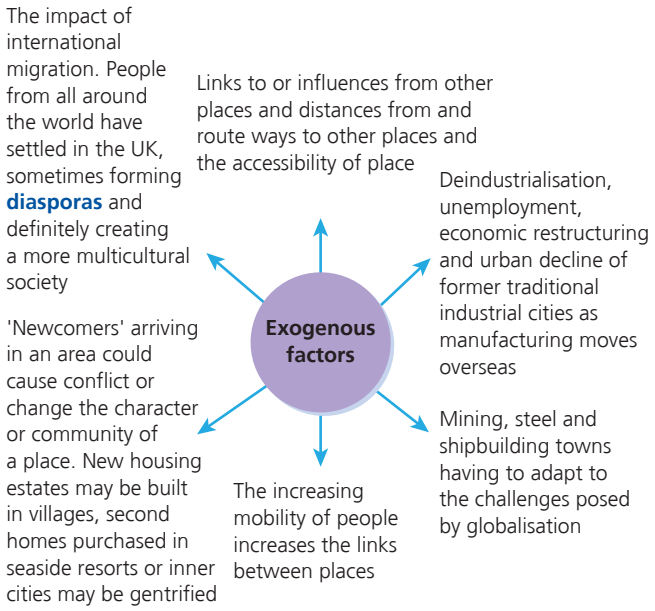


Figure 8.12 Exogenous factors affecting the character of places

Links

There are links between exogenous factors and Chapter 7, Global systems and global governance.

Review questions

1 What factors make your local place important to you?
2 Outline why different groups of people may feel like 'insiders' or 'outsiders' in a place.
3 Explain how endogenous and exogenous factors have affected the character of your local place.

8.2 Changing places: relationships and connections

Demographic change and cultural characteristics

Places change over time. They are dynamic and are shaped by shifting flows of:

- **people**
- **resources**
- **money and investment**
- **ideas**.

Consider, for example, the economic rise and fall of British industrial cities in the nineteenth and twentieth century and the impact this had on their population and environment. British seaside resorts have also undergone significant change in the last sixty years as they have had to adapt to increasing numbers of British people travelling abroad for their holidays.

Economic change and social inequalities

Economic factors tend to have the biggest impact on the character of place but changes may also result from:

- migration and conflict (Damascus, Syria)
- terrorism (New York City, 9/11)
- industrial accidents (Chernobyl)
- natural disasters (New Orleans)
- climate change (the Maldives).

Factors like these affect the way places are perceived and go on to affect the lives of the people living and working there.

External forces can also bring about change, as shown in Table 8.1 (page 362). This can be both positive and negative.

Conflict may arise when people resist changes forced upon their place. The redevelopment of areas of East London for the 2012 Olympic Games was not welcomed by everybody and this can be seen with other redevelopment projects currently planned or taking place in London. Proposed housing estates, landfill sites, wind farms and bypasses can all create tensions between different stakeholders. Think about how your local place has been affected by such developments. Why do people often resist change to their place and how can such conflicts be resolved?

Table 8.1 Some impacts of external forces on place

Agent of change	Example	Impact
Government policies	Regeneration schemes and financial incentives for industries such as subsidies, tax breaks and enterprise zones	These can attract businesses to places and stimulate a positive multiplier effect
The decisions of multinational corporations	In 2019 Honda announced the closure of its Swindon factory to move production to China, Japan and the USA In 2019, Tata Steel announced the closure of its site in Newport, South Wales due to competition from Chinese producers	Job losses for employees and in the supply chain Uncertain future for the 380-acre factory site Job losses for Newport; steel production has been a major characteristic of the town since the 19th century
The impacts of international or global institutions (for example, IMF, World Bank, UN, WHO)	In 2020, the World Bank was running 24 development projects in Haiti Sustainable Development Goals	Post-earthquake and hurricane reconstruction and reaction to COVID-19 Varied levels of success around the world

Examples of continuity and change at a variety of scales

Places can be investigated at a range of scales. In 2012, the BBC in collaboration with the Open University researched the past and present conditions of six London streets (*The Secret History of our Streets*). They used the poverty maps of Charles Booth (c.1886) as a base and investigated the changing social and economic conditions of the residents who had lived there. Portland Road, Notting Hill, considered in 2012, 'the most gentrified street in the UK' and home to some of London's wealthier residents, was in 1899 the worst slum in London. In contrast, Deptford High Street has gone from being the 'Oxford Street of South London' to 'one of the poorest shopping streets in London, marooned amid 1970s housing blocks.'

The following three examples look at changing places on a progressively larger scale. They illustrate how the characteristics of places can be shaped by a very different range of factors, including people, resources, money and investment. They also serve to show how past and present connections within and beyond localities can help shape places and the lives of the people who live there.

Port Sunlight, the Wirral

Port Sunlight is an example of a place shaped almost entirely by the beliefs and ideals of one man, William Hesketh Lever. It was built as a garden village from 1888, on the western bank of the River Mersey, four miles south of Birkenhead on the Wirral. Lever acquired the site to build a new factory for his expanding soap business, with an adjoining model village for his workers. Lever championed the business model of 'prosperity-sharing' and instead of sharing profits directly with his employees he provided them with high-quality, sanitary and spacious housing, with social amenities and welfare provision in picturesque surroundings. Heavily influenced by the Arts and Crafts Movement, over 30 architects were employed to ensure each terrace of houses had a unique design (Figure 8.13).

Figure 8.13 Terrace of workers' houses, Port Sunlight

The open spaces between the houses originally had community allotments (that later became beautiful gardens). Lever also provided uniquely designed public buildings including a church, hospital, post office, fire station, schools, social clubs, dining halls and the now nationally renowned neo-classical Lady Lever Art Gallery. The plan for the village was influenced by the 'garden city' movement and large amounts of open space for parks and recreation were included.

As a company, Lever Brothers became Unilever and over the last 100 years has grown into a global giant producing a huge range of consumer goods. In complete contrast the 'place' of Port Sunlight has changed very little. Although no longer primarily the home of Unilever employees, great efforts are made to maintain the original physical appearance and community feel of the place.

In 1966 almost 1,000 houses and most of the public buildings became Grade II listed and in 1978 the village was designated a Conservation Area. Unilever still has an interest in maintaining the character of Port Sunlight, and in 1999 they set up the Port Sunlight Village Trust (PSVT). Whilst Ellesmere Port and Liverpool, across the River Mersey, have experienced huge amounts of change from deindustrialisation to regeneration, the PSVT and the residents have gone to great efforts to maintain a self-sustaining village that is still true to the original ideals of William Lever.

Devonport, Plymouth

Devonport in Plymouth has changed considerably in the last sixty years, largely as a result of external factors. It was originally established as a naval dockyard town due to its location on the sheltered, natural deep-water harbour known as the Plymouth Sound. By the eighteenth century, it had become one of the fastest growing towns in the country and enjoyed great prosperity.

However, Devonport's fortunes have fluctuated with those of the dockyard. In 1952, the Navy requisitioned Devonport town centre as a storage enclave and enclosed it with a three metre high wall. The community was effectively split in two and the displaced residents were rehoused in flats and apartments ill-suited to family living. Naval jobs continued to decline and this led to significant social and economic problems within Devonport in the 1980s and 1990s.

From 2001–11, Devonport benefited from the New Deal for Communities initiative which provided funding to improve some of the most deprived areas of the UK. The scheme specified place-related outcomes such as addressing crime, community and housing, and people-related outcomes such as education, health and employment. Community groups such as the Pembroke Street Estate Management Board were heavily involved. The physical environment of Devonport has changed considerably. The dividing naval wall was removed, inter-war housing and flats demolished and historical landmarks incorporated into the redevelopment of the area. The regeneration has also tried to attract a wider range of people (Figure 8.14) with options for private or shared ownership and the availability of social housing.

Figure 8.14 The construction of Georgian-inspired homes in the 'Village by the sea' development has attracted people from higher socio-economic groups to Devonport

There are still pockets of deprivation within Devonport, but both quantitative and **qualitative data** show that the New Deal programme and other redevelopment projects have significantly changed and improved the area, with benefits to local residents.

A major redevelopment of Devonport's docks was announced in 2020 to create 600 jobs and future-proof the economy by allowing it to refit the latest nuclear submarines.

Medellin, Colombia

The city of Medellin in Colombia, nicknamed 'The City of Eternal Spring', and the second largest city in Colombia, was for a long period associated with drugs and violence and dubbed 'the most dangerous city in the world.' The notorious drug lord Pablo Escobar wielded enormous power in the city until his death in 1993. Unemployment, crime and poverty were widespread and this in turn created social inequality.

Today, Medellin, with a population of 2.4 million, has become a model for urban regeneration and sustainable city planning through long-term investments in infrastructure and education. As part of a plan to re-brand Medellin, city planners recognised the need to make the city equally accessible to all its citizens, rich and poor. Through this planning philosophy, the city's long-divided social classes are now more able to integrate in everyday economic and educational activities. The city's poorest, many of whom reside in shanty houses in the Aburra Valley, can now access the city's booming economic centre courtesy of a series of outdoor escalators and a gondola system that carries people up and down the valley (Figure 8.15, page 364). Additional innovations include:

- a bus rapid transit system named Metroplus, with dedicated bus lanes
- an extensive above-ground tram system
- a city-wide ride-sharing program.

Emission-free transport has been promoted, and this has been helped by the EnCicla initiative, a free bike-sharing programme that offers an integrated alternative to the city's public and mass transportation systems.

Education, social programmes and the public arts and culture budgets have all been increased to transform the lives of the most underprivileged residents in this city. There are still problems in Medellin. Poverty rates have fallen but inequality between rich and poor has increased and cultural and geographic barriers continue to limit social integration. Crime rates and gang violence remain high. However, change takes time and there is a great feeling of optimism within the city.

Figure 8.15 Medellin's gondola system provides access to the city's booming economic centre

Review questions

4 Summarise how places are shaped by a number of 'shifting flows'.

5 Explain how government policies have affected the character of a place you have studied.

6 With reference to your local place, outline the nature of continuity and change for that place.

8.3 Changing places: meaning and representation

The notion that places change because of the constant movement of people, ideas, wealth and information within them has already been discussed. A further dimension relates to the meaning and representation of place:

- **Meaning** relates to individual or collective perceptions of place.
- **Representation** is how a place is portrayed or 'seen' in society.

Both of these may change over time. They also vary between people and communities as shown by the example of Belfast (see text box opposite) and the place study of Brick Lane and Spitalfields (page 379). In the latter, it is clear that different groups of people attach different values and meanings to the area and hold different views on its present and future development.

This part of the chapter will explore how humans perceive, engage with and form attachments to places and how they present and represent the world to others.

Key terms

Agents of change – These are the people who impact on a place whether through living, working or trying to improve that place. Examples would include residents, community groups, corporate entities, central and local government and the media.

Meaning – Meaning relates to individual or collective perceptions of place.

Perception of place – This is the way in which place is viewed or regarded by people. This can be influenced by media representation or personal experience.

Representation – Representation is how a place is portrayed or 'seen' in society.

Meanings of place

Sense of place

Sense of place is the meaning attributed to a place through our interaction with it. We experience a place by living, working or visiting it. 'Real-life' experiences and memories develop our sense of place. Think about places that were important to you as a child. A favourite park, for example? What memories are attached to these places? What significance do they hold for you? For children, emotional attachment to place is the most important factor shaping place meaning, and this is why 'home' is often their most important place.

Belfast: one place, two representations?

Belfast is the capital city of Northern Ireland and home to 340,000 inhabitants. It is the region's economic powerhouse, an industrial city where the *Titanic* was built and which was once the largest producer of linen in the world. The city suffered greatly during the 30-year period of conflict called 'the Troubles', but has undergone a sustained period of calm and substantial economic and commercial growth since the Good Friday Peace Agreement was signed in 1998.

The city centre has seen large-scale redevelopment and different parts of the city have been rebranded as 'quarters', emphasising their unique history and culture. The Titanic Quarter is one of the largest brownfield redevelopment sites in Europe and includes the Titanic Studios and more than 100 companies, including Citi, Microsoft and HBO. This is the creative and cultural Belfast promoted by the City Council and Tourist Board, looking outwards to build a cosmopolitan city which is open and welcoming.

Figure 8.17 The old: Peace lines or walls were constructed in Belfast to keep neighbouring Protestants and Catholics apart. Many of these walls feature murals, which tourists come to see. In some cases, tourists write their own messages of peace and hope. It is hoped that the walls will be taken down as community relations strengthen

The other representation links to the complex political history of Belfast, the high degree of religious segregation and the image of 'bombs, bullets and balaclavas' which dominated newspaper front pages in the late twentieth century. Peace walls and political parades are an important part of the history of Belfast but some argue that they reflect a city struggling to move beyond the disputes and arguments which have shaped its past.

Figure 8.16 The new: The Titanic Quarter is home to the Titanic Belfast Centre, visited by more than four million people from 145 countries, and the set of the world's most successful television series, *Game of Thrones*, filmed in a former shipyard paint hall

Tourism is now a major part of the Belfast economy. In 2018, visitors spent 5.2 million nights in Belfast and contributed £400 million to the local economy. The evidence suggests that Belfast has successfully managed to change its international image and visitors from all over the world are attracted by the city's culture and vibrancy rather than being put off by events that have gone before.

To find out more about Belfast and the changes which are happening there, go to: www.belfastcity.gov.uk or www.belfasttelegraph.co.uk/ There are useful pages detailing ethnic diversity and segregation in Belfast at: www.geographyinaction.co.uk/

Research suggests that developing a sense of place is important.

- Connecting with one's surrounding environment establishes knowledge of and appreciation for its resources.
- Sense of place supports the development of one's identity.
- A strong sense of place can inspire stewardship and understanding, and nurture empathy.

Is our own experience and enjoyment of place critical to our understanding of place?

Perception of place

Perception of place is the meaning attributed to a place that is developed through what we have heard, seen or read about it.

For example, Dartmoor National Park in Devon is closely linked with ideas of nature and wilderness in different artistic and literary works, including, Arthur Conan Doyle's *The Hound of the Baskervilles* (1902). Many popular images of Dartmoor continue to associate the moor with wilderness, but these hide important human activities and conflicts, including mining, farming, quarrying and military training.

Our perception of place may be shaped by a broad range of sources. Some of these representations of place include:

- advertisements or tourist agency material
- local exhibitions of art, film and photography
- poetry
- song.

Some of these are explored in Section 8.4 below.

Geography fieldwork is relevant here. Fieldwork is an essential component of geography education and studying geography without fieldwork has been likened to studying science without experiments. Practical skills are clearly important in the fieldwork experience, but it is our engagement and interaction with place beyond the classroom which ultimately affects our understanding of that place.

Influence on and manipulation of the perception of place

Links

There are links between management and manipulation of the perception of place and Chapter 7, Global systems and global governance.

Perceptions of **international** places tend to be influenced more by the media than by personal and direct experience. What are your current perceptions of places in countries such as Afghanistan and Syria, for example, and how have these developed? Other influences, such as historical and political relationships or trading links, may also influence how a country is perceived. Governments are keen to attract trade and investment into their countries and a positive place perception is therefore important at an international level. Organisations such as the British Council help to promote the UK through educational and cultural links. The government and monarchy are also seen to play a pivotal role in promoting international relations for the UK.

Perceptions of place are also important at the local scale and increasingly organisations are being employed to promote place, build up a place brand or improve perceptions of place. Investment in place is crucial for its survival and people are more likely to want to live or work in a place with a good reputation and positive image.

Different **agents of change** will aim to manage the perception of place. These may include:

- national and local government
- corporate bodies
- tourist organisations
- community groups.

Government

At both national and local level, strategies have been adopted to manage and manipulate perception of place in order to attract people and investment to a place. These include: **place marketing**, **rebranding** and **re-imaging** (Figure 8.18).

Place marketing

Marketing or public relations (PR) companies may be employed by national and local government to improve or create positive perceptions of place. Using the example of Weston-Super-Mare in Somerset, strategies have included:

- advertising campaigns, including social media marketing through Facebook
- an official Weston-Super-Mare website and newsletter
- a Weston-Super-Mare logo
- creation and promotion of the first ever Love Weston Winter Wonderland, a festive attraction incorporating the annual Christmas lights switch-on. The aim was to increase perceptions of Weston-Super-Mare as a destination point for Christmas shopping.

Figure 8.18 The components of urban rebranding

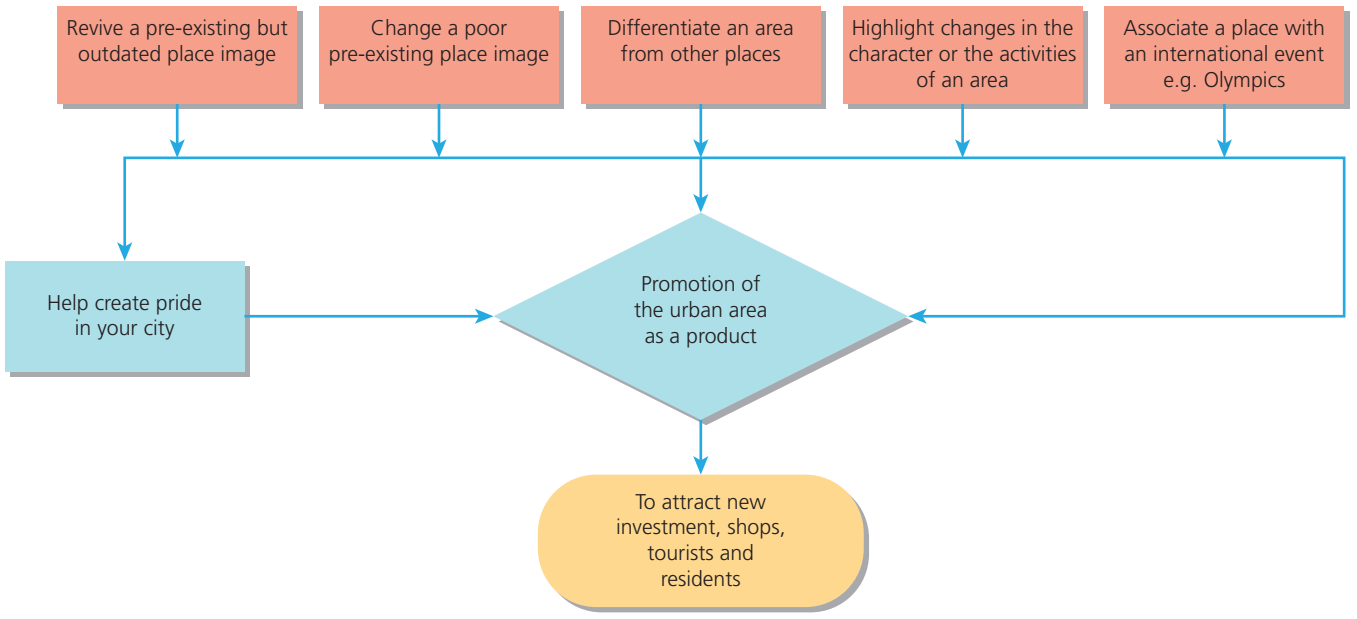


Figure 8.19 The aims of urban rebranding

Think about how your local place is marketed. Who is responsible for this? What are the key objectives of the marketing?

Rebranding

Rebranding is used to discard negative perceptions of a place (Figure 8.19). Its main aim is to try to make a location a desirable place in which to live (as well as to invest and develop industrial activity) and one that people will want to visit for social and recreational purposes and to shop and spend money.

In 2013, the slogan 'People Make Glasgow' was introduced as the new brand name for Scotland's largest city. It was chosen following a **crowdsourcing** social media campaign involving more than 1,500 people from 42 countries. The campaign emphasises that people are the heart and soul of this particular place – 'People Make Glasgow.'

Many argue that place (re)branding must start from the inside and involve local residents with 'insider' experiences. Geographers Edward Relph and Yi-Fu Tuan have emphasised the importance of being inside a place to truly understand it and likewise, it has been argued that without a thorough understanding of place, one would find it difficult to regenerate and rebrand a place.

Rebranding is not without its problems. Different stakeholders may include pre-existing residents, local businesses, potential investors, local government and potential home-owners or visitors and the challenge is to satisfy as many of these groups as possible. Pre-existing residents often want to protect and project their local distinctiveness, while development agencies seek to establish place brands based on government incentives, available technology and an area's international links. Some city regeneration schemes have actually driven out the locals they originally intended to help, as rising property prices and rents have favoured more affluent people.

Re-imaging

Linked to rebranding, re-imaging seeks to discard negative perceptions of a place and generate a new, positive set of ideas, feelings and attitudes of people to that place. This may include the revival of a pre-existing but outdated place image. More commonly, it seeks to change a poor pre-existing image of a place. This has been well documented through the example of Liverpool in the 1980s and 1990s. Deindustrialisation had caused an economic downturn in the city and riots in 1981 dominated newspaper headlines. Large-scale regeneration began and the Tate Liverpool art gallery was one of a number of projects aimed at re-imaging the city's industrial heritage through culture. The Merseyside Development Corporation used the term 'There's life in the old docks yet.' The exterior of the Grade I-listed warehouses of the Albert Docks remained untouched but the derelict interiors were transformed into modern art galleries.

The Tate's presence in the city was seen as a key factor in Liverpool winning the title of European Capital of Culture in 2008, a far cry from the negative imagery of the city in the early 1980s.

Rebranding Amsterdam

Figure 8.20 The motto 'I amsterdam' originated from a photography exhibition in 2004 promoting the city of Amsterdam. It was subsequently adopted for the rebranding and promotion of the city

In the late twentieth century, Amsterdam's reputation as a major international cultural centre had been threatened by a number of factors, including:

- greater competition from other cities both within and outside of the Netherlands
- social and economic decline in some areas
- the city's reputation for being liberal towards soft drugs and prostitution, which was seen as inappropriate for attracting new investors and enterprises
- a failed bid to host the Olympic Games.

A number of rebranding strategies were adopted but the most successful was the 'I amsterdam' slogan, seen to be clear, short, powerful and memorable. The large three-dimensional 'I amsterdam' letters were positioned in front of the city's famous Rijksmuseum in 2005 and the sculpture is now the city's most photographed item, being photographed over 8,000 times a day on sunny days. The use of smartphones and social media has seen the image spread all over the world and Amsterdam has become one of the most successful destination brands on social media. It is argued that the re-imaging has been too successful, with many people beginning to talk about 'overtourism' in the city. Before the COVID-19 pandemic in 2020, tourist numbers were predicted to rise to over 20 million per year.

Other European cities such as Barcelona have undergone successful rebranding programmes in recent decades. Investigate the different strategies used to 'rebrand' these places and evaluate how successful they have been.

Corporate bodies

A **corporate body** is an organisation or group of persons that is identified by a particular name. Examples include institutions, businesses, non-profit enterprises and government agencies.

Many corporate bodies will have an interest in place but some will want to manipulate perceptions of place.

Tourist agencies

Tourist agencies aim to 'sell' place to potential visitors and marketing positive perceptions of place makes this easier. In the UK, tourist organisations range from Visit Britain, the non-departmental public body funded by the Department for Digital, Culture, Media and Sport, to the individuals responsible for promoting a specific tourist attraction. The strategies are similar: to make a place look as good as it can and attract as many visitors as possible. Promotional materials such as brochures, videos, websites, magazine advertisements, slogans and logos are used and places may adopt a unique selling point.

In 2012, the Pembrokeshire Coast National Park Authority put vintage-inspired designs featuring nostalgic images of the Pembrokeshire coast on show at Cardiff Airport, UK railways stations and across the London Underground in order to increase tourism to the area (Figure 8.21). The posters won numerous awards but they were also successful in attracting more people to the area.

Figure 8.21 One in a series of retro-style railway posters used to promote the Pembrokeshire Coast National Park in 2012

Airlines and train companies

Airlines and train companies also seek to manage perceptions of place but they do so in order to get people to use their travel services to visit these places. Figure 8.22 is one of several railway posters marketing a positive image of Torquay in Devon. Artworks like this emerged at the beginning of the twentieth century as railway companies commissioned posters, sometimes by famous artists, to sell the delights of the

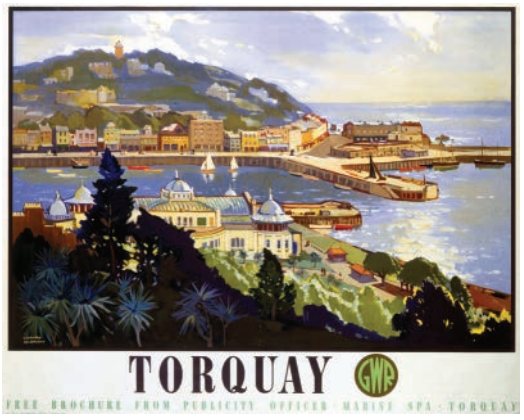


Figure 8.22 Railway poster advertising Torquay c.1947

British coast and countryside and therefore boost the number of train passengers wanting to get there.

The role of community and local groups

Community or local groups may take an active role in managing and improving the perception of their place to attract investment and improve opportunities and services within the area. Regeneration and rebranding strategies have increasingly involved local people, since they have the 'insider' experience of place and will be the people most affected by any changes. Residents' associations and heritage associations play an important role and social media is increasingly being employed to engage and involve local people in planning and place-making schemes.

Review questions

7 Outline the concepts of 'sense of place' and 'perception of place'.

8 To what extent will government and corporate bodies have different reasons for managing the perception of place?

9 With reference to a place you have studied, assess the extent to which attempts to influence place-meaning have been successful?

8.4 Representations of place and the use of quantitative and qualitative sources

Places can be represented in a variety of different forms or media that may give contrasting images. It is therefore important to investigate and present your two place studies using a variety of quantitative and qualitative sources while at the same time acknowledging their limitations. This part of the chapter combines two parts of the specification. Firstly, it looks at how places can be represented using formal and informal sources and, secondly, it introduces the different types of quantitative and qualitative skills you can use to build up your own place studies.

Key terms

Objective – Not influenced by personal feelings or opinions in considering and representing facts.

Qualitative data – Information that is non-numerical and used in a relatively unstructured and open-ended way. It is descriptive information, which often comes from interviews, focus groups or artistic depictions such as photographs.

Some types of qualitative data, such as interviews, can be coded and may then be subject to quantitative analysis.

Quantitative data – Data that can be quantified and verified, and is amenable to statistical manipulation.

Quantitative skills

Statistics

The UK census has been used to detail the social and economic characteristics of the population of the country for over 200 years. Smaller geographical units can reveal basic information about places such as:

- age and gender structure
- ethnicity
- levels of economic deprivation.

The census provides large-scale, **quantitative data**, which has been used by national agencies to better understand and plan for population growth and other demographic changes. The Office for National Statistics (ONS) is the UK's largest independent producer of official statistics. It is responsible for:

- collecting and publishing statistics related to the economy, population and society at national, regional and local levels
- conducting the census in England and Wales every ten years.

Skills focus

The ONS website is a useful starting point for finding out about your local place. You can download data on characteristics such as population, health and crime and use these to produce maps and graphs. Go to www.ons.gov.uk/

The use of quantifiable data such as statistics is not as **objective** as it may first appear. This is because people selectively choose the data they wish to use for their particular purpose. Their use therefore becomes **subjective**. Another criticism of using statistics when

studying place is that they tell us very little about the human experience of a place and what it is like to live there.

Qualitative skills

Maps

Maps have long been used to locate places but they can also influence how we think or feel about a place and as such play a very important role in both our sense of place and perception of place. It is important to cast a critical eye over the reliability and accuracy of maps as throughout history they have distorted reality. For example, early world maps such as the Mappamundi (c.1300) depicted the world as a flat disk with the Holy Land and Jerusalem at the centre. During the period of colonial expansion, maps exaggerated area size and resources (and thus strategic importance). Early cartographers (map-makers) very much 'promoted the cause of empire' (Figure 8.23).

Figure 8.23 This early map of Brazil shows an interior landscape rich in resources but supposedly inhabited by 'less civilised' people. A boat of slaves is shown arriving from Africa

Maps can include hidden bias and influence. Google Maps is a search engine which allows people to search for and find out about places. While undoubtedly an invaluable tool in researching place, it is worth noting that it is not done in a strictly objective way. In fact, Google Maps, like any search engine filters place – directing people towards businesses that have engineered their appearance on the first page of a Google search.

Counter-mapping describes a bottom-up process by which people produce their own maps, informed by their own local knowledge and understanding of places. You could try this for your own local place. Figure 8.48 (page 387) is an example of this: the artist and cartographer Adam Dant illustrated this map of Spitalfields in East London with fifty portraits of the people who make the area distinctive. On the reverse of the map, stories of all those portrayed on the front are provided, plus a guide to the essential Spitalfields landmarks and destinations. The result is that the map not only provides factual information but it also conveys a sense of place.

Analysing the Mercator projection

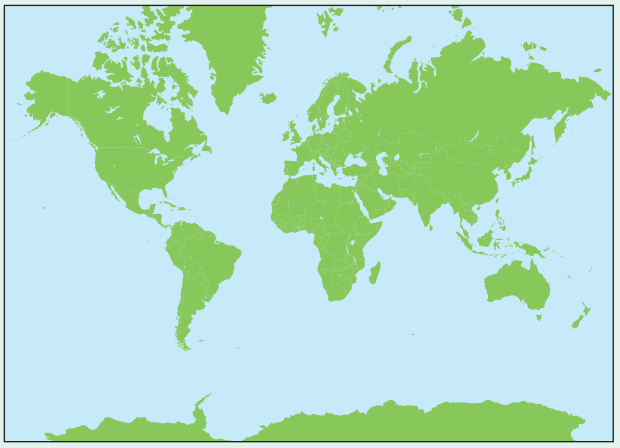


Figure 8.24 The Mercator projection

The Mercator map projection was first used in the sixteenth century as a navigational map for sailors. The northern landmasses look much bigger on the map than they are in reality.

You will commonly see the Mercator map projection in geography classrooms and unthinkingly accept it as an accurate representation of the world. However, educationalists have been denouncing its use in classrooms for a long time because it does not accurately depict the size of the Earth's northern and southern latitudes.

1. It distorts the relative size of land masses. For example, Greenland looks the same size as Africa but in reality Africa is 14 times larger.
2. It is based on the convention that the northern hemisphere is at the top when in truth the world does not come with a label saying 'this way up'.
3. The map is Eurocentric. Not only does it exaggerate the size of Europe but it puts it in the middle of the map.

Skills focus

The expansion of **Geographical Information Systems (GIS)** means that an increasing range of geographical data is available for research and analytical purposes. The Consumer Data Research Centre was established by the UK Economic and Social Research Council and provides a comprehensive set of geographical data which has been mapped and is easy to access. Go to http://maps.cdrc.ac.uk/

Biomapping

One method of trying to measure sense of place is **biomapping** or emotional cartography. This is the mapping of emotions shown by people to certain places through use of a device which records the wearer's Galvanic Skin Response (GSR). This is a simple indicator of emotional response in conjunction with a geographical location. A map can be created which visualises points of high and low feelings.

In a 2015 episode of the series *Coast*, a series of 'stress tests' were carried out on presenter Nicholas Crane by geographer Jon Anderson to measure the electro dermal activity (or nervous energy) of the skin while walking in London, cycling a Boris-bike around Trafalgar Square, and negotiating the London Underground system. The results showed very little variation when he was walking around London but there were notable spikes in nervous energy when he caught the underground train or crossed a busy road. The peak in stress level was recorded as he cycled around Trafalgar Square and had to contend with heavy traffic.

Conducting the same experiment at the coast produced contrasting results. When the presenter took an early evening stroll along the beach at Southend, his stress levels were at their lowest. This was put down to the physical environment of the coast – a combination of natural elements such as the hypnotic and soothing sound of the waves, the feel of the sand under his feet and the sea breeze.

Quantifiable data was produced in this experiment to back up the widely-held cultural notion that going to the seaside is good for you.

Figure 8.25 TV presenter Nicholas Crane preparing to cycle around Trafalgar Square as part of an emotional stress test

Interviews

Interviews can generate detailed insights about a person's sense of place or perception of place.

Advantages of using interviews:

- They are first-hand or direct reports of experiences, opinions and feelings.
- They can be structured or unstructured, the latter being more like an everyday conversation rather than having pre-set questions, and tend to be more informal, open-ended, flexible and free flowing.

Disadvantages of using interviews include:

- Interviewer bias may affect the responses of the interviewee by using leading questions.
- People may not be honest as they like to present themselves in a favourable light.

Ethnography is a research method that explores what people do as well as say. For example, to understand different places as they are experienced and understood from the 'inside', the ethnographer would participate in the daily life of a person or group of people, watching what happens, listening to what is said and generally becoming part of that world.

Media representation of place

Knowledge and understanding of place can be influenced very strongly by a diverse **media** including television, film, photography, art, books, newspapers and the internet. These are increasingly reaching a larger, global audience and are therefore very important in shaping wider perceptions of place. In addition, this has meant that geographical distance has become less of an obstacle for learning about places when people can gain instant knowledge about them from media sources. The notion of near and far places is becoming increasingly blurred as people develop attachments to places they have never visited, through the media.

Photographs

At a time when photo-editing of people is commonplace, so places can easily be 'photoshopped' or filtered to make them appear different and, in the case of tourist areas, more attractive. The difference between a photograph you may have seen of a place and the reality is often down to differences in weather, the time or the season, but with advanced technology, people are increasingly editing photographs to improve the image or perception of that place. Similarly, photographs can be selective in what they show. Marketing images tend to focus on the natural beauty or landscape of places without disturbance from humans; the reality for many tourist places may be very different, as shown in Figures 8.26 and 8.27 on page 372.

Figure 8.26 The Toshogu shrine in the UNESCO World Heritage site of Nikko in Japan

Figure 8.27 The Toshogu shrine in the middle of the day when visitors numbers are at their highest

Figure 8.26 is a tourist image of the Toshogu shrine in the UNESCO World Heritage site of Nikko in Japan. The shrine is set in a beautiful forest and is famed for its ornate architecture including woodcarvings and gold-leaf decoration. Figure 8.27 is taken at the shrine in the middle of the day when visitor numbers are at their highest. The 'reality' of Nikko here contrasts with the representation of the shrine in the tourist image.

Skills focus

Flickr is a photo-sharing platform which can be used to find visual images for your local and distant place. Google Earth is also a source of geotagged photographs. When analysing the value of a photograph, it is important to consider why it was taken and who it was taken for. A photograph can be selective in what it shows. What has been missed out? Why has this not been photographed? The conditions in which it was taken are also important. One photographer has actually created a gallery of his images before and after photo editing – www.peterstewartphotography.com/Before-and-After-Gallery/

Textual sources

Novels may evoke a sense of place – a feeling that the reader knows what it is like to 'be there'. Some places have become so strongly associated with particular authors and stories that they are now promoted or advertised as such, for example, Thomas Hardy's Dorset and Brontë country in West Yorkshire and East Lancashire.

While such associations can play a positive role in representation of place, the reverse can also be true and the role of popular media can be instrumental in creating 'place prejudice'.

The city of Liverpool is one of a number of (mainly northern) British cities that has historically suffered from a negative portrayal in the British media. Unemployment, economic deprivation, gangs, rioting and drugs problems were all characteristic of 'Scouse' depictions in the press in the late twentieth century, and the city authorities adopted an aggressive rebranding of the city to distance it from this poor media image. In 2008 the city was awarded the status of European Capital of Culture, reflecting Liverpool's more positive architectural, musical and sporting heritage.

Poetry

Poetry has long been used to describe and evoke a sense of place. There are many famous poets associated with particular places.

- William Wordsworth is linked to the Lake District.
- Seamus Heaney has written extensively about his Irish roots.
- William Blake famously described the poverty and despair of industrial London.

Poets may refer to specific places in personal and responsive ways but they also enable the reader to sense and imagine what it is like to be in that place.

Daljit Nagra is a contemporary British poet who was born in England to parents who are traditional Sikh

Punjabis. He has written extensively about the Britain where Indians came and settled. His poetry enables the reader to be immersed in the Sikh community and experience Sikh Britain from the 'inside'. Nagra has sought to explore the thoughts, feelings and cultural attitudes of first, second and third generation descendants towards their own community, other minority ethnic communities and the indigenous white population. His poem *Our Town with the Whole of India!* describes the district of Southall in London which migrants had turned into a 'Little India' in the 1970s.

Our town in England with the whole of India
sundering out of its temples, mandirs and
mosques for the customised streets. Our parade,
clad in cloak-orange with banners and tridents,
chanting from station to station for Vasaikhi
over Easter. Our full-moon madness for Eidh
with free pavement tandooris and legless
dancing to boostered cars. Our Guy Fawkes'
Diwali – a kingdom of rockets for the Odysseus-
trials of Rama who arrowed the jungle foe to
re-palace the Penelope-faith of his Sita.
Our Sunrise Radio with its lip sync of Bollywood
lovers pumping through the rows of emporium
cubby holes whilst bhangra beats slam where
the hagglers roar at the pulled-up back-of-the-
lorry cut-price stalls. Sitar shimmerings drip down
the furbishly columned gold store. Askance is
the peaceful Pizza Hut ... A Somali cab joint,
been there for ever, with smiley guitar licks where
reggae played before Caribbeans disappeared,
where years before Teddy Boys jived.

Daljit Nagra, 2014, Extract from *Our Town with the Whole of India!*

The poem describes the changes which have occurred in the community over the years, celebrating the multicultural nature of the area while at the same time highlighting potential threats accompanying it.

Skills focus

To illustrate how the same place can be represented in different ways, compare Blake's poem below with the more romantic, idealised images of London described by Wordsworth. Consider the imagery, the choice of vocabulary and the overall mood of each poem.

London

I wander thro' each charter'd street
Near where the charter'd Thames does flow,
And mark in every face I meet
Marks of weakness, marks of woe.
In every cry of every Man,
In every Infant's cry of fear,
In every voice, in every ban,
The mind-forg'd manacles I hear:
How the Chimney-sweeper's cry
Every black'ning Church appals,
And the hapless Soldier's sigh
Runs in blood down Palace walls;
But most thro' midnights streets I hear
How the youthful Harlot's curse
Blast the new-born Infant's tear,
And blights with plagues the Marriage hearse.

William Blake, 1794

Composed upon Westminster Bridge, September 3, 1802

Earth has not anything to show more fair:
Dull would be of soul who could pass by
A sight so touching in its majesty;
This City now doth, like a garment, wear
The beauty of the morning; silent, bare,
Ships, towers, domes, theatres, and temples lie
Open unto the fields, and to the sky;
All bright and glittering in the smokeless air.
Never did Sun more beautifully steep
In his first splendour, valley, rock, or hill;
Ne'er saw I, never felt, a clam so deep!
The river glideth at his own sweet will;
Dear God! The very houses seem asleep;
And all that mighty heart is lying still!

William Wordsworth, 1802

Music

Music can help to evoke a sense of place. Different types of music may be associated with geographical areas, such as reggae with Jamaica.

Song lyrics may also help to portray particular places. In 2010, Newport rapper Alex Warren and singer Terema Wainwright became an internet sensation after producing a parody of the Jay-Z and Alicia Keys' song about New York. The spoof version of **Empire State of Mind** received almost 200,000 hits within just two days of being placed on YouTube. Like the music video for the original track, the video is shot in black and white with the singers referring to famous landmarks and cultural clichés. However, rather than the Manhattan skyline as a backdrop, the local duo celebrate the sights of the Welsh city. Rather than a grand piano in Times Square, Wainwright plays a battered synthesizer on the riverfront and sings in a thick Welsh accent: 'In Newport / concrete jumble nothing in order / not far from the border / When you're in Newport / Chips, cheese, curry makes you feel brand new / washed down with a Special Brew / Repeat the word Newport, Newport, Newport'.

Similar parodies have been added to YouTube for a number of other UK places.

Television and film

Places are a vital backdrop to most dramas. We associate certain places with different types of stories, such as derelict houses in horror films or space ships in science fiction. However, TV shows and films also play a significant role in representing place and this can be both positive and negative.

The 2015 television dramatisation of the Poldark novels had a positive impact on the Cornish tourist trade as viewers, inspired by the shots of the dramatic coastline and beautiful beaches, flocked to the county to soak up the atmosphere. Hits on Visit Cornwall's website soared by 65 per cent after the first episode alone and the property website Rightmove reported that house-hunting enquiries more than doubled. It is anticipated that the 'Poldark effect' will be felt well into the 2020s.

On a global scale, *The Game of Thrones* TV series has become synonymous with Northern Ireland. It is estimated that 350,000 tourists come to the region annually because of *The Game of Thrones* and spend an average of £50 million a year. The region's film agency suggests that the TV series has brought over £250 million into the economy since 2010. The government and local agencies recognise the importance that the TV series has had on boosting many aspects of the regional economy and supporting the positive representation of the region. Between the 1960s and 1990s, global media representations of Northern Ireland mainly focused on conflict and the troubles, which had a huge impact on investment and visitors to the country. *The Game of Thrones* has created a much more positive representation of place in the media.

These two examples have shown how TV and film can represent place in a positive way. The reverse can also be true. Many crime dramas are located in urban areas, but not all crime occurs in cities. Equally, the sites chosen for filming can portray the same place in very different lights. Compare, for example, the representation of London in *EastEnders* (BBC, 1985–present) with the more glamourous skylines and buildings of *The Apprentice* (BBC, 2005–present) or the settings for the crime dramas *Sherlock* (BBC, 2011–2017) and *Killing Eve* (BBC, 2018–present).

When researching media representations of your chosen place studies, you could have a look at the British Pathé website and YouTube channel. British Pathé is a multimedia resource with film, documentary and newsreel archive material. Over the course of the twentieth century, British Pathé reported on everything 'from armed conflicts and political crises to the hobbies and eccentric lives of ordinary British people.' The resulting archive is a useful audio-visual source which can be used to document changes in place and population over a period of time. The archive is available to view online for free.

Art

Art has long been used to represent place and this is most famously seen in the landscape paintings of the eighteenth and nineteenth century. Painters became synonymous with geographical places:

- Gainsborough and Constable (Figure 8.28) are known for their landscapes of Suffolk, now sometimes referred to as 'Constable Country'.
- French artist Paul Cezanne painted Provence in southern France.
- Japanese artist Hiroshi Yoshida is famous for his iconic wood block prints of Japanese landscapes.

The common criticism of such paintings was that they were pastoral fantasies giving the impression of a rural idyll, which did not exist for the majority of people

living in the countryside at that time. These paintings reflected a romantic vision which still shapes many people's mental images of the countryside and is perpetuated through tourist brochures, chocolate boxes and jigsaw puzzles. Such constructs of rural places are powerful because they shape views on what the countryside is actually like and what it should be like.

Figure 8.28 Constable's painting entitled *The Hay Wain* is based on a site in Suffolk, near Flatford on the River Stour. The location of the painting at Flatford Mill is now a popular tourist location run by the National Trust

Paintings

Paintings may be considered less reliable than photographs because there is more scope for individual interpretation and selection. However, they can also show a deeper understanding of place because they allow the painter to show more of the character of what is there.

An artistic representation of an urban place can be seen in the work of L.S. Lowry, famous for his matchstick human figures and depictions of life in northern mill towns.

Figure 8.29 The iconic Lowry Centre at Salford Quays opened in 2000 on the site of the old docks. It now sits at the heart of Media City UK and has had a major influence on the cultural landscape of the north-west

Lowry became interested in the industrial landscape after he moved to Salford with his family. Although he made many drawings there, his mill scenes (a combination of the real and the imagined) have remained his most famous. Lowry's legacy and association with Salford have continued into the twenty-first century and been used for place regeneration. The Lowry Centre has been a key part of the redevelopment of the Salford Quays in Greater Manchester (Figure 8.29).

Skills focus

Where art has been used to depict place, it is important to ask a number of questions and view it as a cultural product:

- What and who are shown in the picture, and why?
- What and who are not shown, and why?
- Why did the artist paint this scene and who commissioned it?

Sculpture

Art has traditionally represented place but it can also make a place distinctive and be used in the regeneration of an area. An example is the Kelpies sculpture near Falkirk in Scotland (Figure 8.30). The Kelpies are the landmark feature of the Helix Environmental Regeneration Scheme on the Forth and Clyde Canal. They are modelled on heavy horses and represent the industrial history of Scotland and the Falkirk/Grangemouth area. Heavy horses would once have been the powerhouse of the area, working in the foundries, the fields, farms and the canal itself. The Kelpies sculpture has been used to regenerate place here but their national and international prominence has also served to develop a sense of pride and ownership in the local place. Indeed, they were the front cover image of a 2014 book entitled *Scotland: The Best 100 Places*.

Figure 8.30 The Kelpies sculpture

Graffiti

Graffiti is writing or drawing that has been put illicitly on a wall or other surface, often in a public place. It has traditionally been associated with youth cultures claiming ownership of a place but the famous UK graffiti artist Banksy argues that the importance of graffiti is also to give a voice to people who aren't normally heard in the mainstream (Figure 8.31). Many consider graffiti as a type of vandalism and authorities are keen to remove it from public areas. Increasingly however, graffiti is being accepted into mainstream culture and art galleries may now stock graffiti images. In these places, graffiti is seen not as being 'out of place' but as something which can be bought and sold. It is also being used as a type of street art in the regeneration of places.

Figure 8.31 The streets of London have been a canvas for guerrilla artist Banksy's work for over 20 years. Some of the art makes a hard-hitting political point, while other works are intended to be funny

Dismaland, Weston-Super-Mare, Somerset

An unusual type of art installation brought thousands of visitors and greater place recognition to the seaside resort of Weston-Super-Mare in the summer of 2015.

Based in a derelict seafront lido and billed as a 'bemusement park', Banksy's Dismaland featured artworks on themes including the apocalypse, anti-consumerism and celebrity culture.

Unlike other British seaside resorts like Margate in Kent, which has built up a more positive representation following the opening of the Turner art gallery, Weston-Super-Mare relies on its beaches and traditional seaside attractions for visitors.

Banksy's work brought significant benefits to a resort which has experienced significant decline in the last few decades. It is estimated that an additional 150,000 people visited the town as a result of the attraction and they added £20 million to the local economy. More importantly, local tourist chiefs were delighted that it helped to put the place of Weston-Super-Mare more firmly on the map.

Place and architecture

As well as designing buildings, architects are responsible for planning places and this planning is seen as increasingly important in terms of creating sustainable and healthy places. Architecture can also play a pivotal role in the redevelopment and rebranding of a place.

Some redevelopment schemes utilise the existing style or heritage of buildings in an area; others involve more radical change as part of a rebranding process. Compare, for example, the redevelopment of the SouthGate shopping centre in Bath to the redevelopment of the Bull Ring in central Birmingham. The former mimics the city's original and popular Georgian architecture with a Bath-stone façade, while the Bull Ring was a much more radical redevelopment incorporating the construction of the iconic Selfridges department store with a façade of 15,000 aluminium discs mounted onto a blue background. The latter has been seen as one of several architectural landmarks, including the city library, which have made a major contribution to the regeneration of central Birmingham.

Gentrification

Redevelopment of place comes at a cost to some. The process of gentrification has been strongly promoted by local government in many British cities as a method of economic regeneration.

Gentrification is the improvement of housing in an area that was formerly poor and run-down. It is mainly carried out by middle-class residents or newcomers who purchase the properties at cheaper prices and make repairs and improvements which increase the housing value. The nature of the whole area may improve and this leads to greater wealth in this particular place.

Unfortunately, not everyone benefits and the original poorer residents may find themselves 'displaced' as they can no longer afford to live in this area. A similar process has occurred in some British coastal resorts where second-home owners have effectively priced less affluent locals out of the housing market.

Digital or 'augmented' place

The rise in the use of digital technology such as smartphones has led to much discussion about the notion of digital place. Parallel interactive places such as Second Life and Fortnite have existed for some time but these are seen as virtual places rather than real places.

The development of GPS-capable mobile devices has meant that real places have become 'layered' and 'augmented' by technological advances and applications and this can have a huge impact on our sense of place. Smartphones know where we are and are linked to data that knows where other people or things are too. They can provide information and reviews about place in seconds.

Geographical information systems (GIS)

GIS and software developers have been engaging in place for some time.

- Supermarkets employ GIS to map shopping habits and where customers live.
- Police forces want to know about the links between crime and place.
- Politicians want to know about place so that they can focus their time and money on voters.

GIS sources are increasing in number and becoming more readily available as free online resources. They could be used in place studies to illustrate changing demographic and cultural characteristics and economic change and social inequalities.

Review questions

10 What are the differences between qualitative and quantitative data as representations of place?

11 Explain what you have learned from an artistic source about a place.

8.5 Place studies

> The single story creates stereotypes and the problem with stereotypes is not that they are untrue but that they are incomplete, they make one story become the only story ... It is impossible to engage properly with a place or person without engaging with all of the stories of that place or person.
>
> *Chimamanda Adichie,* TED Talk *The Danger of a Single Story,* 2008

The final section of the Changing Places topic asks you to explore the developing character of a:

- place local to your home or study area
- contrasting and distant place.

Factors to consider in choosing the places

- A place for study should be about the size that you could comfortably walk around in a few hours.
- Your local place may be the place where you live or go to school – a village, small town, community or area of a city.
- It can be either urban or rural.
- The contrasting and distant place may be in the same country or a different country but it must show significant contrast in terms of economic development and/or population density and/or cultural background and/or systems of political and economic organisation.
- The geographical distance between your local and distant place may be relatively small.

It is also important to remember that your place studies must incorporate the knowledge acquired from the rest of the specification and thereby further enhance your understanding of the way your life and that of others are affected by continuity and change in these places.

- Before choosing your place studies, it is worth looking at the availability of data and the different sources of information before committing to one distant place.
- For places abroad, charities and non-governmental organisations are useful sources of information, as are government websites and international or global institutions such as the World Bank.

Context for my place studies

Using the examples of central Torquay in South Devon and the area of Brick Lane and Spitalfields in East London, this section will suggest resources and activities which will enable you to build up your own two place studies. I have chosen to present Torquay as my local place study because it is the town in which I work and a place which I have got to know well through fieldwork and recreational activities.

I have less 'insider' knowledge about my chosen distant place of Brick Lane and Spitalfields in East London, but it is a place I have visited and one which I know contrasts significantly with Torquay in terms of both economic development and cultural background.

Both Torquay and Brick Lane have undergone considerable change in the last few decades.

- The seaside resort of Torquay has experienced a decline in tourist numbers since the 1970s resulting in subsequent social and economic problems and a less favourable place identity.
- The Brick Lane and Spitalfields area has undergone social and economic changes of a different nature.

Already contrasting with Torquay in terms of its cultural background (Brick Lane and Spitalfields is home to a large Bangladeshi diaspora), this area of East London is undergoing social and economic change as a result of gentrification and redevelopment projects.

1. Investigating location and locale

Maps

Maps are an important tool for any investigation of place. They can show the location of a place in relation to others and, depending on the type of map chosen, can also display physical and human features of the local area. For a rural place, you could start with an Ordnance Survey map, 1:25,000 scale:

- Look at the physical geography of the area: relief, height of the land, aspect and drainage.
- How have humans impacted the area?
- What are the main land uses?

You could see how the area has changed by looking at older OS maps. Look for changes in the size of the place, types of land use and infrastructure.

Street maps are more suited to urban areas.

- Goad maps are detailed street maps which show individual buildings and their uses. These can be updated or annotated as part of your fieldwork.
- Google Maps also zoom in to street level. Along with Google Earth, these provide geographical coverage of places all over the world. They also provide aerial and satellite images.

A useful exercise is to compare the image of your local place from the OS map to that from Google Earth.

- What characteristics are shown? Which are not?
- How do these images compare with the place as you know it?

For your distant place, Google Street View is a useful tool for seeing what the place is like. The horizontal and vertical street-level images enable you to see land use and building type, which you might not otherwise get to see.

Maps can be used to highlight economic and demographic changes in an area as well as changes in lived experience. In London, for example, Charles Booth's c.1890 'Poverty Map' is viewed as an important step in the development of geodemographics (characterising people based on where they live). Over a period of several years, Booth and his researchers knocked on doors all around London, interviewing and characterising the people they met into seven categories, such as 'Well-to-do' or 'Vicious, semi-criminal'. The houses were then coloured on the map according to their category. The resulting map shows the characteristics of large parts of London at the turn of the twentieth century (Figure 8.32).

Charles Booth's poverty maps for Brick Lane indicate that the area had mixed economic fortunes at that

Figure 8.32 Booth's map of Spitalfields area c.1890. The original notebooks from Charles Booth's survey of life and labour in London (1886–1903) contain interviews with Londoners from all walks of life, eyewitness descriptions of the city, street by street, and the raw data that was later used to compile statistical reports and the famous maps of London poverty

time. Some parts were classed as poor or very poor but equally there were areas of greater affluence too. This contrasted with large areas of nearby Whitechapel and the East End where poverty was much more widespread.

Literary sources such as books, atlases and newspapers

These can be used alongside maps to provide historical information about places. Your local library will be a useful source of information and there may be a Local History society in your area. Investigate how your chosen place has changed over time and consider the role of past connections on current characteristics.

The internet

The **internet** provides a wealth of information about different places. Official government websites provide geographical information such as the census data or local health statistics. An internet search on your chosen places will also result in other statistics, facts, opinions, stories, etc. Remember that this information is created for many purposes (to inform, to persuade, to sell, to present a viewpoint or to create or change an attitude or belief) and unlike more traditional media such as books and magazines, no one has to approve the content before it is made public.

Local place study: Torquay

Figure 8.33 Map showing location of Torquay

Torquay is a seaside town in Devon, part of the unitary authority area of Torbay. It lies 18 miles south of Exeter and 28 miles east-north-east of Plymouth. It is situated on the north of Tor Bay, adjoining the neighbouring town of Paignton on the west of the bay and across from the fishing port of Brixham. Some of the material presented in this place study applies to the area of Torbay, but the main focus is the central tourist area of Torquay.

Constrasting place study: Brick Lane and Spitalfields

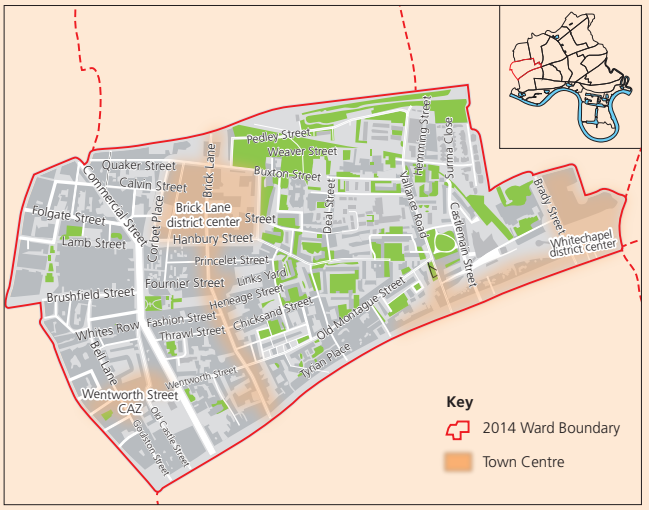


Figure 8.34 Map showing the location of Spitalfields and Banglatown within the borough of Tower Hamlets, London

Brick Lane is a street in the London Borough of Tower Hamlets. It runs from Swanfield Street in the northern part of Bethnal Green, passes through Spitalfields and is linked to Whitechapel High Street to the south by the short stretch of Osborn Street. In the late twentieth century, it became the heart of the city's Bangladeshi-Sylheti community and is known to some as Banglatown. This place study will focus on the southern half of Brick Lane and its immediate surroundings. Some of the data applies to the wider area known as Spitalfields and Banglatown.

Torquay

During the nineteenth century, Torquay grew from being a small fishing village of 800 people into a fashionable seaside resort. John Bartholomew's *Gazetteer of the British Isles,* 1887, described Torquay as '... among the first of English watering places. Its mild and equal climate, its excellent water supply and drainage system, and its bright and pleasant surroundings make it a great resort of invalids and convalescents in winter and spring. It has a museum, public gardens, baths and all the attractions of a great watering place.'

The *Western Morning News* described it in 1864 as 'the most opulent, the handsomest and the most fashionable watering place in the British Isles.' A slight local bias there perhaps but the resort attracted wealthy families, famous writers of the time and royalty.

The arrival of the railway in 1848 made travel more affordable and widespread and Torquay became more popular as a result. The term 'English Riviera' was adopted in the Victorian period as visitors likened the climate and topography of Torquay to the French Riviera resorts. This term has been an important strap line in the marketing of the area ever since. Torquay continued to expand as a tourist resort throughout the nineteenth and twentieth centuries, but like so many British seaside resorts, the 1970s saw the beginning of a decline in tourist numbers.

Figure 8.35

Brick Lane and Spitalfields

Brick Lane was originally called Whitechapel Lane but it is thought that it was renamed because local earth was used by brick and tile manufacturers in the fifteenth century. By the seventeenth century, the street had also become a popular location for breweries.

In the seventeenth century, the area saw an influx of French Huguenots who had been driven out of France. The street and the surrounding area became well known for its weaving and tailoring. Like much of the East End, this area was also a haven for immigrants moving into London to escape persecution abroad or looking for a better life. During the nineteenth and twentieth centuries, it was best known for its Irish and Jewish population.

In the 1990s, the area became known as Banglatown as it was a popular place for immigrants from Bangladesh, particularly Bengalis from the Sylheti region. The weekly Bangladeshi paper, *The Sylheter Dak* even had an office there. In the early 2000s, guidebooks claimed that the street was THE place to go for a curry in London, especially for traditional and authentic Bangladeshi cooking.

Evidence of the people and communities that have given the Brick Lane and Spitalfields area its unique character can be seen all around – a Huguenot church, a Methodist chapel, a Jewish synagogue and Muslim mosque stand among traditional and new shops, restaurants, markets and homes.

In recent times, the area has also developed a vibrant art and fashion scene and the Old Truman Brewery has transformed its vacant and derelict buildings into office, retail, leisure and event spaces.

Figure 8.36

2. Demographic characteristics

A key source of geographic data comes from the census, which is conducted every ten years in the UK. Other socio-economic data on factors such as employment, crime and health are generated on a more regular basis.

- The Office for National Statistics is the UK's largest independent producer of official statistics and its website is a useful starting block for finding out about the social and economic characteristics of your local and distant place. Go to www.nomisweb.co.uk.
- The Local Government Association also provides up-to-date published data about geographic areas and LG Inform Plus can be used to generate reports on places at a census ward level. Go to http://about.esd.org.uk/.

- At an even smaller scale, www.streetcheck.co.uk enables you to check a range of variables for the people living in a particular postcode.
- For geospatial data from the 2011 census, go to http://maps.cdrc.ac.uk/. Type in the postcode for the area you want to investigate in more detail.

Torquay

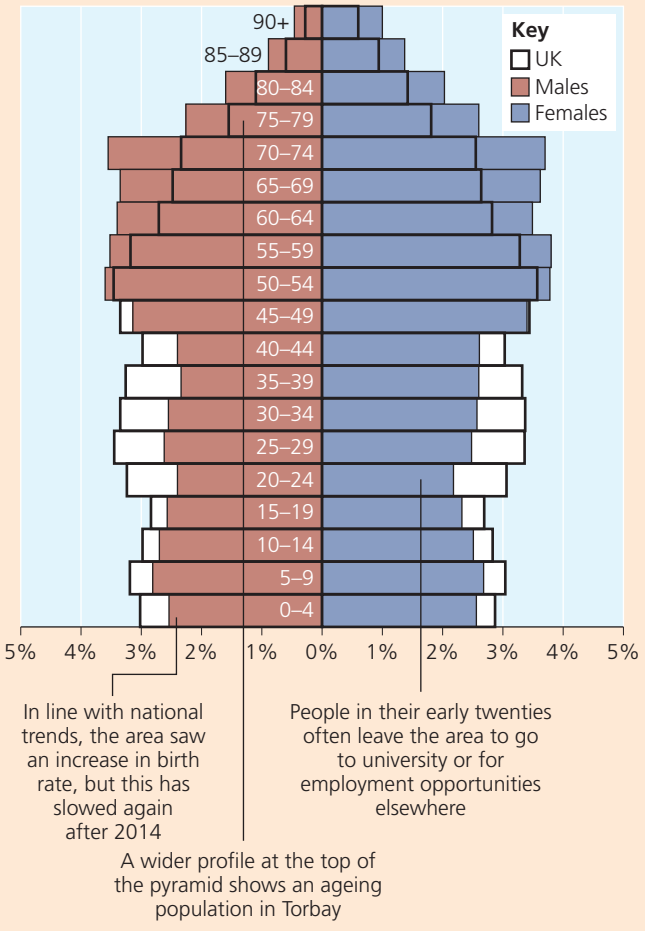


Figure 8.37 Population pyramid for Torbay, 2018. The solid bars represent Torbay's resident population and the hollow bars represent the population structure for England.

According to the ONS, by 2017 Torquay was home to approximately 67,300 people. Today, almost a quarter of these are over the age of 65. This is significantly higher than the national figure of 18 per cent. Torbay is a popular retirement location, which may explain this trend.

There is very little ethnic diversity in Torquay. In the 2011 census, only three per cent of the population identified themselves as not white British. However, Torquay does have a larger BAME community than the surrounding area and the ethnic diversity of the town has increased.

Historically, the tourist-based economy attracted few migrants to Torbay. In the early twenty-first century there was some in-migration of eastern Europeans, but in 2017 there was only an estimated net gain of 200 migrants from overseas. The area still has limited ethnic diversity.

Brick Lane and Spitalfields

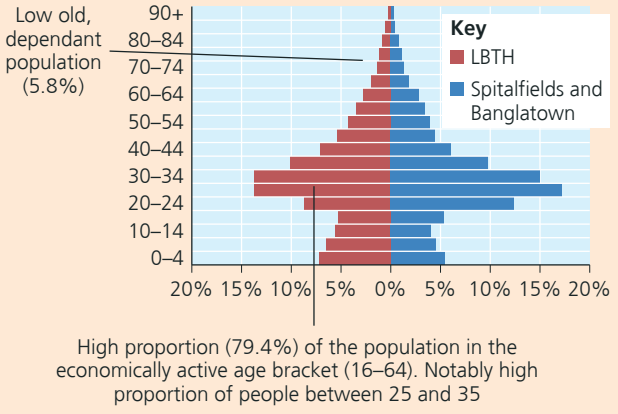


Figure 8.38 Demographic characteristics of Spitalfields and Banglatown ward compared to London Borough of Tower Hamlets (LBTH) in 2017

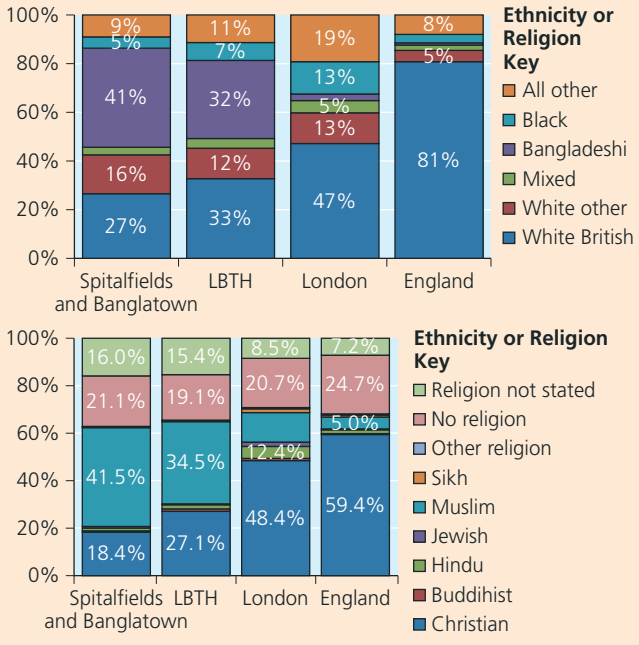


Figure 8.39 Compound bar graphs showing ethnicity and religion in Spitalfields and Banglatown from 2011 census.

Like many areas of inner London, Spitalfields and Banglatown ward has a notably high proportion of its population in the 20–40-year-old bracket. Access to a range of jobs in London attracts economically active people including immigrants. The area is also perceived as a trendy youthful hub.

In 2011 residents of Bangladeshi origin accounted for 41 per cent of the population of Spitalfields and Banglatown, which is higher than the borough average. There is a lower proportion of residents who are white British than the borough average. In terms of religion, Muslims make up the highest proportion of the population at 41.5 per cent. This area has a rich ethnic and religious mix as a result of a series of migratory moves by different groups since the nineteenth century.

3. Economic characteristics

To build up an economic profile of your two places, you could research the following:

- levels of employment and unemployment
- the balance between different economic sectors: primary, secondary, tertiary
- Gross Disposable Household Income (GDHI) estimates
- house prices (property websites and the Land Registry)
- access to services for different economic groups (Go to https://fingertips.phe.org.uk/profile/public-health-outcomes-framework).

The Index of Multiple Deprivation (IMD)

The Index of Multiple Deprivation is a UK-government qualitative study measuring deprivation at small-area level across England. It can be used to show economic inequality between and within different places. The English Indices of Deprivation 2019 are based on 37 separate indicators organised across seven distinct domains of deprivation, which are combined using appropriate weights. The seven different dimensions of deprivation are:

- income
- employment
- health deprivation and disability
- education skills and training
- crime
- barriers to housing and services
- living environment.

Note that these statistics are a measure of relative deprivation, not affluence. Not every person in a highly deprived area will themselves be deprived. Likewise, there will be some deprived people living in the least deprived areas. The IMD data for 2019 has been mapped at http://maps.cdrc.ac.uk/. Select the map showing the Index of Multiple Deprivation data for 2019.

Torquay

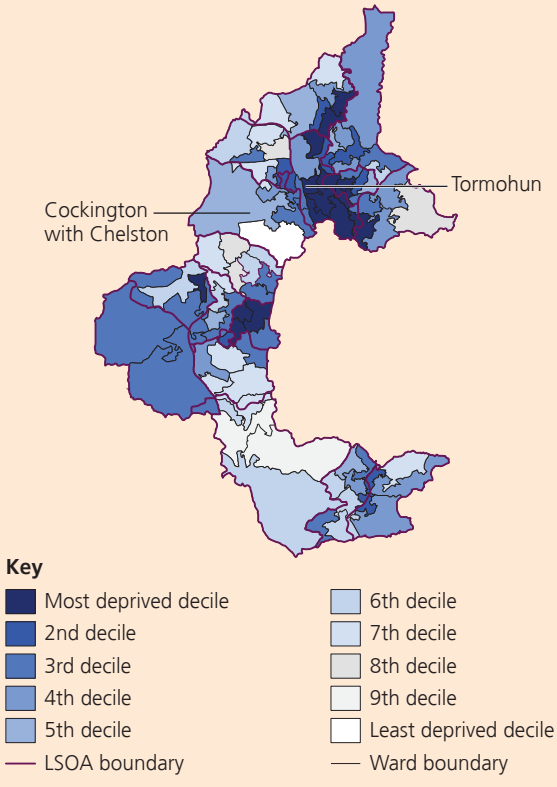


Figure 8.40 Map showing the Index of Multiple Deprivation data for Torbay, 2019

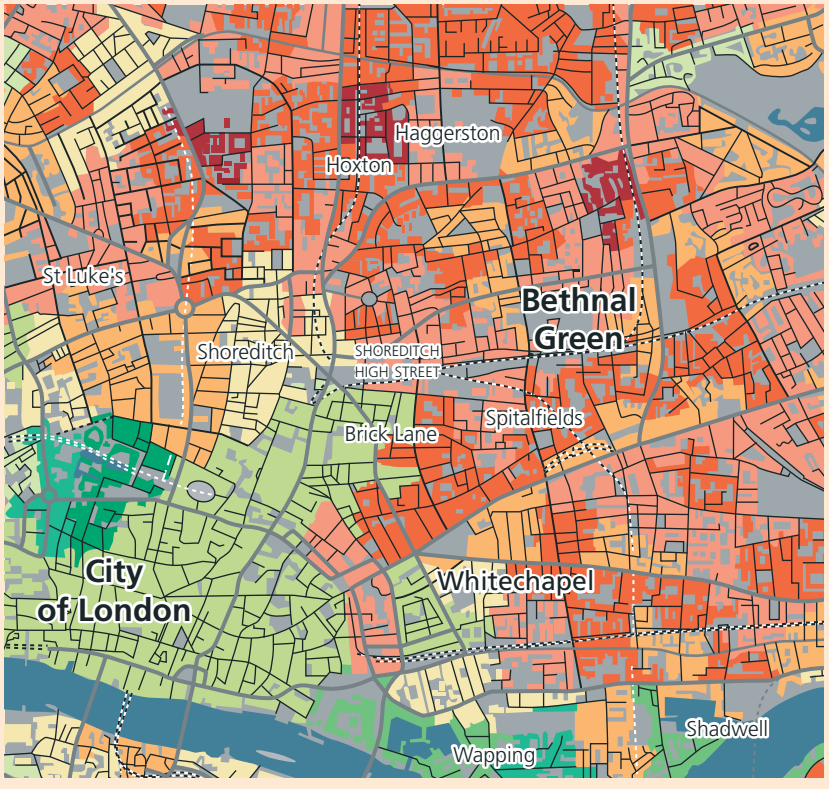
In 2019, Torbay was within the top 20 per cent most deprived local authority areas in England for the rank of average score and the rank of local concentration. A number of areas within Torquay fell into the top 10 per cent most deprived areas, as shown by Figure 8.40. The figure highlights the usefulness of **GIS** and **geospatial data** in identifying the variation in deprivation across the area.

Although the IMD data highlights significant areas of deprivation in Torquay, it is clear that, like most urban areas, it also has areas of great affluence.

The data in Table 8.2 has been collected from a number of sources and highlights the significant income inequality found in Torquay. In doing so, it also emphasises the need to look closely for variation *within* areas as well as between areas.

Strategies such as the redevelopment of Torquay harbour and the designation of Torbay as an urban GeoPark have had some success in reviving the town's fortunes in recent years. Future research could focus on the social and economic changes brought about by strategies like these, as well as looking at the impact of the new South Devon link road which has brought greater accessibility to Torbay.

Brick Lane and Spitalfields



Key
The Indicies of Deprivation for England in 2019 - showing the overall rank, split into deciles

Most deprived decile | 2nd | 3rd | 4th | 5th | 6th | 7th | 8th | 9th | Least deprived decile | Data missing

Figure 8.41 Map showing the Index of Multiple Deprivation data for Tower Hamlets 2019

Like all inner urban areas, there are some areas of significant deprivation in Tower Hamlets, but this has improved over recent years. In 2010, many areas of Tower Hamlets were in the most deprived 10 per cent in England, but in 2019 the majority of the most deprived areas were outside of London. Only 2 of the 20 wards in the borough of Tower Hamlets are now in the 50 least deprived wards in London. Gentrification has accounted for some of this trend, as middle-class families have moved into traditionally poorer parts of London while the poor have moved to boroughs on the outskirts of London where housing costs and rents are lower. The area of Spitalfields and Banglatown has experienced a significant reduction in levels of deprivation. In 2010, it was ranked as the most deprived Lower Super Output Area (LSOA) in all of London, but by 2019 it was ranked 95 out of 633 of the LSOAs in the capital. Figure 8.41, an example of **geospatial data**, shows the considerable variation even within this area.

Table 8.2 Statistical comparison of least and most deprived areas in Torquay

Measure	Least deprived small area (Cockington with Chelston)	Most deprived small area (Tormohun)
Life expectancy at birth–years (2018)	81.2 (male) 84.5 (female)	75.5 (male) 81.6 (female)
Index of multiple deprivation score (2019)	5.4	64.0
Percentage income deprived (2019)	13.3%	26.5%
Percentage employment deprived (2019)	11.7%	24.0%
Jobseekers Allowance/ Universal Credit Claimant Rate (2015–2019)	1.7%	5.1%
Percentage of children in low income families (2013–2016)	16.4%	27.9%

Source: Indices of Multiple Deprivation, 2019, *ONS*

4. Social characteristics and inequalities

The social aspects of the Index of Multiple Deprivation can be looked at in greater detail.

- Local educational data such as school performance can be found at www.education.gov.uk/schools/performance/
- Crime data is available at www.police.uk.
- For health-related data, go to https://fingertips.phe.org.uk/profile/health-profile (see Figure 8.43 and Figure 8.44). This excellent website details the wider determinants of health including social and economic indicators. The data are available at a range of geographical scales and the website enables you to map the data and compare different areas.

Health profiles are also useful. These give a snapshot of health for each local authority in England and are published annually.

Within both Torbay and Tower Hamlets, life expectancy at birth is not evenly distributed across the populations.

Figure 8.42a shows that the gap between the most and least deprived communities in Torbay can be as much as 11 years but averages out to between 7.2 and 7.5 years when comparing gender groups. In Tower Hamlets (Figure 8.42b) the gap between the most and least deprived communities is 12.7 years across both gender groups, whilst the gap for men is up to 10 years and 6.5 years for women.

The 2019 health summary for Torbay shows that there are a number of health indicators in which Torbay performs worse than the England average. Most of these are linked to economic deprivation and inequality.

The health summary for Tower Hamlets gives a more mixed picture. Whilst the district remains among the 20 per cent most deprived in England, some health indicators are worse than the national average and other indicators are better. This perhaps reflects the recent changes to the population in the borough, with some areas experiencing investment and gentrification.

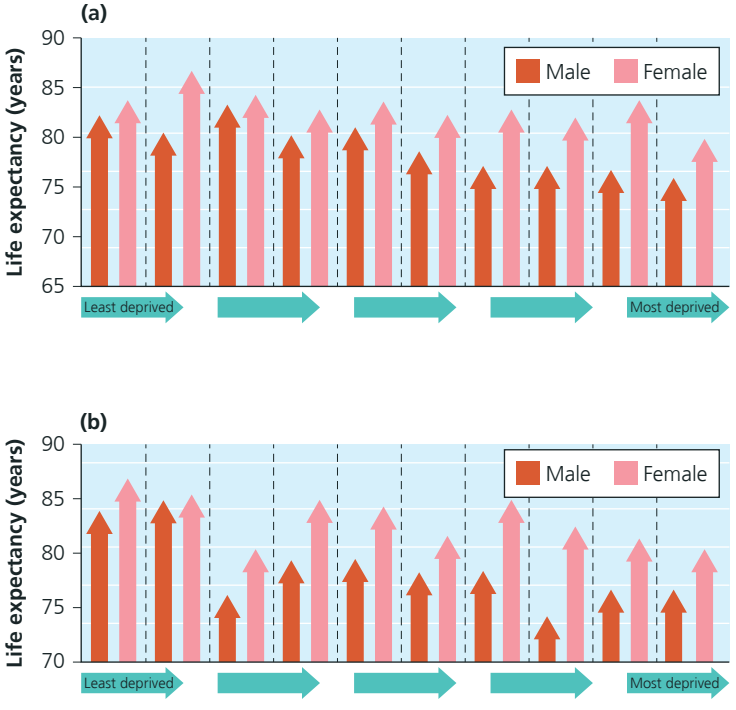


Figure 8.42 Life expectancy at birth by gender and deprivation decile, 2012/14 – Torbay (a) and Tower Hamlets (b)

Compared with benchmark: Better | Similar | Worse | Not compared

Indicator	Period	England	Southwest region	Bath and northeast Somerset	Bournemouth	Bristol	Cornwall	Devon	Dorset	Gloucestershire	North Somerset	Plymouth	Somerset	South Gloucestershire	Swindon	Torbay	Wiltshire
1.01i – Children in low income families (all dependent children under 20)	2016	17.0	13.8	9.8	16.4	19.7	16.3	12.3	12.2	12.3	12.4	19.6	12.7	10.0	13.8	**20.8**	10.2
1.01ii – Children in low income families (under 16s)	2016	17.0	14.0	9.8	16.4	19.7	16.4	12.5	12.5	12.6	12.6	20.0	12.9	10.2	14.1	**21.2**	10.4
B02a – School readiness: The percentage of children achieving a good level of development at the end of reception	2018/19	71.8	72.0	74.9	-	70.6	70.3	72.7	-	71.9	74.7	68.3	71.4	76.8	71.2	**70.8**	72.4
B02a – School readiness: The percentage of children with free school meal status achieving a good level of development at the end of reception	2018/19	56.5	53.0	46.5	-	53.2	51.1	54.7	-	52.4	55.2	54.9	53.0	49.5	54.3	**53.3**	48.8
B02b – School readiness: The percentage of children achieving the expected level in the phonics screening check in Year 1	2018/19	81.8	81.6	81.5	-	81.9	80.7	82.2	-	81.2	83.5	83.0	80.5	84.9	80.1	**84.0**	80.1
B02b – School readiness: The percentage of children with free school meal status achieving the expected level in the phonics screening check in Year 1	2018/19	70.1	67.5	61.6	-	70.3	66.4	65.5	-	65.4	66.2	69.5	67.6	68.8	66.2	**75..4**	63.5
B03 – Pupil absence	2018/19	4.73	4.87	5.09	4.94	5.28	4.98	4.78	4.90	4.53	4.87	5.14	5.11	4.57	4.57	**5.11**	4.72
B04 – First-time entrants to the youth justice system	2018	238.5	241.9	190.4	390.7	402.6	146.3	256.3	312.6	72.3	209.8	305.2	222.6	126.8	382.1	**366.3**	248.8
B05 – 16–17-years-olds not in education, employment or training	2018	5.5	6.4	7.1	6.2	7.7	6.0	5.6	4.5	5.8	5.9	6.4	8.8	5.5	7.5	**6.1**	8.1

Figure 8.43 Snapshot of public health outcome data for a range of places in south-west England. (Data for Torbay has been highlighted)

Compared with benchmark: Better | Similar | Worse | Not compared

Indicator	Period	England	London region	Hammersmith	Harringay	Harrow	Havering	Hillingdon	Hounslow	Islington	Kensington and Chelsea	Kingston upon Thames	Lambeth	Lewisham	Merton	Newham	Redbridge	Richmond upon Thames	Southwark	Sutton	Tower Hamlets	Waltham Forest	Wandsworth	Westminster
1.01i – Children in low income families (all dependent children under 20)	2016	17.0	19.3	21.3	22.2	13.9	16.2	16.3	14.3	31.1	21.1	11.9	23.6	22.7	13.4	21.5	15.5	8.8	23.4	10.0	**32.5**	20.1	18.1	29.0
1.01ii – Children in low income families (under 16s)	2016	17.0	18.8	20.6	21.3	12.9	16.5	16.0	13.8	30.6	20.5	11.7	23.4	22.6	13.1	20.1	14.7	8.5	23.2	9.8	**30.3**	19.4	17.2	27.3
B02a – School readiness: The percentage of children achieving a good level of development at the end of reception	2018/19	71.8	74.1	73.2	74.6	75.0	71.7	74.7	73.8	71.0	70.0	75.9	71.5	76.4	75.5	75.9	75.6	80.6	74.1	73.3	**69.9**	76.4	76.3	70.6
B02a – School readiness: The percentage of children with free school meal status achieving a good level of development at the end of reception	2018/19	56.5	64.1	62.8	64.7	65.8	49.7	63.7	66.5	64.2	60.4	55.8	61.5	66.0	62.3	70.5	63.1	58.2	66.7	56.5	**66.3**	65.1	60.0	67.7
B02b – School readiness: The percentage of children achieving the expected level in the phonics screening check in Year 1	2018/19	81.8	84.2	85.9	84.4	87.0	83.1	84.5	84.8	84.4	85.4	83.6	82.4	81.4	83.4	86.7	84.4	87.1	84.4	84.9	**84.0**	85.5	85.5	85.7
B02b – School readiness: The percentage of children with free school meal status achieving the expected level in the phonics screening check in Year 1	2018/19	70.1	76.2	79.0	76.6	75.2	70.5	75.1	76.1	78.8	78.8	64.7	73.2	69.9	72.4	85.2	75.6	70.8	77.5	72.7	**79.3**	80.8	75.8	80.6
B03 – Pupil absence	2018/19	4.73	4.47	4.95	4.58	4.28	4.57	4.79	4.39	4.90	4.75	4.10	4.52	4.31	4.18	4.51	4.54	3.94	4.26	4.30	**4.41**	4.63	4.39	4.72
B04 – First-time entrants to the youth justice system	2018	238.5	282.5	341.7	322.8	179.0	236.5	221.3	251.9	357.0	165.5	111.4	467.8	410.6	211.2	333.5	233.3	90.0	450.8	114.3	**376.2**	395.8	276.7	258.0
B05 – 16–17-years-olds not in education, employment or training	2018	5.5	4.8	1.5	14.9	2.1	3.1	6.9	3.4	5.5	5.8	2.7	10.0	4.5	2.3	5.2	3.3	2.9	9.9	3.4	**6.8**	3.1	8.8	2.0

Figure 8.44 Snapshot of public health outcome data for the borough of Tower Hamlets (highlighted) compared to other parts of London.

The chart below shows how the health of people in this area compares with the rest of England. This area's result for each indicator is shown as a circle. The average rate for England is shown by the red line, which is always at the centre of the chart. The range of results for all local areas in England is shown as a grey bar. A red circle means that this area is significantly worse than England for that indicator; however, a green circle may still indicate an important public health problem.

Significantly worse than England average
Not significantly different from England average
Significantly better than England average
Not compared

England worst — Regional average* England average — England best
25th percentile — 75th percentile

	Indicator names	Period	Local count	Local value	England value	England worst	England best
Life expectancy and causes of death	1 Life expectancy at birth (Male)	2014–16	n/a	78.7	79.5	74.2	83.7
	2 Life expectancy at birth (Female)	2014–16	n/a	82.7	83.1	79.4	86.8
	3 Under 75 mortality rate from all causes	2014–16	1,488	351.5	333.8	545.7	215.2
	4 Under 75 mortality rate from all cardiovascular diseases	2014–16	345	79.9	73.5	141.3	42.3
	5 Under 75 mortality rate from cancer	2014–16	598	138.9	136.8	195.3	99.1
	6 Suicide rate	2014–16	51	14.1	9.9	18.3	4.6
Injuries and ill health	7 Killed and seriously injured on roads	2014–16	128	32.0	39.7	110.4	13.5
	8 Hospital stays for self-harm	2016/17	426	362.8	185.3	578.9	50.6
	9 Hip fractures in older people (aged 65+)	2016/17	207	568.8	575.0	854.2	364.7
	10 Cancer diagnosed at early stage	2016	459	58.9	52.6	39.3	61.9
	11 Diabetes diagnosis (aged 17+)	2017	n/a	70.4	77.1	54.3	96.3
	12 Dementia diagnosis (aged 65+)	2017	1,526	63.2	67.9	45.1	90.8
Behavioural risk factors	13 Alcohol-specific hospital stays (under 18s)	2014/15–16/17	41	54.2	34.2	100.0	6.5
	14 Alcohol-related harm hospital stays	2016/17	1,167	841.4	636.4	1,151.1	388.2
	15 Smoking prevalence in adults (aged 18+)	2017	16,088	14.8	14.9	24.8	4.6
	16 Physically active adults (aged 19+)	2016/17	n/a	67.1	66.0	53.3	78.8
	17 Excess weight in adults (aged 18+)	2016/17	n/a	62.0	61.3	74.9	40.5
Child health	18 Under 18 conceptions	2016	52	25.7	18.8	36.7	3.3
	19 Smoking status at time of delivery	2016/17	178	15.2	10.7	28.1	2.3
	20 Breastfeeding initiation	2016/17	922	72.0	74.5	37.9	96.7
	21 Infant mortality rate	2014–16	18	4.4	3.9	7.9	0.0
	22 Obese children (aged 10–11)	2016/17	240	19.9	20.0	29.2	8.8
Inequalities	23 Deprivation score (IMD 2015)	2015	n/a	28.8	21.8	42.0	5.0
	24 smoking prevalence: routine and manual occupations	2017	n/a	26.6	25.7	48.7	5.1
Wider determinants of health	25 Children in low income families (under 16s)	2015	4,575	20.2	16.8	30.5	5.7
	26 GCSEs achieved	2015/16	763	56.6	57.8	44.8	78.7
	27 Employment rate (aged 16–64)	2016/17	56,900	76.0	74.4	59.8	88.5
	28 Statutory homelessness	2016/17	58	0.9	0.8		
	29 Violent crime (violence offences)	2016/17	3,393	25.4	20.0	42.2	5.7
Health protection	30 Excess winter deaths	Aug2013–Jul 2016	382	23.0	17.9	30.3	6.3
	31 New sexually transmitted infections	2017	541	690.4	793.8	3,215.3	266.6
	32 New cases of tuberculosis	2014–16	20	5.0	10.9	69.0	0.0

Figure 8.45 Health summary for Torbay, 2018

Source: Public Health England

Representations of place

Particular weight must be given in your place study to the use of qualitative sources. These will vary significantly depending on the place you are investigating but may include novels, poetry, nature writing and travel writing. Use the examples of Torquay and Brick Lane as ideas for possible sources you could investigate.

Artistic representations

Early artistic images of Torquay focused on the attractive coastal scenery of the area and this has been used to great effect in marketing tourism within the area.

Railway posters from the 1930s used idyllic images of sunny seaside holidays to inspire the public to travel to these destinations by train. More recently, the resorts of Torquay, Paignton and Brixham have been actively marketed as the English Riviera, evoking a sense of the Mediterranean, and the brand has used a strong graphic image of a palm tree. The original English Riviera palm tree logo was created by award-winning designer John Gorham in 1982 and quickly became the identifying icon for the English Riviera. The icon has been used in a more recent representation by award-winning Devon artist Becky Bettesworth (Figure 8.46).

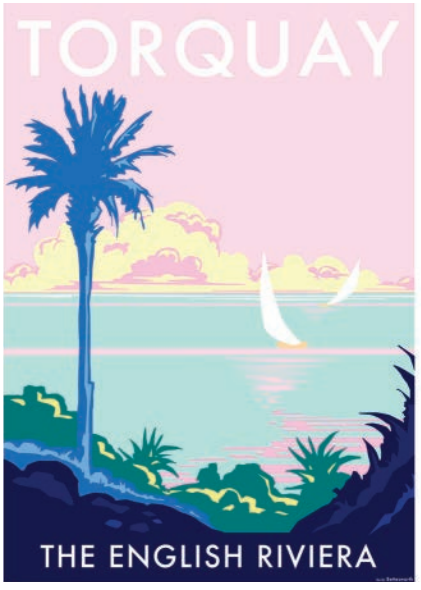


Figure 8.46 This recent representation of Torquay by award-winning Devon artist Becky Bettesworth features the icon of the Torquay palm
Source: www.beckybettesworth.co.uk

Figure 8.47 This doorway off Brick Lane shows the work of street artist STIK, depicting a burka-clad stick figure holding hands with a non burka-wearing stick figure. It is one of the few pieces of street art highlighting the cultural diversity of the area.

The area of Brick Lane and Spitalfields has developed a reputation for fashion and art, particularly street art (Figure 8.47). The *Ravish London* website suggests that this area constitutes 'the Mecca of London street art; a de facto open-air gallery, a square mile of street art to compare with the financial square mile.' The nearby area of Shoreditch saw an influx of artists such as Tracy Emin and Damien Hirst in the 1990s and their success seemed to attract more artists to the area. Brick Lane and Spitalfields have also witnessed a huge increase in the number of galleries opening, helped financially by the gentrification of the area and increasing tourist numbers. In 2011, local artist and cartographer Adam Dant unveiled his Map of Spitalfields Life (Figure 8.48). This map details key locations within the Spitalfields area today, but also the historic background of a variety of places. It includes characters who have made this area special, ranging from artist Tracey Emin to Fred, the chestnut seller. Dant's map provides a very strong sense of place.

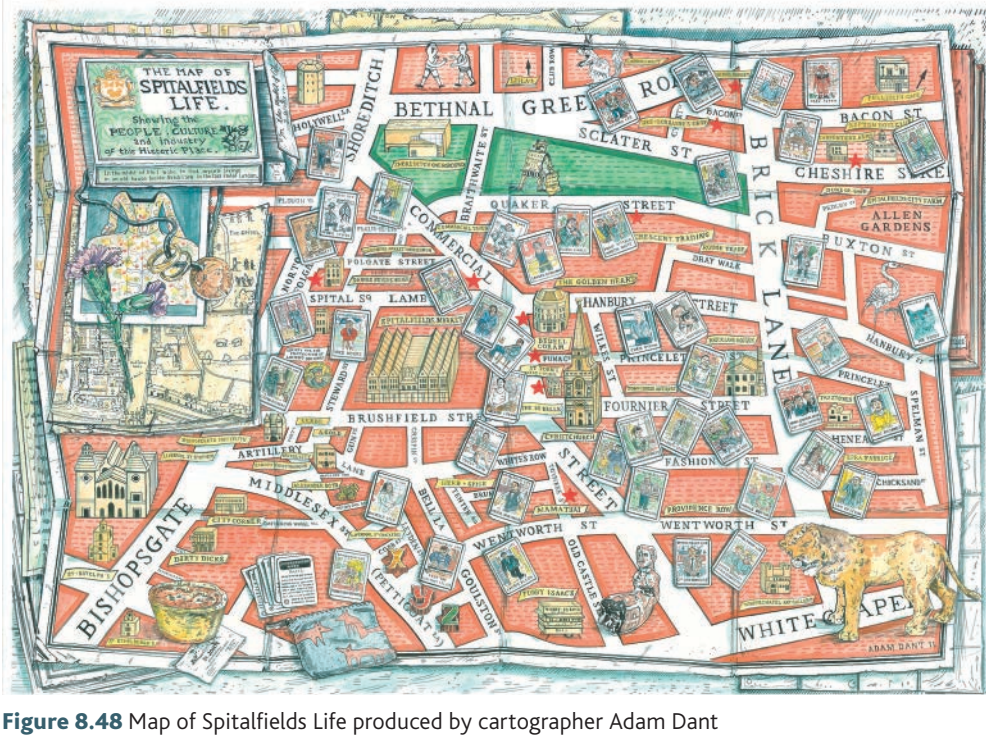


Figure 8.48 Map of Spitalfields Life produced by cartographer Adam Dant

Media representations

Media representations of Torquay have been mixed. For many people, Torquay will always be associated with the 1970s comedy series *Fawlty Towers* (BBC, 1975–1979) loosely based on an experience the actor John Cleese had while filming in Torquay and staying at the Gleneagles Hotel. The twelve episodes of *Fawlty Towers* were not actually filmed in Torquay. A more positive representation of Torquay is that of the crime writer Agatha Christie, who was born in Torquay. The connection with Agatha Christie has been promoted to bring a 'wider sense of glamour and style' to the Riviera brand and, according to the English Riviera website, the International Agatha Christie Festival held in Torquay, is 'fast becoming one of the UK's leading literary festivals.'

Public perceptions of Torquay are mixed. A survey carried out for the *Towards 2015* tourism report concluded that Torquay was perceived as having a 'classy past but an unclear present', while a range of positive and negative comments can be found on online discussion forums. Visitors to Torquay tend to be more complimentary. Typical posts from the *TripAdvisor* website refer favourably to the beautiful coastline, clean beaches and wide range of tourist attractions.

Media representations of Brick Lane are more numerous and more varied. The geographical area of Brick Lane and Spitalfields has numerous websites devoted entirely to it, detailing everything from the best curry houses to walking tours of the street art. The area has also become a popular tourist destination for visitors to London and features prominently in most guidebooks. The *Lonely Planet* writes: 'Full of noise, colour and life, Brick Lane is a vibrant mix of history and modernity, and a palimpsest of cultures. Today it is the centrepiece of a thriving Bengali community in an area nicknamed Banglatown. The southern part of the lane is one long procession of curry and balti houses intermingled with fabric shops and Indian supermarkets.' (Figure 8.49 and Figure 8.50)

In contrast, the website of the Old Truman Brewery located in the northern half of Brick Lane describes itself as being 'home to a hive of creative businesses as well as exclusively independent shops, galleries, markets, bars and restaurants.'

Figure 8.49 Brick Lane road sign written in both English and Bengali

Figure 8.50 View up Brick Lane

Other representations of Brick Lane come from those actually living or working there. Local poet Sally Flood has written extensively about Brick Lane. She was an embroidery machinist for most of her working life and much of her writing was done at the factory, from her seat with its window facing Brick Lane. Sally's memories of life in and around Brick Lane are documented on a number of websites and she has also published a collection of poems relating to this area.

Brick Lane is a mixture
of aromatic spices
Curries, onions and bad drains,
Pakistani restaurants
Jewish trimming shops
and betting shops,
Down-at-heel workers

and hopeful prostitutes,
cars and vans add to the pollution
with exhaust fumes.
Pavements and gutters
are littered with overspill
from dustbins and workshops.
This is where the immigrant
looks for fulfillment!
This is the breeding ground
for discontent,
Where the Meths drinker mixes
with the down-and-out,
Where Workers are exploited
and small-time drug peddlers
sell their dreams!
This is where the thug
dons the crown of King
and bullies thrive,
Where do-gooders
salve their consciences,
This is Brick Lane.

Sally Flood 'The Brick Lane I see', *Window on Brick Lane*, 1980

One of the more famous texts of recent times has been Monica Ali's novel *Brick Lane*, published in 2003 and subsequently made into a film (2007). The story follows the life of Nazneen, a Bangladeshi woman who moves to London at the age of 18 to marry. The fictional couple live in Tower Hamlets and the novel explores her new life in London and adaptations she makes as a Bangladeshi. The following extract gives an impression of the main features of Brick Lane and the changing prosperity of the area as it becomes a more fashionable neighbourhood.

'The bright green and red pendants that fluttered from the lamp posts advertised the Bangla colours and basmati rice. In the restaurant windows were clippings from newspapers and magazines with the name of the restaurant highlighted in yellow or pink. There were smart places with starched white tablecloths and multitudes of shining silver cutlery ... 'You see,' said Chanu. 'All this money, money everywhere. Ten years ago there was no money here.'

Monica Ali, *Brick Lane*

Other sources detailing the lived experience of place

Surveys and questionnaires

Lived experience of place can come from a range of sources.

- In 2008, for example, the Department for Communities and Local Government conducted a place survey to gauge people's perceptions of their local area and the local services they received. In Torbay, the survey found that 82 per cent of Torbay residents were very or fairly satisfied with their local area. This compared favourably with the national average of 80 per cent at this time and was linked to the high quality of life afforded by the natural local environment and geography.
- Since 2011, the Personal Well-being Survey has monitored personal well-being and life satisfaction. Geographical patterns have emerged from this survey but the questions do not relate specifically to place.

You might want to produce your own questionnaires or carry out your own interviews with local people about their lived experience or perceptions of place. Ask your family and friends to take part or you could post a questionnaire online. Social media sources such as Twitter, Facebook, Instagram and Pinterest can provide lived experience of place as people increasingly communicate online about places they have visited. Make sure you acknowledge the strengths and weaknesses of each source.

Blogs

Finally, blogs can be an excellent source of information about an area. In East London, for example, *The Gentle Author* writes a daily blog about the culture of the Spitalfields area and documents the lives of local people and places. 'How can I ever describe the exuberant richness and multiplicity of culture in this place to you?' @The Gentle Author, (www.Spitalfieldslife.com). The *Spitalfields Life* blog was born in 2009 with the words, 'In the midst of life I woke to find myself living in an old house beside Brick Lane in the East End of London.' Today, thousands of people from around the world read each daily posting.

Sources detailing economic and social change

Statistics, maps, newspapers, poems, photos and paintings are just some of the sources which can be used to show change in a place over time. Film or television footage may also be available for your local or distant place.

Photographic sources

Figures 8.51a and 8.51b shows how photographs can be used to document change over time. The photographs from the official Spitalfields Market website show the early market as a collection of sheds and stalls. The second photograph shows the exterior of the market after a major regeneration programme. The website also documents the changes in text, noting the market's evolution from selling fruit and vegetables in the seventeenth century to being a major tourist attraction today, where visitors 'will find designers/makers and artists selling fashions, homewares and accessories or a treasure trove of vintage and antique clothing, furniture and other wondrous oddments.'

Figure 8.51a The early Spitalfields market

Figure 8.51b Spitalfields market, 2015

Photographs have also been used to document gentrification in East London. Older photographs are available from the Tower Hamlets Local History Library and Archive and these have been used to create a series of before and after photographs for the same street. Search for these at http://spitalfieldslife.com/ which is regularly updated and provides fascinating information about the changes in this area.

Text sources

Published media

Gentrification can bring great wealth to an area but can also lead to the displacement of original residents as house prices rise and shops and services change to suit the tastes of the evolving socio-economic composition of the area. National newspapers reported on a series of anti-gentrification protests in and around Spitalfields in September 2015 and a series of letters written to *The Guardian* newspaper highlights major concerns about the impacts of gentrification in East London at this time:

> I was born in Whitechapel in 1974, to a family based in Shoreditch since the 1700s, and the stories and people, the spirit of that 1/4 square mile. Where historically class lines merged, cultures crossed, the rich brushed shoulders with Jack London. The hipsters have my approval for saving the architecture of the area, although I lay that more in the hands of the late 80s British art wave. But it's not the East End: inclusive, classless, tough, oozing character and fear in equal measure, pleasure and peril greeted every turn, never a yard from poverty or riches, sickness or health, life or death.
>
> The current occupants have no verve or spirit, the survival with a smile has gone. It has always been driven by its association with the docks, the clay pits, the silk weavers; a hot bed of immigration and integration. But it no longer has a fraction of the character ... only the scenery remains for us Londoners to easier imagine the feeling of being on the edge of something dangerous or amazing at every turn. Takes a bit more than a beard for that.

Another,

> I'm getting the impression that the 'them and us' mentality is coming from both sides. I live in Hackney with flatmates. None of us are born-and-bred Londoners, but have been here for some 3–8 years. They are British. I'm just a foreigner here for a while. None of us is a 'hipster' either ... I live in

East London because I like the diversity. Not sure where else I'd live, because whenever I go west, it seems so white and it's all coffee chains.

The Guardian website, 24 March 2015

Careful evaluation of sources like this is required but these comments do highlight some of the key issues concerning in-migration and gentrification in this area at this time. In addition, they serve as another record of (some) people's lived experience of place.

Poetry

Finally, we return to the poetry of long-time East London resident Sally Flood. We have already experienced a sense of place in her depiction of Brick Lane. She provides a similar emotional response to place in her writing about the changes occurring in her local landscape as a result of large-scale redevelopment. Many residents welcome such redevelopment schemes as they can improve the physical and economic landscape of a place. Others lament the loss of community and argue that many developments fail to address the capital's housing shortage or help local businesses.

This small street is fast receding
Growing smaller, day by day
Over years I watch it crumble
Watch the houses swept away,
Streets that share my childhood moments
Saw the war years, heard the bombs
Proudly held itself together
Now prepares itself to die.

An extract from 'Mount Terrace', *Tales by Eastenders,* 2015

The sources outlined here are by no means comprehensive but do enable an holistic account of these two places. Take care to evaluate all the sources you use and go through the checklist below to ensure you have used as many different sources of information as possible. Finally, do not forget your own lived experience of place. To be inside a place is to belong to it and identify with it more strongly. Use this to your advantage.

Table 8.3 Data sources

Data source	Examples
Statistics	Census data, employment data
Maps	Ordnance Survey, Google Maps
Geo-located data	Interactive transport app
Geospatial data	Index of Multiple Deprivation, Public Health Determinants of Health
Photographs	Photographs showing change over time
Text, from varied media	Poetry, novels, social media sources including blogs
Audio-visual media	Film archives
Artistic representations	Paintings, sculpture
Oral sources	Interviews, reminiscences, songs

Review questions

12 What qualitative sources were useful for studying the place characteristics of:
   a your local place
   b your contrasting place?

13 What sources of information would you use to investigate the socioeconomic characteristics of a place?

14 Explain the limitations of quantitative data when investigating the lived experience of place.

Further reading

The first three sources are a particularly good starting place for the study of place:

Anderson, J. (2015) *Understanding Cultural Geography: Places and Traces* (Routledge)

Cresswell, T. (2014) *Place: An Introduction* (Wiley-Blackwell)

Phillips, R. (2016) 'Changing Place; Changing Places' A-level subject content overview from www.rgs.org/

Other useful texts include:

Cloke, P. Crang, P. and Goodwin, M. (2014) *Introducing Human Geographies,* Third edition (Routledge)

Freestone, R. and Liu, E. (2017) *Place and Placelessness Revisited* (Routledge)

Horton, J. and Kraftl, P. (2014) *Cultural Geographies: An Introduction* (Routledge)

Jackson, P. (1989) *Maps of Meaning* (Routledge)

Kunstler, J. (1995) *The Geography of Nowhere* (Simon & Schuster)

Massey, D. (1994) *Space, Place and Gender* (Polity)

Oakes, S. (2018) *A-level Geography Topic Master: Changing Places* (Hodder Education)

Rawling, E. (2011) 'Reading and writing place: a role for geographical education in the twenty-first century' in Butt (Ed.) *Geography, Education and the Future* (Bloomsbury)

Relph E. (2008) *Place and Placelessness,* SAGE

Tuan, Yi-Fu (2001) *Space and Place* (University of Minnesota Press)

Fieldwork opportunities

- Investigate how different places are experienced by different people by carrying out a walking tour of a particular place with fellow students. Record your individual feelings and observations as you walk through different areas. Back in class, discuss the reasons why individual students experience places in different ways.
- Geographical areas which have experienced significant immigration or in-migration can be investigated in terms of their changing demographic and cultural characteristics. Follow the example of Doreen Massey and walk down your local high street observing and recording the different groups of people, the shops and buildings within your locality.
- Investigate whether your local place has become a 'clone town'? How has your local place evolved over time and is there evidence of globalisation?
- The ways in which different groups of people experience and perceive places differently can be investigated through the use of surveys and interviews. Examples include the ways in which people with disabilities experience places, but there are other social groups who will have particular perspectives on place.
- Explore the extent to which the views of local people match those of the local government and corporate bodies involved in marketing and regenerating places.
- Compare the ways in which different places are marketed and represented. Investigate the different strategies used to improve perceptions of place.
- Investigate the impacts of industrial decline on different places in the UK. How does the character of these places change? How does it affect people's lived experience of that place?
- There is a wealth of information available for different parts of London and the impacts of the 2012 Olympics could be a focus of fieldwork here. 'The East End of London is a world in itself,' wrote Charles Dickens in the nineteenth century. Little has changed today and the economic and cultural diversity of London lends itself well to further research.

Question practice

A-Level questions

1 Outline how shifting flows of different factors shape the characteristics of places. [4 marks]

2 Figures 8.52 and 8.53 show information taken from the 'Community Life Survey 2018–2019' published by the UK Department for Digital, Culture, Media and Sport. Analyse the data shown in Figures 8.52 and 8.53. [6 marks]

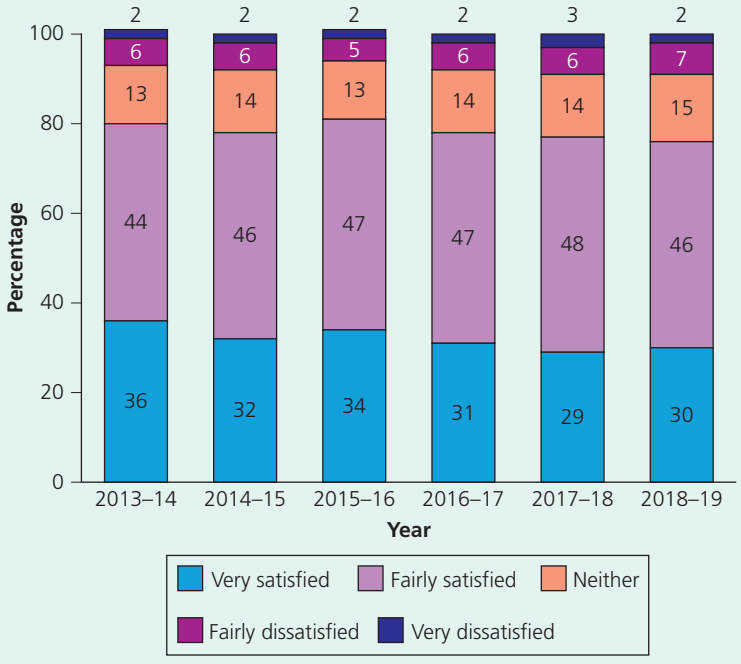


Figure 8.52 Responses to the question: 'Overall, how satisfied or dissatisfied are you with your local area as a place to live?

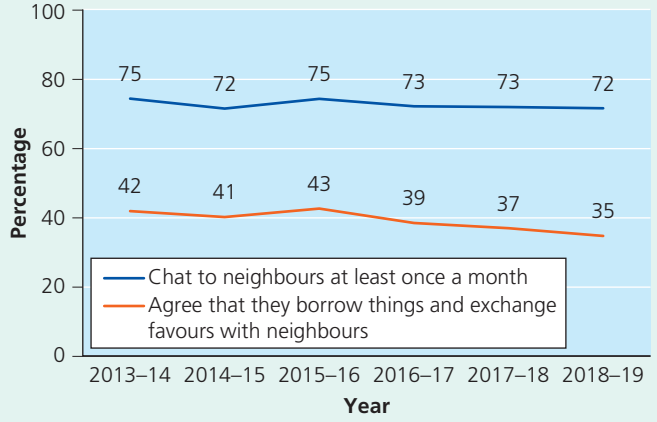


Figure 8.53 Data about the percentage of adults (16+) who chat to their neighbours at least once a month (more than just to say hello), and the percentage who agree they borrow things and exchange favours with their neighbours

Figure 8.54 'Amongst the Nerves of the World' (circa 1930), a painting of Fleet Street, London, by CRW Nevinson

3 Figure 8.54 shows a painting of Fleet Street in London painted by CRW Nevinson around 1930. Using Figure 8.54 and your own knowledge, assess the usefulness of sources such as this in helping to understand the character of a place. [6 marks]

4 'Past processes of development are more important than present processes in influencing the social and economic characteristics of a place.'

Assess the extent to which you agree with this statement with reference to either your local or a distant place. [20 marks]

5 'Globalisation has had a significant impact on shaping the lived experience of place.'

Assess the extent to which this statement is true in both your local and distant place. [20 marks]

Chapter 9

Contemporary urban environments

Population levels are rising, particularly in urban areas. According to the World Bank, in 2018, 55.3 per cent of the total global population were living in urban areas, up from 38 per cent in 1976, although the rate of population growth is slowing – down from 2.46 per cent in 1976 to 1.94 per cent in 2018. By 2050 it is expected that 66 per cent of the global population will be urban with close to 90 per cent of this increase in Asia and Africa. Cities can be efficient. It is easier to provide basic services such as water and sanitation to people living closer together. Likewise, access to health, education and other social and cultural services is more readily available. However, as towns and cities expand, the cost of meeting basic needs increases with a greater strain on the environment and natural resources.

Urban areas are also experiencing the pressures of migration. Globalisation and the impacts of climate change have transformed agricultural practices and in low- and middle-income countries, increasing numbers of people are leaving rural areas to find employment in cities. Unemployment rates are high and the proportion of people living in slum areas has rocketed. Issues of social cohesion are evident as variations in wealth and ethnicity can sometimes lead to hostility.

The future survival of cities depends on sustainable growth and their ability to tackle the major issues such as provision of affordable housing, unemployment, pollution and waste disposal, transport and social inequality.

In this chapter you will study:

- urbanisation and **urban growth**: patterns, reasons and outcomes
- urban forms and the social, economic and environmental issues associated with urban growth
- the impact of the urban environment on local climate, weather and drainage
- features of sustainable urban growth.

9.1 Patterns of urbanisation since 1945

Urbanisation is defined as the process by which an increasing proportion of a country's population lives in towns and cities.

The urban population has grown rapidly from 746 million in 1950 to 3.9 billion in 2014. The most urbanised regions include Northern America (82 per cent living in urban areas in 2018), Latin America and the Caribbean (81 per cent) and Europe (74 per cent). In contrast, Africa and Asia remain mostly rural, with 43 and 50 per cent of their respective populations living in urban areas. This is likely to change over the next 50 years as the fastest growing urban areas are currently found in Asia and Africa. The World Bank predicts that between 2018 and 2050, India, China and Nigeria will account for 35 per cent of the growth of the world's urban population (India: 416 million, China: 255 million and Nigeria: 189 million). India is projected to add 404 million urban dwellers, China 292 million and Nigeria 212 million. The total world urban population is expected to surpass six billion by 2045 and much of the expected growth will occur in low-income countries. The fastest growing urban areas are medium-sized cities and cities with less than one million inhabitants but some reports also highlight the fact that many areas projected to be urban in 2040 have not yet actually been built. In India alone, it is predicted that 70 per cent of cities have yet to be built.

In recent years, some cities have experienced population decline caused by:

- economic contraction (e.g. Detroit, USA)
- low fertility rates (e.g. Nagasaki, Japan and Busan, South Korea)
- emigration (e.g. Poznan, Poland)
- natural disaster (e.g. New Orleans, USA).

Megacities

One striking feature of the last 30 years has been the rapid development of **megacities** – urban areas with a population of more than 10 million people. In 1990, there were 10 megacities; in 2020 there were 30; and by 2030 the UN predicts there will be 43 megacities – housing over 13 per cent of the global population. As Figures 9.1–9.3 show, the development of megacities is largely concentrated in Asia. In 2015, Tokyo was the world's largest city with over 37.4 million inhabitants. There were another eight cities – Delhi, Shanghai, São Paulo, Mexico City, Dhaka, Cairo, Beijing and Mumbai – all with populations in excess of 20 million and rising rapidly. Between 2018 and 2030, the population of Delhi is projected to increase by more than 10 million, while Tokyo is projected to decline by almost 900,000. Such settlements can be defined as **metacities**, the term given to a conurbation (continuous built-up area) of more than 20 million people. The Chinese government has plans to merge nine cities in the Pearl River Delta creating an urban area 26 times larger than Greater London.

Key terms

Megacity – A city or urban agglomeration (urban area incorporating several large towns or cities) with a population of more than 10 million people. According to the UN, London achieved megacity status in 2013. This classification included residents in the Greater London area.

Metacity – A conurbation with more than 20 million people.

Urban growth – An increase in the number of urban dwellers. Classifications of urban dwellers depend on the census definitions of urban areas, which vary from country to country. They usually include one or more of the following criteria: population size, population density, average distance between buildings within a settlement and legal and/or administrative boundaries.

Urbanisation – An increase in the proportion of a country's population that lives in towns and cities. The two main causes of urbanisation are natural population growth and migration into urban areas from rural areas.

The causes of urban growth

The process of urbanisation plays an important role in human affairs. It has historically been linked to other important economic and social transformations

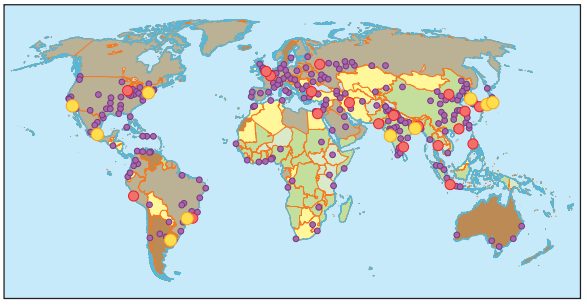


Figure 9.1 Percentage of urban population and city size, 1990

Source: UN Population Division

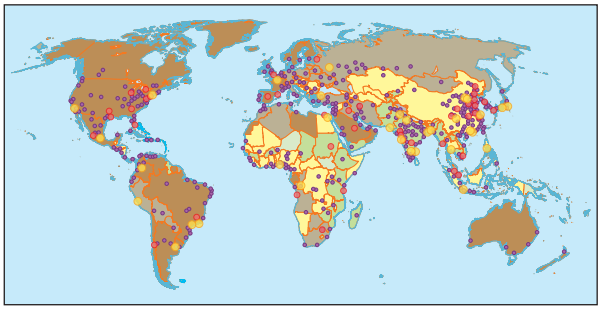


Figure 9.2 Percentage of urban population and city size, 2018

Source: UN Population Division

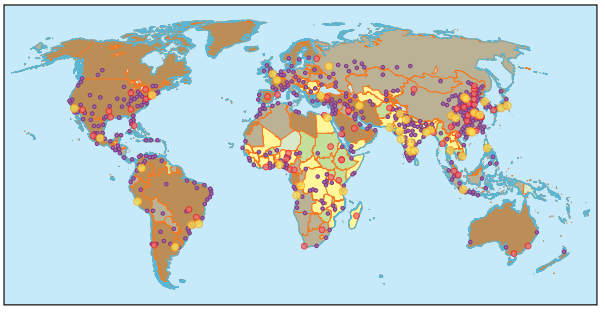


Figure 9.3 Percentage of urban population and city size projected to 2030

Source: UN Population Division

which have brought about greater geographic mobility, lower fertility and longer life expectancy. Cities also play an important role in reducing poverty. They hold much of the national economic activity, government institutions, business and transportation, and have higher levels of education, better health, easier access to social services, and greater opportunities for cultural and political participation. In 2020, São Paulo, Brazil's financial capital, represented 10.7 per cent of all Brazilian Gross Domestic Product (GDP) and was home to 63 per cent of the transnational corporations in Brazil. Likewise, in Kenya, Nairobi, with 8.4 per cent of the country's population, accounted for almost 20 per cent of the country's GDP.

Natural population growth

Urban areas have relatively young age profiles. Across the world, it has traditionally been young adults (15–40 years) migrating for higher paid jobs, better educational opportunities and greater social and cultural diversity. Between 2001 and 2011 the population of large city centres in England and Wales more than doubled, with residents aged 22–29 nearly tripling to make up almost half of their total population. These migrants are in their fertile years – the years during which people have children – and so the rates of natural increase are higher in cities than in the surrounding rural areas. In London, an area stretching from Clapham, south of the River Thames, westwards to Fulham, north of the river has been termed 'Nappy Valley' due to the high proportion of young families living here. In the past, professional couples with young children would have moved out to the suburbs when they could afford it but the rising costs and time involved in commuting has encouraged more young families to remain in the city.

Rural–urban migration

The reasons for rural–urban migration are often divided into 'push' and 'pull' factors. **Push factors** cause people to move away from rural areas, whereas **pull factors** attract them to urban areas. In low-income countries, push factors tend to be more important than pull factors.

Push factors are largely due to poverty caused by:

- population growth, which means the same area of land has to support increasing numbers of people, causing over-farming, soil erosion and low yields
- agricultural problems, including desertification because of low rainfall, systems of inheritance that cause land to be subdivided into small plots, systems of tenure and debt on loans taken out to support agricultural change
- high levels of local diseases and inadequate medical provision
- agriculture is increasingly being organised globally. Land previously used to grow food for local people is now used to produce cash crops for sale to higher income countries. Many traditional rural communities have been driven off their land and into cities
- natural disasters such as floods, tropical storms and earthquakes – people flee rural areas and do not return
- wars and civil strife cause people to flee their land.

Pull factors include the prospect of:

- employment in factories and service industries (for example, hotels) which is better paid than work in rural areas. There is an increasingly high demand for unskilled labour in cities
- earning money from the informal sector, for example, selling goods on the street or providing transport (taxi/rickshaw driver)
- better quality social provisions, from basic needs such as education and healthcare to entertainment and tourism
- a perceived better quality of life in the city, fed in part by images in the media.

Skills focus

Analyse the changes in:

- the percentage of urban population
- city size

between 1990 and 2030 as shown/projected in Figures 9.1 and 9.3.

Consequences of urbanisation and urban growth

Problems with housing, traffic, waste disposal, crime and pollution can be found in cities all over the world irrespective of their economic status. Instead, these issues tend to be linked to the unique geographical circumstances of the city, such as topography, climate

and function. The following section details the key consequences of urbanisation and characteristics of cities in the twenty-first century.

A. Urban sprawl

Urban sprawl is defined as the spread of an urban area into the surrounding countryside. It has been linked to the processes of urbanisation and **suburbanisation** and has traditionally occurred in an uncontrolled and unplanned fashion.

Urban sprawl has many negative impacts:

- Urban sprawl requires more roads and infrastructure such as pipes, cables and wires. It is less economically efficient to service low-density rural areas compared to compact urban developments with the same number of households.
- The reach of urban sprawl into rural areas ranks as one of the main causes of wildlife habitat loss.
- Urban sprawl causes more commuting from the suburbs to the city and thus more fuel consumption and traffic congestion.
- Urban sprawl can increase air pollution since a more car-dependent lifestyle leads to increases in fossil fuel consumption and emissions of greenhouse gases. The areas may also experience higher temperatures in line with the urban heat island effect.
- Sprawl has contributed to the loss of farmland and open spaces, which in turn has led to the loss of fresh local food sources with greater food miles as a result.
- Urban sprawl can have a serious impact on water quality and quantity. Covering the countryside with impermeable surfaces means that rainwater is unable to soak into the ground and replenish the groundwater aquifers. In addition, it can lead to greater water run-off and increased flood risk.
- In addition to the movement of people to the suburbs, one other important component has been the accompanying movement of industry and businesses, including retail companies. This is referred to as **decentralisation** and this outward movement has been blamed for the decline of retail in some city centres and an increasing **homogenisation** of the landscape, where cities become indistinct from one another. American cities in particular have witnessed the huge growth of large edge-of-city complexes including shopping malls and leisure areas. In some cases, new self-contained settlements have developed beyond the original city boundary. These are known as **edge cities**.

B. Shortage of housing in lower-income countries

Population density tends to be high in urban areas and one of the consequences of this is a shortage of accommodation, leading to the presence of large areas of informal and often inadequate housing. These normally develop on the edge of the city or in areas of low land value prone to environmental hazards such as flooding or landslides. They may also be found adjacent to transport networks or in areas suffering high levels of air, noise or water pollution. These settlements tend to have limited access to basic infrastructure such as water, electricity and waste disposal and a lack of services such as health centres and schools.

In 2017, it was estimated that about 900 million people lived in slums. That number rises to 1.6 billion if you include all the different types of informal settlements. This is approximately 25 per cent of the world's urban population. By 2030, it is estimated that one in four people globally will live in a slum or other informal settlement. Mumbai alone had more than nine million slum dwellers in 2020. The number of people residing in slums throughout India is estimated to be 104 million, or nine per cent of the total population.

Informal settlements take on different names and different forms depending on their location. The settlements lining the hills of Rio de Janeiro in Brazil are called favelas; in parts of India they are known as bustees and in West Africa, bidonvilles. The use of the term 'slum housing' has been criticised by those who see it as a political label. In Mumbai, for example, settlements such as Shivaji Nagar and Dharavi are referred to by some as 'homegrown neighbourhoods'. The argument is that they were developed gradually, by the people who live there, with the help of local artisans of construction, and usually with little or no support from the authorities.

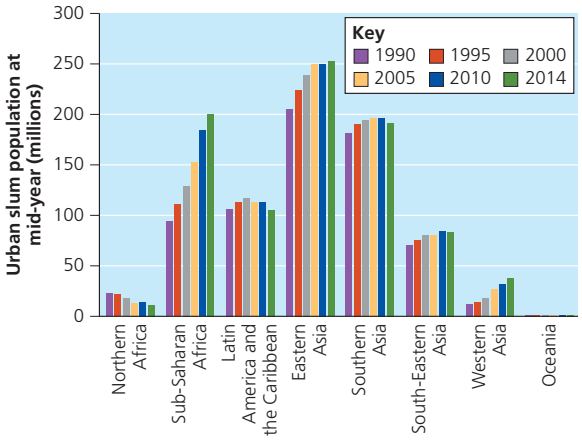


Figure 9.4 The proportion of urban population living in slums, 1990–2014

Source: The New Climate Economy

Figure 9.4 shows that the changes in the proportion of people living in urban slums varied globally. For example, in Sub-Saharan Africa it continued to grow at an ever-increasing rate while in other regions it slowed down or even fell (for example, Northern Africa). This should be viewed in a positive light alongside the relative success of Millennium Development Goal 7 (Ensure environmental sustainability) to improve the lives of slum dwellers. Between 2000 and 2014, for example, the UN reports that more than 320 million people living in slums gained access to improved water sources, improved sanitation facilities or durable or less crowded housing. However, the total number of people living in slums has continued to increase and one of the targets in the post-2015 sustainable development goals is to 'ensure access for all to adequate, safe and affordable housing and basic services, and upgrade slums by 2030.'

In the past, the most extreme strategy adopted by authorities was to eradicate slums. However, this simply moved the problem elsewhere and a more common approach became to acknowledge the presence of slums and provide help in the form of materials or services. These self-help or 'site and service' schemes have proved remarkably successful in some cities but the quantity and/or quality of housing remains inadequate in most urban areas around the world. More recent initiatives have been **slum upgrading programmes.** These seek to improve the slums in partnership with local NGOs and development organisations. They focus on securing rights for dwellers, formalising land tenure rights and providing basic amenities such as electricity, water and waste disposal. **Shack/Slum Dwellers International (SDI)** is an organisation which gives a voice to those living in informal settlements and links up poor communities across Africa, Asia and Latin America. The idea is for slum dwellers to share their knowledge and expertise so that they are not excluded from the economic and political processes happening in their cities.

Dharavi, Mumbai, India: the setting for Slumdog Millionaire

Dharavi, home to more than one million people, is Asia's largest slum. It is located on prime real estate in the centre of Mumbai. Many of the residents are second-generation residents, whose parents moved there many years ago.

The maze of narrow dirty lanes, open sewers and tightly packed huts is a far cry from its roots as a fishing village. Most of the huts have a constant supply of electricity, a gas stove for cooking and many even have a small colour television with a cable connection.

Dharavi is a thriving hub of small-scale manufacturing industries. Small manufacturing units produce leather goods, aluminium bricks and pottery and also recycle plastics for both the domestic and international markets.

The state government has a £1.1 billion plan to redevelop the area and transform it into a modern township, complete with proper housing and shopping complexes, hospitals and schools.

Figure 9.5 Dharavi – Mumbai, India

C. Shortage of affordable housing in higher-income cities

Shortage of affordable housing is a key problem in many wealthy cities. In the UK, the rising demand for accommodation in cities has led to a dramatic increase in both house prices and rental costs. In London, average house prices rose by almost 70 per cent between 2010 and 2019, fuelled partly by in-migration, **gentrification** and by the purchasing of properties by wealthy foreign investors. The city is a major global hub and overseas investors buy properties in London to diversify their international portfolio.

D. Lack of urban services and waste disposal

Financial restraints in lower-income countries can result in a lack of basic services such as water and electricity. Maintenance of infrastructure such as roads, sewers and drainage is also limited and this can result in traffic congestion, polluted watercourses, flooding and the rapid spread of disease. In India, old pipeline infrastructure has not kept pace with urbanisation, resulting in large urban areas devoid of planned water supply and sewage treatment. In 2015, only five per cent of piped water reached the slum areas in 42 Indian cities and towns, including New Delhi.

Waste disposal poses a further problem. The amount of waste increases year-on-year, but there are economic, physical and environmental restraints on its disposal. In Addis Ababa, Ethiopia, the city authority is only able to deal with two-thirds of the rubbish created by the rapidly growing urban population. The rest is left to private contractors to collect or is simply dumped on streets and in rivers, creating a huge health hazard. At Koshe, a huge landfill site on the edge of the city, families live in makeshift housing and search for anything they can use for themselves, or resell (Figure 9.6).

In other cities too, refuse is seen as an opportunity to make money. In Nairobi, recycling is an important part of the everyday economy:

- old car tyres are cut up and used to make cheap sandals
- washing machine doors are used as kitchen bowls, and the drums as storage units
- glass bottles are collected and returned to stores for refilling
- food waste is collected and fed to animals or composted for use on vegetable plots
- tin cans and old oil drums are used to make charcoal stoves, lamps, buckets and metal tips for ploughs.

E. Unemployment and under employment

Since a high proportion of the people who move to cities are relatively young, there is considerable pressure to create sufficient jobs. Unemployment rates are typically high although official data is hard to find and many migrants find employment in informal work such as street hawking. **Under-employment** refers to a situation in which a person is not doing work that makes full use of their skills and abilities. This may occur when a migrant moves to a new city.

F. Transport issues

The processes of urbanisation and suburbanisation have led to increased traffic in cities across the world (Figure 9.7). This has created more congestion and pollution, damaging human health and wasting billions of pounds in lost productivity. The spread of houses into the suburbs and beyond has created surges of morning and evening commuters. Traffic flows for shopping, entertainment and other commercial services add to the problem.

Figure 9.6 Metals, wood, plastic and clothing all have a value at the Koshe dump. Others come to find discarded food which they can eat

Figure 9.7 Kolkata suffers regular gridlock

During the car boom of the 1960s, city planners built more and wider roads as a solution. This didn't work. The more roads created, the more cars they attracted. A 1997 study in California found that new, additional traffic will fill up to 90 per cent of any increase in road capacity within just five years. No matter how much money is spent on traffic infrastructure, congestion and parking problems seem to get worse.

Contemporary urban processes

Higher-income countries have seen a much slower rate of urbanisation in the last few decades and some cities have even witnessed a decline in numbers. Rather than moving into urban areas, as seen in the industrial periods, a more significant trend has been the move outwards.

Key terms

Counter-urbanisation – The movement of people from large urban areas into smaller urban areas or into rural areas, thereby leapfrogging the rural–urban fringe. It can mean daily commuting, but can also require lifestyle changes and the increased use of ICT.

Decentralisation (1) – The movement of population and industry from the urban centre to outlying areas.

Decentralisation (2) – The movement of government institutions or headquarters of businesses away from the central/core areas of a country.

Deindustrialisation – This refers to the loss of jobs in the manufacturing sector, which occurred in the UK in the second half of the twentieth century.

Edge city – A self-contained settlement which has emerged beyond the original city boundary and developed as a city in its own right.

Gentrification – Gentrification is the buying and renovating of properties, often in more run-down areas, by wealthier individuals.

Suburbanisation – The movement of people from living in the inner parts of a city to living on the outer edges. It has been facilitated by the development of transport networks and the increase in ownership of private cars. These have allowed people to commute to work.

Urban resurgence – Urban resurgence refers to the regeneration, both economic and structural, of an urban area which has suffered a period of decline. This is often initiated by redevelopment schemes but is also due to wider social, economic and demographic processes.

Urban sprawl – The spread of an urban area into the surrounding countryside.

Suburbanisation: characteristics, causes and effects

Suburbanisation has resulted in the outward growth of urban development that has engulfed surrounding villages and rural areas. During the mid to late twentieth century, this was facilitated by the growth of public transport systems and the increased use of the private car. The presence of railway lines and arterial roads also enabled wealthier commuters to live some distance away from their places of work.

The towns and cities of the UK demonstrate the effects of past suburbanisation. In the 1930s, there were few planning controls and urban growth took place alongside main roads – this was known as ribbon development. By the 1940s, this growth, and growth between the 'ribbons', became a cause for concern. This led to the creation of **green belts** – areas of open space and low-density land use around towns where further development was strictly controlled.

Since 1950, suburban expansion has increased and has been better planned. During the 1950s and 1960s, large-scale construction of council housing took place on the only land available, which was the suburban fringe. In the 1970s, there was a move towards home ownership, which led to private housing estates being built, also on the urban fringe. Building in these areas allowed people to have more land for gardens and more public open space.

As car ownership grew, the edge of town, where there was more land available for car parking and expansion, became the favoured location for new offices, factories and shopping outlets. In a number of cases, the 'strict control' of the green belt was ignored.

In recent years, new housing estates have been built in suburban areas, along with local shopping centres and

schools (Figures 12.30 and 12.32, pages 667–668). People continue to move to the suburbs because of their desire for a quieter, less congested and less polluted environment. Suburbs are perceived as relatively crime-free environments and they also demonstrate other key benefits of the rural–urban fringe, such as woodlands and parks, golf courses and playing fields. Many are now well-established housing areas, highly sought after in the property market.

The negative effects of suburbanisation relate to urban sprawl and the environmental impacts of this process, discussed earlier. However, suburbanisation can also lead to:

- increasing social segregation within cities as the wealthy move out to the suburbs (for example, from inner London to Richmond) or gather in wealthy enclaves, such as Belgravia. The poor remain in the inner city
- diversion of funding away from inner city areas to the suburbs to pay for new infrastructure and services.

Counter-urbanisation: characteristics, causes and effects

Counter-urbanisation is the migration of people from major urban areas to smaller urban settlements and rural areas. Counter-urbanisation does not lead to suburban growth, but to growth in rural areas beyond the main city. The difference between rural and urban areas is reduced as a consequence of this movement.

A number of factors have caused the growth of counter-urbanisation. One is that people want to escape from the air pollution, dirt and crime of the urban environment. They aspire to the 'rural idyll' – what they see as the pleasant, quiet and clean environment of the countryside, where land and house prices are cheaper. Car ownership and greater affluence allow people to commute to work from such areas. Indeed, many employers have also moved out of cities. Improvements in technology have allowed more freedom of location. The spread of broadband and high-speed internet access means that someone working from a home computer can now access the same global system as a person in an office block in the centre of a city.

At the same time there has been a rising demand for second homes and earlier retirement. The former is a direct consequence of rising levels of affluence. Alongside this is the need for rural areas to attract income. Agriculture has faced economic difficulties and one straightforward way for farmers to raise money is to sell unwanted land and buildings.

Counter-urbanisation affects the layout of rural settlements. Modern housing estates are built on the edges of small settlements, and small industrial units on the main roads leading into the settlement. Former open areas are built on, old properties and some agricultural buildings are converted and modernised.

As with gentrified areas in inner cities, there may be tension between the newcomers and the locals. One of the main areas of conflict is that, despite the influx of new people, local services often close down. Bus services to many rural communities have disappeared, schools and post offices have closed, and churches have closed as parishes are amalgamated into larger units. The main reason for these changes is that the newcomers have the wealth and the mobility to continue to use the urban services some distance away.

The evidence for counter-urbanisation in an area includes:

- an increase in the use of a commuter railway station in the area, including car parking for commuters
- increased value of houses in the area
- the construction of more executive housing in the area, often on newly designated building land, following the demolition of old properties
- conversions of former farm buildings to exclusive residences.

Counter-urbanisation is one of a number of processes contributing to social and demographic change in rural settlements, sometimes referred to as the rural turnaround. This may include:

- the out-migration of young village-born adults seeking education and employment opportunities elsewhere
- the decline of the elderly village-born population, through deaths
- the in-migration of young to middle-aged married couples or families with young children
- the in-migration of younger, more affluent people, which results in increased house prices.

These changes do not take place uniformly within all rural settlements. There are considerable variations between and within parishes. The ones with the most change are key settlements that have a range of basic

services and good access to commuter routes. Such settlements are called **suburbanised villages**.

Urban resurgence: characteristics, causes and effects

Urban resurgence refers to the regeneration, both economic and structural, of an urban area which has suffered a period of decline. An urban resurgence has been seen in many cities in recent years as redevelopment schemes have made city living more attractive. This is particularly the case for the former industrial cities of the UK which suffered from manufacturing decline in the 1970s and 1980s but have reinvented themselves as cities of culture and commerce. London, Birmingham, Manchester and Leeds have all bounced back after severe **de-industrialisation** in the second half of the twentieth century. The cities have revived their fortunes by developing strong financial, business and consumer service industries and have also attracted more university students, young professionals and immigrant workers.

Urban resurgence is evident in the changing landscape of a city. Areas may still contain the industrial architecture of the past, including factories and warehouses, but increasingly these have been converted into housing or commercial use and modern infrastructure and services added. Many urban redevelopment schemes have successfully transformed run-down areas, rebranding them as fashionable districts or 'quarters', which then attract more newcomers, often young professionals, with a high disposable income.

Urban resurgence is often driven by government-led regeneration schemes but there are wider economic, social and demographic processes which are also important. Redevelopment by private companies has led to the wholesale transformation of parts of UK cities in recent years and this has served to attract further investment. City living has also become more attractive as urban areas are improved and people choose to live closer to work, entertainment and leisure facilities, rather than face long and costly commutes. Globalisation and technological change have facilitated the resurgence of some urban areas. Parts of inner East London have experienced a huge in-migration of people, attracted by its reputation for creative and digital start-up businesses. East London Tech City is home to a cluster of independent start-up companies as well as global organisations such as Facebook, Google, Amazon and EE. This area has also become fashionable for its independent shops, galleries, markets, bars and restaurants.

Major sporting events can act as a catalyst to changing the fortunes of an area. The London Olympics brought much needed investment to former industrial parts of East London, while the 2014 Commonwealth Games encouraged urban regeneration and business investment in parts of Glasgow depressed by the

The resurgence of central Birmingham

The Jewellery Quarter in Birmingham is one example of an area which has experienced bust and boom. In the early 1900s, it employed over 20,000 people in jewellery making, metalworking and hallmarking but a combination of foreign competition, reduced demand and the bombing of the area during the Blitz led to a sharp decline in the area's fortunes. Decline continued throughout the twentieth century and in spite of a small number of regeneration attempts, it was not until the early 2000s that large-scale improvements began to be seen. Former warehouses and factories have been converted into loft-style apartments and townhouses, and more than 30 restaurants, bars and cafés have helped create a vibrant hub for young businesses and professionals (Figure 9.8).

Further evidence of urban resurgence in Birmingham can be found in the Gas Street Basin canal area. Close to the city centre, this area was once the hub of a thriving canal transport network moving heavy goods such as coal and glass. Today, the canals have been cleaned up and are part of an attractive area where bars and restaurants line the waterways and traditional narrowboats navigate past large arts and entertainment complexes.

Figure 9.8 The Birmingham Mint redevelopment is found on the edge of the Jewellery Quarter in Birmingham. The Grade II listed former minting factory was converted into flats with many of its traditional features still intact.

decline of shipbuilding, steelmaking and heavy engineering. Resurgence has a positive multiplier effect, initiating further improvements and attracting greater investment into an area. However, as more people are attracted back into a city, greater pressure is put on the urban infrastructure and some people may find themselves displaced as house prices rise in line with demand. There are also concerns that not everyone benefits from resurgence and this has led to increasing inequality between rich and poor.

Many American cities including New York, Boston and Los Angeles have experienced an urban resurgence in recent decades. The second half of the twentieth century saw a huge population decline in urban areas as families increasingly moved out to the suburbs. This was accompanied by the loss of manufacturing and retail businesses and soon the traditional 'downtown' areas – the major retail centre of the city – were in decline along with neighbouring residential areas. The term '**dead heart syndrome**' was used to describe this process.

The 1990s saw the beginning of a resurgence. A sustained period of national economic growth, successful regeneration schemes and more attractive urban design has helped revive the fortunes of many American cities. In addition, young people are remaining in cities even when they start a family, while many older people whose children have left home are moving back to the city to be closer to urban services. This population revival increases demand for services and has fuelled a prosperous urban economy in many cities.

Urban change

De-industrialisation, decentralisation and the rise of the service economy

De-industrialisation refers to the loss of jobs in the manufacturing sector, which occurred in the UK in the second half of the twentieth century. Prior to this, the Industrial Revolution and the rise of manufacturing industry had been a key development in the growth of many urban areas. Cities became synonymous with particular types of industry, such as textiles (Manchester), iron and steel (Sheffield) and shipbuilding (Glasgow). Thousands of jobs were created and people migrated into urban areas. By the 1980s, many of the older industrial cities were experiencing severe economic problems associated with the decline of manufacturing. This has been attributed to three main factors:

- mechanisation: most firms can produce their goods more cheaply by using machines rather than people
- competition from abroad, particularly the rapidly industrialising countries of the time such as Taiwan, South Korea, India and China
- reduced demand for traditional products as new materials and technologies have been developed.

Table 9.1 shows an overall downward trend in manufacturing employment in the UK. Urban areas bore the brunt of these job losses but unemployment figures varied significantly between cities and depended upon the size of the city, the composition of the urban economy and the actions of local government. Cities in the manufacturing heartlands of northern England

The New York City High Line

Urban resurgence is evidenced in New York with the changing nature of the New York City High Line – a 1.5 mile-long section of elevated rail track built originally to carry goods to and from Manhattan's largest industrial district. It was abandoned in the 1980s as the Lower West Side underwent a period of manufacturing decline but has been successfully redeveloped in the 2000s as an elevated park and walkway lined with trees, grasses and shrubs.

The walkway has become a site for artistic commissions and cultural events and the five million annual visitors have increased spending in local shops and cafes, as well as encouraging real estate development in the neighbourhoods that line its route.

The High Line has given new life to a piece of industrial infrastructure as a public green space and functions essentially like a green roof. Porous pathways contain open joints so water can drain between planks, cutting down on the amount of storm-water that runs off the site into the sewer system. The planting design is inspired by the self-seeded landscape that grew on the elevated rail tracks during the 25 years in which they were derelict.

Figure 9.9 New York City's High Line

Table 9.1 Trends in manufacturing employment in the UK, 1980–2020

Year	Number of employees (thousands)
1980	6,403
1985	5,071
1990	4,908
1995	4,201
2000	4,099
2005	3,306
2010	2,815
2015	3,007
2020	3,011

Source: ONS

such as Manchester, Liverpool and Sheffield suffered more extensive job losses than cities with more diverse economies.

There was also significant variation *within* cities. Inner city areas contained many of the old types of workplace most likely to be closed – old plants with the oldest production techniques, lowest productivity and most unionised workforces. Inner city areas also lacked suitable land for the expansion of existing manufacturing and, as a result, new investment tended to focus on the edge of urban areas or more rural locations. This movement of industry away from the inner city was part of a wider process known as decentralisation, which also affected residential and retail land use in the late twentieth century.

The decline in manufacturing employment in the late twentieth century was accompanied by the rise of the service economy in urban areas. The **service economy** covers a wide range of activities including:

- **tertiary activities** such as financial services (for example, banking, accountancy and insurance), retailing, leisure, transport, education and health
- **quaternary activities** where knowledge or ideas are the main output, such as advertising, computer programming and software design.

Population growth fuels the service sector but it has also grown because:

- financial services are needed to support manufacturing industries, which are still important in many cities today
- as societies become more technologically sophisticated, they need a larger range of specialised services to keep them running
- as societies become wealthier, they demand more leisure and retail services.

The expansion of the service sector has been evident in cities all over the world and for many urban areas, there has been a dramatic shift in their economic core from manufacturing to service-based activities. The major financial centres are all located in world cities, while corporate headquarters cluster in urban areas where they can access national and international markets, a highly skilled labour force and specialist support services.

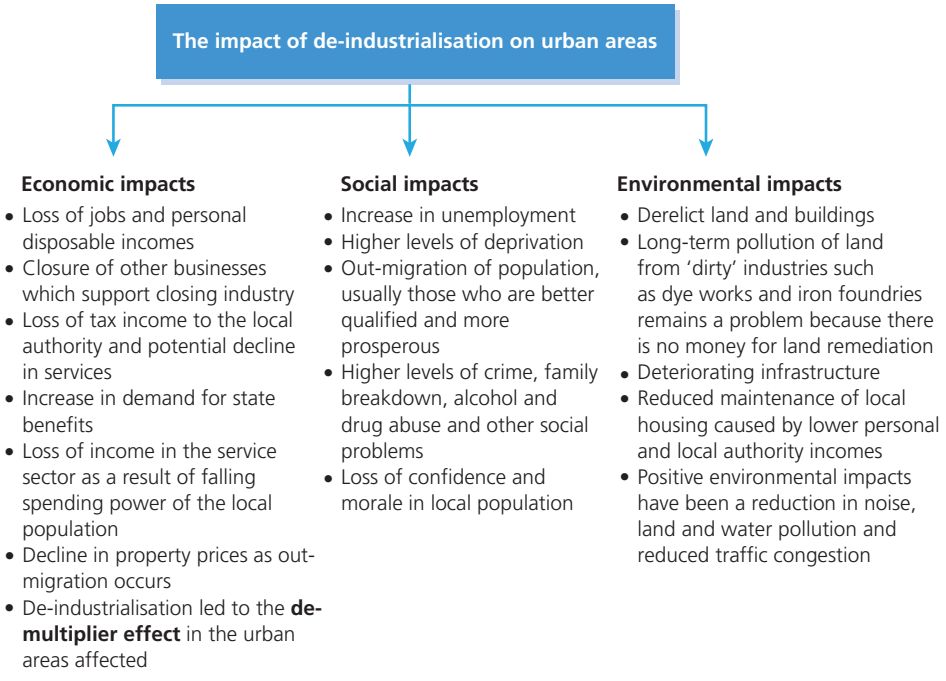


Figure 9.10 The impacts of de-industrialisation on urban areas

Source: Adapted from Drake and Lee (2000) The Urban Challenge *(Hodder Education)*

Table 9.2 A summary of urban policy in the UK since 1979

Urban policy	Details	Examples
1979–1991 Emphasis given to property-led initiatives and the creation of an entrepreneurial culture	Greater emphasis placed on the role of the private sector to regenerate inner city areas. Coalition boards were set up with people from the local business community and they were encouraged to spend money on buying land, building infrastructure and marketing to attract private investment	Enterprise Zones Urban Development Corporations Urban Land Grants Derelict Land Grants City Action Teams
1991–1997 Partnership schemes and competition-led policy	A greater focus on local leadership and partnerships between the private sector, local communities, voluntary sector and the local authority. Strategies focused on tackling social, economic and environmental problems in run-down parts of the city, which now included peripheral estates	City Challenge City Pride Single Regeneration Budget
1997–2000s Area-based Initiatives	Many strategies in the 2000s focused upon narrowing the gap in key social and economic indicators between the most deprived neighbourhoods and the rest of the country. Local authorities were set targets to improve levels of health, education and employment opportunities and funding was allocated to assist them in delivering government objectives	Regional Development Agencies (RDAs) New Deal for Communities National Neighbourhood Renewal Strategy The Housing Market Renewal Programme
2011–present Devolving responsibilities to cities in England	Bespoke packages of funding and decision-making powers negotiated between central government and local authorities, and/or Local Enterprise Partnerships (LEPs), have enabled them to take greater responsibility for business support, planning, transport and health	City Deals

Table 9.3 An evaluation of three regeneration policies which have taken place in the UK, 1980–present

Regeneration policy	Details	Successes	Failures	Case studies
Urban Development Corporations (1980s)	Urban Development Corporations (UDCs) were set up in the 1980s primarily to regenerate inner-city areas. The boards of UDCs were mostly made up of people from the local business community and they were encouraged to spend money on buying land, building infrastructure and marketing to attract private investment. Funding came direct from central government	Effective in attracting new businesses to run-down areas and improving the environment of UDC areas. By the mid-1990s, they had attracted over £12 billion in private-sector investment and created 190,000 jobs nationally	The property-led approach did little to tackle social problems Local people complained they had little involvement in the process and, in the London Docklands, locals did not tend to benefit from the new housing and jobs created	London Docklands Development Corporation (LDDC) Read: 'London Docklands: Post LDDC Developments' in *GeoActive,* 18:373 (OUP)
City Challenge (1990s)	This was a scheme where cities had to compete with each other for government regeneration grants. The cities with the 'best' schemes were awarded the grants. This was a local authority led scheme which formed partnerships between the private sector, local communities and the local authority. Strategies focused on tackling social, economic and environmental problems in run-down parts of the city	The fact that local authorities had to bid for funding was judged to have resulted in more successful regeneration schemes City Challenge gave equal importance to buildings, people and values 1997 data revealed that City Challenge had improved over 40,000 houses, created 53,000 jobs and reclaimed 2,000 ha of derelict land	Resources were thinly spread over large areas Areas which had previously received government funding based on need, no longer received funding because their bid was unsuccessful Money was lost preparing bids by local authorities who did not win funding	Hulme City Challenge partnership, Manchester Read: Barker, Redfern and Skinner, *A2 Geography* (Hodder) p. 157

Table 9.3 (Continued)

Regeneration policy	Details	Successes	Failures	Case studies
City Deals (2011–present)	Between 2012 and 2014, 26 City Deals were agreed The first wave, completed in 2012, covered the eight largest English cities outside London The second wave, completed in 2014, covered the next 14 largest English cities and their wider areas. They allow cities to: • take responsibility for decisions that affect their area • design their own strategies to help businesses grow • create economic growth; and decide how public money should be spent	The City Deal approach to infrastructure funding has proven successful in delivering increased economic activity, infrastructure funding and development certainty in a number of cities and will continue to be rolled out across the United Kingdom	Many employers feel that adult employment and skills funding streams remain too fragmented and project-specific to allow cities to integrate policy and delivery across sectors and between partners A number of barriers to further local integration of employment and skills delivery remain. This is due to the payment-by-results nature of the deals This has resulted in funding for programmes and specific projects rather than creating more flexible, non-ringfenced funding sources	The Manchester 'Working Well' pilot was implemented in 2014. It is built around a payment-by-results mechanism which rewards service providers for achievement of employment-related outcomes Read: www.greatermanchester-ca.gov.uk/what-we-do/work-and-skills/working-well/

City bars, restaurants and clubs provide important environments for networking, which is heavily drawn upon in business.

While the growth in the service sector has gone some way to reducing the unemployment caused by de-industrialisation in urban areas, a number of problems still exist.

- Many of the men who lost jobs through de-industrialisation have continued to suffer from long-term unemployment.
- Many of the service jobs created are part-time or temporary.
- The number of service jobs created has not always made up for the loss of manufacturing jobs.
- Inner city locations have been avoided by both service industries and newer manufacturing companies leading to continued inner city decline.

Urban policy and regeneration in Britain since 1979

Urban policy relates to the strategies chosen by local or central government to manage the development of urban areas and reduce urban problems. Regeneration has been a key element of urban policy in the UK since the 1980s and while early strategies focused on 'top-down' economic regeneration, subsequent policies have recognised the need to adopt a more holistic approach, tackling economic, social and environmental problems from the 'bottom-up'.

Table 9.2 (page 406) provides a summary of 'explicit' urban policies adopted in the UK since 1979. The use of the term 'explicit' is important because other more general strategies concerned with education, health and employment have helped in the regeneration of urban areas. Table 9.3 looks at three of the main regeneration strategies in more detail.

Review questions

1 What are the key issues caused by urbanisation in the twenty-first century?

2 How do these issues vary between low- and high-income countries?

3 To what extent has there been a change of focus in the aims of various urban regeneration schemes in the UK since 1979?

9.2 Urban forms

Urban form refers to the physical characteristics that make up built-up areas, including the shape, size, density and organisation of settlements. It can be considered at different scales: from regional to urban, neighbourhood and street and is evolving continually in response to social, economic, environmental, political and technological developments. A government report on 'Urban form and infrastructure in the UK' (2014), reported that the UK's urban form is characterised by 64 'primary urban areas', including one built-up megacity region (London, and the Greater South East), six large

metropolitan areas (Birmingham, Leeds, Liverpool, Manchester, Newcastle-Upon-Tyne and Sheffield), and 56 towns and cities with more than 125,000 people.

This section of the chapter will focus first on the largest urban forms: those of megacities and world cities found in different parts of the world. It will then go on to explore patterns of land use within cities and finally it will consider how urban landscapes are changing in the twenty-first century.

The characteristics of megacities and world cities

Globalisation and economic competition between countries and cities has led to the rapid rise of the megacity. Faced with poorer economic prospects in rural areas and the perception of a better life in the city, mass migration has fuelled large-scale population growth. Added to this, government policies, such as the establishment of Special Enterprise Zones in Chinese cities, can encourage greater financial investment. Problems occur because this growth in population is not matched by a growth in resources and infrastructure. Megacities in lower-income countries have tended to sprawl in a haphazard fashion and this has inevitably led to the challenge of providing employment, housing and basic services. In addition, there are concerns about how city authorities can effectively govern such large cities.

Megacities can be a force for good. On average they produce two to three times more GDP than other cities and a United Nations report on urbanisation in 2014 highlighted a number of other characteristics and benefits:

1. They offer opportunities to expand access to services, such as healthcare and education, for large numbers of people in an economically efficient manner.
2. It is less environmentally damaging to provide public transport, housing, electricity, water and sanitation for a densely settled urban population than a dispersed rural population.
3. Urban dwellers have access to larger and more diversified employment markets.
4. Better levels of education and healthcare can improve the lives of the poor and empower women in countries where they do not have equal status.
5. Megacities are centres of innovation where many solutions to global problems are being trialled.

The growth of large urban areas can also fuel political pressure for change. Political protests are more common in urban areas where large numbers of younger people are brought together.

Key terms

Fortress landscapes – This term refers to landscapes designed around security, protection, surveillance and exclusion.

World city (or global city) – These are cities which have great influence on a global scale, because of their financial status and worldwide commercial power. Three cities which have traditionally sat at the top of the global hierarchy – New York, London and Tokyo – are now being joined by the likes of Beijing, Shanghai and Mumbai. These cities house the headquarters of many transnational corporations (TNCs), are centres of world finance and provide international consumer services.

Historically, the greatest global cities were generally the largest. However, today size is not so important. Of the world's most populous cities, only Tokyo, New York and Beijing are in the top rankings of the world's most important cities. It has been argued that what matters today is influence and the term **world city** (or **global city**) has been given to the cities which have greatest influence on a global scale. This can be measured in a number of ways and it is worth casting a critical eye over the different methods adopted (Figure 9.11).

One of the more commonly cited rankings is carried out by the Globalization and World Rankings Research Network (GaWC) based at Loughborough University. The Alpha, Beta and Gamma rankings, shown in Table 9.4, were based initially on the connectivity of cities through four 'advanced producer services': accountancy, advertising, banking/finance and law. New indicators were added in 2004 but economic factors are still deemed more important than political and cultural indicators.

Below these rankings are cities with 'sufficiency of services'. These are not world cities but have sufficient services so as not to be dependent on world cities. They tend to include smaller capital cities and traditional centres of manufacturing regions.

Figure 9.11 The characteristics of world cities

Table 9.4 GaWC city rankings, 2020

City ranking	Description	Examples from 2020 ranking
Alpha ++ cities	More integrated than all other cities and constitute their own high level of integration	London and New York
Alpha + cities	Other highly integrated cities that complement London and New York, largely filling in advanced service needs for the Pacific Asia	Tokyo, Hong Kong, Paris, Shanghai, Singapore, Beijing, Sydney, Dubai
Alpha and alpha – cities	Very important world cities that link major economic regions and states into the world economy	Chicago, Milan, Mumbai, Moscow, São Paulo, Frankfurt, Toronto, LA, Madrid and more
Beta level cities	These are important world cities that are instrumental in linking their region or state into the world economy	Bangalore, Copenhagen, Guangzhou, Zagreb, Oslo, Cairo and more
Gamma level cities	These can be world cities linking smaller regions or states into the world economy, or important world cities whose major global capacity is not in advanced producer services	Lahore, St Petersburg, Durban, Islamabad, Bristol and more

Spatial patterns of land use in urban areas

The term **urban morphology** refers to the spatial structure and organisation of an urban area. Traditionally this may have been affected by physical factors such as relief and drainage. Early industrial areas developed close to rivers where they could harness the power of water for energy and transportation. Flat land was also important for transportation of goods via road or rail. Relief still plays an important role today because flat land is easier to build on and may attract a higher land value. Conversely, flat land close to rivers may pose a flood risk. In poorer cities, informal settlements are often found on undeveloped steep land. Brazil's largest shanty town, Rocinha, is built on a steep and rugged hillside overlooking Rio de Janeiro. The poorer parts of the shanty town are found higher on the hilltop, with many houses only accessible by foot.

Skills focus

To what extent do physical factors still affect the pattern of land use in towns and cities today? Investigate different cities around the world to consider the impacts of relief, drainage and other factors such as soil and rock type and natural hazards on their land use.

As humans have been able to overcome the limitations imposed by physical factors, urban form today is more strongly influenced by human factors.

The main factor affecting land use in high-income countries is **land value** and this is traditionally higher in the centre of a city where accessibility is greatest.

The **Peak Land Value Intersection (PLVI)** (Figure 9.12) is the point with the highest land value and from here, land prices decline in line with the theory of **distance-decay**.

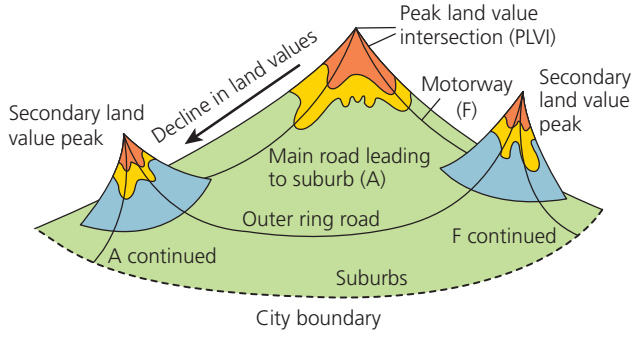


Figure 9.12 Typical land values for a city in a high-income country
Source: Stephen Codrington

It is usually only very profitable businesses such as large retailers that can afford the high prices of the PLVI. In UK cities, the site has often been occupied by the likes of Marks & Spencer.

Other retailers and commercial enterprises tend to occupy most of the **central business district (CBD)** but cannot all afford to pay the high prices required for the most accessible locations. Consequently, smaller retailers and businesses are more likely to be found towards the edge of the CBD.

Taking a transect from the CBD to the suburbs, land values fall significantly as the different land users are less reliant on accessibility and unable to pay the higher prices associated with this. This is known as the **bid-rent theory**. As Figure 9.13 shows, traditionally there has been a move from retailing to industrial and commercial and then to residential areas.

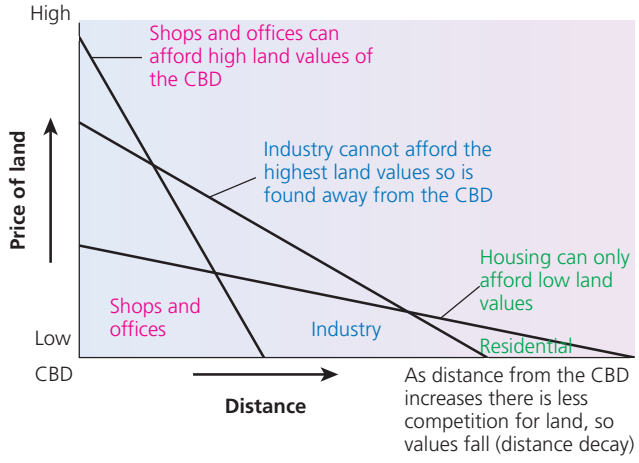


Figure 9.13 The bid-rent theory

The 1980s trend towards out-of-town shopping centres affected land values in some cities and this can be seen by the presence of secondary land-value peaks shown in Figure 9.12. An increase in traffic meant that the city centre was no longer always the most accessible part of the city and a lack of space for expansion encouraged some of the large retailers to look elsewhere for potential new sites. For the consumers, the growth of edge-of-town developments provided free parking and other land uses such as cinemas and restaurants, which proved to be highly attractive. Land values subsequently rose in such areas.

New urban landscapes

Moving from the centre outwards, there have been a number of significant changes to the urban landscape in recent years.

Changes in the city centre

Some city centres in the UK have experienced decline in the last 30 years. In the 1980s and 1990s this was largely due to the development of out-of-town retail parks and the decentralisation of business and residential areas, which served to pull people away from the CBD. High parking costs, congestion and perceptions of the city centre as dirty and unsafe were further disincentives for shoppers. More recent competition has come from the phenomenal growth of internet shopping.

A number of strategies have been devised to help reverse the decline of city centres including the provision of a more attractive shopping environment, the construction of all-weather shopping malls, improvements in public transport links and the establishment of business and marketing teams to co-ordinate management of the CBD and run special events. In addition to this, the 2000s have seen two other notable strategies.

Town centre mixed development

Many cities are encouraging the development of functions other than retailing to increase the attractions of the city centre. These include:

- a wider range of leisure facilities including cinemas, theatres, cafés, wine bars, restaurants and other cultural and meeting places. Where these offer services of different kinds and at varying prices and degrees of quality, a greater range of people will be attracted
- the availability of spaces, including gardens and squares or plazas, to enable people-watching and other activities

Main land-use zones in a city

Central Business District (CBD): This central area contains the major shops, offices and entertainment facilities.

Inner city: This is an area of old housing and light manufacturing industry. This area dates back to the Industrial Revolution when it was comprised mainly of factories and terraced housing providing accommodation for the factory workers. Many British cities have witnessed regeneration within these areas in the last three decades.

Residential: These areas consist of housing from a range of periods, which has traditionally increased in both size and price as one moves towards the outskirts. However, urban regeneration schemes and gentrification have meant that some of the most expensive property can now be found in traditional 'low class residential' areas, while council estates on the edge of cities are now among some of the most depressed parts of British cities.

Green areas: Such as parks tend to be dotted throughout an urban area. They range from large botanical gardens of national importance down to playgrounds within a housing estate.

Out-of-town retail developments: Originally developed by large supermarkets, these spaces soon expanded to include non-food retail units and entertainment complexes. They have had a negative economic impact on some town and city centres. In 1994 the UK government started to actively discourage their development.

Business or science parks: These tend to be found on the edge of urban areas where there is good access to major roads. Some science parks are also located near to universities.

Industry: Manufacturing industries often require large areas of land and tend to locate towards the edge of cities where cheaper land is available. De-industrialisation brought about the large-scale decline of manufacturing industry in British cities and former factory sites have either been demolished or converted into other land uses. The latter may still occupy their historical location in the inner city.

Informal settlements: Also known as slums or shanty towns, these are a feature of cities in low-income countries. They have traditionally developed on the edge of cities although they are also found adjacent to transport routes or in areas of the city unpopular with residents such as rubbish dumps. Physical factors such as steep slopes, unstable land and areas prone to natural hazards may also encourage their development.

Exeter: Mixed-use city centre redevelopment

In Exeter in 2007, the Princesshay mixed-use city centre redevelopment was opened. It replaced an unattractive, dated post-war facility. It contains retail units, flats, cafes and restaurants. It attracts tourists by showing new public art pieces and hosting events and festivals, as well as being the entrance to the city's medieval underground passages. It was hoped to make the area a secure and safe setting for city centre living and a night-time economy.

In late 2019, a new plan for a nearby development, called CityPoint, was unveiled as well as plans for the redevelopment of the old bus station into hotels, restaurants and bars, residential accommodation, office space and the civic centre.

Figure 9.14 shows part of the City Centre Quarter of Exeter. The Roman walk signposted recognises the heritage of the city and runs along the original Roman city wall. In this photo, you can see shops, a restaurant and two storeys of flats with cedar-clad walkways. These residential units help to create a sense of place, community and safety in the city centre, as well as meeting a local housing need.

- the promotion of street entertainment such as at Covent Garden in London
- developing nightlife, such as clubbing. (There are negative issues associated with this, including the high level of policing that is necessary)
- developing flagship attractions such as We The Curious and M Shed in Bristol
- constructing new offices, apartments, hotels and conference centres to raise the status of the CBD for business and encourage tourists to remain near the city centre
- encouraging residential areas to return to city centres by providing flats, redeveloping old buildings (a form of gentrification) or building new upmarket apartments.

The combination of strategies above and the stricter planning controls placed on out-of-town developments has meant that many large cities in the UK have successfully attracted shoppers and visitors back to the city centre. However, decision makers are still worried about the decline of the CBD in smaller cities and urban areas.

The development of cultural and heritage quarters

Many cities across the UK have initiated the planning and development of cultural or heritage quarters as a deliberate model for urban regeneration of declining inner urban areas. Culturally-led urban development first began to appear in the 1980s and early UK examples are the Sheffield Cultural Industries Quarter and the Manchester Northern Quarter. A prerequisite for a cultural quarter is the presence of cultural production (making objects, goods, products) or consumption (people going to shows, visiting venues and galleries). Heritage quarters focus more on the history of the area based around small-scale industries. The most successful quarters tend to be those actually making something or those associated with a product, such as the Birmingham Jewellery Quarter. Areas like these have built up a regional, and in some cases national, reputation which attracts visitors and tourists from further afield and brings financial benefits to the wider area.

Some critics have argued that not all towns and cities need cultural quarters and that in some areas they have simply created higher property values. Experiences of different 'quarters' has shown that some are more successful than others. However, as a tool for regeneration, improving perceptions of place and preserving history and culture, 'quarters' tend to be viewed in a positive light.

Gentrified areas

Gentrification is defined as the buying and renovating of properties often in more run-down areas by wealthier individuals. It is an important process of housing improvement supported by groups such as estate agents and local authorities and it has helped to regenerate large parts of British inner cities in the last few decades. Unlike the regeneration schemes described in Table 9.3 (page 406–7), gentrification involves the rehabilitation of old houses and streets on a piecemeal basis and is carried out by individuals or groups of individuals rather than large organisations. Gentrification can happen for a number of reasons:

- The Rent Gap: This refers to the situation when the price of property has fallen below its real value, usually due to lack of maintenance or investment, and there is a 'gap' between actual and potential price. Such properties are attractive to builders, property developers or individuals who can afford to renovate the properties and then sell them on to make a profit.
- Commuting costs: Commuting can be time-consuming, expensive and stressful. Moving closer to the city centre can eliminate the need to commute.
- The 'pioneer' image: This refers to the trend of creative individuals such as artists and designers moving into more 'edgy' neighbourhoods. These groups are not interested in the conformity of suburban living but are drawn to the diverse cultural opportunities of the urban centre. The gentrification of areas such as Hoxton and Shoreditch in London and SoHo in New York City has been linked to their reputation as the location of vibrant arts scenes.
- The support of government and local decision makers: Both groups are keen to improve the economy and environment of inner city areas and gentrification is seen as an important part of this.
- Changing composition of households: Many cities have seen the growth of single or two-person households without children. These households are more likely to see the benefits of inner-city living.

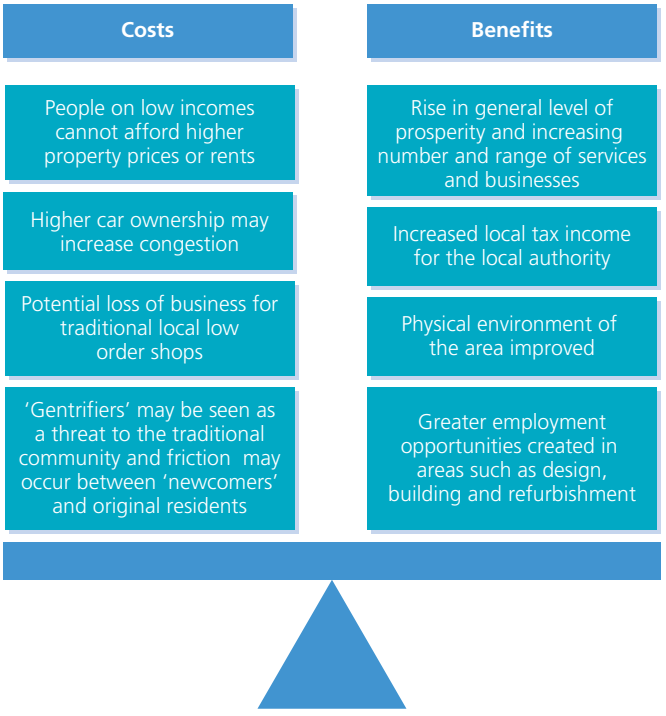


Figure 9.15 The costs and benefits of gentrification

Gentrification has both costs and benefits, as shown in Figure 9.15. Some commentators have emphasised the importance of gentrification in inner-city regeneration; others have raised concerns about the displacement of low-income families and small businesses. In many parts of London, gentrification has contributed significantly to the lack of affordable housing and prices have been pushed up even beyond the level of well-paid professional workers. Anti-gentrification protests are becoming more common.

Homes sold for a pound in Liverpool

One innovative way of regenerating inner city estates without displacing low-income families has been a scheme adopted in a number of British cities where run-down housing is sold off to individuals or families at very low prices. In 2013, Liverpool City Council sold 20 derelict homes in Kensington, Granby and Picton for £1 each. In order to prevent people abusing this scheme, there were a number of basic rules. First, buyers had to show they were capable of doing the house up to a reasonable standard and secondly, they had to sign an agreement to live in the property for five years and not sublet it. There were about 1000 applicants for each house. The scheme has now been extended to other houses and shops in the area with the overall aim to improve the built environment and make it a thriving community again.

Fortress landscapes

The term **fortress landscapes** refers to landscapes designed around security, protection, surveillance and exclusion. In the UK, a number of relatively simple strategies have been adopted to reduce crime in urban hotspots such as the city centre and inner city estates. These include:

- greater use of closed-circuit television (CCTV)
- railings and fencing around private spaces
- 'mosquito' alarms, which emit a high-pitched sound heard only by young people, to discourage loitering around certain buildings
- effective use of street lighting
- speed bumps to prevent joyriding.

More recent strategies have focused on the concept of 'designing out crime' through better urban architecture. In Manchester, for example, the redevelopment of housing in parts of the Greenheys and Wythenshawe estates has included more windows to provide better natural surveillance, provision of front gardens with fences or hedges to mark a clear boundary between private and public space and bins in gated compounds rather than open alleyways. Features which have been avoided are recessed doorways for people to hide in, projecting window sills or exposed rainwater downpipes, which would make it easier for anyone to climb on to the roof and dark alleys and dead ends associated with muggings and drugs deals.

There is also evidence in some UK cities of the exclusionary tactics adopted by North American cities to segregate people from 'others' perceived as threatening or undesirable. The use of anti-homeless spikes fitted into the ground in shop doorways or outside upmarket apartments has been heavily criticised but the high pitched 'mosquito' alarm, sloped bus shelter seats and special benches which deter skateboarders raise few eyebrows. Gated communities are also becoming a feature of some urban landscapes. These are not as common in British cities as they are in places like the USA and South Africa but the use of guards or the electronic control of access into housing complexes is increasing.

Links

The notion of Insider and outsider perspectives on place, discussed on pages 358-9 in Chapter 8, is relevant here.

Fortress LA

Surveillance and exclusion measures are seen to be at their most extreme in American cities. Los Angeles has developed a reputation as a city built on fear. The author Mike Davis has detailed how paranoia and fear of gangs, minorities and the homeless have led to a refashioning of the urban landscape which includes:

- gated communities
- armed-response security units in residential areas
- shopping malls surrounded by staked metal fencing and an LAPD surveillance tower.

Edge cities

Edge cities are self-contained settlements which have emerged beyond the original city boundary and developed as cities in their own right. They are associated with the urban landscape of North America and have been viewed by some as a feature of postmodern urban living. Edge cities are largely the result of urban sprawl. This has happened on a much larger scale in the USA as a result of higher car ownership, greater willingness to travel long distances for work, shopping and entertainment and limited planning restrictions in the suburbs. Los Angeles is the classic example of a sprawling urban settlement. In 2015, the core city of Los Angeles which is about 30 kilometres wide, had a population of just under 4 million. However, this is surrounded by a metropolitan area of nearly 18 million people which is more than 100 kilometres at its widest point. Within this area there are more than 20 edge cities (for example, Pasadena and Century City).

Edge cities develop close to major roads or airports and tend to be found in close proximity to shops, offices and other businesses which decentralised from the original city. Edge cities may lack a clear structure but they do have a wide range of amenities including schools, shopping malls and entertainment facilities. Residents may rarely go back to the original core city.

Edge cities have been linked to extreme **social segregation**, where the wealthy have moved to the new suburban settlements leaving only the poor and disadvantaged sections of society in the original city boundary.

The concept of the postmodern western city

The term **postmodernism** is used to describe the changes that took place in Western society and culture in the late twentieth century. It mainly concerned art and architecture and marked a departure from the conformity, uniformity and brutalist nature of modernism. Postmodernism is characterised by the mixing of different artistic styles and architecture.

Both fortress landscapes and edge cities have been viewed as features of the **postmodern western city** but other key features include:

- a more fragmented urban form comprising independent settlements (such as edge cities), economies, societies and cultures
- a greater emphasis on producer services and knowledge-based industries rather than industrial mass production
- eclectic and varied architecture as seen in the London cityscape
- spectacular flagship developments such as the Guggenheim museum in Bilbao, Spain
- greater ethnic diversity but heightened economic, social and cultural inequalities and polarisation.

While elements of postmodernism can be seen in cities all over the world, it is important to acknowledge that the notion of the postmodern western city has been based on the experiences of a small number of (mainly American) cities. Los Angeles is often referred to as the archetypal postmodern city but the experiences of cities like Los Angeles are clearly not representative of cities elsewhere in the world.

Review questions

4 How have urban processes changed over time and how do these vary across cities?

5 Why are many cities in high-income countries now seeing a resurgence in their urban areas?

6 Produce a timeline that shows how urban land use in the UK has changed over time.

7 What are the main factors influencing land use in different cities around the world?

9.3 Social and economic issues associated with urbanisation

Economic inequality

Some argue that a key feature of cities in Asia, Africa and South America is the increasingly large wealth gap between rich and poor residents, known as **economic inequality**. In Mumbai, this stark contrast can be seen where the world's most expensive home towers over one of the largest slum areas of the city. Antilia is the 27-storey home of Mukesh Ambani, chairman of a global energy conglomerate and estimated to be worth over $35 billion. Within a short distance of the house is Byculla, an area inhabited by some of Mumbai's nine million slum dwellers. Most people here survive on less than $2 a day.

Key terms

Cultural diversity – The existence of a variety of cultural or ethnic groups within a society.

Economic inequality – The difference between levels of living standards, income, etc., across the whole economic distribution.

Social segregation – When groups of people live apart from the larger population due to factors such as wealth, ethnicity, religion or age.

Urban social exclusion – Economic and social problems faced by residents in areas of multiple deprivation.

Figure 9.16 Valued at $1 billion, Antilia, the world's most expensive home, is made of steel and glass, has spectacular ocean views, swimming and gym facilities and three helipads. It is seen here towering above a more modest part of Mumbai

Inequality exists in all urban areas. Enormous contrasts in wealth can be found over relatively small distances as seen in Figure 9.16. When you do fieldwork in an area, you can identify which areas are more affluent and you can sense if a neighbourhood is improving or deteriorating. The wealthy and the poor seem to concentrate spatially – a form of **social segregation**. There are a number of reasons for this:

- **Housing**: Developers, builders and planners tend to build housing on blocks of land with a particular market in mind. The requirement to include a proportion of 'affordable housing' may affect housing value in some areas but wealthier groups can choose where they live, paying premium prices for houses well away from poor areas, with pleasing environments and services such as quality schools and parks. Poorer groups typically have less choice.
- **Changing environments**: Housing is only a partial explanation for inequality since housing neighbourhoods change over time. In the UK, houses that were built for large families in Georgian and Victorian times are now too big for the average family. Many have been converted into houses of multiple occupation (HMOs), often occupied by those on low incomes or students. Conversely, former poor areas are being gentrified. The 'right to buy' legislation of the 1980s transformed many council estates, as houses were bought by their occupants and improved.
- **The ethnic dimension**: Ethnic groups originally come to the country as new immigrants. When they first arrive they may suffer discrimination in the job market and may be either unemployed or employed in low-paid jobs. They are only able to afford to buy cheap housing or they have to rent privately (for example, Moss Side in Manchester or St Paul's in Bristol). Therefore, newly arrived migrants concentrate in poor areas in the city, often clustered into multicultural areas. Such ethnic groupings tend to persist into later generations.

Measuring poverty and inequality

The Index of Multiple Deprivation is a UK government qualitative study measuring overall deprivation for every neighbourhood in England. It is based on 37 separate indicators, organised across seven distinct domains of deprivation. These domains are combined,

using appropriate weights, to calculate the Index of Multiple Deprivation 2019 (Figure 9.17).

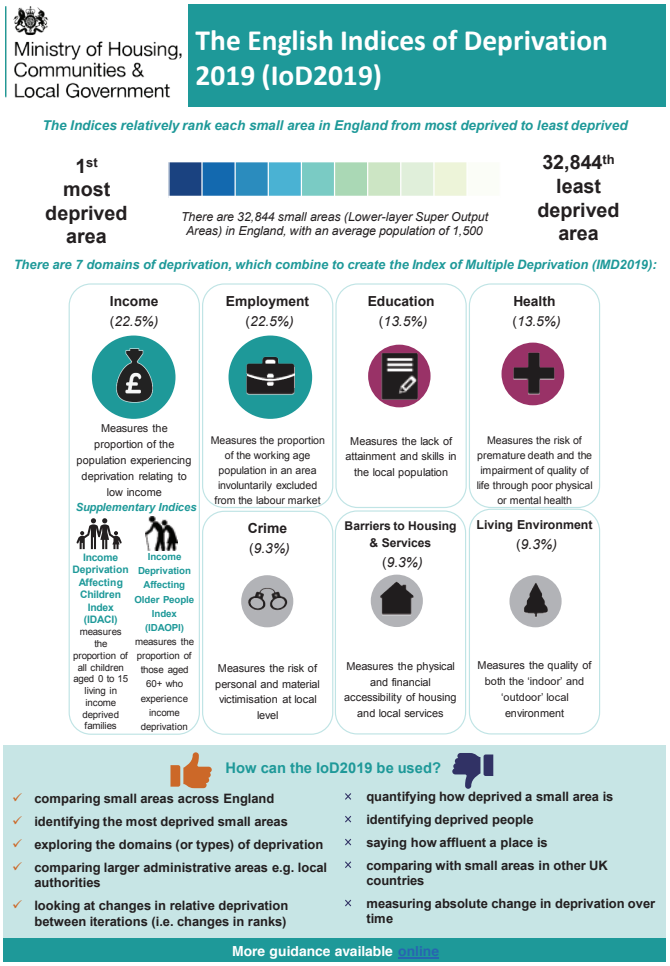


Figure 9.17 The English Indices of Deprivation, 2019 (IoD2019)

Source: Ministry of Housing, Communities & Local Government

As was the case in previous versions of the Indices, the IoD2019 reveals concentrations of deprivation in:

- large urban conurbations
- areas that historically have had large, heavy manufacturing industry and/or mining sectors (such as Birmingham, Nottingham, Hartlepool)
- coastal towns (such as Blackpool or Hastings)
- parts of east London.

There are also pockets of deprivation surrounded by less deprived places in every region of England.

Of the most deprived 10 per cent of neighbourhoods in England (3,284), 137 rank as highly deprived on six of the seven domains. These neighbourhoods are not evenly distributed across England: 88, or 64 per cent of them, are located within just eight local authority districts.

It is important to note that these statistics are a measure of relative deprivation, not affluence, and important to recognise that not every person in a highly deprived area will themselves be deprived. Likewise, there will be some deprived people living in the least deprived areas.

GIS focus

Go to http://maps.cdrc.ac.uk/ and select the map showing the Index of Multiple Deprivation data for 2019. Type in the postcode for the area you want to investigate in more detail. The website also maps geodemographic data from the 2011 census and classifications for different types of urban dweller.

Urban social exclusion refers to the problems faced by residents in areas of multiple deprivation. These people are excluded from full participation in society by their social and physical circumstances. They cannot access a decent job because of poor education or obtain decent housing because of poverty. They often suffer from poor health and from high levels of crime in an unattractive physical environment. In a city, inequality can cause lack of social cohesion and in extreme cases lead to civil unrest.

Inner city areas have traditionally been the most deprived urban neighbourhoods. The characteristics of this decline, shown in Figure 9.19 (page 418), were caused by de-industrialisation in the second half of the twentieth century when unemployment became a major problem. Population loss followed and the movement of younger, more affluent and more skilled residents left behind a population who were older, less skilled and poor.

There is no clear geographical pattern to urban poverty today. It clearly exists but the image of a troubled inner core surrounded by affluent suburbs is outdated. Some of the highest levels of urban poverty are now found in peripheral estates while many inner city areas have been transformed by regeneration schemes. Rich and poor areas today are found across city and suburb alike.

Inequality remains a major challenge for cities around the world in the twenty-first century. Indeed, the gap between rich and poor can be far greater *within* a city than across a whole country. The inequalities tend to exist in terms of access to job opportunities, education, housing and basic public services such as water and sanitation. In many poorer cities, for example, adequate water and sanitation services are primarily channelled to upper- and middle-class neighbourhoods, while

Skills focus

It is possible to measure the quality of life in an area using primary data, such as the quality, density and condition of housing and the nature of the physical and social environment (Figure 9.18). It is also possible to use secondary data from a census to assess deprivation levels. This may include poverty in terms of low income, or shown by poor health or the lack of possessions, such as cars. It is common for the poorest parts of an urban area to suffer from **multiple deprivation** (a combination of social, environmental and economic deprivation).

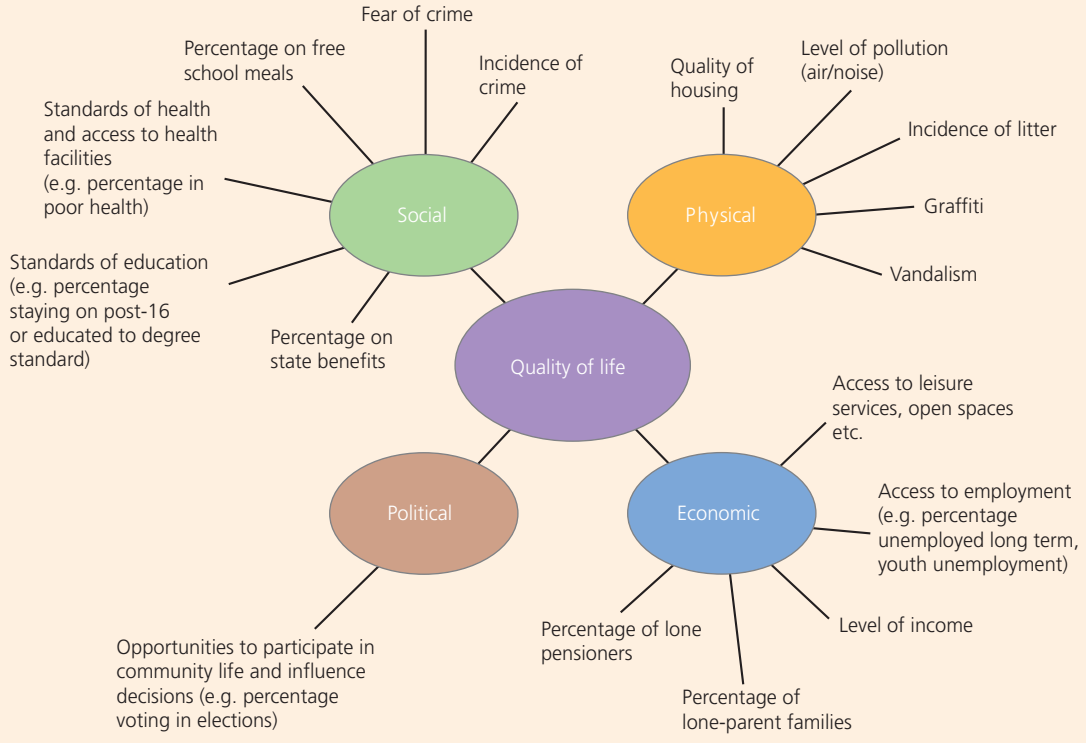


Figure 9.18 Measures of quality of life

low-income neighbourhoods often depend on distant and unsafe water wells and lack any form of waste disposal. The poor also tend to live in overcrowded informal accommodation lacking basic infrastructure and services. The knock-on impacts of this are poorer health, higher unemployment and a lack of social mobility. The poor get stuck in a cycle of poverty from which it is hard to escape.

It is important to distinguish between poverty and inequality. Poverty is an absolute term, referring to a level of deprivation that does not change over time. Inequality is a relative term referring to the differences between people, usually economic, over a geographic distribution. In practice, poverty and inequality often rise and fall together but this need not necessarily be the case. Inequality can be high in a society without high levels of poverty due to a large difference between the top and the middle of the income spectrum.

Tackling poverty and inequality in cities

The main driver of urban inequality is wealth. Richer cities like London tend to have much higher rates of inequality, while the more equal cities are those that are smaller, have lower average wages and are coming out of the end of years of industrial decline. On the national stage, a redistributive tax system can help to reduce inequality, but at the local level the effects of this are less clear. General strategies to reduce poverty and inequality are described below.

- Enforcing a living wage or providing an urban subsidy: In China, minimum wages have grown steadily. In 2020, Shanghai continues to have the highest minimum wage in China at US $358 per month, followed by Shenzhen and Beijing, both at US $318 per month. In some rural areas wages can be as low as US $166 per month. Despite this, China

Figure 9.19 The web of inner city decline, despair and deprivation

still has some of the highest income inequality in the world.

- Provision of schools: Education may be taken for granted in many high-income countries but schooling beyond primary school level helps lift millions of people out of poverty in lower-income countries.
- Supporting low-skilled workers who want to develop their abilities: Cities need to ensure labour markets are inclusive, giving all residents the skills and opportunities needed to enter the workforce.
- Access to affordable housing: The presence of slum housing in poorer cities should not hide the fact that access to affordable housing is a major problem in wealthy cities too. In London, the lack of affordable housing has been blamed for breaking up families, reducing employment prospects and mobility and slowing the economy. The London Living Rent scheme offers affordable rented accommodation to middle-income households who currently rent and want to build up savings to buy a home. It is one of three schemes introduced by the Mayor of London. The others are London Affordable Rent for households on low incomes and London Shared Ownership.
- Greater provision of public transport: Affordable public transport schemes improve mobility for the city poor, enabling them access to employment, education and services that could improve their lives. Medellín, Colombia, has a cable car system to help its citizens get around (Figure 9.20). Opened in 2004, it has four separate lines and allows people from previously isolated neighbourhoods to travel into the centre of the city. The cable car system connects to the city's metro system and offers some great views of the city. It has helped transform a once very dangerous city into a top

global tourist destination. Further lines are under construction.

Figure 9.20 The Medellín Metrocable system

- Enforce minimum environmental standards: Poor health is strongly linked to poor environmental conditions. This can be improved through effective legislation.
- A number of British cities, including Liverpool and Sheffield, have established 'Fairness Commissions' looking at how local areas can address inequality.

Cultural diversity

Cultural diversity refers to the existence of a variety of cultural or ethnic groups within a society. Culture can relate to nationality, race, age and traditions. Immigration is a key influence. Urban areas, especially large cities, are places where cultural diversity tends to flourish. Cities like London, New York and Amsterdam have received migrants from all over the world and this has led to the creation of multicultural urban societies. London is considered to be one of the most diverse cities in the world. More than 300 languages are spoken by the people of London, and the city has at least 50 non-indigenous communities with populations of 10,000 or more.

Globalisation has increased movement around the world and it is now common to find people from different parts of the world living in major urban areas. The term **diaspora** is commonly used to describe a large group of people with a similar heritage or homeland who have settled elsewhere in the world. In some global cities, these diasporas make up a larger proportion of the population than the indigenous residents. The United Arab Emirates state of Dubai is one such example. The majority of the city's population is made up of expatriates. Two-thirds of the population is of Asian descent, with people coming from India, Pakistan, Bangladesh and Sri Lanka. Arabic is recognised as the official language, but many languages are readily spoken, including English, Urdu, Punjabi, Tagalog, Bengali, Hindi, Persian and Chinese. In Los Angeles, 57 per cent of the population is multilingual. Cultural enclaves such as Chinatown, Koreatown and Thai Town are just some of the areas reflecting its cultural diversity.

Why is there greater cultural diversity in cities?

Cultural diversity results largely from immigration and immigrants are more likely to choose to live in urban areas. There are a number of reasons for this:

- cities tend to offer a greater range of employment opportunities
- cities are the first point of entry into the country for many immigrants
- cities tend to house earlier immigrant groups with the same ethnicity
- established cultural diversity in cities means there are specialist shops and religious centres located there
- urban populations tend to be more tolerant of immigrants.

In the UK, there have been a number of significant migrations which have shaped the cultural diversity of major cities (Figure 9.21, page 420). The port city of Liverpool attracted many Irish migrants in the nineteenth and twentieth centuries and there are large minority ethnic Indian communities in cities such as Leicester and Greater Manchester, where labour-intensive industries such as clothing were traditionally located. More recently, the influx of Eastern European migrants has led to changes in the demographic make-up of many British cities. In Southampton, for example, more than 10 per cent of the population are now Polish and specialist Polish supermarkets and restaurants have opened to cater for them.

Cultural diversity brings many benefits to an urban society. The most commonly cited advantages relate to the greater exposure people get to different foods, music, language and religion. Events such as the Notting Hill Carnival in London and Mela in Newcastle have become part of the British calendar and serve to illustrate an acceptance of and interest in different cultures.

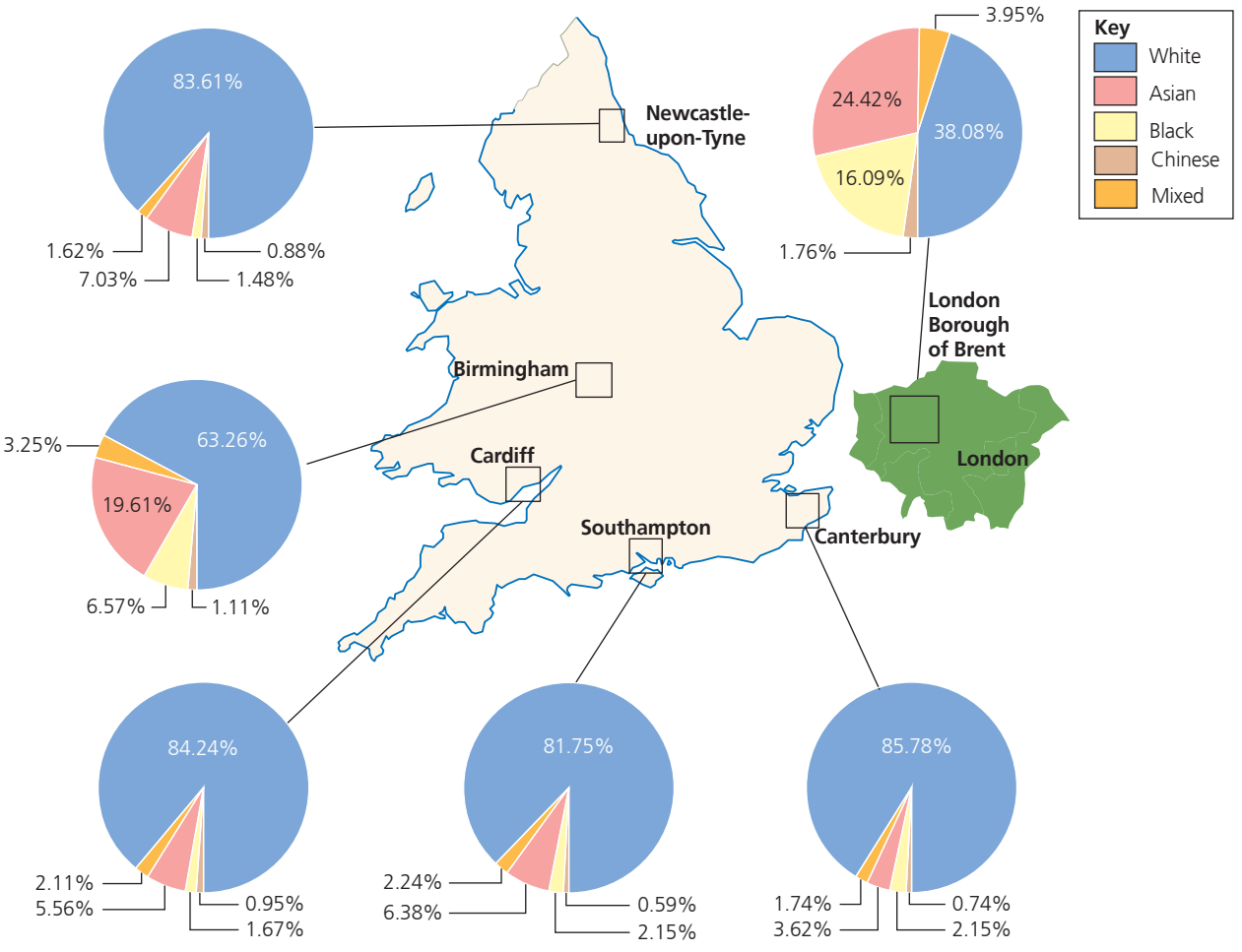


Figure 9.21 The ethnic breakdown of population for a selection of British cities, 2009

Issues associated with cultural diversity

Cultural diversity can put extra pressure on already stretched urban services. Where language differences exist, local authorities may need to provide English lessons or bilingual literature. Hospitals may need to cater for specific requirements and schools may alter their curricula and holiday patterns to cater for different ethnic groups. Variations in educational attainment have also been noted and it is the responsibility of local authorities to ensure that all children have the same opportunities.

GIS focus

You can use the 2011 census data to investigate cultural differences in your local urban area. Go to www.datashine.org.uk and select the Origins & Beliefs data. Compare the patterns shown with some of the other statistics available, such as employment rates, educational attainment, housing and health indicators.

Many countries have adopted a multicultural policy to protect and celebrate cultural diversity but some argue that at an urban level, this can encourage culturally and spatially distinct communities leading 'parallel lives'. This is known as segregation and there is evidence of this in cities around the world. In Dubai, for example, British migrant workers tend to live in expat (sometimes gated) communities and integrate little with local society. London has also witnessed an influx of very wealthy immigrants in recent years and these have tended to cluster together. The more affluent boroughs of Mayfair and Knightsbridge are home to a multitude of millionaires from the Middle East and the former Soviet Union and the price of housing in these areas now far exceeds the reach of even wealthy British people.

Social segregation based on ethnicity

In some European and North American cities, different ethnic communities have become isolated from wider

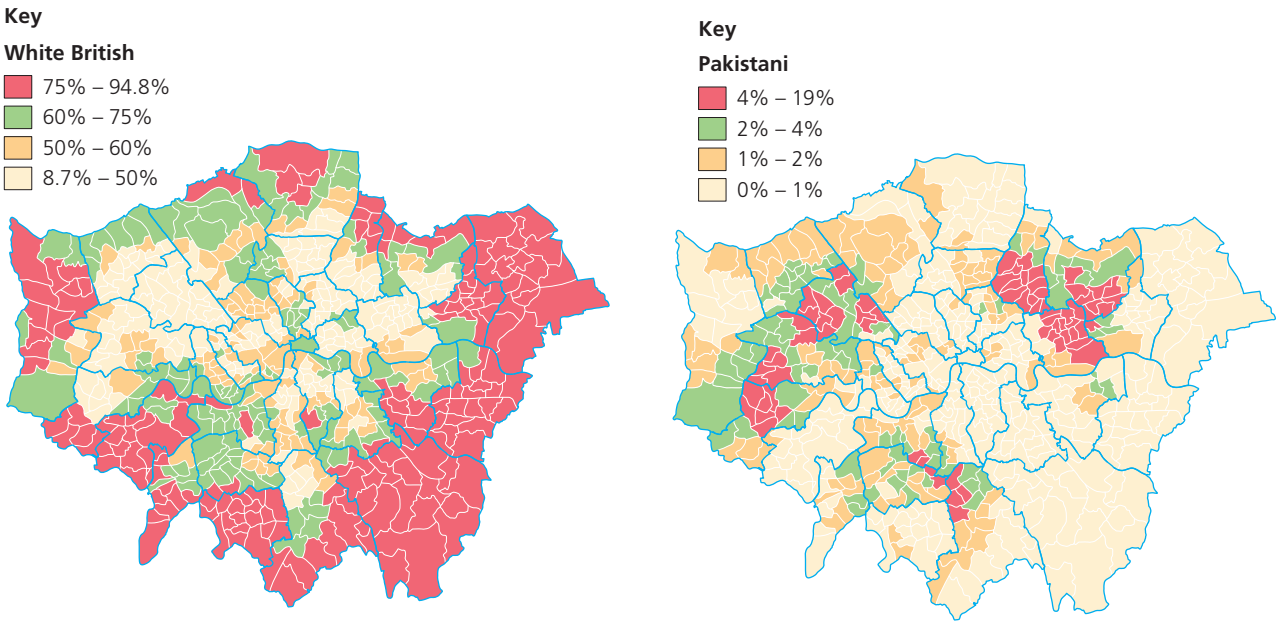


Figure 9.22 The distribution of white British populations shows a concentric pattern around the edge of London while those of Pakistani descent are concentrated in three main areas of the city (2011 census)

society as they have maintained their own language and beliefs and limited their interaction with others. Local schools can sometimes become dominated by a particular group and this can lead to suspicion and hostility as younger people from different backgrounds do not get to know each other.

In American cities, the term **ghetto** has been used to describe an area of a city where the population is almost exclusively made up of an ethnic or cultural minority. As in many European cities, these are often located in the poorer parts of the city, which wealthier residents have left and where unemployment rates are high.

The place study on Brick Lane and Spitalfields in Chapter 8 highlights the large concentration of people of Bangladeshi descent in East London. There are geographical patterns for other ethnic groups in the capital too (Figure 9.22).

Figure 9.23 Reasons for ethnic segregation in urban areas

Such geographical patterns tend to be the result of self-segregation, but there are also external factors which have encouraged minority ethnic groups to live in particular areas. These are shown in Figure 9.23 (page 421).

In spite of rising tensions between ethnic groups in some cities, analysis of the 2011 census data revealed that many large cities in the UK, including Leicester, Birmingham, Manchester and Bradford, recorded a decrease in segregation for most ethnic groups between 2001 and 2011. Figure 9.24 shows the changes for Indian and Chinese groups across England and Wales. In London too, even the most diverse wards of Brent and Newham experienced a further decrease in segregation.

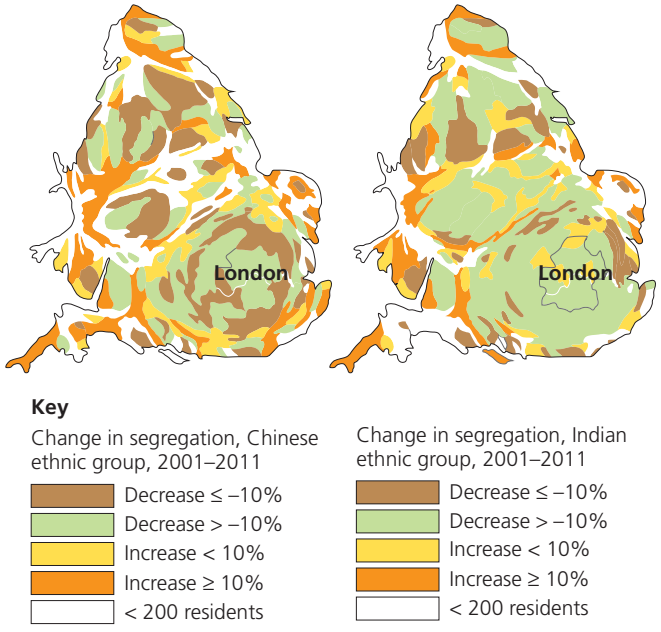


Figure 9.24 Changing rates of segregation for British Asian populations 2001–2011

Source: Guardian News & Media Limited

To tackle the negative issues associated with ethnic segregation, a policy of **interculturalism** has been promoted by some. This emphasises interaction and the exchange of ideas between different cultural groups. Areas of mutual interest are found and community engagement is conducted in 'intercultural spaces' such as libraries, schools, sports clubs and community centres.

In the UK, there is legislation on anti-racism, employment rights and opportunities to combat discrimination, prejudice and racism. Encouraging greater political involvement of different cultural groups can also encourage greater integration and provide a voice for those who may feel under-represented.

Review questions

8 Outline the different ways that social and economic issues vary across and within cities.

9 Using an example of a city, evaluate how successful it has been in tackling issues such as poverty and segregation.

9.4 Urban climate

Urban areas create their own set of local atmospheric conditions (temperature, precipitation, relative humidity, visibility, air quality and wind speed) that differ from those in the surrounding rural areas. This is referred to as an **urban microclimate** (Table 9.5). The affected area extends both upwards and downwind of the city.

Key terms

Albedo – The reflectivity of a surface. It is the ratio between the amount of incoming insolation and the amount of energy reflected back into the atmosphere. Light surfaces reflect more than dark surfaces and so have a greater albedo.

Particulate air pollution – A form of air pollution caused by the release of particles and noxious gases into the atmosphere. Emissions of particles can occur naturally but they are largely caused by the combustion of fossil fuels.

Photochemical pollution – A form of air pollution that occurs mainly in cities and can be dangerous to health. Exhaust fumes become trapped by temperature inversions and, in the presence of sunlight, low-level ozone and other noxious gases form. It is associated with high-pressure weather systems.

Temperature inversion – An atmospheric condition in which temperature, unusually, increases with height. As temperature inversions do not allow convection, they trap pollution in the lower layer of the atmosphere.

Urban heat island – The zone around and above an urban area, which has higher temperatures than the surrounding rural areas.

Urban microclimate – The small-scale variations in temperature, precipitation, humidity, wind speed and evaporation that occur in a particular environment such as an urban area.

Warm air forms a dome over a city. The urban canopy layer (UCL) is the layer of air closest to the surface, extending upwards to the mean building height. Above this lies the urban boundary layer (UBL). This may be over one kilometre in height during the daytime, but

just a few hundred metres at night. It is the UBL that forms a dome of warmer air that can extend as a plume downwind of the city (Figure 9.25).

Table 9.5 The effects of urban areas on local climate

Element of climate	Effect of urban area (compared to nearby rural areas)
Temperature	
Annual mean	0.5–3°C increase
Winter minima	1.0–2°C increase
Summer maxima	1.0–3°C increase
Precipitation	
Quantity	5–10% increase
Days with less than 5 mm	10% increase
Snowfall – inner city	5–10% decrease
Relative humidity	
Annual mean	6% decrease
Winter	2% decrease
Summer	8% decrease
Visibility	
Fog in winter	100% increase
Fog in summer	30% increase
Wind speed	
Annual mean	20–30% decrease
Calms	5–20% increase
Extreme gusts	10–20% decrease
Radiation	
Ultraviolet in winter	30% lower
Ultraviolet in summer	5% lower
Total on horizontal surface	15–20% lower
Pollution	
Dust particles	Up to 1,000% increase

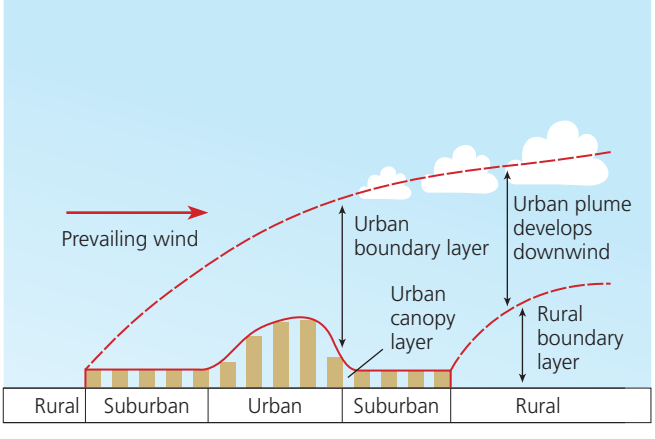


Figure 9.25 The urban climate dome

Temperatures: the urban heat island effect

Urban and suburban areas experience higher temperatures than the surrounding rural areas and this difference is what constitutes an **urban heat island (UHI)** (Figure 9.26). The annual mean air temperature of a city with one million or more people can be 1 to 3°C warmer than its surroundings and on a clear, calm night, this temperature difference can be as much as 12°C. Temperatures will fluctuate in urban areas dependent on season, weather conditions, sun intensity and ground cover. Surface urban heat islands are typically largest in the summer. Smaller urban areas will produce heat islands, but the effect tends to decrease as city size decreases.

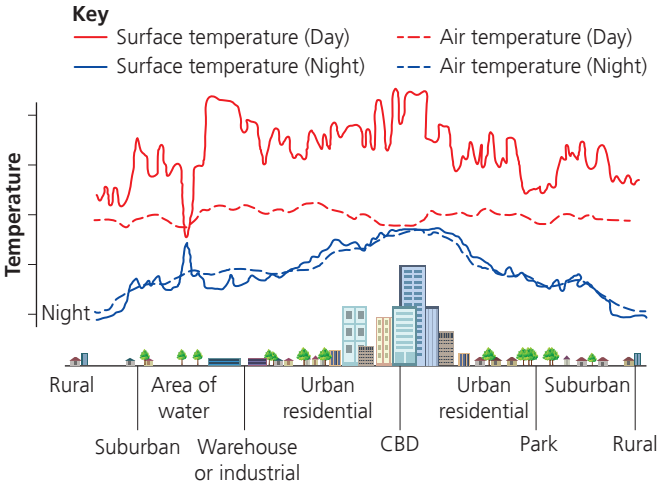


Figure 9.26 The urban heat island effect

Source: Adapted from US Environmental Protection Agency

Figure 9.26 shows a transect from the rural–urban fringe through to the high-rise characteristics of the central business district and back to the rural–urban fringe. It can be seen that typically the urban temperatures are at their highest in the mid-afternoon over the CBD. Secondary peaks of high temperature appear over other built-up areas such as the suburban residential areas. The temperature range from rural fringe to city centre is often greatest at night due to the high heat storage capacity of building materials compared to vegetation. There is very little variation in surface temperatures over areas of water. Air over water is often relatively cool; a combination of the absorption of heat by the water itself and the evaporation at its surface, lowers the temperature.

Figure 9.27a Landsat satellite image of Atlanta, Georgia, 2000 – aerial photo

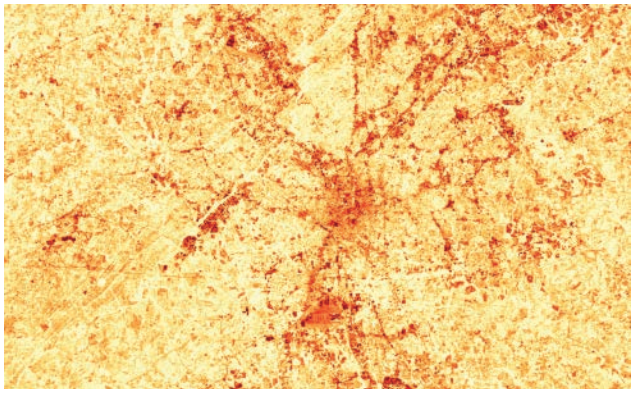


Figure 9.27b Landsat satellite image of Atlanta, Georgia, 2000 – land surface temperature map

The Landsat satellite images shown in Figures 9.27a and 9.27b provide two views of urban Atlanta, Georgia, USA on 28 September 2000. The first image is an aerial photo-like view of the area, where trees and other vegetation are green, roads and development appear grey, and bare ground appears tan or brown. The second image is a land surface temperature map, in which lower temperatures are yellow and higher temperatures are red. Because vegetation cools the surface through evapotranspiration of water, the most densely vegetated areas (darkest green in top image) are the coolest areas (palest yellow in bottom image). Where development is densest, the land surface temperature is near 30°C.

Cities tend to be warmer than rural areas for the following reasons:

- Surfaces in the city tend to be much less reflective than those in rural areas. Building materials such as concrete, bricks and tarmac have a much lower **albedo** and act like bare rock surfaces, absorbing large quantities of heat during the day (Figure 9.28). Much of this heat is stored and slowly released at night. Some urban surfaces, particularly buildings with large windows, have a high reflective capacity, and multi-storey buildings tend to concentrate the heating effect in the surrounding streets by reflecting energy downwards. In winter, rural areas keep snow for a much longer period and therefore have a greater albedo ranging from 0.86–0.95.
- Air pollution from industries and vehicles increases cloud cover and creates a 'pollution dome', which allows in short-wave radiation but absorbs a large amount of the outgoing radiation, as well as reflecting it back to the surface.
- Urban areas are designed to dispose of surface water quickly. This reduces the potential for cooling by evaporation. Reduced vegetation means less evapotranspiration.
- Heat comes from industries, buildings and vehicles, which all burn fuel (Figure 9.29). Although they regulate the temperature indoors, air conditioning units release hot air into the atmosphere. Even people generate heat and cities contain large populations in a small space.

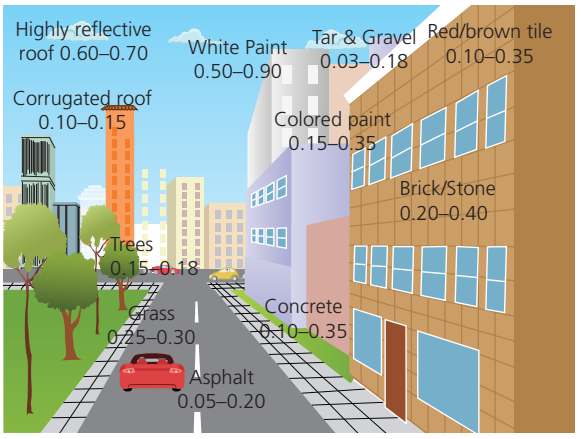


Figure 9.28 Urban environment albedos tend to be much lower than those of rural areas. Rural surface albedos may include deciduous forest (0.17), coniferous forest (0.14) and sand (0.37)

Source: Adapted from US Environmental Protection Agency

Figure 9.29 This infrared thermal image shows that buildings and roads emit much more heat than lighter surfaces and vegetation.

Fieldwork opportunities

Urban temperatures can be mapped using isotherm maps as shown in Figure 9.30. In this particular case, the temperature in central London rises to around 3°C higher, while Richmond Park (dark blue area below left) is about 1°C cooler than its surroundings. It is possible to investigate the presence of an urban heat island in your local urban area by taking temperature readings for a transect from the rural–urban fringe through the urban centre and out to the other side of the settlement. Some UHI studies have been done by individuals on a bicycle but it is better to have a larger group of students who can record the temperature at allotted points at exactly the same time of day or night.

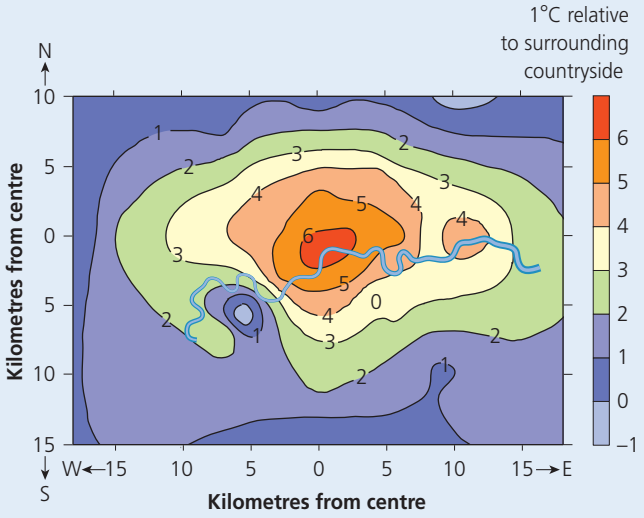


Figure 9.30 This isotherm map shows temperature differences between a rural reference climate station and a number of urban climate stations located across London, under calm and dry conditions at 02.00–03.00 hrs for six urban heat island events during the summer of 2000 (1 July to 30 September).

Source: Greater London Authority

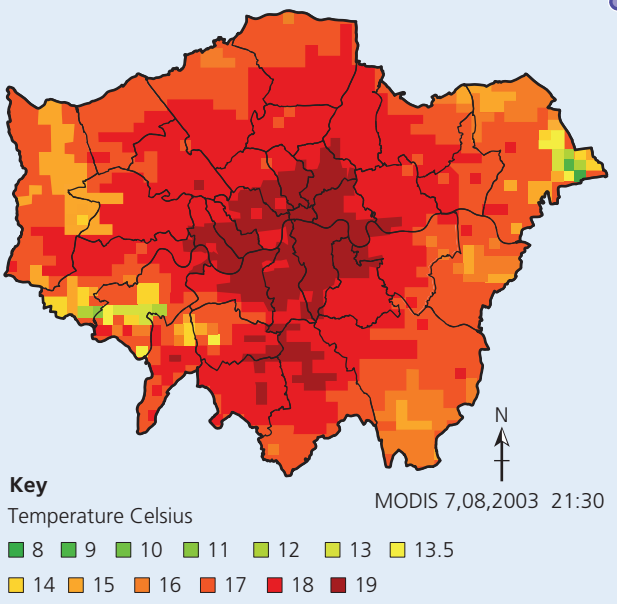


Figure 9.31 Surface temperatures across London recorded by the MODIS satellite on 7 August 2003 at 21:30 hrs during the August 2003 heat wave. During this heat wave, the UHI intensity reached 9°C on occasions

Source: International Proceedings of Chemical, Biological and Environmental Engineering

An alternative view of London's UHI can be gained from the analysis of surface temperatures as measured by infrared cameras located on satellites. Figure 9.31 shows there is a positive correlation between high surface temperatures and high density urban areas. The relatively cool areas to the southwest coincide with the large open and green spaces of Richmond Park.

Why is the urban heat island a matter of concern?

Considerable work has been done on the UHI in London in recent years and a number of concerns have been highlighted:

- As temperatures rise in the summer months, conditions can become uncomfortable in buildings and on city transport systems. During extreme heat island events, the cases of heat stroke, asthma, organ damage and even death increase. During the heatwave of 2019, Paris recorded a high of 42.6°C. Most of the 1435 heat-related deaths in France that year occurred in Paris.
- The hot and still, anticyclonic weather conditions responsible for intense urban heat island events also produce higher air pollution levels. This is because the chemical reactions that produce ozone and smog are accelerated by high temperatures, while the lower wind speeds keep the heat and pollution trapped in the city.
- Excessive heat puts an increased strain on the supply of energy for cooling and air conditioning.
- In warmer periods, the added heat from the urban heat island can lead to increased water consumption by residents and businesses. This places extra strain on the water supply infrastructure and can lead to water-use restrictions. Evapotranspiration rates will also be higher, with the result that plants and trees will potentially extract water from the soil at greater rates than normal.

- The earlier flowering times of plants and trees in cities and a prolonged growing season may cause discomfort for city residents who face a longer allergy season.
- The prolonged survival and higher reproduction rates of some animal and insect pests can be problematic and there is a greater potential for algal blooms in water courses as a result of rising temperatures.
- There is an increased risk of the deterioration of historical monuments and buildings through increased rates of temperature-related chemical weathering.
- Climate change is expected to increase the intensity of the urban heat island effect in most urban areas.

Strategies for managing the urban heat island

In light of the largely negative consequences of extreme urban heat island events, urban planning and design have increasingly focused on strategies to reduce the urban heat island effect:

- **Cool surfaces**: Cool roofs built from materials with high albedo absorb and store less solar energy during the day and thus are not major emitters of heat into the urban atmosphere at night. Cool roads and pavements with reflective coats and seals are being trialled around the world.
- **Green roads**: Green roads are roads with a more porous surface that allows water to seep in and grass to grow through. This reduces the amount of heat absorbed by the road surface.
- **Green roofs**: These consist of a growing medium planted over a waterproof membrane and can reduce rooftop temperatures by 20–40°C on a sunny day (Figure 9.32). In addition, green roofs reduce rainwater run-off, act as insulators and increase urban biodiversity by providing habitat space for birds and small animals.
- **Urban greening**: Planting trees and vegetation provides shade (surface peak temperature reductions of 5–20°C may be possible) and can have a natural cooling effect as seen by lower temperatures in urban parks around the world. Urban trees also act as carbon stores, can reduce urban flooding by intercepting rainfall and filter pollutants from the air.
- **Sky view factor**: Sky view describes the relative openness between buildings in an urban area. A restricted sky view, as found for narrow streets and

Figure 9.32 The curving green roof structure of the School of Art, Design and Media at Nanyang Technological University in Singapore

London skyscraper responsible for melting car parts

Figure 9.33 The 'walkie-talkie' building, 20 Fenchurch Street, London. The *brise soleil* is evident at the top of the building

Individual buildings can sometimes have unexpected impacts on their immediate surroundings. In September 2013, the 525-ft tall building at 20 Fenchurch Street in London, nicknamed the 'walkie talkie', was blamed for melting car parts. The concave design and mirrored glass caused the Sun to shine intense rays of light on to the pavement below and this led to damaged wing mirrors and plastic panelling on parked cars. A special architectural sunshade known as a *brise soleil* was subsequently fitted to prevent the problem.

tall buildings, will reduce the escape of heat from street and building surfaces. This can contribute to the accumulation of heat within 'street canyons' and lead to the increase of air temperatures. In addition, if streets are angled perpendicular to the prevailing wind, this will reduce the chances of ventilation and removal of heat and pollutants that accumulate between the buildings during intense urban heat island events.

- **Cool cars**: Due to its higher reflectivity, the interior of a light coloured car does not heat up as much as a darker one. This reduces the need for air-conditioning and reduces the amount of heat that is emitted into the surrounding air.

Precipitation

Rainfall can be higher over urban areas than rural areas. This is partly because higher urban temperatures encourage the development of lower pressure over cities in relation to the surrounding area. Convection rainfall tends to be heavier and more frequent, as does the incidence of thunder and lightning. There are several possible reasons for this:

- The urban heat island generates convection. As ground surfaces are heated, rapid evapotranspiration takes place and can result in cumulus cloud and convectional weather patterns.
- The presence of high-rise buildings and a mixture of building heights induces air turbulence and promotes increased vertical motion.
- Due to the low pressure caused by rising air, surface winds are drawn in from the surrounding rural area. This air then converges as it is forced to rise over the higher urban canopy. A similar process occurs as the prevailing wind moves over the city. Friction from the urban boundary creates an orographic process similar to a mountain barrier but the moving air may split apart due to the barriers created by high-rise buildings. As the air comes back together downwind of the high buildings, they are thought to converge and rise upwards forming rain clouds.
- City pollution can increase cloud formation and rainfall. Pollutants act as hygroscopic (water attracting) nuclei and assist in raindrop formation. There is also some suggestion that city pollution enhances the chance of lightning as the cloud droplets take on different electrical charges.
- Cities may also produce large amounts of water vapour from industrial sources and power stations.

Studies have shown that rainfall downwind of major urban areas can be as much as 20 per cent greater than it is in upwind areas. The heating of the surface and the overlying air creates instability in the atmosphere that encourages air to rise. As it rises, it cools, and water vapour condenses into rain that falls downwind of the city.

Fog

In cities, the occurrence of fog increased along with industrialisation. Records of London weather show that in the early 1700s, there would have been about 20 days of fog every year but by the end of the 1800s this had risen to over 50 days. It was discovered in the 1950s that the average number of particles in city air in the more developed world was much greater than in rural areas. The particles acted as condensation nuclei and encouraged fog formation at night, usually under high-pressure weather conditions.

In the UK the Clean Air Acts of the 1950s resulted in a dramatic reduction in smoke production and particulate emissions, and a decrease in the number of foggy days. In contrast, cities undergoing more recent industrialisation events are experiencing more fog.

Figure 9.34 People wearing face masks in Beijing. In December 2015, the first ever air pollution red alert was declared in the city as heavy winter fog combined with polluted air.

Cities such as New Delhi and Beijing (Figure 9.34) suffer regular winter fogs and the term 'airpocalypse' has been used to describe the high death toll which occurs when they trap pollutants to create a toxic smog.

Thunderstorms

Thunderstorms develop in hot humid air and are characterised by violent and heavy precipitation associated with thunder and lightning. In urban areas the chance of thunderstorms is increased, particularly during the late afternoon and early evening in the summer months.

Thunderstorms are produced by convectional uplift under conditions of extreme instability. Cumulonimbus clouds may develop up to the height of the tropopause, where the inversion produces stability. The updraught of air through the central area of the towering cloud causes rapid cooling and condensation. This leads to the formation of water droplets, hail, ice and super-cooled water, which coalesce during collisions in the air. During condensation, latent heat is released that further fuels the convectional uplift. As raindrops are split in the updraught, positive electrical charge builds up in the cloud. When the charge is high enough to overcome resistance in the cloud, or in the atmosphere, a discharge occurs to areas of negative charge in the cloud or to Earth. This produces lightning. The extreme temperatures generated cause a rapid expansion of the air which develops a shock wave. This is heard as thunder.

Wind

Key terms

Channelling – Wind redirected down long, straight, canyon-like streets where there is less friction. These are sometimes referred to as urban canyons.

Venturi effect – The phenomenon in which wind velocities increase when air flows through an increasingly narrow gap (for example, between two buildings).

In simple terms, buildings artificially increase the friction on urban winds and reduce wind speed. In practice, it is much more complex and urban designers and architects strive to reduce the negative impacts of effects of wind. There are four main effects of wind in urban areas:

- On calm and clear nights, when the urban heat island effect is at its greatest, convectional processes can draw in strong localised winds from cooler surrounding areas.
- In many cities, particular districts have become dominated by clusters of high-rise buildings that have reshaped the skylines (for example, La Défense, Paris). This has caused an increase in the **channelling** effect where wind at street level may be artificially increased or reduced based on the orientation of the surrounding street. This is good for reducing pollution but is associated with pedestrian discomfort when wind is directed down long straight streets. The **Venturi effect** intensifies the wind speed as wind is funnelled through small openings. It can cause severe gusts at street level and may make walking difficult for some, especially those with mobility challenges.
- An individual building modifies the air passing over and around it (Figures 9.35 and 9.36). On the windward side, the overpressure, which increases with height, is deflected both upwards over the building and downwards. The descending flow on the windward side can be very strong, forming a vortex when it reaches the ground which can sweep around the windward corners. In the lee of the structure, low pressure draws in air, forming complex vortices downwind.

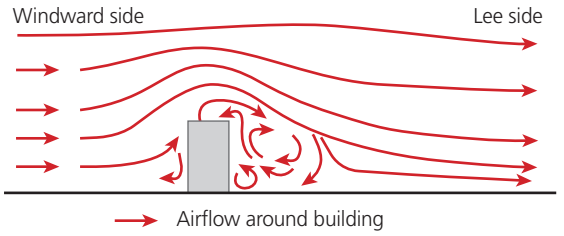


Figure 9.35 The effects of large urban buildings on winds

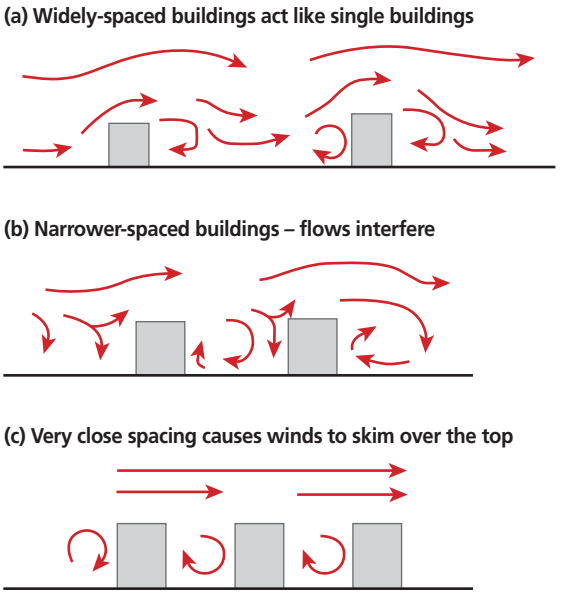


Figure 9.36 Airflow in urban areas modified by more than one building

The spacing of buildings also affects the air flow (Figure 9.37). If they are widely spaced, each building will act as an isolated block. In the case of closely spaced buildings, the wake of one building interferes with the airflow around the next. The closer they are together, the greater the skimming effect, leaving 'dead-air' zones between them where pollution and litter can build up.

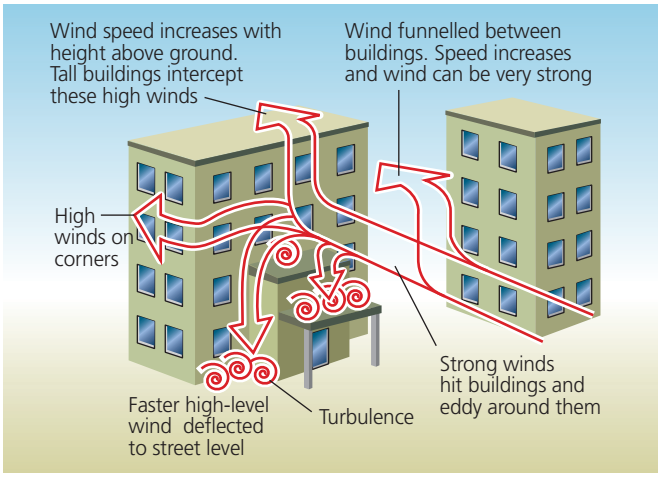


Figure 9.37 Airflow modified by a single building

Building design and city planning can mitigate the potential problems caused by urban winds (Table 9.6).

Table 9.6 How design features can impact on urban winds and climate

Design feature	Impact(s) on urban winds/climate
Aligning main streets with prevailing wind	Pollutants are flushed out Reduction in UHI effect
Building on stilts	Allows the wind to flow through the base of a building
Building on a podium	Downdraughts do not reach ground level
Porches built over doorways	Downdraughts do not reach ground level
Barriers built in potential Venturi effect hotspots (e.g. La Grande Arche de la Défense, Paris)	Reduction in high-speed gusts between buildings

When buildings are designed it is important that pollution emitters (chimneys) are high enough to ensure that pollutants are released into the undisturbed flow above the building and not into the lee eddy or the downward-flowing air near the walls.

Air quality

Air quality in urban areas is often poorer than in rural areas (Table 9.7, page 430). **Particulate air pollution** is caused by the release of particles and noxious gases

Coping with the wind: the Burj Khalifa, Dubai

Figure 9.38 The Burj Khalifa in the United Arab Emirates state of Dubai is the world's tallest building

At over 828 m (2,716.5 feet) high, the Burj Khalifa is the tallest building in the world. At such heights, a major concern for any structure is wind stress and, in particular, wind vortexes pulling the building from side to side. The Burj Khalifa has to withstand gusts of over 240 kilometres per hour and a series of aerodynamics improvements had to be made to ensure the building could withstand these. The softened edges of the building deflect the wind around the structure and prevent it from forming whirlpools or vortices, and the entire tower was orientated relative to the prevailing wind direction. The 206-storey Burj Khalifa may still sway slowly back and forth by about 2 metres at the very top but it does not suffer any of the problems experienced by skyscrapers in the past. In the 1970s, for example, people working in the 240-metre tall John Hancock building in Boston, USA suffered motion sickness even in moderate winds and over 5,000 glass panels blew out, shattering on to the pavements below.

Utilising the wind: Masdar City, Abu Dhabi

The reliance on oil in the Middle East is not sustainable and the Masdar Initiative in Abu Dhabi therefore focuses on renewable energy in an attempt to diversify the future energy mix. It is hoped that the zero-carbon, zero-waste Masdar City will be home to 40,000 residents when it is completed in the early 2020s, providing an ultra-modern example of sustainable urban living.

The power of the Sun has been harnessed to provide plentiful solar energy in Masdar City but the local winds have also been used to save energy. In the central atrium of the Masdar Institute, a barajeel, or wind tower provides cool air and a fine water mist, reducing the need for air conditioning. Meanwhile, the Siemens headquarters is positioned to face the direction of the prevailing wind, making use of a Venturi tunnel underneath the building to maintain a cooling airflow through specially designed wind channels. Thanks to the Venturi effect, a breeze flows up to the roof of the building through atria in the buildings structure, cooling public spaces without any energy costs.

Figure 9.39 The orientation of the buildings at the Masdar Institute was designed to optimise street shading. The city grid was also angled to help capture and funnel prevailing winds to aid cooling

Table 9.7 Types of urban air pollution

Pollutant	Cause	Impacts
Carbon monoxide (CO) is a colourless, tasteless, odourless and poisonous gas produced by incomplete combustion of fuel	It is estimated that road transport is responsible for almost 90 per cent of all carbon monoxide emissions in the UK. Concentrations tend to be highest close to busy roads	Carbon monoxide affects the transport of oxygen around the body by the blood. Breathing in low levels can result in headaches, nausea and fatigue
Nitrogen oxides (NO_x), when exposed to sunlight, react with hydrocarbons to produce the components of petrochemical smog – a mixture of ozone, nitric acid, aldehydes, peroxyacetyl nitrates (PANs) and other secondary pollutants	Road transport is estimated to be responsible for about 50 per cent of total emissions of nitrogen oxides	Nitrogen oxides can inflame the lining of the lung and impacts are more pronounced in people with asthma. They can also cause accelerated weathering of buildings and acid rain. The PANs also damage plants
Particulate matter (PM) – tiny bits of solids or liquids suspended in the air	Particles originate mainly from power stations and vehicle exhausts. Other particulate matter includes small bits of metal and rubber from engine wear, dust, ash, sea salt, pollens and soil particles	Particles smaller than about 10 micrometers are referred to as PM_{10} and can settle in the airway and deep in the lungs, causing health problems
Sulphur dioxide (SO_2) is a colourless gas with a sharp odour. It reacts easily with other substances to form harmful compounds, such as sulphuric acid, sulphurous acid and sulphate particles	In the UK the major contributors are coal and oil burning by industry, such as power stations and refineries	Short-term exposure may cause coughing, tightening of the chest and narrowing of the airways. Sulphur dioxide can also produce haze, acid rain, damage to lichens and plants and corrosion of buildings

into the atmosphere. Emissions of particles can occur naturally but they are largely caused by the combustion of fossil fuels. A combination of dust, soot and gases are produced from vehicles and industrial processes and this has had a negative impact on human health in many cities around the world.

Air pollution varies with the time of year and with air pressure. Concentration of pollutants may increase five- or six-fold in winter because **temperature inversions** trap them over the city (Figure 9.40).

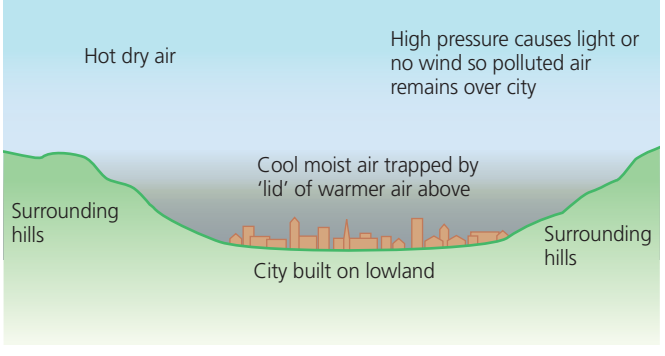


Figure 9.40 An urban temperature inversion

The mixture of fog and smoke particulates produces **smog**. This was common in European cities through the nineteenth and first half of the twentieth centuries because of the high incidence of coal burning. Britain suffered particularly badly, many of the smogs being so thick they were known as 'pea-soupers'. In the month of December 1952, smog in London was responsible for more than 4,000 deaths.

Smogs remain a feature of many cities around the world today and headlines regularly focus on the negative impact they have on human health. A particular concern has been the increase in **photochemical smog**. Photochemical oxidants (ozone and peroxyacetyl nitrate (PAN)) are associated with damage to plants and a range of discomforts to people, including headaches, eye irritation, coughs and chest pains. The action of sunlight on nitrogen oxides (NO_x) and hydrocarbons in vehicle exhaust gases causes a chemical reaction which results in the production of ozone. (Do not confuse this low-level ozone with the high-level ozone in the atmosphere, which protects the Earth from damaging ultraviolet radiation.) Los Angeles has had a serious problem with photochemical smog for decades because of its high density of vehicles, frequent sunshine and basin topography that traps photo-oxidant gases at low levels (Figure 9.41). The chemical PAN, linked to vehicle emissions, has been deemed particularly hazardous and linked to the famous Los Angeles 'eye-sting'. Numerous attempts have been made to improve public transport in the city but Los Angeles is a large, sprawling city and it is much easier to get around with a car. Consequently more recent strategies have included stricter city vehicle emission standards and the Clean Air Action Plan (CAAP) for the port of Los Angeles.

Photochemical smog is a particular hazard during anticyclonic conditions because once the air has descended, it is relatively static owing to the absence of wind. Such weather systems tend to be stable and can persist for weeks during the summer months.

Pollution reduction policies

Reducing air pollution in cities has become increasingly important for national and local government and a number of different strategies have been adopted.

Figure 9.41 The skyline of Los Angeles, USA, with a layer of brown haze or smog hanging over the city

Clean Air Acts

After the catastrophic London smog of 1952, the British government decided legislation was needed to prevent so much smoke entering the atmosphere. The Act of 1956 introduced smoke-free zones into urban areas and this policy slowly began to clean up the air. The 1956 Act was reinforced by later legislation. In the 1990s, for example, tough regulations were imposed on levels of airborne pollution, particularly on PM10s. Local councils in the UK are now required to monitor pollution in their areas and to establish Air Quality Management Areas where levels are likely to be exceeded.

In London, measures to clean up construction sites have been introduced; these sites are responsible for around 12 per cent of the city's NOx emissions. The use of dust suppressants at industrial sites has also been increased.

Vehicle control and public transport

Figure 9.42 (page 432) shows the type of strategies adopted by different cities to reduce the number of vehicles in urban areas. Greater provision of public transport and general restrictions on polluting vehicles can be very effective.

In London, a variety of measures have been introduced:

- 280,000 trees have been planted; their leaves and bark absorb pollution – for example, when the bark of the London plane tree flakes off, it sheds pollutants.
- The world's first Ultra Low Emission Zone came into being in 2019.
- Increased use of hybrid and zero emission engine buses.

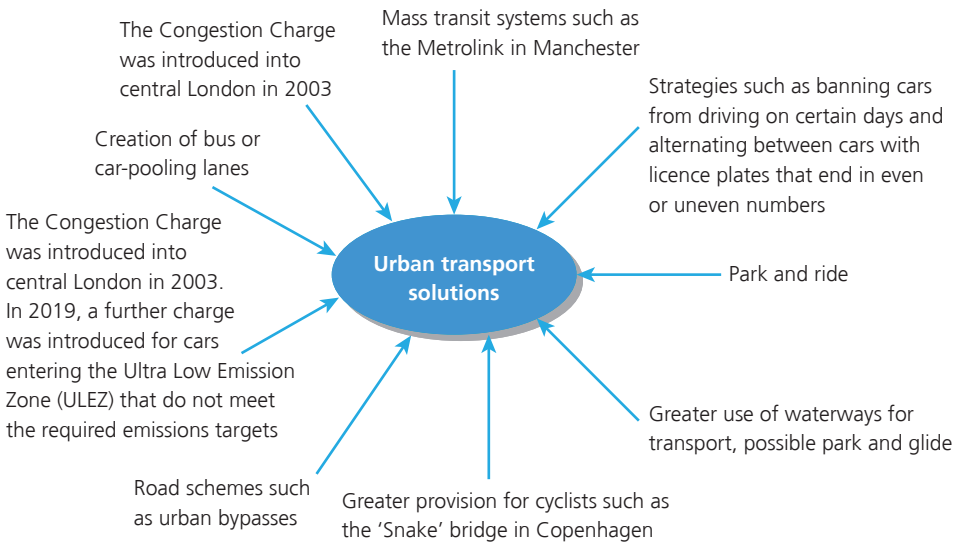


Figure 9.42 Strategies to reduce the number of cars in urban areas

- 1500 electric vehicle (EV) charging points had been installed by 2020, with more to come.
- A £48 million scrappage fund for polluting cars is in place.

Air quality standards have improved. Before the first of these measures was implemented in 2016, London's air exceeded the hourly legal limit for nitrogen dioxide for more than 4,000 hours. In 2019, this had fallen to just above 100 hours, a reduction of 97 per cent.

Zoning of industry

Industry has been located downwind in cities if at all possible and planning legislation has forced companies to build higher factory chimneys to emit pollutants above the inversion layer.

9.5 Urban drainage

Built-up areas need to be drained to remove surface water run-off. Traditionally this has been achieved by using underground pipe systems to convey the water away as quickly as possible. They were not designed to take into account the amenity aspects of drainage systems, such as water resources management, community facilities, landscaping potential and provision of varied wildlife habitats. Water quality issues have become increasingly important; pollutants from urban areas are being washed into rivers or into the ground and once polluted, groundwater is extremely difficult to clean up.

Modern approaches such as **sustainable urban drainage systems (SUDS)** deliver a more holistic approach to managing surface water and wherever possible mimic natural drainage.

Urban hydrology

As already stated above, precipitation falls in greater amounts and with greater intensity in towns and cities than in the surrounding rural area. Natural landscapes like forests, wetlands and grasslands trap precipitation and then allow it to infiltrate slowly into the ground. In contrast, impermeable urban surfaces like roads, car parks and rooftops prevent precipitation from infiltrating; most of it remains above the surface where it runs off rapidly in unnaturally large amounts.

Urban areas are also designed to shed water quickly. Sloping roofs, smooth, rounded guttering and cambered roads all contribute to the rapid movement of water away from the surface. Water is then gathered in smooth storm sewer systems which act like a high-density drainage system. It gathers speed and erosional power as it travels underground. When the water leaves the storm drains and empties into streams, they fill rapidly. Figure 9.43 shows how the water cycle is affected when a rural landscape is altered to an urban one.

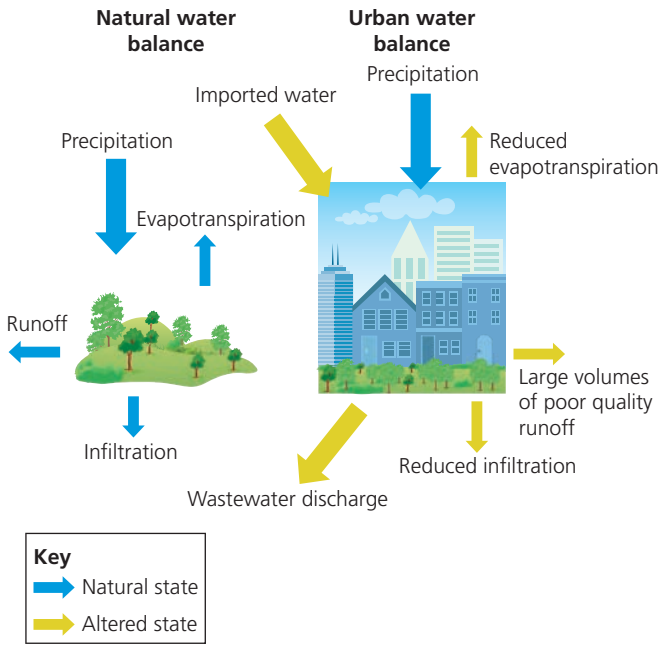


Figure 9.43 The effect of urban surfaces on the water cycle

Because much of the precipitation is unable to infiltrate impermeable urban surfaces, groundwater and soil water storage levels are reduced. Since this is the water that feeds streams during dry periods, base level flows are reduced. The resultant storm hydrograph for an urban river (Figure 9.44) shows a river with a **flashy** discharge but low base flow.

The result of this is that urban areas are more likely to have flooded rivers after heavy rainfall.

The combination of population and urban growth, along with the predicted increase in the occurrence of severe weather as a result of climate change, has meant that many more people are now at risk from flooding.

According to the Asian Development Bank (ADB), the Asian population that is vulnerable to inland flooding is expected to reach 350 million by 2025. The sprawling and growing cities are sinking under their own weight, combined with the effects of groundwater extraction by their residents. In parts of Jakarta, a city of 9.6 million people, the ground has sunk 2.5 metres in less than a decade.

Catchment management in urban areas

While flooding is important when managing catchment in urban areas, it is not the only thing to be taken into consideration by planners.

Figure 9.45 (page 434) summarises these issues. Pollutants can harm fish and wildlife populations, kill native vegetation, foul drinking water supplies, and make recreational areas unsafe and unpleasant. Sediment from erosion can fill spaces between rocks on the stream bottom, thus reducing living space or habitat for the biological communities.

One way to deal with the flood problem is shown by the way the city of Los Angeles has managed its river. The Los Angeles River channel (Figure 9.46, page 434) is

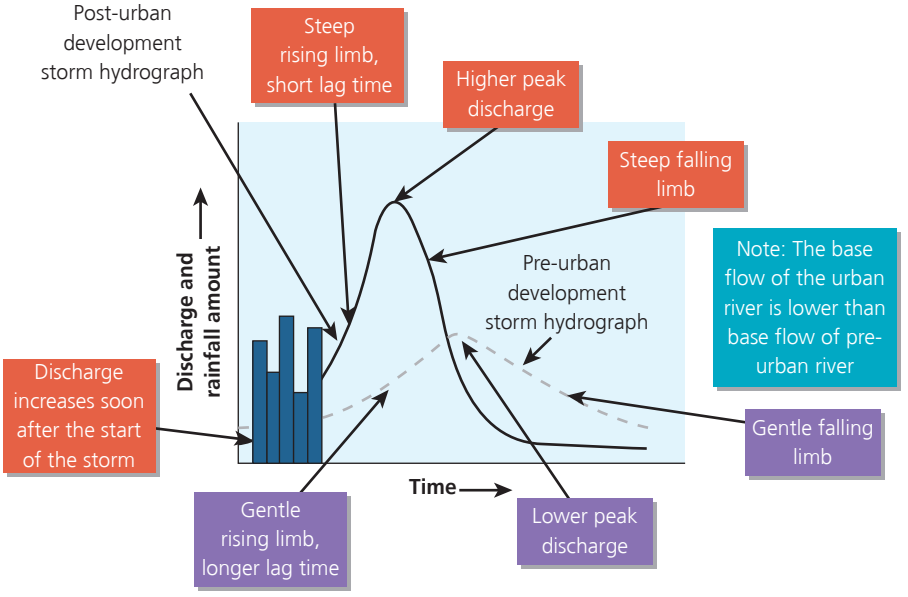


Figure 9.44 Changes in a storm hydrograph for a stream following urban development of its catchment

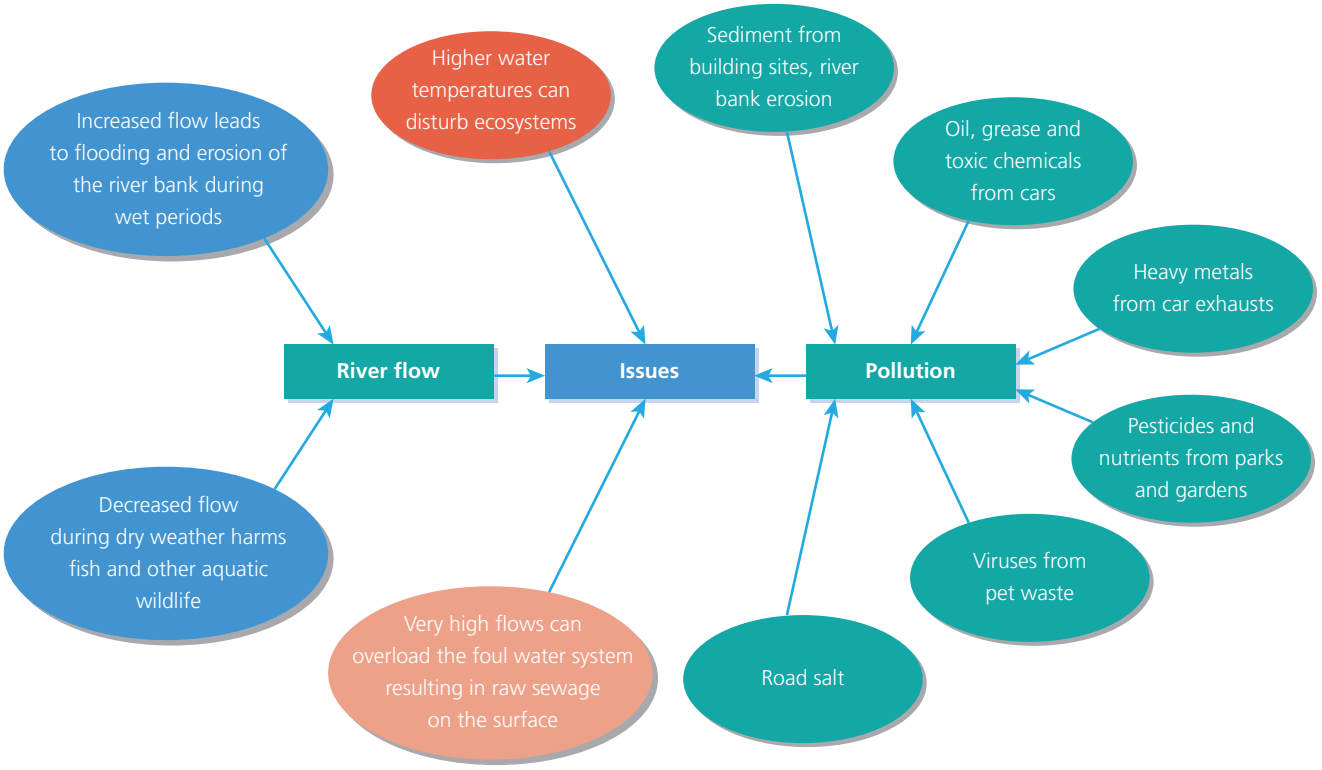


Figure 9.45 Issues associated with urban river catchment management

designed to be **fail-safe**. Devastating flash flooding of the LA River led the city authorities in the 1930s to convert it from natural and meandering to cement and controlled.

While directing the river through this built channel has helped control flooding, it has removed the **ecosystem services** that a river typically provides. In addition, the entire system could be paralysed if one part of the structure sustains significant damage, such as from an earthquake.

Figure 9.46 The channelised Los Angeles River

An alternative to this hard engineering is a series of measures that are described as **safe-to-fail.** They fall under a general title of (SUDS):

Sustainable urban drainage systems (SUDS)

SUDS is a relatively new approach to managing rainfall by using natural processes in the landscape to reduce flooding, control flooding and provide amenity for the community (Figure 9.47).

Roofwater is collected in water butts for use in gardens or flows to grass channels called swales. It then travels on to grass basins where it is stored before being released into local ditches.

Figure 9.47 This bioswale rain garden is a sloped retention area designed to capture and convey water while allowing it to infiltrate the ground slowly over a 24- to 48-hour period. Some of the water is soaked up by the vegetation, thereby reducing flooding by natural means

The Cheonggyecheon River project, Seoul, South Korea: an example of river restoration and conservation in a damaged urban catchment

Seoul is one of the world's largest cities and has a population of over 10 million people – 24.5 million in the larger metropolitan area. It is also one of the most densely populated, with over 17,000 people per square kilometre. One central area in the city, Cheonggyecheon, was built over the course of a small seasonal stream and had become overcrowded and blighted by traffic.

The Japanese colonial administration had first begun dredging the Cheonggyecheon stream in 1918, aiming to cover it over – it was seen as a sanitation and flood risk. The Second World War and the Korean War halted all conversion and maintenance work. Refugees and returning populations then established squatter communities along the stream's banks. The stream was fully covered between 1958 and 1961, and a four-lane overpass was built in 1971. This continued until 2003 when an ambitious two-year, $281-million scheme was launched to restore the river. This was transformed into a 5.8 km ecologically sensitive green pedestrian corridor.

Reasons for and aims of the project

- The city of Seoul is in the process of changing from a development-oriented urban landscape dependent on the car to one that values the quality of life of its people and the importance of functioning ecosystems. This mirrors changes in South Korea's planning priorities. In the second half of the twentieth century the country strove for accelerated industrialisation and modernisation. Now, however, there is a different emphasis from both city authorities and residents on health, sustainability and social responsibility.
- The ageing elevated freeway and concrete deck covering the Cheonggyecheon stream posed safety risks and needed to be repaired or removed.
- The government wanted to improve connectivity between the city's north and south sides, which the freeway divided.
- The areas next to the overpass housed a dense warren of over 100,000 small businesses. The overpass had a serious impact on their competitiveness. By the end of the twentieth century, buildings were disintegrating, the area was polluted and there was little green space.
- The project aimed to create both ecological and recreational opportunities along a 5.8 km corridor in the centre of Seoul.

Description of work carried out

- The elevated freeway and concrete deck covering the stream were dismantled.
- Twenty-two bridges – 12 pedestrian bridges and 10 for cars and pedestrians were built to improve movement north-south across the area.
- Car use was discouraged in the area, rapid bus lanes were added, and improved loading and unloading systems were implemented.
- To address the variable flow of the stream, water from the nearby Hanang River is pumped to the area to create a consistent flow with an average depth of 40 cm. This water is treated to make sure that it is not polluted.
- The corridor runs from Seoul to an ecological conservation area outside the city, and is split into three zones which mark the transition from an urban landscape to a natural environment:
  - In the central historic zone, underground waterways were redirected to create a new stream bed with landscaped banks, and the foundations and stones of earlier bridges were included as decorative elements. There is seating throughout to encourage the public to use the space.
  - The middle zone has recreation areas, waterfront decks and stepping stones that bridge the two banks. Its design focused on using environmentally friendly materials. Other features include fountains and waterfalls.
  - The stream widens as it enters the final zone, which is designed to look overgrown and untamed. Sections of the pier and overpass were left as industrial mementoes. Where the stream flows into the Hanang river is a wetlands conservation area.

Figure 9.48 Restoration of the Cheonggyecheon area

Figure 9.49 The Cheonggyecheon area following restoration

The attitudes and contributions of parties involved

- South Korean president Lee Myung Bak included the restoration of the stream in his successful bid to become the Mayor of Seoul in 2001, and it became a priority for his administration.
- Transportation experts were concerned that removing the elevated highway would increase traffic congestion and chaos in the northern end of the city since it carried 169,000 vehicles per day.
- The Cheonggyecheon Restoration Citizens' Committee of professional and citizen groups took responsibility for gauging public opinion. It communicated the project's goals and organised public information sessions.
- Local businesses saw the project as either an interruption or threat to their livelihoods. The Citizens' Committee surveyed the markets in detail, carried out interviews with those likely to be affected, and held regular consultation meetings. To minimise inconvenience and stimulate business activity, the Seoul Metropolitan Authority:
  - provided extra parking
  - reduced parking fees
  - improved the loading and unloading systems
  - promoted Cheonggyecheon businesses
  - provided support, subsidies and grants for business restructuring
  - struck special arrangements with those businesses that had to relocate.

Evaluation of the project

Since the restoration, Cheonggyecheon has become popular with residents and visitors alike for rest and relaxation. The stream has become a tourist attraction, drawing an estimated 18.1 million visitors by the end of 2008.

Public facilities, such as the Dongdaemun plaza and various seating schemes, have encouraged a diversity of cultural programming. Parts of the project have become well known as lively and accessible venues.
The Cheonggyecheon Museum, situated on the embankment, has permanent and temporary exhibitions on the area's history and reconstruction.

Economic sustainability

Despite its central location the area had fallen behind the rest of Seoul in terms of economic viability. Since the end of the project, development capital has been invested in residential construction and property prices have risen at double the rates elsewhere in the city. The number of businesses in the areas closest to the restoration work has risen, compared with decreasing trends in most other city districts.

Traffic

Speeds in the central business district slowed by only 12.3 per cent. Bus and subway train usage has increased. The stream-side walk is an attractive alternative to street-level journeys, and there is more pedestrian activity in the area.

The removal of large numbers of cars from this artery has created a 2.5°C reduction in average temperatures. Air flows freely along the path of the stream, creating a cooling wind corridor. A 2005 public survey showed respondents overwhelmingly noticed improvements in air and water quality, noise and smells.

Environment

The stream has re-established lost habitats, plant and animal species have reappeared, and it has become an urban wildlife haven. Schoolchildren have access to a valuable educational resource through the ecology embedded seamlessly within their urban environment.

Inclusive design

The design was non-inclusive. There was limited consideration of certain user groups, for example, older people, people with visual impairments and people with mobility problems. In September 2005, a group led a protest march demanding the right to access the new pathways alongside the stream.

In response, lifts were provided at seven locations, together with free wheelchairs for users with mobility problems. Despite this, irregular surfaces are uncomfortable for people who use a wheelchair and poorly-lit, congested tunnels are difficult for people with visual impairments.

Rain falling on roads or paths soaks through a permeable block paving where it is filtered and stored in the stone below, or it flows into grass channels, which have a stone filter drain underneath, before it joins the rest of the SUDS system.

Usually there will only be a little water in the detention ponds and the swales when the rainfall is light. If it rains heavily, the swales and basins fill for a short period, protecting those areas downstream. Water is collected, cleaned and stored in the local landscape, providing an attractive place for play and wildlife.

One example of SUDS being put into place is at the Moor Park Centre in Bispham, North Blackpool.

A case study of this can be found at www.susdrain.org/case-studies/case_studies/moor_park_blackpool.html. Susdrain, the community for sustainable drainage, have many other, mainly English, case studies. A lot of work has also been carried out in the USA. The link below contains more information on the use of such systems in Phoenix, Arizona, which has a desert climate with typical infrequent but heavy rainstorms. www.epa.gov/sites/production/files/2015-10/documents/phoenix_gi_evaluation.pdf

Key term

Sustainable urban drainage systems (SUDS) – This is a relatively new approach to managing rainfall by using natural processes in the landscape to reduce flooding, control flooding and provide amenity for the community.

Review question

10 How do climatic conditions vary across a city and what are the key factors influencing this?

9.6 Urban waste and its disposal

Waste creates a number of big issues for urban authorities and planners. Inadequate waste disposal can be linked to air and water pollution, both of which have a negative impact on human health. It is also becoming increasingly expensive to deal with waste as space for landfill is running out and treatment of waste through incineration is costly. Methane (CH_4), a potent greenhouse gas, is emitted when biodegradable waste, most importantly food and paper/card waste, decomposes anaerobically. If not addressed now, refuse sites will produce 8 to 10 per cent of the global polluting greenhouse gas emissions by 2025. Recycling rates are increasing but it is clear that much more needs to be done to tackle the growing waste crisis in cities all over the world. Some argue that the first step in waste management is to *stop* calling it 'waste'.

Waste generation

Globally, waste increases by about 7 per cent, year on year. Population growth accounts for much of this increase but economic development also plays a role, since greater personal wealth increases consumption of goods and services and this leads to more waste. In urban areas, where there are large concentrations of people, the amount of **municipal solid waste (MSW)** is particularly high and is set to increase significantly over the next decade as a result of urbanisation and rising living standards (see Figure 9.50). In 2002, there were 2.9 billion urban residents who generated about 0.64 kg of MSW per person per day. In 2012, this had increased to about 3 billion residents generating 1.2 kg /capita/day. By 2025 this is predicted to increase to 4.3 billion urban residents generating about 1.42 kg/capita/day of municipal solid waste. Solid waste is very much seen as an 'urban issue' because urban residents produce about twice as much waste as their rural counterparts. Globally, rural dwellers tend to be poorer, purchase fewer store-bought items (which results in less packaging), and have higher levels of reuse and recycling.

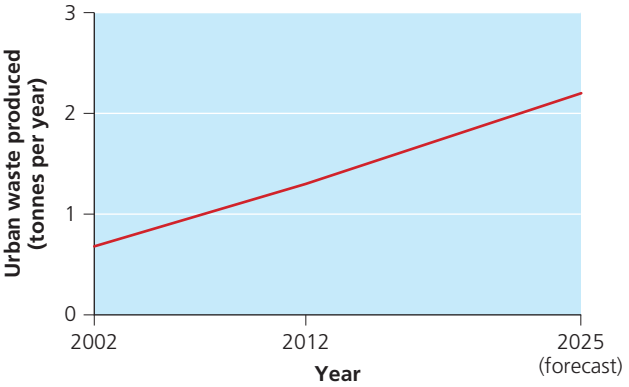


Figure 9.50 Urban waste generation 2002–2025

Source: 'What a Waste: A Global Review of Solid Waste Management', (2012), World Bank

Waste generation varies significantly between cities. As Figure 9.51 (page 438) shows, 2010 rates of waste production were much higher in cities in high-income countries. This is because waste generation tends

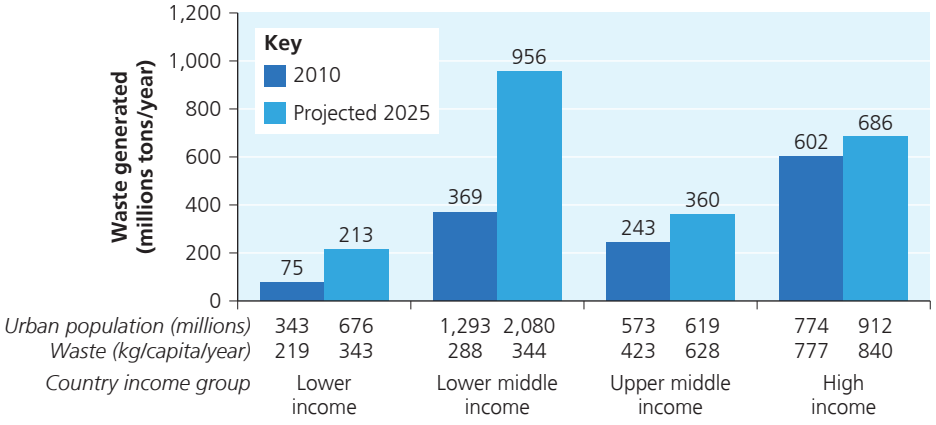


Figure 9.51 Urban waste generation by income level and year

to be greater where disposable incomes and living standards are higher. However, it is the cities in low- and middle-income countries which are set to see the biggest increase in waste generation over the next decade as a result of rapid urbanisation and continued industrialisation. The amount of municipal solid waste is growing fastest in China. It is already the world's largest generator of waste and by 2030, its volume of MSW is projected to be double that of the USA.

A tipping point?

One interesting finding to come out of studies on urban waste management is that both richer and poorer cities in a region often outperform middle-income ones. The average resident of Kuala Lumpur, for example, uses 497 litres of water per day and produces 816 kilograms of waste per year. This contrasts starkly to richer Singapore – 309 litres of water and 307 kilograms of waste, but also poorer Delhi – 209 litres of water and 147 kilograms of waste. Somewhere in the midst of economic growth, consumption appears to overtake sustainability and it is not controlled until a city becomes comparatively wealthy. The 'tipping point' in Asia appears to be a per capita GDP of around US $20,000.

Sources of waste in urban areas

Urban waste is made up of millions of separate waste items, the sources of which are detailed in Table 9.8. Some of the larger waste items such as organics (food and horticultural waste) and paper are easier to manage, but wastes such as multi-laminates, hazardous (for example, syringes) and e-waste pose disproportionately large problems for disposal. This is a particular concern for low-income countries where they may not have the facilities to properly dispose of them safely. Annually, 77,000 tonnes of hazardous waste is deposited in Cairo.

The composition of waste varies according to a number of factors such as level of economic development, cultural norms, geographical location, energy sources and climate. As a country urbanises and populations become wealthier, consumption of inorganic materials (such as plastics, paper and aluminium) increases, while the relative organic proportion decreases. Generally, low- and middle-income countries have a high percentage of organic matter in the urban waste stream, ranging from 40 to 85 per cent of the total.

Impacts of increasing waste generation

- The costs of collecting and treating waste are high. In lower-income countries, solid waste management is usually a city's single largest budgetary item and it is common for urban authorities to spend 20–50 per cent of their budget on solid waste management.
- Environmentally, waste is a large source of methane, a powerful greenhouse gas. Waste also contributes to water, ground and air pollution.
- Untreated or uncollected waste can lead to health problems such as respiratory ailments, diarrhoea, cholera and dengue fever.
- Many city authorities are struggling to collect increasing quantities of urban waste. The 2012 World Bank report on waste found that 30–60 per cent of urban solid waste in lower-income countries is uncollected. Egypt generates about 80 million tonnes of solid rubbish per year, according to government figures. The country's inadequate waste infrastructure is supplemented by the Zabaleen, Cairo's informal

Table 9.8 Generators and types of solid waste

Source	Waste generators	Types of solid waste
Residential	Households	Food wastes, paper, cardboard, plastics, textiles, leather, yard wastes, wood, glass, metals, ashes, special wastes (e.g. bulky items, consumer electronics, white goods, batteries, oil, tyres) and household hazardous wastes (e.g. paints, aerosols, gas tanks, waste containing mercury, motor oil, cleaning agents) and e-wastes (e.g. computers, phones, TVs)
Industrial	Light and heavy manufacturing, fabrication, construction sites, power and chemical plants	Housekeeping wastes, packaging, food wastes, construction and demolition materials, hazardous wastes, ashes, special wastes
Commercial	Stores, hotels, restaurants, markets, office buildings	Paper, cardboard, plastics, wood, food wastes, glass, metals, special wastes, hazardous wastes, e-wastes
Institutional	Schools, hospitals (non-medical waste), prisons, government buildings, airports	Same as commercial
Construction and demolition In some cities this can represent as much as 40 per cent of the total waste stream	New construction sites, road repair, renovation sites, demolition of buildings	Wood, steel, concrete, dirt, bricks, tiles
Urban services	Street cleaning, landscaping, parks, beaches, other recreational areas, water and wastewater treatment plants	Street sweepings; landscape and tree trimmings; general wastes from parks, beaches, and other recreational areas, sludge

Source: Adapted from 'What a Waste: A Global Review of Solid Waste Mangement' (1999), World Bank

waste collectors. Each morning, across Cairo they collect and bring home huge amounts of rubbish for the women and children to sort through.

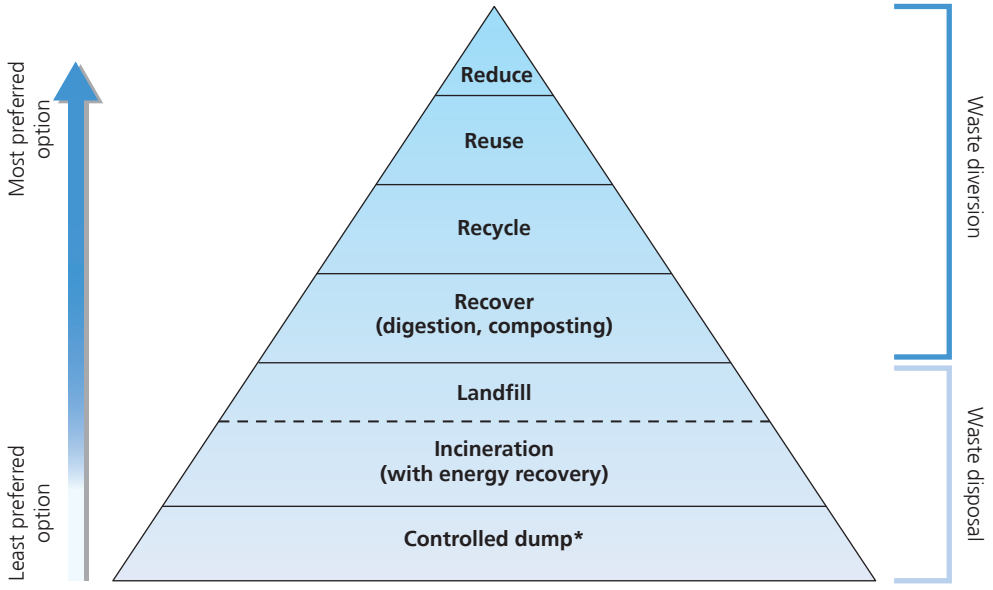
- Cities are running out of landfill space. The Borj Hammoud landfill site in Beirut, Lebanon, was planned to be shut down by 2019. This has been delayed because of a nearly 30 per cent reduction in daily waste production following mass protests and the COVID-19 lockdown. In fact, permission has been granted to raise it by a further 2 m. Following a waste crisis in 2015, when rubbish was left piled up in the streets, the government has proposed expanding existing sites. Local residents fear that the odours being emitted by the site are toxic pollutants. Long-term exposure to these strong odours can be linked to respiratory diseases, allergies and the spread of bacteria. In addition, water that has leached out from the present site is being dumped into the sea, polluting the water and making the sea in surrounding areas dangerous to swim in.

Approaches to waste management

Waste management generally follows an accepted hierarchy, as shown in Figure 9.52 (page 440). The key target is to reduce the amount of waste produced in the first place and this can be done through a combination of waste-related legislation, education and financial incentives. At the other end of the spectrum, waste disposal occurs through incineration or landfill.

Methods of waste management and disposal

Landfilling and thermal treatment of waste (incineration) are the most common methods of waste disposal in high-income countries. Most low- and lower middle-income countries dispose of their waste in open dumps. Some of this disposal may be **unregulated.** This means it is not controlled or supervised by regulation of law. Solid waste that is not properly collected and disposed of can be a breeding ground for insects, vermin and scavenging animals, and can thus pass on air- and water-borne diseases. A survey conducted by UN-Habitat in 2009 found that in areas where waste is not collected frequently, the incidence of diarrhoea is twice as high and acute respiratory infections six times higher than in areas where collection is frequent. Environmental threats include contamination of groundwater and surface water by leachate, as well as air pollution from burning of waste that is not properly collected and disposed of.



*A controlled dump is one in which there has been site selection, controlled access and possibly compaction of waste. The bottom ash, non-combustibles and bypassed waste from incineration all go to landfill.

Figure 9.52 The waste hierarchy

Recycling and recovery

Resource recovery is the selective extraction of disposed materials for a specific next use, such as recycling, composting or energy generation. **Recycling** is carried out when materials from which the items are made can be reprocessed into new products. In recent years, the global market for recyclables has increased significantly. The world market for post-consumer scrap metal is estimated at 400 million tonnes annually and around 175 million tonnes annually for paper and cardboard (UN-Habitat 2009). This represents a global value of at least $30 billion per year. Recycling, particularly in low- and middle-income countries, occurs through an active, although usually informal, sector. Estimates suggest that about 1 per cent of the urban population – at least 15 million people – survive by salvaging recyclables from waste.

Manufacturing new products using recycled materials can save significant energy. For example, producing aluminium from recycled aluminium requires 95 per cent less energy than producing it from virgin materials. **Urban mining** is the name given to the process of recovering compounds and elements from products, buildings and waste which would otherwise be left to decompose in landfills. By collecting and salvaging valuable components to be reused and recycled, there is a greater chance of reducing landfill waste.

The key advantages of recycling and recovery are reduced quantities of disposed waste and the return of materials to the economy. However, there are some negative environmental issues. Notably, energy may be required for the operation of material recovery from waste and this leads to greenhouse gas emissions. Informal recycling by waste pickers will have little greenhouse gas emissions, except for processing the materials for sale or reuse, which can be relatively high if improperly burned (for example, metal recovery from e-waste).

Trade

Waste may be moved between and within countries. The **global waste trade** is the international trade of waste between countries for further treatment, disposal, or recycling. Toxic or hazardous wastes are often exported from high to low-income countries as seen in the example of e-waste in the textbox.

Guiyu: The town of electronic waste

Electronic waste, also known as e-waste, refers to discarded electrical or electronic devices. A rapidly growing surplus of electronic waste around the world has resulted from quickly evolving technological advances and an estimated 44.7 million tonnes was generated in 2016. According to the UN Step report (2013), e-waste is now the world's fastest growing waste stream. Much of this e-waste comes from the USA and Europe but it is shipped to poorer countries in Asia and Africa to be processed and recycled, often informally. Electronic goods often contain toxic substances such as lead, mercury, cadmium, arsenic and flame retardants. Once in landfill, these toxic materials seep out into the environment, contaminating land, water and the air. Those who work at these sites suffer frequent bouts of poor health.

The town of Guiyu in Guangdong province in China has long been known as one of the world's largest electronic waste dump sites. At its peak, some 5,000 workshops recycle 15,000 tonnes of waste daily including mobile phones, computer screens and computers shipped in from around the world. Large quantities of pollutants, heavy metals and chemicals are released into the air and local water supplies are heavily contaminated by toxic metal particles. Locals suffer respiratory and stomach problems and children living in the area have abnormally high levels of lead in their blood.

However, these countries often do not have safe recycling processes or facilities and hazardous wastes are not properly disposed of or treated, leading to contamination of the surrounding environment. International laws such as the Basel Convention have been introduced to prevent transboundary movement of hazardous waste but evidence suggests it still happens.

Reduction (incineration)

Incineration of waste (with energy recovery) can reduce the volume of disposed waste by up to 90 per cent. General waste can be safely burned at high temperatures and under carefully controlled conditions to produce electricity and heat. This is referred to as **energy from waste** and an increasing number of cities are processing their waste in this way. Incineration without energy recovery is still common but it is not a preferred option due to costs and pollution. The open-burning of waste, which is common in poorer countries, is particularly discouraged due to severe air pollution associated with low temperature combustion.

Burial (landfill)

Burial is the placement of waste in man-made or natural excavations, such as pits or landfills. Landfill sites are a common final disposal site for waste from urban areas. In lower-income countries there may simply be a hole in the ground where open dumping occurs. In higher-income countries, there are much stricter regulations and the types of material that can be sent to landfill are often defined by law. In the UK, most landfill sites now control and collect the gas that is released by the decomposing waste, often using it to generate electricity through turbines.

The environmental problems caused by landfills can be numerous. The greenhouse gas methane is produced by rotting organic matter and other chemicals like bleach and ammonia can produce toxic gases that negatively impact the quality of air in the vicinity. Dust and other forms of non-chemical contaminants can also make their way into the atmosphere. Landfills can also affect groundwater and river quality because toxic chemicals can leach out and contaminate the water.

Since 1999, Singapore has been incinerating much of its waste and sending the incineration ash and non-incinerable waste to an offshore landfill called Semakau Landfill. The Semakau site has been formed by the construction of a dyke in the shallow sea. It is lined with an impermeable membrane to stop leakage into the sea below. Each cell is covered with a layer of earth once it has been filled to ground level. Subsequently, grass and trees take root to form a green landscape. The actual **submergence** of waste in oceans is banned by international convention but, according to the United Nations, some companies have been dumping radioactive waste and other hazardous materials into the coastal waters off Somalia, taking advantage of the fact that the country lacks strong governance.

Table 9.9 A comparison of incineration and landfill

Type of disposal	Advantages	Disadvantages
Landfill	Facilities are properly sited with necessary controls Different types of waste accepted and ordered	Unsightly Often opposed by neighbouring residents Potential leaching of chemicals threatens groundwater supply Decaying matter produces methane, a strong greenhouse gas which is also explosive Landfill takes up a lot of space High transportation costs
Incineration	Can reduce volume of waste needing disposal by 90 per cent Can inactivate disease agents Can reduce toxicity of waste Can be used to produce energy Incinerator bottom ash can be recycled as a secondary aggregate	Expensive Not all waste is combustible Poses challenges of air pollution and incinerator bottom ash disposal Capacity limitations Unpopular with local residents

Landfill vs incineration in Amsterdam

Amsterdam is the cultural capital of The Netherlands – a densely populated, low-lying country which has a growing population and a diminishing amount of spare land. In the late twentieth century, rising prosperity in the country had led to the greater production and consumption of goods and a consequent increase in waste generation. Lack of space and growing environmental awareness forced the Dutch government to take measures to reduce their reliance on landfill.

The Dutch approach is simple: avoid creating waste in the first place, recover the valuable raw materials from it, generate energy by incinerating residual waste, and only then dump what is left, but do so in an environmentally acceptable way. This approach – known as 'Lansink's Ladder' – was incorporated into Dutch legislation in 1994.

Landfill

- The increasing level of material consumption and the significant lack of physical space, together with environmental deterioration of the land, forced the Dutch government to take measures early on to reduce the landfilling of waste.
- There were increasing numbers of objections to waste disposal sites from the public due to smell, soil pollution and groundwater contamination.
- In 1995, the government introduced a landfill tax on every tonne of material landfilled. This gave waste processing companies the financial incentive to look for other methods such as recycling and incineration. Landfill tax was increased year on year until 2012 when it was repealed because the low level of landfilling had rendered the tax unnecessary.
- A landfill ban covering 35 waste categories was introduced in 1995.
- The amount of waste sent to landfill decreased significantly in the late 1990s and early 2000s. By 2006, the country had already reached the targets of the Landfill Directive set for 2016.

Waste to energy incineration in Amsterdam: the Afval Energie Bedrijf (AEB) incineration plant

- In a strategy referred to as waste-to-energy (W2E), Amsterdam has created the Afval Energie Bedrijf (AEB) incineration plant capable of producing 1 million MWh of electricity annually. Beyond the energy factor, the plant is also being used to create heating for several communities around Amsterdam, and produces 300,000 gigajoules of heat annually.
- Annually, 1.4 million tons of waste is brought to the W2E plant. This amounts to 600 trucks and one freight train per day of refuse from the Amsterdam metropolitan area.
- 64 per cent of the waste that ends up at the plant is recycled.
- Aware of the environmental effects of the gases from this process, the AEB plant has installed a complex process of scrubbing the flue gases. Attempts are being made to close the loop for other by-products so that the material can be used in other industrial processes – from trace elements for manufacturing, to fly ash for construction. Whatever material is left becomes landfill.
- Next door to the W2E plant is the Waternet water treatment plant. The two plants work together: the incineration plant supplies energy and heat for water treatment processes; the water treatment plant injects its sludge and biogas into the incineration plant as an additional fuel source.

- As the plant performs several functions (elimination of waste, generation of electricity and heat) simultaneously, it compares positively to other disposal methods, actually avoiding 438 kilotons of CO_2 per year.

The environmental comparison

If the equivalent amount of waste was put into landfill, the CO_2 emissions per year would be 1036 kilotons. This would mainly be as a result of methane gases developing at the landfill. If these gases were either captured or burned off, it would reduce the total emissions to 404 kilotons – still higher than the 438 kilotons of CO_2 'saved' by the AEB incineration plant.

Conclusion

The waste to energy strategy in Amsterdam is considered both economically and environmentally better than landfill. However, incineration is still a controversial waste disposal strategy and in 2014, the Dutch government set the target to reduce waste-to-incineration by 50 per cent and focus on improving rates of reduction and recycling. Like other twenty-first century cities aiming for greater sustainability, the movement away from waste disposal to overall waste reduction is the ultimate goal.

Further research

Even though Singapore has an efficient waste collection and disposal system, the government, as part of a new sustainable programme to meet future environmental challenges, has called for a change in the way people view waste management. The government is encouraging everyone to consume less and conserve more. Go to: www.towardszerowaste.sg/zero-waste-masterplan/

Review questions

11 What are the key issues surrounding urban waste in the twenty-first century and how do these vary in cities around the world?

12 What are the best strategies for tackling urban waste disposal?

13 What do you understand by the term 'waste stream'?

9.7 Other contemporary urban environmental issues

Environmental problems tend to be worse in poorer cities experiencing rapid growth and impact most severely on the more vulnerable groups within that urban society. The main problems concern pollution of the air and water and waste disposal but urbanisation is also leading to a loss of land and therefore habitats. By 2030, it is estimated that the urban land area will have expanded by as much as 3.3 million square kilometres. The Cities and Biodiversity Outlook (CBO) project states that 'urban expansion is occurring fast in areas adjacent to biodiversity hotspot areas and faster in low-elevation, biodiversity-rich coastal zones than in other areas'. Noise pollution can also be an issue. In Shenzen, China, three-quarters of the complaints about pollution concern noise pollution rather than air or water pollution. Noise pollution is linked mainly to traffic and industry.

This section of the chapter focuses on three of the biggest environmental threats facing cities in the twenty-first century: atmospheric pollution, water pollution and dereliction.

Atmospheric pollution

Atmospheric pollution is caused by the release of particles and noxious gases into the atmosphere and this can have a negative effect on human health. In 2014, the WHO found that urban air pollution was 2.5 times higher than the recommended levels in about half of the urban populations being monitored. Air pollution accounts for an estimated 4.2 million deaths per year due to stroke, heart disease, lung cancer and chronic respiratory diseases. Around 91 per cent of the world's population lives in places where air quality levels exceed WHO limits. An estimated 3,799 Londoners died prematurely from air pollutants in 2019, principally from exposure to nitrogen dioxide and fine PM2.5 particles; down from 9,400 in 2010.

Dhaka's air pollution

Dhaka, the capital of Bangladesh, is the world's fastest growing megacity with as many as 400,000 new migrants arriving each year. It stands on the bank of the Buriganga River and developed as an important trading centre. It has a diverse economy, a high literacy rate and rapidly developing infrastructure, but the city faces major urban challenges such as poverty, crime, congestion and political violence. However, one of the biggest problems the city faces is air pollution. In 2020, Dhaka topped the list of cities with the worst atmospheric pollution. On many occasions it is categorised as 'hazardous'. The pollution is caused mainly by vehicle emissions and industrial practices. The two-stroke engines used in auto rickshaws (baby-taxis), tempos, mini-trucks and motorcycles have been identified as highly polluting and there are estimates that about 90 per cent of the vehicles that ply Dhaka's streets daily are faulty and emit smoke far exceeding the prescribed limit. Industrial pollution comes from factories in and around the periphery of Dhaka, manufacturing products including textiles, fertilisers, pesticides and pharmaceuticals. However, according to a 2011 World Bank report, it is the brick-making kilns which account for about 40 per cent of Dhaka's fine-particle air pollution (Figure 9.53).

The kilns cause 750 premature deaths a year from cancer and cardiopulmonary disease. In 2010, the national government ordered a shutdown of fixed-chimney kilns by July 2013. Facing opposition from kiln owners, the government has extended that deadline several times.

Figure 9.53 Air pollution from the brick kilns of Dhaka. The biggest culprit is the 'fixed-chimney' kilns, which use outdated technology to burn coal imported from India and firewood chopped from nearby forests.

Human activity produces the pollutants but it is primarily the weather that determines what will happen once they are released into the air. During wet or windy conditions pollution concentrations remain low, either blown or washed away. During periods of still, hot weather, pollution is able to build up to harmful amounts, leading to what are known as **pollution episodes**.

Managing air pollution

A number of strategies used to improve air quality are described on pages 431–2. Further measures include ensuring that:

- houses are energy efficient
- urban development is well served by public transport routes
- street design is appealing and safe for pedestrians and cyclists
- waste is well managed.

More park and woodland areas also help. The environmental company AECOM has calculated that London's 8.3 million trees provide £95 million worth of air filtration every year in terms of health costs avoided. Together, these strategies serve as a catalyst for local economic development and the promotion of healthy urban lifestyles.

Figure 9.54 Air pollution monitoring station in London

GIS focus

City pollution levels are monitored continuously (Figure 9.54, page 444) and you can use this data to examine geographical patterns. Go to www.londonair.org.uk/london/asp/nowcast.asp for a 'nowcast', which is a service to show current pollution levels in detail across London in comparison with the government's Air Pollution Index. The Daily Air Quality Index (DAQI) details levels of air pollution in the UK and provides recommended actions and health advice. For European cities, the Urban Air Quality Index is updated every hour and allows comparison of air quality over a 24-hour period. Other forms of social media are increasingly used to convey information to the general public. For example, UK-AIR provides automated tweets about current and forecast air quality including episodes of poor air quality.

There are greater concerns about air pollution in low- and middle-income countries which lack the finances and legislation to adequately tackle the problem. Over 90 per cent of air pollution in these countries is attributed to vehicle emissions due to the high number of older vehicles, poor vehicle maintenance and low fuel quality. More recently, schemes such as the UN Partnership for Clean Fuels and Vehicles aim to improve air quality; there has been greater investment in improving road quality; and commitments have been made to promote non-motorised journeys.

Water pollution

Water pollution is the contamination of water sources including rivers, lakes, oceans, aquifers and groundwater. It occurs when pollutants are directly or indirectly discharged into water without adequate treatment to remove harmful compounds. Indirect sources of water pollution include contaminants that enter the water from soils or groundwater and from the atmosphere via rain. The high concentration of impermeable surfaces in urban areas increases run-off from roads and can carry numerous pollutants such as oils, heavy metals, rubber and other vehicle pollution into waterways and streams. The reduction in water percolation into the ground can also affect the quantity and quality of groundwater. Stormwater run-off in urban areas can overwhelm combined stormwater and wastewater treatment systems when high volume flows exceed treatment capacities. Reports suggest that urban stormwater can be as polluted as untreated domestic wastewater.

Causes of water pollution in urban areas

- Surface run-off from streets carrying oil, heavy metals and other contaminants from motor vehicles.
- Industrial waste.
- Untreated or poorly treated sewage which is low in dissolved oxygen and high in pollutants such as nitrates, phosphorus and bacteria. Treated sewage can still be high in nitrates.
- Rubbish dumps, toxic waste, chemical and fuel storage, which can all leak pollutants.
- Intentional dumping of hazardous substances.
- Air pollution can lead to acid rain, nitrate deposition and ammonium deposition, which can alter the water chemistry of an area.

Consequences of water pollution

Over 1.2 billion people lack access to clean water and waterborne infections account for 80 per cent of all infectious diseases. Increased water pollution creates breeding grounds for malaria-carrying mosquitoes as well as damaging ecosystems, leading to species extinction. For humans, different forms of water pollutants affect health in different ways:

- Heavy metals from industrial processes can slow development, result in birth defects and may be carcinogenic.
- Industrial waste often contains toxic compounds that damage the health of aquatic animals and those who eat them. They can cause immune suppression, reproductive failure or poisoning.
- Microbial pollutants from sewage often result in infectious diseases that infect aquatic life and terrestrial life through drinking water. Microbial water pollution causes diseases such as cholera and typhoid fever, which are a major cause of infant mortality in low-income countries.
- Organic matter and nutrients can cause an increase in aerobic algae and deplete oxygen from the water. This can lead to the suffocation of fish and other aquatic organisms.
- Suspended particles in freshwater reduce the quality of drinking water for humans and the aquatic environment for marine life. They can also reduce the amount of sunlight penetrating the water, disrupting the growth of photosynthetic plants and micro-organisms.

Water problems in India

India suffers chronic water problems. In 2019, it was estimated that around 70 per cent of surface water in India was unfit for consumption. A primary cause of this was the quantity of domestic sewage generated by rapidly expanding towns and cities. A survey of 27 cities found that untreated sewage flowing in open drains was causing serious deterioration of groundwater quality with knock-on impacts to human health. Vector-borne diseases such as cholera, dysentery, jaundice and diarrhoea are widespread and water pollution was also found to be a major cause of poor nutrition and under development in children. According to the World Health Organization, more than 87 per cent of people in India's cities (compared with 33 per cent in rural areas) now have access to a toilet, but it is the leaking and incomplete sewage systems that are leading to the contamination of rivers and lakes. The World Bank believes that such a release of pollution upstream lowers economic growth in downstream areas, reducing GDP growth in these regions by up to a third.

Steps are being taken to tackle India's water problems. The Prime Minister Narendra Modi has made cleaning the Ganges, the river that is holy to Hindus, a key policy goal, while there are also plans for infrastructural improvements and wastewater recycling in cities across the country. In the city of Ahmedabad, the state council introduced a scheme in which children are being paid to use public toilets. Open defecation is a problem where people do not have access to proper sanitation. The number of public toilets has also been increased to prevent human waste running into wells and streams, contaminating water that may then be used for drinking or bathing.

Figure 9.55 The Dharavi slum area in Mumbai – severely polluted with solid and liquid wastes generated by Mumbai

Managing water pollution

Improving water quality requires strategies to prevent, treat and remediate water pollution. Ideally pollutants are prevented from entering watercourses in the first place but the reality is that potential pollutants are treated before they are discharged. The final and often most expensive strategy is when polluted watercourses are restored through remediation.

In most high-income countries, water quality improvement focuses on two approaches:

1. the construction of water-treatment facilities and wastewater plants
2. regulations aimed at 'point source' polluters such as industries which discharge water pollution into receiving waters or sewer systems that flow into treatment plants.

The most difficult water-quality challenge is dealing with 'non-point source' pollution which is the result of precipitation run-off from a wide range of sources including fertilisers and pesticides from agriculture, and chemicals and toxins from urban settlements. These pollutants are difficult to regulate.

Lack of money and inadequate technology in low-income countries have resulted in much lower water quality standards. Effective legislation is often absent and enforcement of pollution controls limited.

Key water pollution strategies

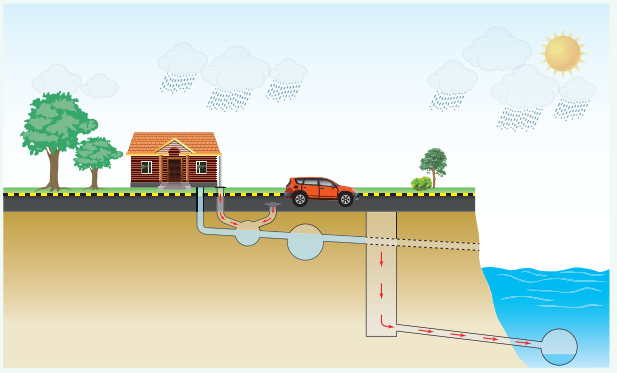
- **Low-impact development (LID)** is a stormwater management approach that can help to reduce stormwater run-off. This is done primarily through the use of vegetation and permeable surfaces to allow infiltration of water into the ground. Permeable streets and pavements, 'green' roofs, rain gardens and more urban parks allow water to infiltrate into soils rather than flow directly into sewers. Filtering stormwater through vegetation and soil has been shown to reduce organic pollutants, oils and heavy metals by more than 90 per cent.
- **Legislation, regulation and enforcement:** There are many different anti-pollution laws and agreements in operation worldwide. However, these laws need to be enforced. Some cities have adopted incentive-based approaches, charging polluters per unit. Charges start low but are increased if pollution continues, creating an incentive to reduce discharges and purchase wastewater treatment technologies. Regulation works

in a similar way. Factories are allowed to discharge only limited amounts of carefully controlled pollutants. By slowly reducing the levels of permitted discharges, year by year, pollution levels are reduced.

- **Education and awareness:** The more people know about the causes and effects of pollution, the more likely they will be to avoid adding to the problem. In 2014, Wessex Water used mobile billboards in hotspot areas urging its customers to get behind a campaign encouraging people to bin wet wipes rather than flush them down the toilet. Wet wipes are a common cause of sewer blockages since they do not decompose like toilet paper.
- **Improvements in sewage and wastewater processing**.
- **Appropriate technology:** One example of this is the Janicki omniprocessor – a small-scale innovation aimed at providing clean water in low-income countries. The omniprocessor first boils raw 'sewer sludge', during which the water vapour is separated from the solids. Those solids are then put into a fire, producing steam that drives an engine producing electricity for the system's processor and for the local community. Finally, the water is put through a cleaning system to produce drinking water.

The Thames Tideway Tunnel

Much of London is served by a combined sewerage system, collecting sewage from homes as well as rainwater runoff from roads, roofs and pavements. The interceptor sewers, constructed by Sir Joseph Bazalgette following the 'Great Stink' of 1858, are still the backbone of London's sewer network today, but they are struggling to cope with the expanding population and the demands of modern-day living. During periods of heavy rain, the Victorian drainage system discharges raw sewage directly into the River Thames, killing large quantities of fish and other aquatic life and threatening public health. In 2013 alone, 55 million tonnes of raw sewage was washed into the river. The Thames Tideway Tunnel is under construction to upgrade London's sewerage system (Figure 9.56). A specialised tunnel boring machine (TBM) has been used to excavate a circular tunnel, seven metres in diameter, using a rotating cutterhead, while simultaneously creating a tunnel wall using concrete segments. The completed tunnel will be 25 kilometres long and up to 65 metres deep. The tunnel will take the sewage to the Becton sewage treatment works in the Essex marshes and, when thoroughly cleaned, it will be released into the Thames. For more information, go to www.tideway.london/the-tunnel. It is hoped that the whole scheme will be completed in 2025.



Now:
The low level interceptor sewers fill up and overflow into the River Thames.

After:
The overflow will be diverted into the tunnel instead of going into the river.

Figure 9.56 By intercepting the sewage before it enters the river, the Thames Tideway Tunnel will help prevent the tidal River Thames from being polluted with untreated sewage, which can stay in the river for up to three months before the tide finally takes it out to sea

Dereliction

Dereliction refers to the state of having been abandoned and become dilapidated. In urban areas, derelict buildings are often associated with former industrial sites or run-down housing estates. In the UK, de-industrialisation led to many people leaving the inner city and industrial buildings being abandoned. Alongside this, services such as public houses and shops may also have become vacant as areas became subject to urban decline.

Dereliction has a negative impact on the surrounding area. Crime and vandalism rates tend to be higher, house prices fall and out-migration of residents takes place. The high costs involved in urban renewal often discourage authorities and individuals from renovation or rebuilding. In addition, the presence of listed buildings, which are subject to considerable planning regulation, can significantly increase the investment needed.

Derelict land can pose a risk to human health. Contamination from industrial processes lives on in an environment long after the industry that produced or used them is gone. In Glasgow, research has linked proximity to contaminated derelict sites with health problems such as low birth weight infants, heart disease, cancer and respiratory disease. Many of the industries prevalent in Glasgow during the industrial period were 'dirty' ones, with high usage of toxic and dangerous chemicals and environmental degradation

of the surrounding areas. These industries included shipbuilding, dye works, tanneries, distilleries, cast iron foundries and chemical manufacturing.

Derelict London

For some, derelict buildings provide a fascinating look at our architectural past. The *Derelict London* website includes over 3,000 photographs of London sights not normally seen by tourists. The author Paul Talling clearly states that not all of the photographs are of derelict buildings but, as the *New Statesman* review suggests, the appeal of the website is 'in how it seems to trace the skeleton of a dead city while it is still in apparently rude health.'

The website is a useful resource for looking at how buildings have changed over time. Some properties have fallen into even greater disrepair but there are also photographs taken before and after regeneration has taken place.

Figure 9.57 Lambeth Hospital buildings, Kennington 2008 (left) and 2013 (right)

Lambeth Hospital shut in 1976. Much of the original workhouse was demolished, although the water tower survived but lay derelict. Some of the more modern hospital buildings remained in a derelict condition until being cleared for a new housing development. The former workhouse administrative block is now a cinema museum.

Source: www.derelictlondon.com

Key terms

Brownfield site – A term used in urban planning to describe land previously used for industrial purposes or some commercial uses.

Dereliction – The state of buildings having been abandoned and become dilapidated.

Greenfield site – An area of undeveloped land.

Land remediation – The removal of pollution or contaminants from the ground, which enables areas of derelict former industrial land to be brought back into commercial use.

Strategies to deal with derelict urban sites

One of the most common strategies for tackling urban dereliction is through **regeneration schemes**. UK government-led schemes such as Urban Development Corporations, City Challenge and City Deals have been well documented and have had varying levels of success in tackling derelict areas. The government focus on using **brownfield sites** for new building developments rather than **greenfield sites** has also met with limited success. According to government figures, the proportion of homes built on the green belt doubled, from 2 to 4 per cent of all new homes built between 2016 and 2017, while the share built on brownfield land fell from 61 to 56 per cent. London had the greatest proportion of dwellings (including conversions) built on previously-developed land (98 per cent). The advantages of using brownfield sites in urban areas are numerous: it improves the physical environment; it revives older urban communities; existing infrastructure can reduce costs and encourage faster occupancy; it preserves historical landmarks and heritage architecture; it reduces urban sprawl and it preserves greenfield land. The disadvantages tend to focus on the greater costs of clearing contaminated land and the fact that most brownfield sites are in inner city areas, which have higher levels of traffic congestion and noise.

Land remediation

Land remediation is the term given to the removal of pollution or contaminants from the ground. This enables large areas of derelict former industrial land to be brought back into commercial use. In preparation for the 2012 London Olympic Games, a 350-hectare area of East London was 'cleaned up' and has now become the largest new urban park in Europe, with 100 hectares of open land and 45 hectares of new habitat. Over 2.2 million square metres of soil was excavated, of which nearly half was treated by soil washing, chemical stabilisation, bioremediation or sorting. Eighty per cent of the excavated material was re-used on site. A total of 235,000 m^3 of contaminated groundwater was successfully treated.

Community action

Community action has been at the heart of Detroit's recent strategy to tackle its derelict land areas. Once the centre of global car production, Detroit's population has fallen from nearly two million in the 1950s to less than one million today. The city now has 40,000 vacant and derelict lots, comprising around one-third of its area. Between 1970 and 2000, over 150,000 buildings were demolished and large commercial developments built in order to try and revitalise the area, but the decline continued. Detroit community groups are now using areas of the city for a range of small, community-based activities, including urban farming. By 2010, they had converted over seven hectares of unused land into more than 40 community gardens and microfarms yielding over six tonnes of produce a year, including hay, alfalfa, honey, eggs, milk, beef, flowers, vegetables and herbs. Much of it is produced by volunteers and students, and sold on to other community organisations such as soup kitchens. Derelict buildings have been converted into community centres, cafés and greenhouses. The chief executive of Recovery Park Farms, an organisation creating the urban farms, says that as the city has so much vacant land, it could farm up to 2000 hectares commercially indoors, and create up to 20,000 jobs directly and indirectly.

Review questions

14 Which aspects of urban sustainability do you think are the most important and why?

15 How realistic is it to expect all cities, rich and poor, to adopt sustainable practices?

9.8 Sustainable urban development

Cities pose a threat to both the local and global environment. Cities consume three-quarters of the world's resources and generate a majority of the world's waste and pollution. We have already seen how the consequences of waste disposal, pollution and dereliction are felt most severely in the cities in which they are created but, increasingly, the effects of urban growth are also being felt globally. Cities now rely on energy and resources from all over the world, while the pollution and waste they generate is also dispersed globally.

Ecological footprint

Ecological footprint is a measure of the environmental impact of communities or individuals. It is defined as the total area of productive land and water required to produce the resources a population consumes and absorb the waste it produces. To live within the means of our planet's resources, the world's ecological footprint would have to equal 1.7 hectares per person; in 2019 humanity was using natural resources 1.75 times faster than our planet's ecosystems can regenerate.

The variation both between countries and within countries is huge. Cities tend to have a higher ecological footprint than rural areas and wealthier cities have a higher ecological footprint than poorer ones. London's ecological footprint (Figure 9.58) covers an area almost twice the size of the UK.

Figure 9.58 The ecological footprint of London in 2011

Research by the Global Footwork Network (2011) found that, for example, San Francisco's footprint was about six per cent higher than the average American's. The higher than average footprint is attributed to increased consumption and waste production by residents and services within the city. Therefore, reducing the ecological footprint of cities is a priority to achieve greater global sustainability.

Dimensions of sustainability

One of the post-2015 Sustainable Development Goals is to make cities inclusive, safe, resilient and sustainable and the city as an entity is viewed as a key factor in building a more sustainable world. The idea of sustainability was first brought to wider public

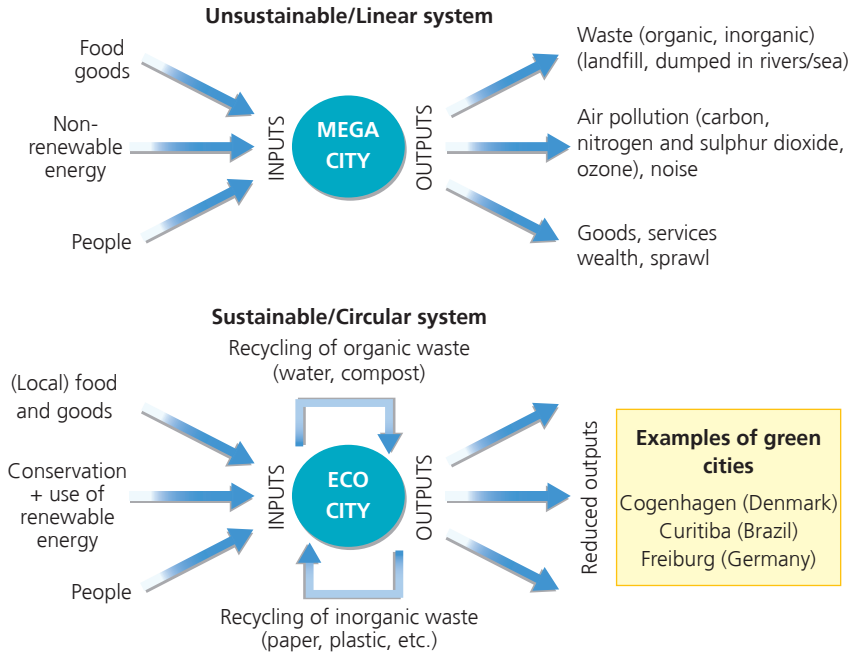


Figure 9.59 The city as a system

Source: Rogers, R. (1997). Cities for a small planet

awareness in the 1987 Brundtland Report entitled 'Our Common Future' by the UN World Commission on Environment and Development (WCED). **Sustainable development**, it stated, was 'meeting the needs of the present without compromising the ability of future generations to meet their own needs.' The term incorporates social, economic and environmental concerns and a **sustainable city** is therefore one which provides employment, a high standard of living, a clean, healthy environment and fair governance for all its residents, as shown in Figure 9.60. People-centred planning has increasingly been incorporated into sustainable urban design and the notion of **liveability** is important here. Liveability may mean different things to different people. For some, it is tied to natural amenities such as parks and green space; for others, to cultural offerings, career opportunities, economic and political stability or some degree of safety within which to raise a family. In the context of the global liveability rankings, it relates to which cities provide the best or worst living conditions for their residents.

Features of sustainable cities

Thinking of the city as a system enables a comparison between the characteristics of a typical megacity and that of a green city. The linear system of the megacity is unsustainable. Uncontrolled use of inputs and outputs leads to resources becoming exhausted and extremely high pollution and waste levels. The circular nature of the sustainable city system shows how some of the outputs are recycled, which reduces both the demand for new input resources and pollution and waste levels (Figure 9.59).

Key terms

Ecological footprint – The total area of productive land and water required to produce the resources a population consumes and absorb the waste produced.

Liveability – The characteristics of a city which improve the quality of life for the people living there.

Sustainable city – A city which provides employment, a high standard of living, a clean, healthy environment and fair governance for all its residents.

Sustainable development – Meeting the needs of the present without compromising the ability of future generations to meet their own needs.

Urban resilience – The capacity of individuals, communities, institutions, businesses and systems within a city to survive, adapt and grow, no matter what kinds of chronic stresses and acute shocks they experience.

Sustainability can be measured in a number of ways and different organisations will use different criteria to rank cities in terms of their sustainability. Terms such as 'eco-city' and 'green city' are often used to describe cities with a good environmental record, but the range of indicators used for measuring urban sustainability needs to be far greater and incorporate the social, economic and political elements highlighted in Figure 9.60 and Figure 9.61.

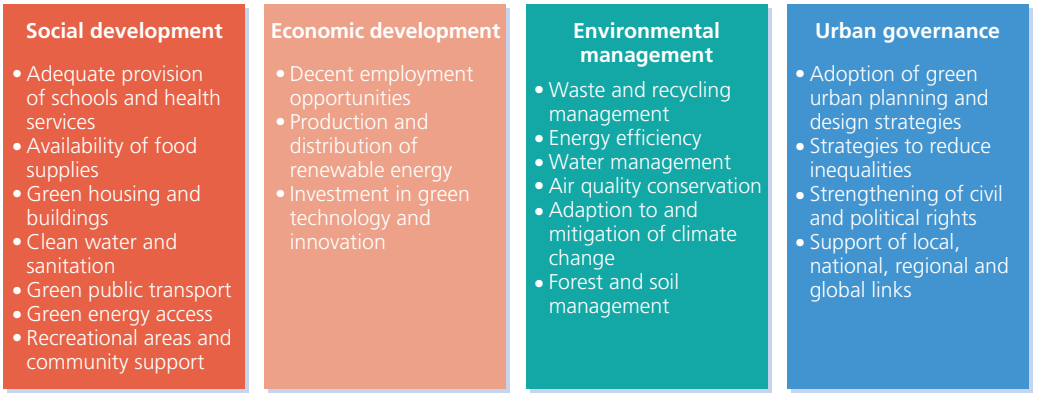


Figure 9.60 Pillars for achieving urban sustainability

Source: United Nations

Another twenty-first century buzzword which can be linked to the idea of sustainability is that of resilience. **Urban resilience** is defined as the capacity of individuals, communities, institutions, businesses and systems within a city to survive, adapt and grow, no matter what kinds of chronic stresses and acute shocks they experience.

Table 9.10 Urban resilience

Chronic stresses (day-to-day challenges)	Acute shocks (sudden threatening events)
High unemployment	Earthquakes
Inefficient public transport	Floods
Endemic violence	Disease outbreaks
Chronic food and water shortage	Terrorist attacks

A sustainable city is one which can cope with the physical, social, and economic challenges that are a growing part of the twenty-first century.

Go to https://resilientcitiesnetwork.org/ to look at some of the proposed ideas for developing greater urban resilience.

Contemporary opportunities and challenges in developing more sustainable cities

The main obstacle to cities becoming more sustainable tends to be financial. In the context of rapidly growing urban populations and limited budgets, city authorities often choose short-term 'cheaper' solutions over long-term planning. Likewise, in cities where large numbers of people still lack access to water and electricity, providing basic infrastructure tends to take priority over investment in environmental projects and sustainable development. Further challenges to developing sustainable cities include poor infrastructure, weak institutions and lack of enforcement of planning regulations.

Table 9.11 on page 453 highlights some of the main challenges facing cities. Drawn from the United Nations World Economic and Social Survey 2013, the table also suggests possible opportunities for overcoming these problems.

Figure 9.61 Key features of sustainable cities

Copenhagen: Europe's most sustainable city?

Copenhagen, the capital city of Denmark, regularly tops the list of the world's most sustainable cities and was awarded the title of European Green Capital in 2019. The climate action plan (2011) for the city stated that the city aims to be a green, smart, carbon-neutral city by 2025. It aims to do this by:

- adding 100 new wind turbines
- reducing both heat and commercial electricity consumption by 20 per cent
- encouraging 75 per cent of all journeys to be by bike, on foot or by public transport
- the biogasification of all organic waste
- installing 60,000 m^2 of new solar panels
- meeting 100 per cent of the city's heating requirements with renewables.

Social development

- Only two per cent of employees in Copenhagen work more than 40 hours a week. This allows more time for family, friends, hobbies and recreational activities, all of which are associated with lower stress levels.
- The city has 249 miles of cycle lanes, which makes cycling an easy and safe option. More than half of commuters travel to work or school by bike each day. It is estimated that the residents of Copenhagen have one million fewer sick days simply due to the fact that they cycle so much.

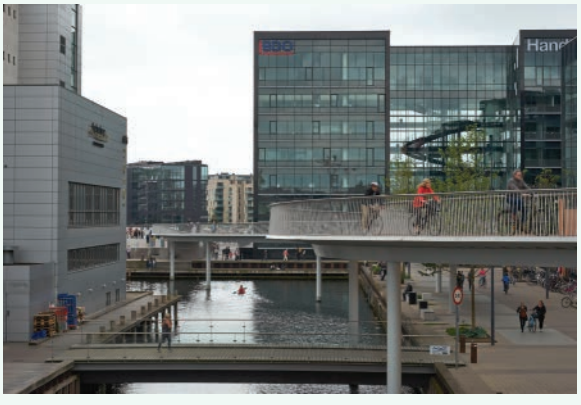


Figure 9.62 The elevated *Cykelslangen,* or Cycle Snake opened in Copenhagen in 2015. It is a two-way 220-metre bridge. You can get across the harbour in less than a minute, creating a less stressful atmosphere for cyclists, pedestrians and vehicle users alike

Economic development

- Copenhagen is the economic and financial centre of Denmark. It is home to a number of international companies and has successful business clusters in IT, biotechnology and pharmaceuticals.
- Copenhagen has some of the highest wages in the world and unemployment rates are low.
- Danes have an entrepreneurial spirit. In Copenhagen they have held on to their independent shops and cafes. Small boutiques and local coffee shops are the norm rather than global chains.
- Reasonable rents in the city allow young start-up companies to flourish.

Figure 9.63 Added to its green credentials, Copenhagen regularly tops the lists of the most desirable cities in which to live

Environmental management

- The harbour that bisects the city has been transformed from an industrial zone into a cultural and residential hub. Where once 100 overflow channels carried wastewater into Copenhagen harbour after heavy rains, investment in reservoirs and conduits has meant that the harbour water is now so clean that residents can go swimming in it.
- Copenhagen is served by an integrated public transportation network: a driverless Metro, regional trains and buses.
- In 2001, a large offshore wind farm was built just off the coast of Copenhagen at Middelgrunden. It produces about 4 per cent of the city's energy.
- Copenhagen's Finger Plan 2007 includes protection of its green belt and limits urban sprawl through better use of existing city land.
- In 2019, the Amager resource centre opened. It is a waste incinerator building that also incorporates hiking and running trails, a street fitness gym and an 85-metre climbing wall – the world's highest outdoor climbing wall.

Urban governance

- As a country, Denmark has one of the highest taxation levels in the world but this provides very generous state welfare provision which reduces inequality.
- The Citizen's Dialogue Project, a publicly financed scheme, involves direct public participation in planning.
- Denmark boasts very high income equality and gender equality and is perceived to be a very fair and egalitarian society.

Table 9.11 Challenges to and opportunities for building sustainable cities

Main urban trends	Developing countries		Developed countries	
	Challenges	Opportunities	Challenges	Opportunities
Social				
By 2025, urban population will live mainly in small (42 per cent) and medium-sized (24 per cent) cities	Improve access to housing, water and sanitation; improve public infrastructure; foster institutional capacity	Investment in public infrastructure (including transportation); construction of compact buildings in middle-income cities; strengthen links between cities and rural areas	Social cohesion	Investment in compact urban development and decentralisation
Number of urban people living in slums continues to grow	Reduce number of urban poor and disease risk; improve social cohesion; reduce youth unemployment	Investment in universal access to affordable water and sanitation; public transportation, and creation of jobs to reduce growth of slums; employment of the 'youth' dividend in low-income countries.	Reduce urban unemployment due to economic crises (of youth in particular); provide adequate housing in poor neighbourhoods	Strengthening and widening of social safety nets; upgrading investments in social protection for an effective response to crises and their aftermath
Inefficient use of public services (water, electricity)	Improve waste and recycling management; support consumption of local produce; change overconsumption patterns of high-income households	Subsidies to households and small firms to reduce non-saving water systems and waste; incentives to local communities to improve recycling systems	Change overproduction and overconsumption styles; improve waste and recycling management	Investment in retrofitting in buildings of water- and energy-saving devices; upgrading of public infrastructure
Ageing	Create productive employment for older persons	Investment in universal pensions; extension of working age; support for family networks	Fiscal pressure to reduce health costs; improve productivity	Investment in retraining older persons, and extending the working age
Economic				
Inequality and financial fragility	Create policy space for inclusive development; reduce underemployment; promote economic diversification	Investment in green industry, adaptation to climate change, structural economic change (industrial and service 'leapfrogging' for least developed countries); strengthening regional cooperation	Reduce unemployment; boost economic growth; strengthen international cooperation	Investment in green infrastructure; policy coherence and coordination
Food insecurity	Improve access to food; increase productivity	Investment in urban agriculture, local crops, storage facitities; R&D	Reduce food waste	Investment in storage infrastructure; reducing food subsidies; policy coordination
Environmental				
Energy access	Provide access to clean energy and reduce use of 'dirty' energy in poor households (e.g. least developed countries); discourage high energy consumption in high-income households)	Investment in capacity development, energy-saving devices, production and use of renewable sources of energy; subsidies and incentives for efficient energy use and water use for middle- and high-income households	Reduce overproduction and overconsumption to sustainable levels	Investment and incentives to produce and use renewable energy sources; decentralisation of energy production
Climate change	Reduce impact on livelihoods; reduce carbon emissions; generate financial resources for adaptation	Investment in health and education infrastructures and facilities; adaptation and mitigation technology and early warning systems; green public transportation; strengthen regional cooperation for green technology transfer	Upgrade disaster risk prevention systems; reduce carbon emissions to sustainable levels	Investment in mitigation, industrial green transformation; retrofitting of buildings; policy coordination

Source: UN/DESA, Development Policy and Analysis Division

Skills focus

There are several well-documented examples of cities facing challenges. Research one of these cities. Outline the challenges it faces and the methods it is using to outline these challenges.

Strategies for developing more sustainable cities

There is no 'one scheme fits all' approach towards achieving sustainability because the challenges faced by individual cities are diverse and depend on their population size, economic status, technological capacities and development priorities. However, some of the key strategies are summarised below:

- **Investment in infrastructure such as roads, water, sewers and electricity and services such as schools and healthcare**: In Brazil, Curitiba's credentials as a sustainable city are well documented. Its integrated bus system is just one part of the Curitiba Master Plan introduced in the 1970s. An extensive network of dedicated bus lanes provides a service comparable to underground or subway systems but at a cost estimated to be about 200 times less. The associated reduction in traffic has led to a significant reduction in carbon emissions.
- **Green investment in low-income countries can help poorer cities 'leapfrog' from high-carbon energy use to a zero-carbon development path**: This could provide employment for the 'youth bulge' within these cities.
- **Investment in the production and use of renewable energy sources, as well as the renovation of infrastructure, retrofitting of buildings and improved electricity and water efficiency**: In the city of Freiburg in Germany, solar investment subsidies are given to residents installing solar panels and 'plus energy' homes have been constructed which create more energy than they consume. Low-energy construction standards have also been introduced in a bid to reduce CO_2 emissions by 40 per cent by 2030.
- **Investment in the reduction of waste production and improvement of waste collection and recycling**: The 'Garbage that is not garbage' scheme in Curitiba promotes recycling through the separate collection of different waste components, while the 'Garbage purchase' programme encourages residents in the favelas to sell their rubbish back to the city in exchange for food, bus tokens and football match tickets. The scheme has helped clean up densely populated areas that the rubbish vans cannot reach.
- **Provision of more 'green' areas**: A study carried out by the European Centre for Environment and Human Health concluded that living in an urban area with green spaces has a long-lasting positive impact on people's mental well-being.

  Meanwhile, another study in 2019 by Fields in Trust found that parks delivered an estimated £34.2 billion of health and well-being benefits each year.
- **Investment in more sustainable and affordable housing**: Low-carbon housing developments include the experimental BEDZED development and Greenwich Millennium Village in London. Environmentally sustainable, recycled and local materials were used in their construction and the accommodation comprises a mix of social housing and private units.
- **Adoption of a local currency**: Local currencies such as the Bristol Pound serve the needs of local people because they keep money within the local economy. Research by the New Economics Foundation has found that for every local currency pound spent in a local business, £1.73 is generated through the multiplier effect. In contrast, for every pound spent in a chain store, only 35 pence is re-spent in the local economy. Despite research suggesting that the Bristol Pound venture is too small to have a significant local impact, by late 2017, five million Bristol Pounds had been spent.
- **Active participation of different city stakeholders including government, residents and local businesses in urban planning.**
- **Disaster risk reduction:** Schemes such as tidal barrages and early warning systems can help mitigate the impacts of floods, storm surges and other hazards to which some cities are vulnerable.

Finally, it has been argued that greater investment in rural areas is important to reduce the rural–urban migration that has put increasing pressure on cities in the last few decades.

Review questions

16 Which are the most effective strategies for tackling inequality in urban areas?

17 What are the main influences on climate and drainage in a city? How can humans limit their impact on both of these?

Case study of an urban area: London

Figure 9.64 Aerial view of the City of London's skyline

London is the capital and most populous city of the UK. In 2019, the Greater London Authority reported that the population had topped 8.9 million. This figure is expected to reach 11 million by 2050. London has a history stretching back to Roman times when it was named Londinium. From then it developed as a port around the navigable River Thames and eventually became the seat of political power and government. Despite leaving the EU, in 2018 the Institute for Urban Strategies ranked London first in the Global Power City Index. This is due to a number of factors:

- Economically, London is a global financial centre with a growing reputation as a technological hub and top rankings for software and multimedia development. Its reputation as an economic powerhouse means that it is particularly successful in attracting direct foreign investment.
- London performs very well globally in terms of the number of people in higher education, the quality of universities and access to libraries.
- More than 300 languages are spoken by the people of London, and the city has at least 50 non-indigenous communities with populations of 10,000 or more. Virtually every race, nation, culture and religion in the world can claim at least a handful of Londoners.
- According to the World Cities Culture Forum, London is seen as one of the most cosmopolitan and tolerant cities in the world, attracting a large diversity of people – from activists to business leaders, intellectuals to fashionistas. The city is also a major centre for art forms including music and dance, while its leading museums and galleries are among the most visited in the world.
- London is one of the most visited cities in the world due to a combination of history, heritage, art and culture.
- The hosting of the 2012 Olympic Games further raised the profile of London as a world city and led to the huge redevelopment of formerly run-down areas.

Cultural diversity in London

London is often referred to as 'the world under one roof' because of its multicultural population. Its function as a port has resulted in a long history of immigration and this has led to a rich ethnic and cultural diversity. Many see the arrival of hundreds of West Indian men aboard the *Empire Windrush* in 1948 as the start of mass immigration into the UK. Many of these men were answering adverts offering employment and intended only to stay a few years. In fact, the people of the *Windrush* and subsequent immigrant groups have played a vital role in creating a new concept of what it means to be a Londoner.

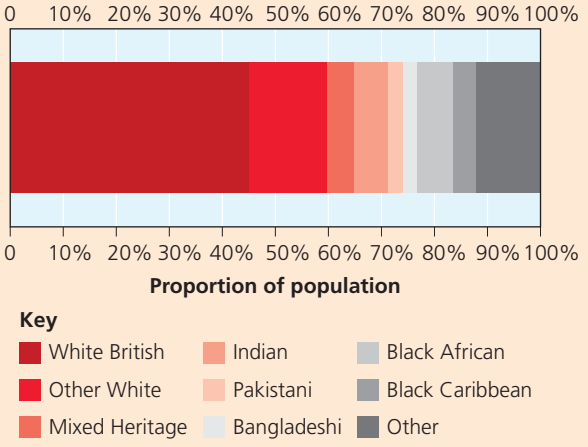


Figure 9.65 Ethnicity in London, 2011

Source: Trust for London & New Policy Institute

In 2015, just over 3.8 million of London's residents (44 per cent) were of a black and minority ethnicity origin. This is expected to increase to 50 per cent by 2038. In spite of some instances of hostility towards immigrant groups, London is regarded as a welcoming city and residents generally recognise the positive social and economic contributions immigrants make to life in the capital.

Economic and social well-being

London is the economic powerhouse of the UK and home to a third of the country's entire wealth. In parallel, however, it is also home to some of the country's highest poverty rates, with a reported 2.1 million people living in poverty.

In 2015, the London Fairness Commission reported that for every £1 of wealth owned by the bottom 10 per cent of London households, the top 10 per cent own £172.

Inability to get on to housing ladder (ever) due to low income and inflated London house prices

Long commutes on buses because of inability to afford the tube fare

Jobs, sometimes paying below the minimum wage

Multimillion pound houses

Fine dining and theatre experiences

Holidays abroad

Twice as likely to die from chronic lower respiratory illness

Lack of leisure time due to holding down more than one job to make ends meet

Overcrowded high-rise flats

Fear of crime and gang violence

Well-paid jobs with six-figure bonuses for some

Use of private schools and hospitals

Large proportion of monthly income spent on renting low-cost accommodation a long distance from work

Figure 9.66 A city of two halves

Key points from London's Poverty Profile, 2017

- The percentage of Londoners living in poverty (after taking into account housing costs) has decreased from 29 per cent to 27 per cent over the last six years.
- There are nearly one million people in poverty living in the private rented sector, more than any other housing sector. The number of children living in poverty in the private rented sector has tripled over the last ten years.
- Wealth inequality is more marked than income inequality in London.
  - The top 10 per cent of Londoners receive nearly 30 per cent of income but own just over 50 per cent of the total wealth.
  - The bottom 50 per cent of Londoners receive nearly 25 per cent of income but own only 5 per cent of the total wealth.
- 21 per cent of employees are paid less than the London Living Wage, compared with 22 per cent the previous year. This was the first year, since 2009, with no increase.
- Eight per cent of adults of working age receive out-of-work benefits, a decrease of 12 per cent from 2013.
- The net increase of 6,700 affordable homes in 2015/16 is less than 40 per cent of the target set out in the London Plan. Only one in four new homes was affordable.
- Within England, seven in ten households in temporary accommodation are in London. Over 80 per cent of these have children.
- London's infant mortality rate has fallen by around 40 per cent in just over a decade. London has gone from having an infant mortality rate that was higher than that for the whole of England to a rate that is lower.

Source: https://trustforlondon.fra1.digitaloceanspaces.com/media/documents/LPP_2017_key_findings.pdf

Skills focus

To see how these figures have changed over time and how different boroughs compare, go to www.londonspovertyprofile.org.uk/key-facts/ It is also worth looking at the 2015 Index of Multiple Deprivation data for London which measures relative deprivation. This can be viewed at http://maps.cdrc.ac.uk/.

Inequality impacts on everyday life. Compare the map showing average incomes in Figure 9.67 with life expectancy in Figure 9.68. The figure varies by about 5 years between the extremes. The difference between Hackney and the West End is the same as the difference between England and Guatemala.

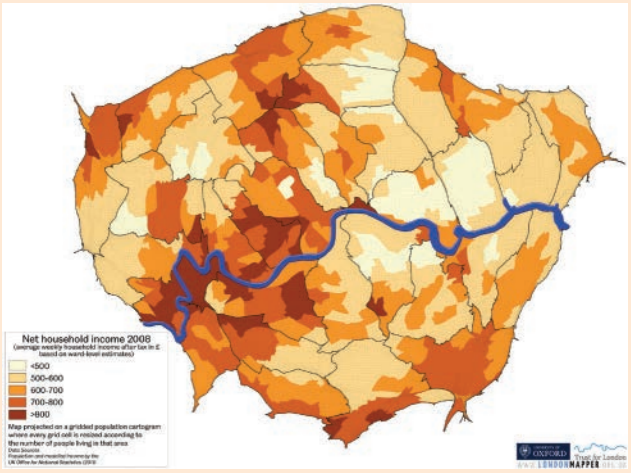


Figure 9.67 Average household income after tax (before housing costs) based on ward-level estimates for 2008 projected on a gridded population cartogram (Data source: Office for National Statistics, 2011)

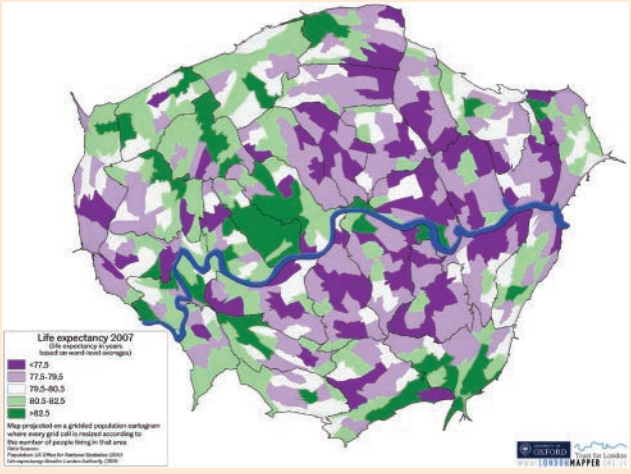


Figure 9.68 Average life expectancy based on ward-level estimates for 2007 projected on a gridded population cartogram (Data source: Greater London Authority, 2011)

The nature and impact of physical environmental conditions

In the first half of the twentieth century, 'smog' events became so strongly linked to the city of London that the term 'London particular' was and still is used in reference to smogs. Since then, air quality has improved significantly. The introduction of the Clean Air Act in 1956 and more recent legislation, along with the introduction of greener buses, taxi age limits and the Ultra Low Emission Zone (ULEZ), have reduced CO_2 emissions in particular. However, nitrogen dioxide levels still breach EU legal limits and Figure 9.69 shows that London still suffers higher levels of air pollution than surrounding rural areas. The 2014 Public Health England report on air pollution says that 5.3 per cent of all deaths in people age 25 and over are now linked to air pollution. These figures vary geographically but the highest percentage of deaths linked to air pollution are in London. More specifically, the boroughs of Kensington & Chelsea and Westminster have the highest rates, where 8.3 per cent of deaths can be attributed to air pollution.

Linked to air quality is the impact of London's urban heat island, shown in Figure 9.30 (page 425). With the centre of London already up to 10°C warmer than the surrounding rural areas and average summer temperatures predicted to rise further, summer heat waves pose a threat to homes, workplaces and public transport. They have a negative effect on health, particularly that of vulnerable people, and lead to greater consumption of water and energy.

The London Climate Change Partnership has suggested that extremely high demands on London's power supply network may lead to 'brownouts', due to the high cooling demand, and increases in electricity demand for cooling could negatively affect London's sustainability.

One further concern linked to London's physical environment is the threat of flooding. This comes from five sources – tidal, fluvial (from rivers and tributaries), surface (from rainfall), sewer and groundwater flooding. Climate change will bring wetter winters and more frequent heavy downpours, as well as rising sea levels and higher tidal surges, all of which pose a major threat to London. Fifteen per cent of London is on the flood plain, protected by flood defences. Residential areas are located within this area but it also includes much of the infrastructure Londoners rely on daily: 49 railway stations, 75 underground stations and 10 hospitals. The Thames Estuary 2100 (TE2100) project is one response to the risk of flooding but other strategies are being adopted. Go to http://london.gov.uk for more information on these.

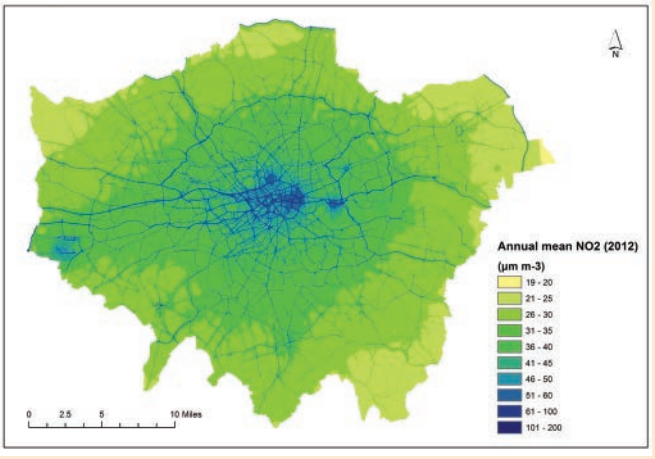


Figure 9.69 Concentrations of nitrogen dioxide are higher along the roads in London and at key transport hubs such as Heathrow Airport and City Airport. Over the course of the day, NO_2 levels are particularly high during rush hour as they predominantly come from vehicle emissions

Moving towards greater urban sustainability

The London Sustainable Development Commission was established in 2002 to advise the Mayor of London on making London a 'sustainable world city'. The Commission updated its aims in 2015 in line with Sustainable Development Goals (SDGs) agreed by world leaders. SDGs are a 'blueprint to achieve a better and more sustainable future for all' by 2030. Various strategies have since been implemented to tackle key issues such as inequality, air pollution and urban deprivation. A snapshot of strategies is shown below. There are many more to investigate and evaluate.

As the urban area of London continues to expand, greater pressures will be placed on the services and infrastructure. Strategies encouraging greater urban sustainability will be key to the city's future viability. Proposals to make London into the world's first **National Park City** are being discussed.

Further research

There are numerous areas of London's geography to research. The Olympic legacy, inequality and the success of strategies to tackle environmental problems are just three areas which merit more attention. Look also at the place study of East London in Chapter 8 for information on the Bangladeshi diaspora, and the impacts of gentrification and urban resurgence. Useful websites include www.london.gov.uk and www.tfl.gov.uk. The London mapper website visualises data in a variety of different ways: www.londonmapper.org.uk/ Likewise, the 2014 book *London: The Information Capital* includes 100 maps and graphics on all aspects of London living. Both sources are excellent examples of GIS. Finally, a useful resource for looking at the extent and impact of inequality in London is geographer Danny Dorling's *The 32 Stops: the Central Line*.

Table 9.12 Urban sustainability strategies adopted in London

Social developments	Economic developments
Urban renewal has brought economic investment and growth to parts of London previously suffering urban decline. For example, Olympic regeneration in Newham has provided additional social housing, leisure facilities and greater investment in education. The Queen Elizabeth Olympic Park provides parklands, waterways and leisure activities. From 2016, the Olympic Stadium became home to West Ham United FC A focus on improving education in the capital through the London Challenge initiative has continued to improve outcomes for pupils in London's primary and secondary schools at a faster rate than nationally. Pupils in London do better in school than the rest of the country In Newham, the Every Child programme offers children the chance to learn a musical instrument or take part in cultural events. Children in this area also benefit from universal free school meals, which have been linked to greater attainment and better health. Similar schemes operate in other boroughs	The London Living Wage is a calculated hourly rate of pay which gives the wage rate needed for a worker in London to provide their family with the essentials of life, including a cushion against unforeseen events. It is not compulsory but a voluntary commitment made by employers. It has been estimated that over 10,000 London families have been lifted out of working poverty as a direct result of the Living Wage When Crossrail opens (planned date 2022), it will increase London's rail-based transport network capacity by 10 per cent. It is hoped this will cut journey times across the city, ease congestion and encourage regeneration At a borough level, schemes such as Workplace in Newham have been introduced to help people find jobs. Workplace provides residents with free advice about employment options and training. Since 2007, Workplace has supported over 20,000 Newham residents into employment, and helped more than 900 businesses fill their vacancies
Environmental developments	**Urban governance**
The Congestion Charge was introduced in 2003 to discourage driving in central London to reduce congestion and pollution More recent 'green' transport policies have included Routemaster buses with green electric hybrid engines, a fleet of eight hydrogen fuel buses, more than 1,400 charge points to support the use of electric vehicles, zero emission taxis from 2018 and a central London Ultra Low Emission Zone (ULEZ) from 2020 with the objective of reducing air pollutant and CO_2 emissions from road transport Cycle superhighways and the cycle hire scheme more commonly known as 'Boris Bikes' have helped to increase cycling in London by 173 per cent since 2001. The aim is for a 400 per cent increase by 2025, although there are continued concerns about cycling safety In response to the UHI effect in London, targets have been set to increase green cover in central London by 5 per cent by 2030	The role of the Mayor of London was created under the Greater London Authority Act 1999 as part of the government's commitment to restore a city-wide government for London Mayors are important to ensure that a city has a strong voice and can attract investment from home and abroad. Devolved powers include planning, transport, employment, economic development, health and policing

Case study of a contrasting urban area: Mumbai

Although not its capital, Mumbai is India's largest city with a population estimated to be in excess of 25 million in 2020. It is the financial and commercial centre of India as well as home to the popular Bollywood movie industry.

Figure 9.70 The location of Mumbai

Mumbai is located on the west coast of India and is the capital of the state of Maharashtra. It developed as a trading centre selling local goods such as gold, jewellery and textiles. The arrival of the East India Company in the early seventeenth century led to the establishment of the British Raj (Empire) and goods such as raw cotton were regularly shipped to Britain for manufacturing. Originally, Mumbai was a series of seven islands separated by swamps, but by 1845 these had been filled in and Mumbai occupied one large island. Mumbai has a natural deep-water harbour and has been the main port in the Arabian Sea since the opening of the Suez Canal in 1869.

Figure 9.71 Skyline of Mumbai and Marine Drive, known as the Queen's Necklace, Mumbai

Following India's independence from the British in 1947, Mumbai developed rapidly. High-rise, modern architecture, the Bombay Stock Exchange, tarred roads and a boom in manufacturing and services have changed the city's status and brought it on to the world stage. In 2015 it accounted for 33 per cent of India's income tax, 6.16 per cent of GDP (the largest single contributor in India), 25 per cent of industrial output and 40 per cent of foreign trade. Mumbai is now a megacity with the highest population density of any city in the world. Such numbers place an inevitable strain on the urban infrastructure.

Economic and social well-being

Mumbai's population has more than doubled since 1991 and this is largely due to the influx of migrants from other parts of India seeking better employment opportunities. The resulting population is very diverse and 16 major languages of India are spoken here. Poverty and inequality are two of the big issues, however. It is estimated that around 60 per cent of Mumbaikars live in 'slums' and the average Indian would need to work for three centuries to pay for a luxury home in Mumbai.

Links

See the case study on page 399 for further information on Dharavi.

Dharavi slum

There are many informal slum areas in Mumbai but Dharavi in central Mumbai is perhaps the most famous, brought to the attention of many by the Oscar-winning film *Slumdog Millionaire* in 2008. Until the late nineteenth century, this area was mangrove swamp inhabited by Koli fishermen. When the swamp filled in (with coconut leaves, rotten fish and human waste) the Koli people lost their fishing ground, but there was more land area for others. The Kumbhars came from Gujarat to establish a potters' colony, Tamils opened tanneries and thousands more travelled from areas such as Uttar Pradesh to work in the rapidly expanding textiles industry. The result was a very diverse neighbourhood in a very diverse city.

Cottage industries have thrived in Dharavi. It is home to thousands of micro-industries, including garment-makers, tanners, welders and potters, which produce over US$650m worth of goods annually. However, the living and working conditions remain very poor. Years of government neglect have resulted in inadequate hygiene standards and it is said to have the highest population density in the world at over 300,000 people per square kilometre. Housing quality is poor and areas within the slum lack basic infrastructure. In some parts of the slum, toilets are shared by as many as 1,000 people and services such as water and electricity are not always available.

Due to the northward expansion of the city of Mumbai, Dharavi found itself occupying an area of prime land in the new business district of India's richest city. This made it a key target for developers eager to make money from the construction of luxury apartments, while the government has been keen to improve the appearance and reputation of the area. The government-led Dharavi Redevelopment Project will see all residents who can prove residency since 2000 provided with a new, 300-square-foot house for free. The scheme has not been without controversy. The main

concern has been the potential loss of the community networks and businesses which have built up there.

The nature and impact of physical environmental conditions

Mumbai has a tropical climate. The south-west monsoon brings heavy rainfall to the city between June and September and although Mumbaikers adapt to the wet conditions, the rains can be devastating for the city, which is mainly built on low-lying land. In 2019, Mumbai received more rainfall than usual during the monsoon season. More than 250 mm fell on one day in August and 403 mm in just four days in early September. This is compared with the monthly expected average rainfall of 341 mm. Several people were washed away and others were killed by collapsing walls. Electricity, water supply, communication networks and public transportation were totally shut down; more than 400 people died; and over 10,000 homes were destroyed. The city suffered losses running into billions of pounds. Urban growth was of course partly to blame. There was nowhere for the rainwater to go as rapid and often uncontrolled development had replaced most of the public parks, private gardens, beaches, mangrove swamps and wetlands with a built environment.

Figure 9.72 Impact of monsoon rainfall on the daily commute in Mumbai

The Greater Mumbai Disaster Management Action Plan was created in response to flooding in 2005. It identified the risks and vulnerabilities the city could face in the future, including earthquakes and cyclones; it created the Disaster Management Cell to co-ordinate relief and rescue efforts; and it widened and deepened the Mithi River which drains out into the Arabian Sea. The plan is clearly a positive strategy for Mumbai and shows the government is keen to address the major flood issue, but environmentalists are concerned that there are still too many factors which make Mumbai vulnerable to flooding again: continued construction on the floodplain, removal of mangrove forests and the clogging of storm drains and waterways with plastic rubbish. In addition to this, the Intergovernmental Panel on Climate Change (IPCC) has predicted that the increase in rainfall, heat, humidity and sea-level rise associated with climate change will make Mumbai the second most at-risk city in the world.

The future?

Economic growth has clearly brought wealth to many in Mumbai and new housing projects are trying to address the shortage of housing. However, this megacity is struggling to cope with its rapid growth. Services are stretched or non-existent, air and water supplies are polluted, inequality is growing and over half the population still lives in slums. The Indian government has pledged to make cities like Mumbai 'smarter' in terms of the economy and environment. In addition, non-government organisations have been working towards improving the lives of slum dwellers in Mumbai. However, as Figure 9.73 shows, Mumbai faces considerable environmental pressures and much greater investment and long-term planning are essential if the city is to become truly sustainable in the twenty-first century.

Further research

Mumbai has been well-documented in geographical magazines. Further useful information can be gathered from the World Bank (www.worldbank.org/) which funds a number of projects in Mumbai. In March 2015, *The Guardian* newspaper's Guardian Cities Team published a five-day feature on Mumbai.

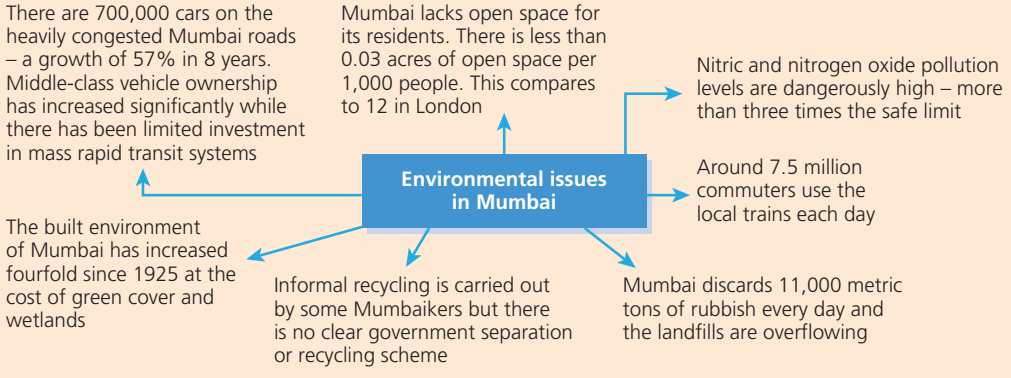


Figure 9.73 Environmental issues in Mumbai

Fieldwork opportunities

The Office for National Statistics (www.ons.gov.uk) provides a range of data for different wards within urban areas. These statistics can be used alongside primary data on housing, land use, provision of services and environmental quality to compare different parts of an urban area. The Field Studies Council (FSC) has numerous useful documents and resources on their website to help you investigate the quality of urban areas and inequalities. Go to www.geography-fieldwork.org

It is also worth looking at the 2019 Index of Multiple Deprivation data which measures relative deprivation. This can be viewed at http://dclgapps.communities.gov.uk/imd/iod_index.html

The redevelopment of King's Cross Station in London provides a new fieldwork venue in the UK capital. Disused railway land on the 27-hectare site is being transformed into a mixed-use site complete with homes, shops, offices, galleries and restaurants. The website below provides both historical and current information about the site as well as brochures and videos. Visit www.kingscross.co.uk/discover-kings-cross

Changes in temperature and/or pollution levels can be measured and mapped across an urban area. Traffic counts at strategic points may show a correlation with such data.

Infiltration rates can be measured across a variety of urban surfaces. This could then be linked to a Sustainable Urban Drainage System.

Further reading

Two relatively recent urban geography books:

Montgomery, C. (2015) *Urban Geography* (Routledge)

Hall, T. and Barrett, H. (2011) *Happy City: Transforming Our Lives Through Urban Design* (Penguin)

Pacione, M. (2009) *Urban Geography: A Global Perspective* (Routledge)

See also Drake, G. and Lee, C. (2000) *The Urban Challenge* (Hodder)

The United Nations publishes a number of useful reports on cities. In 2014, a report entitled *World Urbanisation Prospects* was published outlining the key issues resulting from urban growth.

UN-Habitat – an organisation concerned with global urban issues has also published a number of reports entitled *The State of the World's Cities* which focus on sustainability issues. A *World Cities Report* was published by the UN in 2016 (www.unhabitat.org/)

The Country Activities Report 2019 provides an evolving picture of the portfolio on UN-Habitat's projects and programmes from 2018 onwards. It is an up-to-date window into the most recent activities. (https://unhabitat.org/sites/default/files/2019/05/un-habitat_country_activities_report_-_2019_web.pdf)

The Urban Geography Research Group (UGRG) of the Royal Geographical Society/Institute of British Geographers has a website with useful links and resources. (www.ugrg.co.uk)

The late Hans Rosling's brilliant Gapminder website contains a wealth of data on socio-economic trends and variations around the world (www.gapminder.org/)

Water Supply for Urban Poor in India – produced by Water Aid, a charity that is committed to helping both the rural and urban poor gain access to safe water supplies. (https://smartnet.niua.org/sites/default/files/resources/water-supply-for-urban-poor-in-india-web.pdf)

Improving Life for Low-income Londoners: How the next mayor can tackle inequality (www.trustforlondon.org.uk/publications)

For current urban trends and processes it is worth reading newspaper websites such as *The Guardian* and *The Independent*, which often have up-to-date features on cities. *The Guardian* has a twitter feed @guardiancities. It also has a resilient cities page supported by the Rockefeller Foundation.

Magazines such as *Geographical* magazine and *Geography Review* regularly include articles on specific cities and urban issues.

Question practice

A-level questions

1 Outline the issues associated with the disposal of urban waste. [4 marks]

2 Study Figure 9.31 (page 425). Analyse the extent to which Figure 9.31 supports the statement that 'temperatures rise towards the centre of a city during a heatwave.' [6 marks]

3 Assess the extent to which the natural hydrological cycle can be affected by urban growth. [9 marks]

4 Assess the success of one named urban regeneration policy on Britain since 1979. [9 marks]

5 Assess the extent to which environmental problems in cities affect residents' lived experience within those cities. [20 marks]

Chapter

10 Population and the environment

Human population growth is the driver of many environmental, socio-economic and political challenges that we face in the world today. This chapter will examine the relationship between some key aspects of the physical environment and population numbers as well as people's health and well-being. It will in turn consider the impacts of population change on the environment.

The ability of a place to support a population depends upon whether it can provide an adequate food and water supply. Local food production is determined by climate, soils and other aspects of the environment such as relief and topography. The relationship between population change and the environment also has a bearing on the economic development of a place and its ability to trade with other places. It is essential therefore to understand the reciprocal relationship between the physical environment and human populations at all scales of study in order to understand how the future of our planet will look.

In the chapter you will study:

- the relationship between the physical environment, food production systems and food security
- the relationship between the physical environment and human health and well-being
- natural and migration population change
- population ecology and the relationship between population and resources
- global population futures – varying possible scenarios of future population growth.

Links

There are links between this chapter and Chapter 1, Water and carbon cycles; Chapter 7, Global systems and global governance and Chapter 8, Changing places.

10.1 Introduction

The relationship between population and the environment

Appreciating the many facets of the relationship between population and the physical environment is pivotal to understanding the geographical characteristics of **population distribution** and change, as well as the impact of people on the environment.

Features of the physical environment such as the climate, geology and topography determine available water supply, the nature of soils, drainage patterns and potential environmental hazards. In turn, these features affect settlement patterns, food production, energy supplies and human health and so contribute to population size, distribution and its success in developing the resources it has at its disposal.

The size of population and its rate of growth will have a reciprocal impact on the physical environment. The degree and nature of the impact will depend upon the level of resource consumption, standard of living and mitigation put in place to reduce any negative impacts. Recent trends in population growth have increased resource consumption, leading to many negative impacts on the physical environment, including:

- climate change and resulting increases in extremes of temperature, floods, droughts and sea level rises
- pollution of water and land as well as atmospheric pollution on a local and regional scale
- ozone depletion

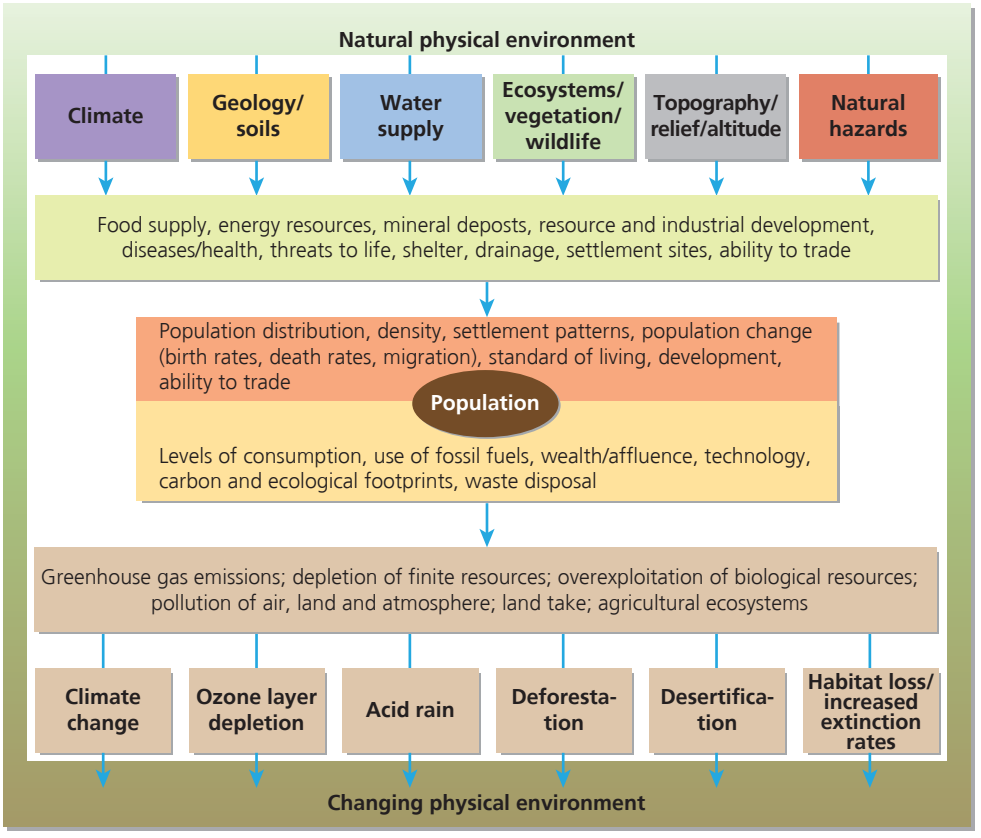


Figure 10.1 The inter-relationship between human population and the physical environment

- depletion of finite natural resources
- damage to wildlife and their habitats leading to increased extinction rates and consequent threats to species interdependence.

Figure 10.1 outlines the different facets of this reciprocal relationship between population and resources. This is further examined in the population ecology section of this chapter, particularly by the Population, Resources and Pollution Model.

Elements in the physical environment

Climate

Of all the elements in the physical environment, this is arguably the most important. Rainfall, temperature, wind velocity and levels of solar insolation all determine food productivity, the type of farming system adopted and which species are selected for cultivation. For example, many crops need temperatures of at least 5°C to grow and pasture-fed livestock need to be supplemented with fodder crops in temperatures below 5°C. High concentrations of population are determined by adequate rainfall and temperatures that are suitable for growth of crops and rearing of livestock, with sufficient sunlight for photosynthesis. Characteristics of climate can also drive the level and nature of diseases. For example, tropical diseases such as malaria, yellow fever and Ebola will directly affect death rates and life expectancy in the human population. High infant and child mortality rates in tropical and subtropical areas can also influence fertility rates as families seek to compensate for their loss. Climate change arising from global warming is the aspect of the environment giving greatest concern to the human population at the start of the twenty-first century.

Links

There are links to Chapter 1, Water and carbon cycles, in particular to Water, carbon, climate and life on Earth on page 41.

Soils

The most important feature of soil is its fertility. This depends upon soil structure, texture, acidity (pH), organic matter and nutrients. These will determine agricultural output and the type of farming system employed. Fertility can be maintained with artificial chemical fertilisers in areas of high population density

but this may have unsustainable outcomes, such as contributing to water pollution, eutrophication and increased greenhouse gas emissions. Areas with fertile soils are often associated with densely populated areas. Some, however, such as rich volcanic or alluvial soils, are also prone to hazards (such as volcanic eruptions or floods), which may negatively affect the population.

Water supply

People need access to water to survive. It is important for human hydration but its key use is for irrigation to maintain food production. It has other key uses including hygiene and sanitation as well as in many industrial processes. Its importance in determining population distribution is well illustrated in Egypt, where 95 per cent of its population of 102 million people live on about 4 per cent of the country's land area, within 12 miles of the River Nile.

Geology and other resource distributions

Concentrations of other resources such as fossil fuels or other valued minerals have given rise to industrialisation and consequently densely populated conurbations in Europe (such as the Ruhr Valley), parts of the USA and increasingly now China and India. Even when these resources become depleted, industrial inertia leaves a legacy of large, dense populations with new tertiary industries emerging to serve them.

Population parameters

Distribution, density, numbers and change

Elements of the physical environment clearly determine **distribution** patterns and **population density**. For example, on a national scale, Egypt's population is very unevenly distributed, determined primarily by water supply and specifically by access to the River Nile. Thus the Nile Valley has areas with a population density of 2,000+ people/km² while the average for Egypt as a whole is 103 people/km² because the desert areas have fewer than 10 people/km². Most countries in the world have uneven distributions due to a combination of the factors in the physical environment and the level of urbanisation as urban areas tend to be more densely populated than surrounding rural areas. Figure 10.2 shows global population density estimates for 2020. Even at this scale, the densely populated areas relating to water supply or resource distribution are clearly visible along the Ganges Valley in India, in eastern China, South East Asia and north west Europe. Areas of population sparsity can be correlated to deserts, continental interiors and where there are extremes of cold temperatures in Siberia and the Canadian Shield.

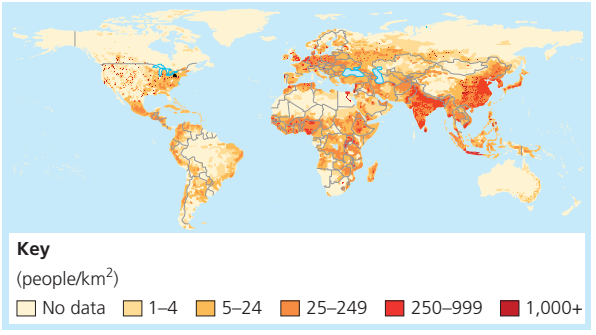


Figure 10.2 Global population density estimates, 2020
Source: Center for International Earth Science Information Network

Population numbers are recorded by census data, which in most countries is collected every five or ten years. Numbers are constantly changing so population size between census dates is estimated using information on rates of population change such as birth rates, death rates and net migration.

With numbers and size being dynamic phenomena, **population change** considers population from a temporal as well as a spatial perspective. Population change can be measured by rates of growth (or decline) but there are many other aspects of change that can be observed. These are covered in more detail in Section 10.4 on page 506.

Key terms

Population density – The average number of people living in a specified area, usually expressed as the number of people per km².

Population distribution – The pattern of where people live. This can be considered at all scales from local to global, in an area or country.

Role of development processes

The process of development has been a narrative associated with the human ability to acquire and make use of the resources made available by the physical

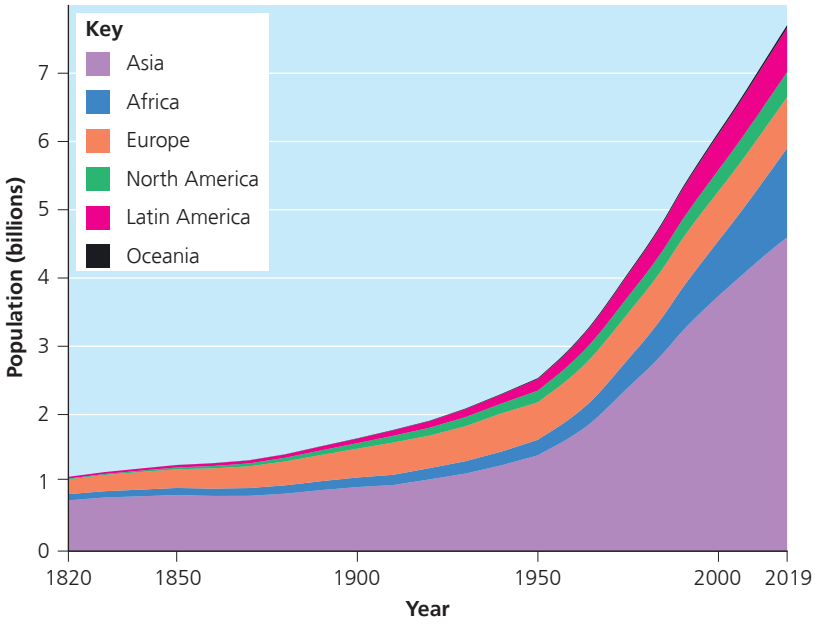


Figure 10.3 Global patterns of population numbers

Source: Our World in Data

environment. Over time people have migrated to areas of fertile soils and resource abundance, where they have thrived and population has grown. Development surges have resulted from some trigger. The most notable of these have been the Neolithic Agrarian Revolution in Mesopotamia, around 12,000 years ago and the Industrial Revolution in Europe in the eighteenth and nineteenth centuries. These have sparked technological developments which have enabled specific areas (and the Earth as a whole) to support a larger human population. Also important has been the increasing ability of the human race to control many of the infectious diseases which have threatened survival and kept death rates high. This has been one of the main reasons for the rapid rise in world population from around 1 billion in 1800 to 7.7 billion by 2019 (See Figure 10.3).

Rapid population growth and the process of development have had a largely negative impact on the environment itself. The impact of individuals will depend on where they live on the planet, their individual lifestyles and the resources and services available to them. These themes are further examined in Section 10.5 on population ecology in this chapter, which considers ecological footprints and the global challenges resulting from continued population growth.

Global patterns of population numbers, densities and changes

At the time of writing, the world's total population is around 7.8 billion and according to a UN report published in 2019, it is expected to increase by nearly two billion in the next 30 years to reach 9.7 billion by 2050. Asia is by far the most populated continent, with around 60 per cent of the world's population and the two most populated countries, China and India. India is expected to surpass China as the world's most populous country by the end of the 2020s. Figure 10.3 shows the pattern of population numbers by continent.

Global patterns of population density are shown in Figure 10.2 (page 464), in which coastal areas and seaboards, major river basins and areas of industrialisation and urbanisation are the most densely populated areas.

The world's population continues to increase but the rate of population change varies greatly across regions.

The population of sub-Saharan Africa is projected to double by 2050 while other regions are expected to experience lower rates of population growth until 2050, including North Africa and West Asia (46 per cent increase), Australia/New Zealand (28 per cent), Central and Southern Asia (25 per cent), Latin America and the Caribbean (18 per cent), Eastern and South East Asia (3 per cent) and Europe and Northern America (2 per cent).

Some countries in Europe and Asia, notably Italy, Poland, Russia, Japan and South Korea, are already experiencing a population decline, which seems likely to continue (Figure 10.4).

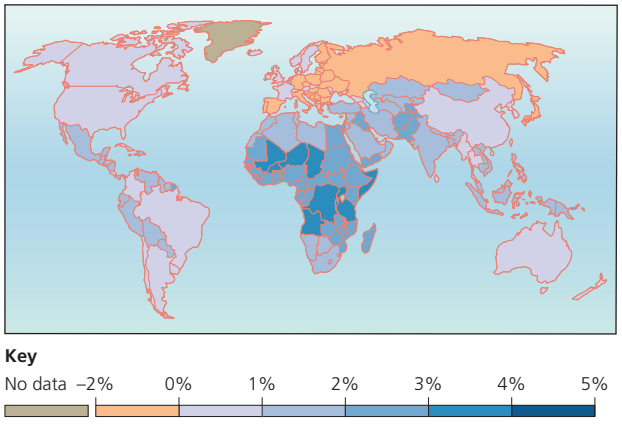


Figure 10.4 Global pattern of population rate change, 2020
Source: Our World in Data

Review questions

1 Explain how elements in the physical environment influence population distribution.
2 Identify the main impacts of human population growth on the physical environment.
3 Describe the global pattern of population density.
4 Outline the key features of population change rates.

10.2 Environment and population

The amount and type of food produced varies from place to place. That variability depends upon both environmental and human variables. The two key environmental variables are **climate** (temperature, growing season, water availability) and **soils**. The variation in production can occur at all scales from local to global.

Key terms

Agricultural productivity – The ratio of agricultural outputs to agricultural inputs.

Agriculture – The science or practice of farming, including cultivation of the soil for the growing of crops and the rearing of animals to provide food and other products such as wool, cotton or rubber.

Climate – A region's long-term weather patterns. This is measured in terms of average precipitation, maximum and minimum seasonal temperatures, sunshine hours, wind speeds, humidity and the frequency of extreme weather events.

Food security – Food security exists when all people at all times have access to sufficient, safe, nutritious food to maintain a healthy and active life.

Greenwashing – a term used to describe exaggerated environmental claims, usually made to promote the green credentials of big corporations, i.e. a form of environmental propaganda.

Salinisation – The build-up of salts in soil, eventually to toxic levels for plants.

Soil – The upper layer of earth in which plants grow, a black or dark brown material typically consisting of a mixture of organic remains, clay and rock particles.

Topography – The relief and drainage of an area.

Zonal soil – A categorised soil group, which covers a wide geographical region; its profile and properties are determined primarily by the actions of the regional climate and natural vegetation on the parent rock from which it is formed.

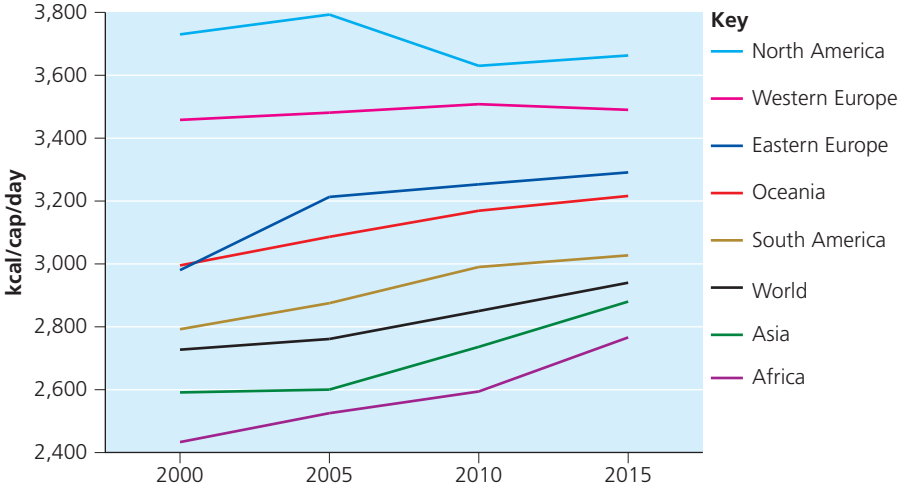


Figure 10.5 Daily per capita food supply by region, 2000–2015

Global and regional patterns of food production and consumption

Food production

Global food production is not evenly distributed. Food production is determined by two main drivers:

- **Environmental capacity** – influenced by the physical environment, climate, water availability and soil fertility and so on.
- **Human capacity** refers to population size (human resources) and levels of skills, technology and capital investment in agricultural infrastructure.

In the early 1960s, average global food supplies for humans stood at only 2,300 calories per person per day; this was very unevenly distributed. In high-income countries the average was 3,030 calories per day, whereas in low-income countries of the developing world it was below 2,000. Probably more than half the people in these low-income countries suffered from chronic under-nutrition.

By 2015, despite the fact that global population had increased by two and half times, the world could produce enough food to provide every person with more than 2,940 calories per day. This should have been enough to ensure that all had an adequate and balanced diet. Despite some successes achieved by the UN's **Millennium Development Goal** (MDG) of halving poverty and hunger by 2015, unfortunately food production and availability remained uneven and 820 million people still suffered from under-nutrition in 2018. This, however, represents a fall from 15 per cent to 10.5 per cent in a 15-year period.

These gains in production resulted from a number of factors:

- The package of technologies referred to as the **green revolution**, including increased use of new, high-yielding crop varieties and technologies. Irrigation, fertiliser and pesticides increased yields for millions of farmers, although with some negative social and environmental impacts.
- An increase in globalised trade that addresses seasonality as well as promoting specialisation of production where it is most efficient.
- Investment in production in low-income countries by TNCs and other organisations (governments and NGOs) to increase productivity.

Figure 10.5 (page 466) shows how food supply has changed in regions of the world between 2000 and 2015. The rates of change vary.

Links

There are links here to Chapter 7, Global systems and governance, including Flows of ideas and technology, page 300; International trade and access to markets, page 306; Nature and role of TNCs, page 314; and the case study on the banana trade, page 321.

Skills focus

Describe the changes in food supply for the six regions of the world shown in Figure 10.5 (page 466). Suggest reasons why these changes vary from region to region.

In low-income countries, production and consumption of the main agricultural products have been growing at higher rates than in the higher-income countries. This is a result of higher population growth rates, increasing wealth and a greater responsiveness of demand to income growth in these poorer countries. In contrast, a slower growth of demand has occurred in the richer world because high per capita consumption and slow growth of population have had a dampening effect on the growth in demand for many commodities. Countries such as China and Brazil have high growth rates, while North America and Europe have lower growth rates and, in some cases, production has declined.

Figure 10.5 also shows that the global food production increased from 2,727 kcals per capita per day in 2000 to 2,940 kcals per capita per day in 2015, resulting in the global pattern of food supply shown in Figure 10.6.

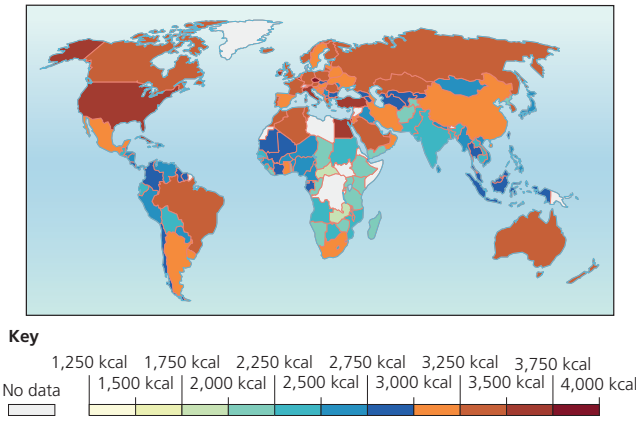


Figure 10.6 Global food supply by calories available, 2013

Source: Our World in Data

Food consumption

Food consumption expressed in kilocalories (kcal) per capita per day is a key variable used for measuring changes in global and regional food availability. Analysis of **United Nations Food and Agriculture Organization (FAO)** data shows that dietary energy has been steadily increasing on a worldwide basis; availability of calories per capita from the mid-1960s to 2015 increased by nearly 600 kcal per capita per day and by 800 kcal per capita per day in developing countries (see Table 10.1). In short, the world has made significant progress in raising food consumption per person over the past 50 years but there are still problems of undernourishment and hunger in some regions.

Table 10.1 Global and regional per capita food consumption (kcal per capita per day)

Region	1964–1966	1974–1976	1984–1986	1997–1999	2015	2030 (est)
World	2,358	2,435	2,655	2,803	2,940	3,050
Developing countries	2,054	2,152	2,450	2,681	2,850	2,980
Near East and North Africa	2,290	2,591	2,953	3,006	3,090	3,170
Sub-Saharan Africa	2,058	2,079	2,057	2,195	2,360	2,540
Latin America and the Caribbean	2,393	2,546	2,689	2,824	2,980	3,140
East Asia	1,957	2,105	2,559	2,921	3,060	3,190
South Asia	2,017	1,986	2,205	2,403	2,700	2,900
Industrialised countries	2,947	3,065	3,206	3,380	3,440	3,500
Transition countries	3,222	3,385	3,379	2,906	3,060	3,180

Source: WHO

According to the FAO, undernourishment, or hunger, is dietary intake below the minimum daily energy requirement. Dietary energy requirements differ by gender and age, and for different levels of physical activity. Accordingly, minimum dietary energy requirements vary by country and from year to year depending on the population structure but generally range between 1,700 and 2,000 kcals per day. Figure 10.7 shows the global pattern of countries with a prevalence of undernourishment as a percentage of the total population. Those regions suffering most from undernourishment are sub-Saharan Africa, south, central and eastern Asia and Latin America.

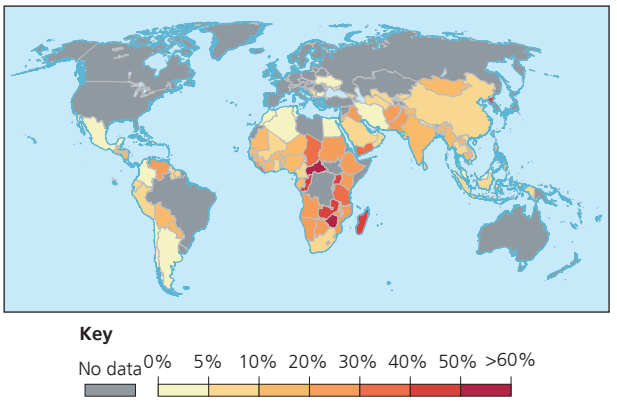


Figure 10.7 Global pattern of undernourishment in per cent, 2017
Source: Our World in Data

In the past twenty-five years, food consumption has been growing at a faster pace than world population growth. Over the same period, rising incomes in industrialising countries and changes in consumer preferences in higher income countries have influenced the following changes in food consumption patterns:

- an increase in consumption of higher value meat and dairy products in emerging economies
- lower red meat consumption in many developed economies because of health and environmental concerns.

Agricultural systems

Agriculture can be studied in a systematic way and farms can be considered as open systems (see Figure 10.8). The inputs include physical, cultural, economic and behavioural. Generally, as an area develops the physical factors become less important as the human inputs increase.

Figure 10.9 summarises how types of farming systems can be categorised. Within that general pattern there are the following types:

- **Commercial farming** is carried out so that the majority of the produce is sold and the income generated can provide a livelihood for the farm workers as well as be invested back into the farm.
- **Subsistence farming** means that the majority of the produce is consumed by the landowner and the farm workers, though a little surplus may be sold to buy other living requirements and/or be invested in the farm.
- **Intensive farming** is a system requiring a high level of inputs per unit area of land being farmed. The nature of these inputs can vary, which means that intensive farming can either be:

Figure 10.8 Aspects of the farming system

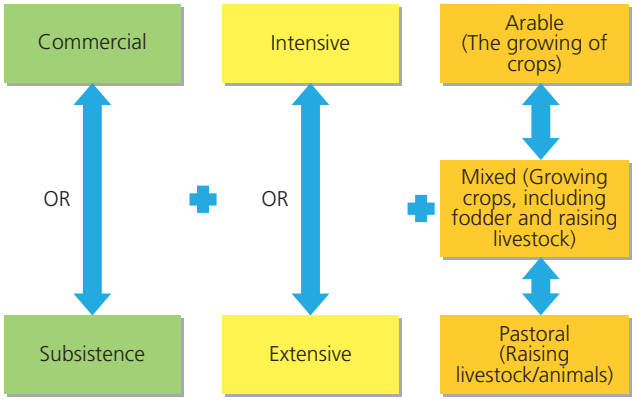


Figure 10.9 Farming types

– **capital intensive:** money is invested in soil improvement, machinery, buildings, pest control, high-quality seeds/animals. There are few people employed and so output is high per hectare and per worker, for example, tomato production in The Netherlands.

– **labour intensive:** the number of farm workers is high and so there is a high output per hectare but a low output per worker, for example, rice cultivation in the Ganges Valley.

- **Extensive** farming is carried out over a large area with relatively fewer capital or labour inputs per unit area of land being farmed. This varies greatly.

There are areas where, although the labour force is minimal, there is a high capital input. Inputs include: quality seeds/animals, pesticides and insecticides, such as for wheat farming in the Canadian Prairies. Other areas still have a relatively small labour force but rely on the sheer amount of land they are farming to provide sufficient output for their needs, for example, sheep ranching in Australia.

Extensive farming systems usually produce a high total output because of the large area being farmed, but the output per unit area of land will be lower than in intensive farming systems.

Another way of distinguishing between farming types is whether inputs of artificial chemicals are used or not. In organic farming of crops or livestock, no artificial chemical additives are involved. Other farming systems use artificial inputs to increase productivity, such as chemical fertilisers for crops or hormone treatment of livestock.

Agricultural productivity

Productivity is a key measure of the economic performance of agriculture and an important driver of farm incomes. It represents how efficiently the agricultural industry uses the inputs that are available to convert into outputs. It is a relative measure which enables economists to see if improvements are made by comparing one year to another. Understanding and assessing the potential for greater productivity is a critical first step on the pathway to producing more with less.

Those looking at the results need to consider the longer-term trends as changes from year to year are often shaped by factors outside the farmer's control. Factors such as weather, animal disease, policy interventions and general economic conditions can have short-term effects on productivity.

Agricultural productivity is typically measured in terms of yield: how many kilograms of grain per hectare, kilograms of meat per animal or litres of milk per cow, etc. The most commonly used measurement of productivity is **total factor productivity (TFP)**.

TFP is the ratio of agricultural outputs (gross crop and livestock output) to inputs (land, labour, fertiliser, machinery and livestock). In ecological terms, this can also be considered to be the '**energy efficiency ratio**', and as outputs increase relative to the level of inputs, so productivity is improved. As producers use inputs more effectively and precisely, or adopt improved cultivation and livestock raising practices, their TFP grows while using a fixed or even reduced amount of inputs.

For crops, TFP improves with:

- higher yielding, disease resistant and drought or flood tolerant crop varieties
- more efficient and timely cultivation and harvesting practices
- using technologies that indicate precisely when and how much water and fertiliser to apply.

For raising livestock:

- breeding animals for favourable genetic qualities and behaviour
- using better animal care and disease management practices
- adoption of high quality feeds contribute to greater productivity.

Over the past 55 years, agricultural productivity has increased at an average of between 2.5–3 per cent each year in developing countries and between 2–2.5 per cent each year in more developed countries. Figure 10.10 shows that increases in productivity over this period can be attributed to a number of factors:

- **Expansion of land being used for agriculture (extensification)** – this is especially the case in lower income developing countries as most cultivable land is already used in developed countries. To a more limited extent, additional irrigation has been needed to bring land into production.
- **Intensification** – additional economic inputs such as machinery, fertilisers, pesticides and higher yielding seeds have had a major impact in increasing productivity. Intensification is used to increase yields in both the developed world, where it was associated with the 'industrialisation' of agriculture, and in the developing world, where it was associated with Green Revolution technology.
- **Total factor productivity** – better and more precise use of inputs, particularly based on scientific research and development. This has been the more dominant factor in increasing productivity in the last 20 to 30 years.

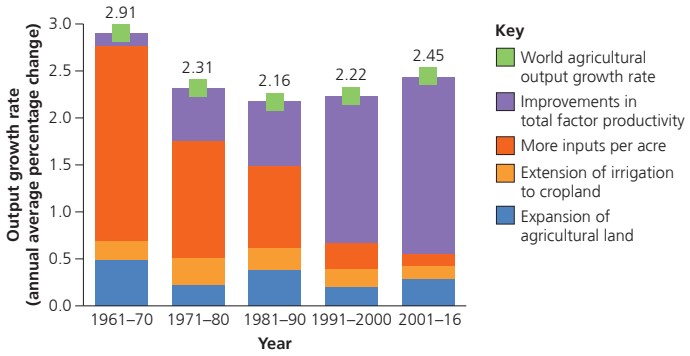


Figure 10.10 Growth in global agricultural output, 1961–2016
Source: U.S. Department of Agriculture

Projections for agricultural productivity forecast continued growth until 2050 but at a slightly lower rate of 2.3 per cent.

In summary, most parts of the world are getting smarter at producing agricultural products more efficiently. Growth in productivity coming from expansion of farming land or large irrigation projects is less important than in the past. Figure 10.10 also shows that accelerated intensification plays a lesser role now than in the latter part of the twentieth century. Increased output is mostly a result of scientists and farmers making more efficient use of agricultural inputs to make farming systems more productive and extracting more output from each unit of input.

Key physical environmental variables

As we have seen from Figure 10.8 (page 469) farmers make decisions on how to use their land based on:

- their own strengths, knowledge and skills
- cultural factors
- economic variables
- physical factors.

Two of the most important physical factors are the climate and the soils; they are closely related because soils are often determined by the climate and the natural vegetation of an area.

Climate and soils

Climate and soils are known as 'limiting factors' in agricultural production because specific crop types can only survive and be economically productive in certain climatic conditions and in suitable soil types. McCarty and Lindberg's Optima and Limits Model (see Figure 10.11) suggests that:

- each crop type will have optimum growing conditions where climate and soils (and other physical conditions) are ideal and will give maximum productivity
- there are physical limits beyond which a specific crop type cannot be viably grown
- as you move away from the optima towards the limits then costs of production increase (for example, more inputs such as fertilisers are needed) and/or yields decrease
- crops are not always grown in optimum conditions, but to be successful they are grown within their environmental limits.

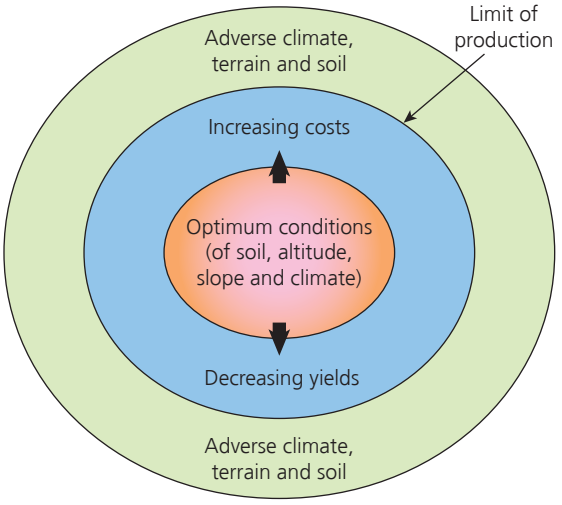


Figure 10.11 McCarty and Lindberg's Optima and Limits model
Source: Waugh, D. (2000). Geography: An Integrated Approach

Many crop types have been hybridised or modified by agricultural scientists to produce new varieties of the crop which are more tolerant of a wider range of climatic and soil conditions. New varieties have extended the spatial limits within which many crop types can be grown and some have been developed specifically to grow optimally in certain conditions.

Livestock are generally more resilient than crops, but even so, certain climatic and soil conditions suit some animals more than others and specific breeds are optimally more productive in some areas than in others.

Characteristics of major climate types

By studying Figure 10.12 (page 472) and Figure 10.2 (page 464), a correlation can clearly be seen between the patterns of world population density and types of climate.

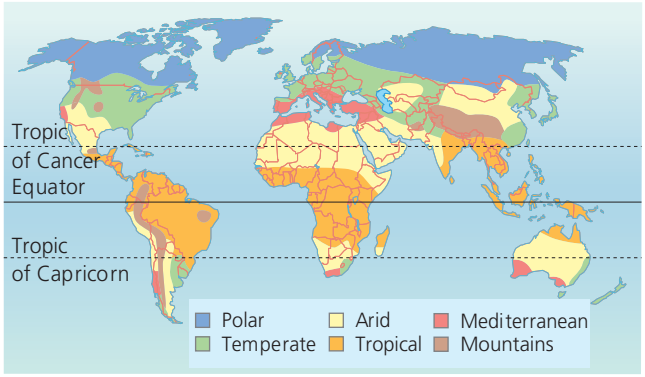


Figure 10.12 Climate zones of the world

Skills focus

When asked to analyse data that appears in map form, you must be able to recognise and describe patterns. This is not the same as simply naming places/regions. To what extent do the patterns on Figure 10.2 (page 464) and Figure 10.12 match?

Climatic type 1: Polar climate

Characteristics and distribution

The polar climate is one of long and intensely cold winters, with very low temperatures (falling below –40°C) accompanied by snow and strong winds. The cold is such that:

- there is a permanent covering of the land by glacial ice with a surface layer of snow, or
- it freezes the land from the surface downwards to produce permanently frozen ground (permafrost). In summer, the top layer thaws for a short time but the ground underneath remains frozen. In winter, there is a covering of snow. This area is known as the tundra.

According to the Arctic Human Development Report, the polar areas are occupied by approximately 13.1 million people spread between eight countries with a population density of less than four people per square kilometre. Most of these people live in the tundra climatic area of continental North America and Eurasia. Glacial areas are mostly devoid of permanent human settlement. They are visited by nomadic hunters and oil and mineral prospectors.

Relationship to human activities

In the second half of the twentieth century, the number of people living in the Arctic region started to grow rapidly because of improved healthcare for indigenous populations and the discovery of vast natural resources, which led to a large influx of immigrants. Recently population growth has slowed down in general and in the north of Russia, the total population has been declining. It is estimated that two-thirds of the total Arctic population lives in relatively large settlements, however the indigenous peoples tend to be found in small, widely scattered communities.

The only form of arable farming in such a place is one where the farmer creates an artificial environment. For example, Tim Meyers farms seven hectares of land in the Yukon-Kuskokwim Delta near Bethel, Alaska (Figure 10.13). The nutrient-rich soil is frozen and so Meyers first had to develop a method for thawing the ground. It was a two-year process of clearing vegetation and then spreading manure, composted tundra and a 'slurry' of salmon waste. Added to this was some lake water and a small amount of dry molasses to 'increase biological life'. Raised beds and high polytunnels help mitigate the low temperatures and short growing season, and a vast underground root cellar stores and extends the lifespan of the harvest. He grows and raises root vegetables, brassicas, gourds, strawberries, poultry and eggs.

Figure 10.13 Tim Meyer's farm in the Alaskan tundra

In North America, some Indian and Inuit groups hunt caribou (the American reindeer) but in Eurasia, many Arctic tribes herd them. The reindeer are not totally domesticated, as they generally roam free on the Arctic pastures. A single owner can possess hundreds or even thousands of animals. Traditionally, the herders followed the animals northwards to their breeding grounds. They return south in autumn to escape the harsh polar winter. Reindeer have never been bred in captivity, and individuals are just tamed for milking and as draught animals. The reindeer is the source of everything from milk, meat and skin (clothing and tents) to antlers and the fatty marrow of the bones.

Further details about human activities in cold climates can be found in Chapter 4 (page 170).

Climatic type 2: Tropical monsoon climate

Characteristics and distribution

The monsoon climate occurs in much of India and Bangladesh. It involves the seasonal reversal of winds. In the winter, they are from the north and north east, blowing outwards from central Asia. These winds are known as the north east or retreating monsoon. They are very dry and most rainfall is concentrated in the state of Tamil Nadu in the south east, while much of the remainder of India experiences a winter drought.

In summer (from May–June to September) the land in and around the Thar Desert in Rajasthan heats up rapidly, creating low pressure. This causes a reversal in the wind direction as the low pressure draws in warm, moist air from the south west that originated over the equatorial area of the Indian Ocean. These south-west monsoon winds bring high and intense amounts of rainfall that is increased by the uplift of the air over the Western Ghats mountain range of south-west India as well as by intense convection caused by the hot land surface. The monsoon rains cause flooding to many towns and cities along the west coast. An example is Mumbai, which has an annual mean rainfall of 2,386 mm, with all but 120 mm falling in the summer months (June–September). Temperatures range from 32°C in summer to 18°C in winter. Some of the south-west monsoon air flow continues around the eastern side of the Western Ghats and continues north-eastwards, picking up more moisture from the Bay of Bengal. Further uplift of this moist air over the foothills of the Himalayas brings seasonal monsoon rainfall and intense flooding to the lower Ganges valley and Bangladesh until mid-October.

Relationship to human activities

Rice seedlings are grown in nurseries until the summer monsoon rains begin and then they are transplanted into flooded fields (10–12 cm deep water). The fields are level and have low mud walls to retain water. Although rice is best suited to lowland flood plains with deep fertile alluvial soils, it can be grown in terraces cut into the slopes of hilly areas where the flood waters are retained by earth bunds. Rice growing in India is predominantly a labour-intensive farming system where a large, cheap labour force using simple tools and oxen does all the work (Figure 10.14).

The Indian economy gains due to good monsoon rains in the country. It can support large numbers of people in both rural areas and the rapidly growing cities of the Indo-Gangetic Plain. In good years, the surpluses produced have enabled India to become a leading exporter of both rice and wheat, from which the economy gains. On the other hand, weak monsoon rains result in crop failure, which affects the economy in a negative manner due to lower production. This translates into rising prices, fewer exports, less food security and rising debts for farmers, leading to increases in suicide among farmers in some regions.

Climate change is affecting the monsoon climate and increasing its unpredictability. There have been more fluctuations in the timing and the amount of rainfall, making it difficult for farmers to plan, resulting in crop failure. There has been a drive by the Indian government to increase the amount of land that is irrigated, especially by more efficient micro-irrigation, but it still stands at less than half the total land under cultivation.

Figure 10.14 Labour intensive rice farming in the monsoon area of India

Climate change and agriculture

According to the FAO, climate change is not hypothetical and the effects of global warming are already becoming obvious. The world is beginning to witness floods, droughts, shifting of monsoon patterns and more frequent and intense weather events. These will have wide-ranging effects on the environment, especially on agriculture and **food security**.

More than any other region, the people in the sub-Saharan Africa region are likely to be hardest hit as a result of climate change. Agriculture is likely to be

the most vulnerable economic sector because of its dependence on climate and weather. People in sub-Saharan Africa are mainly agrarian, with over 60 per cent of the population living in rural areas and working on small farms. Agricultural produce accounts for 23 per cent of the sub-Saharan African economy. The region has experienced more frequent and more intense climatic extremes over the past few decades and temperature increases in the region are projected to be higher than the global average increase. Over one billion people will face the direct impacts of climate change, which even at a 1.5°C global temperature rise will include:

- longer and more frequent heatwaves
- rapid drying out of soils, especially in the western Sahel
- less rain, longer and more frequent droughts, especially in West Africa and southern African countries
- reduced crop yields, especially in West Africa, which has been identified as a climate change hotspot.

With sub-Saharan Africa's population expected to increase by a further one billion by 2050, climate change poses a real threat to food security in the region.

Climate change will also impact adversely on the Asia-Pacific region. Coastal areas of south and South East Asia are likely to face the combined threats of changing rainfall patterns, resulting in severe water shortages or floods, along with rising temperatures and sea levels. This will affect:

- growing conditions in the south-west Pacific, causing changes to growing seasons and reductions in yields
- people's ability to live on some of the low-lying Pacific islands.

Links

There are links here to Chapter 7, Global systems and governance, particularly

- Interactions of governance between different scales, page 330, and,
- Unequal flows of ideas and Unequal flows of technology, page 300.

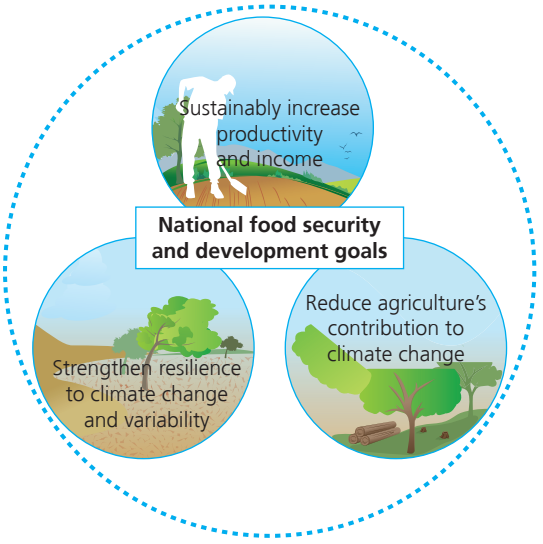


Figure 10.15 Climate Smart Agriculture (CSA)

At its global conference on Agriculture, Food Security and Climate Change, the FAO introduced the concept of **Climate Smart Agriculture (CSA)** (Figure 10.15).

CSA is an integrated approach to address the interlinked challenges related to food security and climate change through:

- **increased productivity (economic)** – to improve food and nutrition security and to boost the incomes of the world's rural poor who are reliant on agriculture for their living
- **enhanced resilience (social)** – to reduce vulnerability to drought, pests and disease and improve capacity to adapt to challenges such as shortened seasons or erratic weather patterns
- **reduced emissions (environmental)** – to pursue lower emissions for each unit of food produced and reduce/remove greenhouse gas (GHG) emissions from agriculture, including avoiding deforestation.

Climate Smart Agriculture has a clear focus on addressing the challenges presented by climate change by considering both the synergies and trade-offs that exist between productivity, adaptation and mitigation. The ideas behind CSA have been welcomed by many global organisations, including the World Bank, which has committed to supporting countries or specific projects to deliver CSA.

There are currently 455 members of the Global Alliance for Climate Smart Agriculture (GACSA) and many of them are inter-governmental groups, NGOs, research institutions or private sector organisations. However, there are relatively few state members of GACSA from regions most at risk; only eight African countries are currently members. There is also concern in civil society groups that GACSA's promotion of CSA is greenwashing a plan to produce more with less.

Links

There is a link here to Chapter 1, Water and carbon cycles, particularly The carbon cycle, page 24, and The major stores of carbon, page 25.

Characteristics of zonal soils

Soil is made up of a mixture of solids (minerals and organic matter), liquids and gases that occur on the land surface and is characterised by one or both of the following:

- horizons, or layers, that are distinguishable from the initial material as a result of additions, losses, transfers and transformations of energy and matter
- the ability to support rooted plants in a natural environment.

Soil performs five essential functions:

- **Cycling nutrients**: Carbon, nitrogen, phosphorus and many other nutrients are stored, transformed and cycled in the soil.

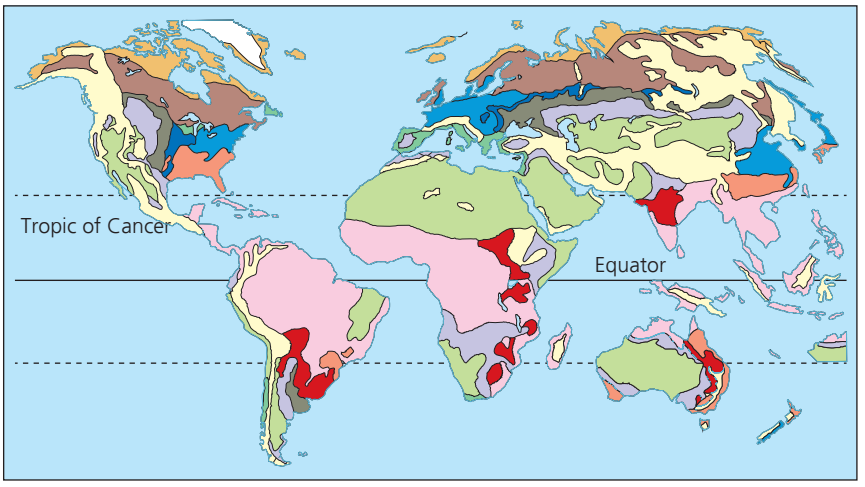
Example of adaptation to climate change

Changing climatic conditions and extreme weather events are particularly treacherous for small, resource-poor farmers who do not have any savings, alternative income or insurance. The **Indian Agricultural Research Institute (IARI)** carried out **Climate Change Adaptation Projects** to test technologies and strategies for sustainable livelihood security in rural communities that are vulnerable to climate risks. One such community is the Mewat district of the Indian state of Haryana, about 100 km south of Delhi.

Mewat is a low rainfall, drought-prone area. Analysis of past weather data showed that the mean minimum temperature increased at the rate of 0.18°C every 10 years during the monsoon (kharif) season, and by 0.47°C every 10 years during the dry (rabi) season. Between 2020 and 2050, temperature change is expected to accelerate, with minimum temperatures increasing by 1.87°C during the monsoon season and by 2.73°C during the dry season.

In recent years, the district has also experienced unseasonal hailstorms and extreme rainfall patterns that used to occur once in every ten years but are now more frequent. These storm events have negatively impacted farmers and production. In 2017, the winter crop was destroyed by hailstorms and then the summer crop was destroyed by intense rainfall leading to flooding, when 250 mm fell in just five hours. The annual average rainfall for the district is only 600 mm. Based on these findings, IARI developed a customised set of interventions to conserve resources, adapt to climate changes and improve livelihoods:

- Superior seed varieties were tested and the successful ones were made available through village seed banks.
- Heat-stress tolerant varieties of the major crop, wheat, were introduced, increasing yields by 12 to 18 per cent.
- Agro-forestry is encouraged as certain tree-crop interactions are beneficial. For example, cowpea and pearl millet perform better when planted alongside Prosophis trees. Trees offer shade and a windbreak that creates a micro-climate to conserve soil moisture. Prosophis trees also produce a bonus yield of dry leaves and twigs, which act as a mulch and add organic matter as well as a source of fuelwood from annual loppings.
- Farmland was levelled using lasers and this improved water-use efficiency by about 15 to 20 per cent.
- Underground pipelines were laid in farmers' fields and used for delivering high-tech drip irrigation; this resulted in an additional 40 per cent saving. Overall, the irrigated area increased by 45 per cent while the labour hours required for irrigating crops were reduced by 28 per cent. Rainwater harvesting and storage in covered tanks is also encouraged to maximise water efficiency.
- Greater crop diversification was introduced with high yielding varieties of chilli, aubergine, tomato and onion, accompanied by improved production technologies (for example, raised-bed planting, starting seedlings in a nursery, and micro and sprinkler irrigation) which increased household incomes by up to 86 per cent higher than households that maintained conventional cropping.
- Farmers were also given access to an information and communications technology platform, **mKRISHI®**, which is operated by **Tata Consulting Services**. Using their mobile phones, farmers and farmers' groups were connected to weather forecasting and agricultural production advisory services.



Key

- Soils of the tundra
- Podzols and related soils of the boreal forest
- Grey forest soils of the forest-steppe transition
- Brown earth and leached soils of the deciduous forest
- Chernozems of the temperate grasslands
- Chestnut soils and brown soils of the semi-arid grasslands
- Red and grey soils of the deserts
- Red and brown soils, cinnamon soils of the 'Mediterranean' woodlands
- Red-yellow podzolic soils of the subtropical woodlands
- Soils of mountainous areas
- Dark grey and black soils of the tropics and subtropics (vertisols)
- Red and yellow tropical rain forest and savanna soils (ferrallitic, ferruginous, ferrisols)

Figure 10.16 A simplified world soil map

- **Regulating water**: Soil helps control where rain, snowmelt and irrigation water go.
- **Sustaining plant and animal life**: The diversity and productivity of living things depend on soil.
- **Filtering and buffering pollutants**: Minerals and microbes in the soil filter, buffer, immobilise, degrade and detoxify potential organic and inorganic pollutants.
- **Physical stability and support**: Soil structure provides a medium for plant roots.

A **zonal soil** is a major soil group often classified as covering a wide geographic region or zone and embracing soils that are well developed and mature, having taken a long time to develop from the parent material. They are in a state of dynamic equilibrium with the climate, vegetation and parent matter, forming the base for the climatic climax community (see Chapter 6 pages 237–8). Within the area of a zonal soil there are variations caused by local factors.

Figure 10.16 shows the distribution of the major zonal soil types.

Zonal soil 1: Chernozem

Chernozems are deep black soils rich in organic matter, their name deriving from the Russian for black earth (Figure 10.17). Their mineral content mainly derives from wind-blown (loess) sediments. Chernozem has a black colour and contains a high percentage of organic matter as well as high natural percentages of nutrients such as phosphorus and ammonia. It is over one metre deep and has a clay-like structure which is good for retaining water.

These soils are found in regions with a continental climate (cold winters and hot summers) and flat or gently undulating plains with a natural climax vegetation of tall grasses. They cover an estimated 230 million hectares worldwide, mainly in the middle latitude steppes of Eurasia and North America.

Figure 10.17 A chernozem soil agricultural landscape in late autumn in the Ukrainian steppes

The people who inhabited these regions (for example, Native Americans on the prairies of North America and Eurasian Nomads on the steppes) were nomadic herdsmen or hunter-gatherers. The high natural fertility of chernozems and their favourable **topography** attract modern agriculture because they permit a wide range of agricultural uses, including arable cropping (with supplemental irrigation in dry summers) and cattle ranching.

The black earth region of the steppes of Ukraine has a climate that allows planting in both winter and spring. The single harvest is carried out between July and October. This time range allows planting and field work periods to be spread out, as well as helping with crop rotation.

Crops include cereals (wheat, barley and maize), oilseed plants (sunflower, oilseed rape and soybean), as well as potatoes.

Zonal soil 2: The red/yellow latosols of the tropical rainforest.

Figure 6.46 (page 268) describes latosols. They are the product of a hot, wet climate and thick forest cover. Although they appear to be very fertile, most of the organic nutrients are stored in the vegetation and not the soil.

One successful way of farming in this area is by shifting cultivation. Indigenous peoples (such as the Quicha and the Kayapo of the Amazon Basin) clear small areas of vegetation. They then burn it because the ash provides nutrients for the infertile soil. The land is then farmed for two to three years before they move on to another area of the rainforest. This allows both the forest and the soil to recover. This only supports very low numbers of people and the rainforests have typically been sparsely populated.

In more recent times both poor settlers and ranchers have moved into the rainforest areas and vast tracts of forest have been cleared. The causes and impacts of this clearance on both the environment and the population are covered in Chapter 6 (page 272).

Soil problems and their management

Soil erosion

Soil erosion is defined as the wearing away of the top layer of soil (**topsoil**). It is the most fertile layer because it contains the most organic, nutrient-rich materials. It is this layer that farmers want to protect for growing their crops and for their animals to graze on. Soil is eroded by both water and wind (Figure 10.19, page 478, and Figure 10.20, page 479). Water erosion is judged to be the more serious of the two types. Reports suggest that around 36 billion tonnes of topsoil are lost each year due to water erosion and deforestation. The FAO estimates that five to seven million hectares of productive land is lost through erosion and other forms of degradation each year.

Figure 10.18 (page 478) summarises the ways in which soil is eroded and transported away from fields.

Repeated erosion reduces the fertility of the soil by:

- removal of topsoil that is rich in nutrients and organic matter
- reduction of the depth of soil available for rooting and water storage for crop growth
- reducing infiltration of water into soil, thereby increasing run-off and erosion.

This in turn can lead to:

- loss of seeds, seedlings, fertiliser (nitrates) and pesticides
- young plants being 'sandblasted' (wind erosion)
- increased difficulty (therefore fuel consumption and man-hours) of farming operations.

Damage to the environment can include:

- deposition of sediment onto roads, neighbouring land and into drains and ditches
- damage to the quality of water courses, lakes and rivers through excess inputs of sediment and

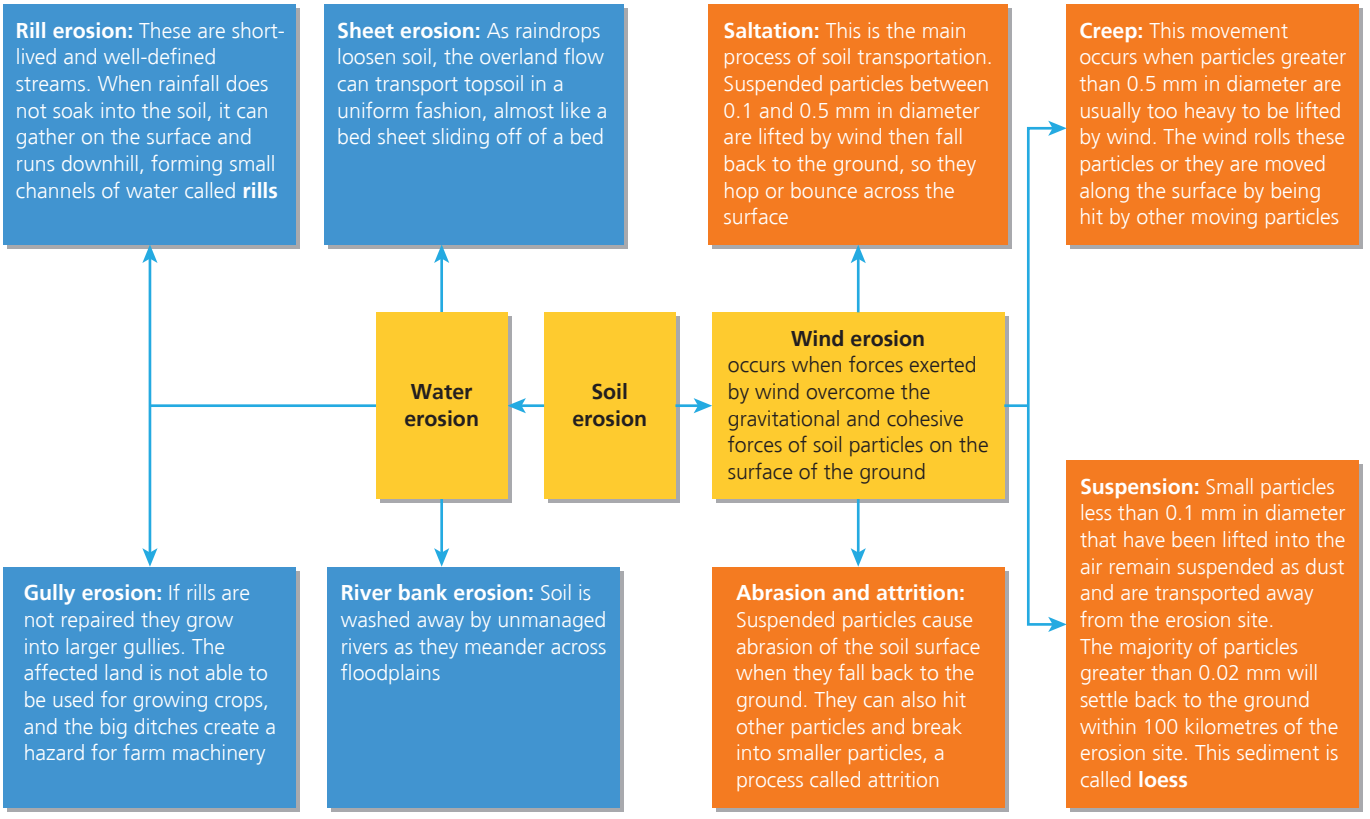


Figure 10.18 Characteristics of soil erosion

increased chemical loading, which in turn may lead to **eutrophication**.

- increased run-off and sedimentation causing a greater flood-hazard downstream
- sediment in rivers damaging the spawning grounds of fish.

Water erosion control

Figure 10.19 Sheet erosion on red sandstone soils in Norfolk

The key to reducing soil erosion by rainwater is to reduce the amount of surface flow of water. This is done by:

- installing and maintaining field drains (see Figure 1.33, page 21) and ditches. Sediment should be removed from ditches and replaced in the fields where it came from
- reducing the amount of water running off roads and farm track onto fields (again by drains and ditches)
- the judicious use of farmyard manure to stabilise the topsoil
- protecting soil in winter by early sowing or the use of cover crops
- work across slopes whenever possible. Contour ploughing reduces overland flow and the formation of rills and gullies.

Links

There are links here to Chapter 1, specifically Drainage basin systems, page 10, and Examples of how land use changes can affect the water cycle, page 20.

Wind erosion control

Figure 10.20 Wind erosion of sandy soil in East Anglia

Wind erosion control is carried out by:

- increasing soil cohesion by applying organic matter (for example, farmyard manure) to the soil; this improves its structure
- increasing the roughness of the soil surface or by leaving crop residues or stubble in fields and not ploughing them into the soil. In Burkina Faso, for example, millet and sorghum stubble is left at one metre high so that it stabilises the soil and also where it traps dust and leaves blown by tornadoes; this material eventually adds to soil depth and fertility
- increasing plant cover so that surface wind speed can be cut

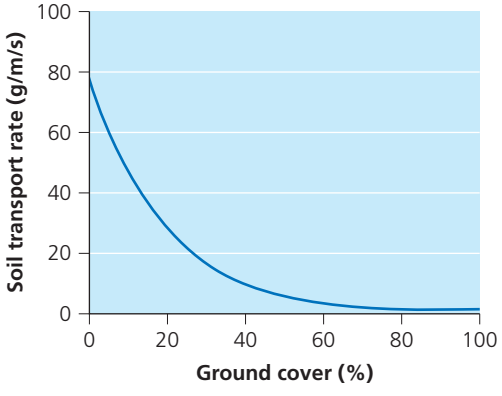


Figure 10.21 Relationship between ground cover and wind erosion

- increasing plant cover to about 50 per cent gives fields adequate protection from wind erosion (Figure 10.21)
- planting lines of trees or hedgerows as windbreaks cuts wind speed which reduces both evaporation (by up to 20 per cent) and wind erosion. A tree line has a wind speed reduction effect for up to 12 times the height of that tree line, both before and after the barrier. This means that the cropped area between windbreaks can be as wide as 100 m if the trees are over 5 m high. This can be done on a large scale and is one of the reasons for the Great Green Wall project across the Sahel in Africa, to prevent further desertification of the region.

Waterlogging of soil

Soil is considered to be waterlogged when enough of the pore spaces in the soil are occupied by water rather than air for there to be insufficient oxygen (**anaerobic conditions**) for plants to respire. Plants differ in their demand for oxygen and so there is no universal level of soil oxygen that can identify waterlogged conditions for all plants. In addition, a plant's demand for oxygen in its root zone will vary with its stage of growth.

Lack of oxygen in the root zone of plants causes their root tissues to decompose. The consequence is that a plant's growth and development are stalled. If the anaerobic circumstances continue for a considerable time the plant eventually dies. Fortunately, waterlogged conditions often do not last long enough for the plant to die. Once a waterlogging event has passed, plants recommence respiring. As long as soil conditions are moist, the older roots close to the surface allow the plant to survive, though it may be stunted with low yield.

Waterlogging occurs under two conditions:

- Surface-fed waterlogging happens when precipitation, irrigation water or river floodwater exceeds the combination of evapotranspiration and percolation so that the water stays in and on the surface of the soil.
- Groundwater-fed waterlogging is caused when the rate of rising groundwater is not matched by the rate of evapotranspiration. This may be a natural rise in groundwater or be caused by seepage from irrigation canals.

Links

The methods and impacts of soil drainage are discussed in Chapter 1, Water and carbon cycles, on pages 21–2.

Salinisation

Salinisation refers to a build-up of salts in soil, eventually to toxic levels for plants (3,000–6,000 parts per million (ppm) of salt causes problems for most cultivated plants). Salt in soils can be directly toxic to plants but its biggest effect is that it decreases the osmotic potential of the soil water and reduces

the rate of root water uptake. If the soil has greater concentrations of solute than the root, then the plants cannot get water from the soil.

Problems with salinisation are most commonly associated with excessive water application, rather than with too little. All irrigation water contains dissolved salts derived as it passed over and through the land. This is particularly true of water that has come from groundwater sources. Rainwater also contains some salts. Evaporation of this water from the dry surface of the soil leaves the salts behind. It is also likely to become a problem on soils where the groundwater is within three metres or less of the surface. In such cases, water rises to the surface by capillary action rather than percolating down through the entire soil profile. It then evaporates from the soil surface, again leaving the salts behind on the surface.

Figure 10.22 Salinisation of soil in the Indus Valley, Pakistan

Salinisation of soils is a problem in more than one hundred countries worldwide and irrigation is often a major contributing factor. Estimates of how much land is affected globally vary but scientists suggest that 10 per cent of all arable land is affected, but this rises to about 25 per cent of all irrigated land being salt-affected. It is particularly acute in hot, semi-arid areas which are poorly drained and use a lot of irrigation water; there is never sufficient rainwater to flush the salt out of the soil. These conditions are found in China's Northern Plain, Central Asia, the San Joaquin Valley of California and the Indus Plain of Pakistan (Figure 10.22). Salt concentrations are high enough in much of this irrigated land to decrease yields significantly. In extreme cases, land is abandoned because it is too salty to farm profitably.

The 'treatment' for salinisation is to flush the soil with lots of water. However, this results in the salinisation of rivers and groundwater where the flush water goes. For example, following the flushing of soil in the Lower Colorado River valley, the river became too salty downstream for the Mexican farmers to use for irrigation. The Mexican government forced the USA to construct a desalination plant near the Mexican border so the water would be useable.

In extreme cases, when the salt crust is too thick, it cannot be flushed, as water just runs off the salty surface.

Structural deterioration of soil

Soil structure is the arrangement of soil particles into groupings. These groupings are called **peds** or aggregates, which often form distinctive shapes typically found within certain soil horizons. For example, granular soil particles are characteristic of the topsoil. Soil aggregation is an important indicator of the workability of the soil. Soils that are well aggregated are said to have 'good soil tilth.'

Soil structures can be classified as follows:

- **Granular and crumb structures** are individual particles of sand, silt and clay grouped together in small, nearly spherical grains. Water circulates very easily through such soils. They are commonly found in the top, A-horizon of the soil profile.
- **Blocky structures** are soil particles that cling together in nearly square or angular blocks having more or less sharp edges. Relatively large blocks indicate that the soil resists penetration and movement of water. They are commonly found in the lower (B-horizon) layers of the soil where clay has accumulated.
- **Prismatic and columnar structures** are soil particles which have formed into vertical columns or pillars separated by vertical cracks. There is some permeability but water moves more slowly and generally drainage is poor. They are also found in the lower, more clay-rich layer of soil.
- **Platy structures** are made up of soil particles aggregated in thin plates or sheets piled horizontally on one another. Plates often overlap, greatly impairing water circulation. They are commonly found in forest and high latitude soils.

Under natural conditions, the breakdown of soil structure is rare but where agriculture is practised, changes in soil structure are common. The causes of structural deterioration by agricultural practices, resulting problems and how these can be managed are shown in Table 10.2.

Table 10.2 Structural deterioration of soil: problems and management

Cause of deterioration of structure	Resulting problems	Management of soil structure deterioration
Reduction in soil organic matter: • Cultivation causes **physical fracturing and mixing** of soil and increased aeration; this stimulates breakdown of soil organic matter • Most of the **plant matter is removed** when harvesting crops and burning stubble; so less organic material is returned to the soil in the form of decaying plants	• The bonds holding aggregates become weak because of the low organic matter content • They can no longer withstand disruptive forces such as raindrop impact • Individual particles of soil break off from the aggregates causing breakdown in structure • When exposed to heavy rain, soils may form a surface crust that impedes infiltration of water and/or emergence of seedlings	• Reduce use of chemicals and multi-harvesting that puts soil under pressure • Replace lost organic material with compost or manure (artificial fertilisers restore fertility but do nothing to restore soil structure) • Plant trees (agro-forestry) as litter and debris from trees replenish organic matter • Leave land fallow or rotate crops to allow some natural stabilisation of the weakened soil
Compaction of soil created by a combination of pressure and sliding forces from: • **farm machinery** – tractor wheels, plough soles, rotary blades etc. • **grazing livestock** – when hooves of livestock compress the soil (also known as poaching) The way soil reacts to application of pressure depends on texture, how hard and deep the soil is and the shape of the contact area. Compaction is worsened in wet conditions	• The concentrated weight of farm vehicles repeatedly moving over the same ground compacts the soil into ruts, which become impermeable to water • Rainwater gathers in ruts and runs off downslope causing, in extreme cases, gully erosion • Plough soles smear the soil immediately below the depth of a plough furrow creating an impermeable layer called a plough pan • Hooves of livestock compact soil breaking structure and creating impermeable pads, prone to erosion	• Using existing drive lines to minimise the compaction over a wider area • Low inflation of tyres reduces compaction • Adopt conservation agriculture with zero tillage • Using additional tool attachments behind tractor tyres to loosen the compacted soil • Avoid working soils when wet • Management of stock levels to sustainable rates

Strategies to ensure food security

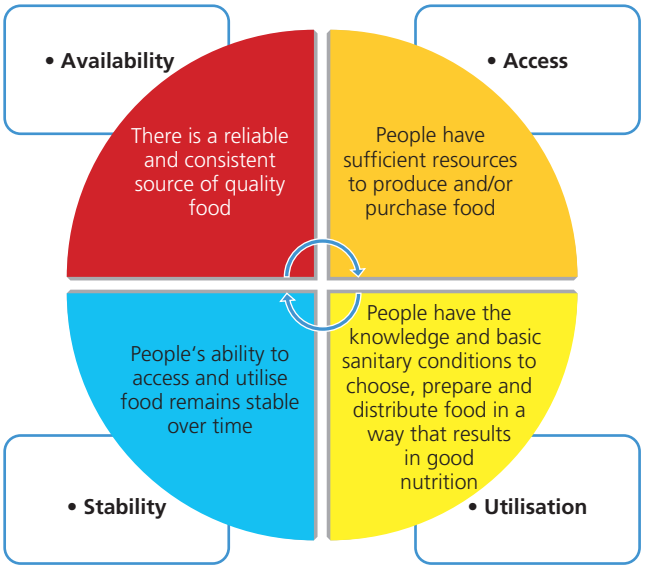


Figure 10.23 The four main components of food security

Food security depends upon (Figure 10.23):

- **Food availability:** Sufficient quantities of food available on a consistent basis.
- **Food access:** Having sufficient resources to obtain appropriate foods for a nutritious diet.
- **Food use:** Appropriate use based on knowledge of basic nutrition and care, as well as adequate water and sanitation.
- **Food stability:** That the availability, access and utilisation components remain sustainable for future generations.

Opinions about food security vary. Discussion points include:

- There is enough food in the world to feed everyone adequately, but it is not necessarily in the places where it is needed.
- Future food needs can – or cannot – be met by current levels of production.
- National food security is very important – or no longer necessary – because of global trade.
- Globalisation may – or may not – lead to the persistence of food insecurity and poverty in rural communities.
- Households get enough food, but whether that food fulfils the nutrition needs of all members of that household is unclear.

These are all discussion points that will influence the strategies that are taken to ensure food security to feed

a global population of nearly 10 billion people by 2050. The discussions centre around three main issues:

- **Production** – strategies to address this will focus on increasing production to meet the demands of a growing population (supply management).
- **Distribution** – this approach accepts the Food and Agriculture Organization (FAO) view that sufficient food is already being produced, even to feed the predicted 2050 population, but it is unevenly distributed. Strategies addressing this issue will focus on reducing post-production losses including wastage at both the production and consumption end of the chain and ensuring better mobility of food produced through free trade or aid.
- **Demand** – the uneven distribution is caused by over-consumption of food in wealthier countries that can afford to purchase it, especially high calorific value food such as meat, which requires more resources to produce. Strategies addressing this issue are likely to focus on reducing the demand for certain types of food to support more even distribution strategies (demand management).

There is considerable overlap between these different strategies to ensure food security and the likelihood is that an integrated approach will be needed, which contains elements of all three. To achieve food security sustainably, other elements, including protecting natural ecosystems and reducing greenhouse gas emissions, need to be incorporated into this integrated approach (Figure 10.24).

Links

There are links here to Chapter 7, Global systems and governance – Global governance, page 327.

Strategy 1: Increasing food production

The **Global Agriculture and Food Security Program (GAFSP)** believes that the world will need to produce about 56 per cent more food by 2050 to feed everyone. GAFSP provides an intermediary fund hosted within the World Bank to complement and reinforce existing international efforts to tackle food security. The Program's mission is to strengthen resilient and sustainable agricultural systems in low-income countries. By investing in smallholder farmers and their communities, GAFSP is helping progress towards the UN Sustainable Development Goals (SDGs), particularly SDG2 (Zero Hunger).

GAFSP receives donations from wealthy governments and from the Gates Foundation; it pools these donations to allocate funding where it is most needed and likely to make an effective impact.

By 2018, GAFSP had allocated $1.7 billion to projects that would tackle hunger and malnutrition and improve food security across 45 countries, reaching and supporting over 13 million smallholder farmers. For more details of individual projects visit the GAFSP website at www.gafspfund.org/projects.

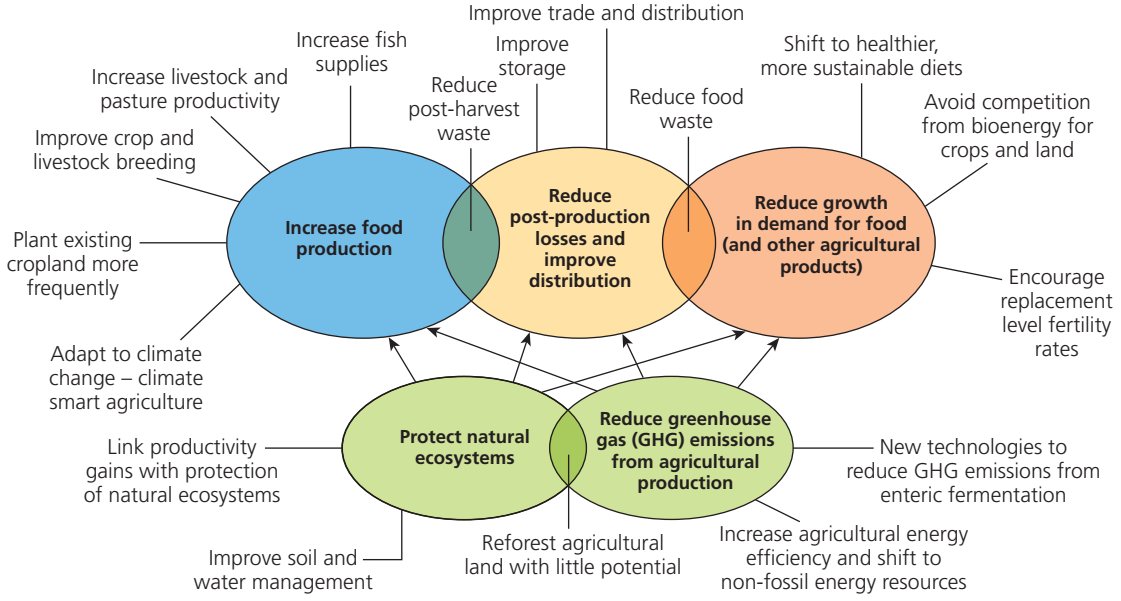


Figure 10.24 Overview of sustainable approach to ensuring food security

The work of GAFSP and partner organisations, such as the FAO, can be very targeted but contribute to more general measures to increase food supply without necessarily expanding agricultural land. These include:

- **Increasing livestock and pasture productivity** – suggested actions for farmers include improving pasture fertilisation, feed quality and veterinary care, as well as raising improved breeds and using rotational grazing, which will increase yields without agricultural land expansion.
- **Improve crop breeding** – the use of biotechnology to develop, through hybridisation, high-yielding variety (HYV) seeds, suited to specific environmental conditions could be advanced and extended further, particularly in sub-Saharan Africa. New advances in molecular biology offer great promise for additional yield gains, by making it cheaper and faster to genetically code plants and purify crop strains. Genetic modification (GM) may be dismissed by environmental protesters in Europe, but it might provide a solution to hunger in Africa.
- **Plant existing cropland more frequently** – food production can be boosted by reducing fallow land or by increasing double cropping. This puts pressure on land fertility, but carefully managed crop rotation accompanied by research conducted to determine where cropping intensity increases are most feasible will close some of the production gap.
- **Increase fish supplies** – fish is an important source of protein and nutrition to populations that otherwise lack these elements in their diet. An important aspect of increasing fish supply is the development of **aquaculture** (fish-farming), particularly in LDEs where the creation and stocking of pools with fast-breeding species, such as tilapia, can increase food and nutrition supply in a sustainable way. Productivity from aquaculture can be improved by selective breeding to improve growth rates of fish, improving feeds and disease control and using water recirculation and other pollution controls. However, this so-called '**Blue Revolution**' is not without its challenges. It competes with agricultural land for food production; it places a high demand on freshwater supplies (which may be limited) and requires, in most cases, wild fish to be caught to supply feed for the farmed fish.

Strategy 2: Reducing post-production losses

According to the FAO, the rate of global food production has increased more than the demand for food from population increase. However, approximately 25 per cent of food produced for human consumption goes uneaten, with loss or waste occurring all along the food chain from field to fork.

Figure 10.25 (page 484) shows that food losses around production are more prevalent in developing regions, while food waste around consumption is more prevalent in developed regions. Reducing losses during production and post-harvest practices is a way of increasing food supply, while reducing food waste is a way of reducing demand. Addressing both types of loss would help to increase and balance food security in the future. Measures taken include improving food storage in developing countries, streamlining expiration labels and setting targets to reduce food waste.

Improving post-harvest practices

After a crop has ripened a variety of functions are performed, including harvesting, drying, threshing, storage, transportation and marketing. Inefficient practices at each stage may allow crops to be exposed to contamination by micro-organisms, chemicals, excessive moisture, fluctuating temperature extremes, mechanical damage or ineffective storage. These practices compromise food safety and contribute greatly to food losses.

Improving post-harvest management has the potential to reduce losses, as well as to impact positively on the health and well-being of people living in the region. For example, in Uganda the most serious of food-related health risks is the constant threat of food poisoning caused by aflatoxin contamination. Aflatoxins are naturally occurring, highly carcinogenic poisons produced by a fungus that is particularly prominent in maize. It can develop when produce comes into contact with soil during harvesting, threshing and drying. The problem has become so widespread in East Africa that aflatoxin poisoning has become an epidemic.

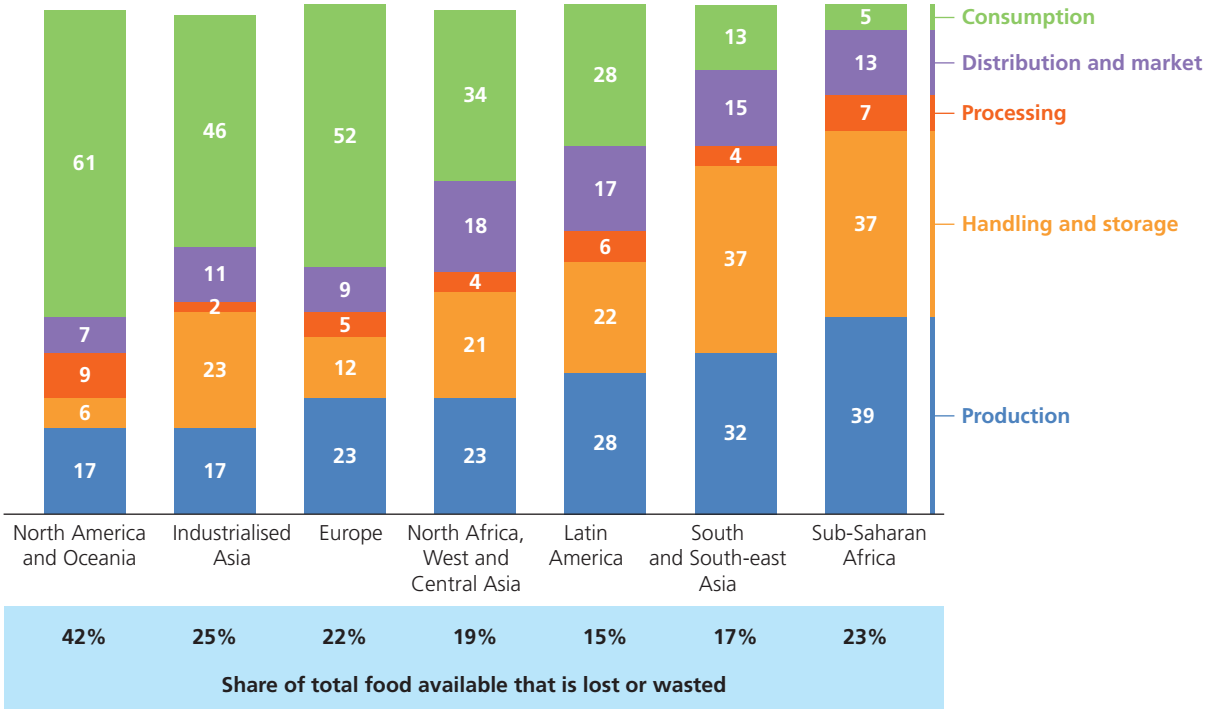


Figure 10.25 Where food loss and waste occur along the food supply chain
Source: 2018 World Resources Institute

The solutions

Training and guidance given to farmers in developing countries, such as Uganda, can reduce the crop's susceptibility to aflatoxins and other contamination. This includes instruction in:

- controlling moisture content and avoiding direct crop contact with exposed soil during drying
- understanding the impact of weather conditions at the time of harvest, so that harvesting is undertaken at the optimum time (for example, rain dampens crops leading to mould growth)
- drying crops by reducing the moisture content below the level required for mould to grow
- precautions to avoid damaging grain during threshing
- killing all life stages of insect pests prior to storage by placing the grain into a solar oven for 1–5 hours.

Contamination can also occur after grain has been placed into storage, due to pest infestation and poor storage conditions that lead to accelerated growth rates of fungi. Farmers are:

- advised against stockpiling harvested crops either directly in baskets or in polypropylene bags on the floor of their homes (for fear of theft)
- supported with improved methods of low humidity storage and new technologies such as hermetic (airtight) storage units or metallic silos to protect crops from insects, rodents, birds, rain and temperature fluctuations. The FAO has built more than 45,000 small, metal storage silos (just big enough for use by a single farmer) in 16 different developing countries.

Redistributing food

Some perfectly good food just never gets eaten. It might be because a farmer cannot afford to harvest an entire field, or because a grocer has ordered too much of an item and cannot sell it all. One way to reduce loss and waste is to simply redistribute food by giving it to food banks or NGOs for redistribution to communities in need. This is difficult to organise and achieve effectively beyond a local level, especially with fresh food. Redistribution on an international level (other than in emergencies) can lead to the problem of food 'dumping' in poorer countries which undermines the livelihoods of local farmers.

Reducing food waste close to the fork

- **Improve food labelling**: confusion around 'use-by', 'sell-by', 'best-before' and other date labels can lead people to throw out food that is still perfectly good

to eat. For example, one survey conducted in the UK found that one-fifth of food thrown out by households was incorrectly perceived as being out of date due to confusing labels. For food safety and to avoid wastage the focus of date labelling should be 'use by'.

- **Launch consumer awareness campaigns**: campaigns revealing how much food people actually waste may address consumers' consciences. By providing constructive solutions for cutting down on that waste, such as providing storage tips for fruit and vegetables, grocers or local government officials could play a part in these initiatives to reduce waste.
- **Set food loss and waste reduction targets**: setting time-bound targets inspires action by raising awareness, focusing attention, and mobilising resources. For example, the European Union has announced a target of reducing food loss and waste by 50 per cent by 2050.

Strategy 3: Reducing growth in demand

Shift to healthier, more sustainable diets

As societies in emerging economies have become more affluent, the demand for meat products in many countries, such as China, has increased enormously. Consumption of beef, lamb and goat is projected to nearly double by 2050 at current rates of growth. Meat is very resource intensive to produce and requires at least 10 times more land per gram of edible protein than common plant proteins, such as peas and beans, as more land is required to grow fodder.

Greater numbers of people eating less meat and consuming more plant-based food as part of their diet would free up more land for crop production for human consumption. This may require better marketing of plant-based food, improvement of meat substitute products or implementing policies to disincentivise meat consumption. There would undoubtedly be a reaction against this by pastoral farmers, especially in wealthier countries. Arguably a lot of farmland, especially steeper or marginal land, is only suitable for pasture and cannot be used for crops, so it is already being put to its most effective use.

Encouraging replacement level fertility rates

One of the most obvious strategies to ensure food security is to control population growth. Most of the world's nations are close to or below replacement level fertility of approximately two children being born per woman (see Section 10.4, pages 507–8). This level of fertility stabilises population growth and in the longer term will reduce pressure on food supplies. Sub-Saharan Africa is an exception globally, with fertility rates of above five children per woman of childbearing age in many countries. This rate is falling but may need to fall more rapidly to ensure food security in this region by 2050. Coercive policies are difficult to implement in poor regions with high infant and child mortality. Measures to encourage lower fertility, such as family planning initiatives, increasing the awareness and provision of contraceptive methods and compulsory education for girls (to give better career prospects), may be the best option. Many demographers and economists argue that the only way to secure lower fertility rates in poor areas is through economic development that can contribute to improved healthcare.

Review questions

5 Explain the difference between intensive and extensive farming systems.

6 Assess the extent to which climate change is likely to impact on agricultural activity.

7 Outline two types of soil problem that affect agriculture and suggest how each can be managed to mitigate the impacts on soil fertility and agricultural production.

8 Evaluate different strategies that might be used to ensure food security.

10.3 Environment, health and well-being

Few things matter more to people than their **health**, **mortality** and **well-being** and these are directly affected by the environment in which they live. As Figure 10.26 (page 486) shows, this can be the physical environment where factors such as air and water quality are important, the economic environment which affects the type of housing they live in and the jobs that they do or the social environment determining where an individual is educated or lives, and the people and institutions with whom they interact.

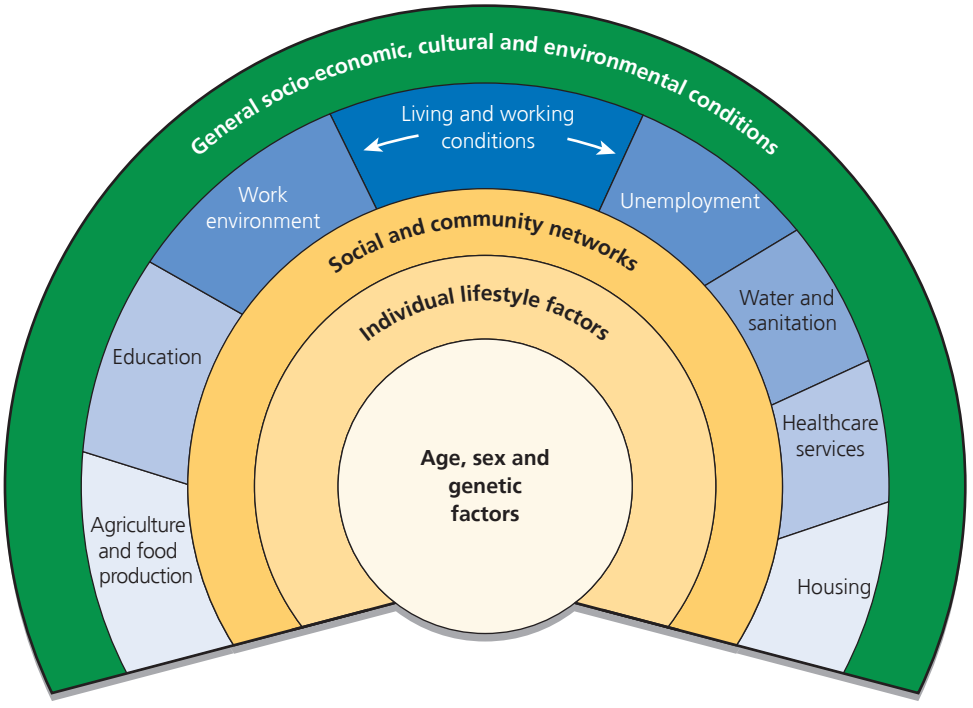


Figure 10.26 Determinants of health

Source: Dalgren G and Whitehead M (1991) Policies and strategies to promote social equity in health, Institute for Future Studies, Stockholm

Key terms

Disability-adjusted life years (DALYs) – A measure of morbidity within a society. They measure the number of years of healthy life lost by being in poor health or a state of disability.

Epidemiological transition – This describes the process by which the pattern of disease and mortality changes over time. The pattern is transformed from one of high infant/child mortality and infectious epidemics to one with a high prevalence of chronic degenerative and lifestyle-based diseases.

Health – Defined by the World Health Organization (WHO) as a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity.

Morbidity – Morbidity refers to being ill or unhealthy with a disease or the symptoms of a disease; it can also describe the incidence of disease within a population. Some diseases are so infectious that by law they must be reported. These include malaria, rubella and tuberculosis.

Mortality – Mortality relates to death. It can be measured by death rate, infant mortality rate, child mortality rate and case fatality rate.

Non-communicable disease – A medical condition or disease that is, by definition, non-infectious and non-transmissible among people. Non-communicable diseases include heart disease, the largest cause of deaths globally (see Figure 10.28), and cancer.

Well-being – The state of being comfortable, healthy or happy.

Global patterns of health, mortality and morbidity

Health is unevenly distributed. If all the people who died in 2019 were imagined as a group of 1,000 people, Figure 10.27 shows how **mortality** rates are considerably higher in low- and middle-income countries. More than half (540) of these 1,000 deaths

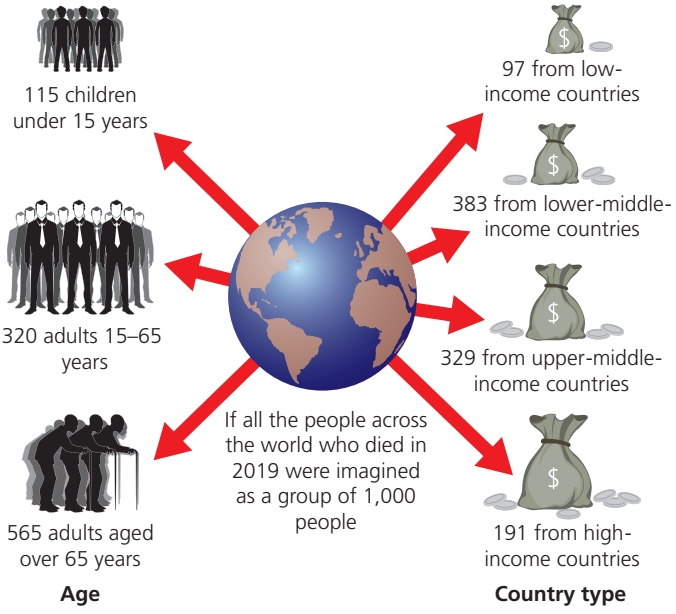


Figure 10.27 Distribution of mortality across age and country type in 2019

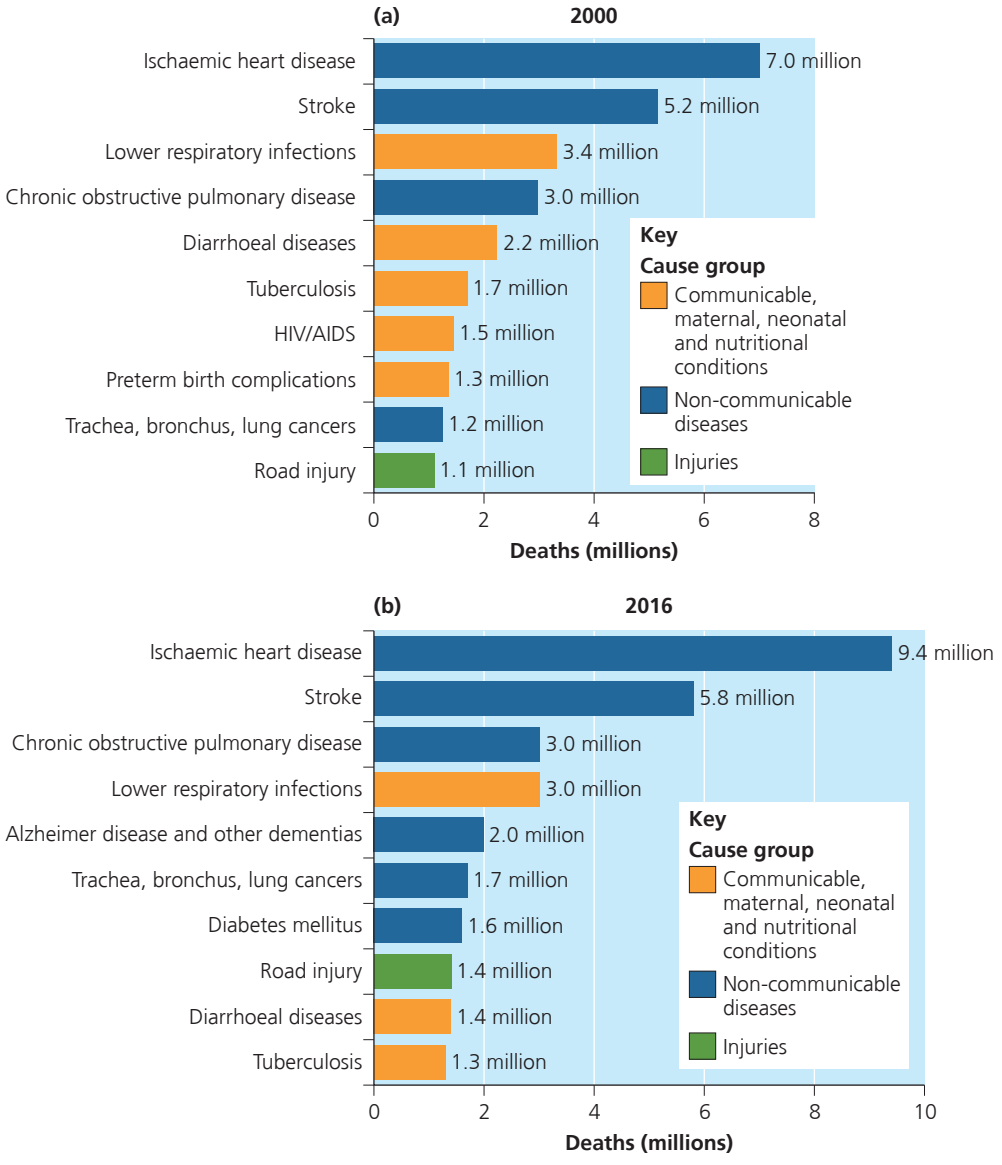


Figure 10.28 Top 10 causes of death globally, changes from 2000 (a) to 2016 (b)

Source: WHO

would have been caused by just 10 conditions, shown in Figure 10.28.

Non-communicable diseases (NCDs) were responsible for 71 per cent of all deaths globally in 2018, up from 60 per cent in 2000. The four main NCDs are cardiovascular diseases, cancers, diabetes and chronic respiratory diseases. In comparison, deaths from infectious diseases are declining. Historically, deaths from NCDs are associated with older age groups and in high-income countries. In 2016, however, WHO evidence suggested that of the 41 million deaths caused by NCDs, 15 million (36.5 per cent) were people between the age of 30 and 69 and that low- and middle-income countries were disproportionately affected with 32 million (78 per cent) of all NCD deaths.

Figures 10.28a and 10.28b show that the number of people dying globally from NCDs increased between 2000 and 2016, while deaths from communicable diseases have fallen. This is the result of improvements in sanitation, diet, treatments and medication. For example, HIV/AIDS, the cause of one million deaths in 2016, is now outside the top ten causes.

There is likely to be a spike in deaths caused by communicable diseases in 2020–21 because of the Coronavirus pandemic. The highly infectious COVID-19 virus was the cause of over 1.7 million deaths by the end of 2020. This will lift it into the top 10 causes of death for the year. With further waves of infection occurring in both HDEs and developing countries in the winter of 2021, the

death toll from the disease has already exceeded 2.2 million globally.

Economic and social development

As societies and nations develop, the economic and social conditions in which people live should improve. As part of this development, improvements in health and healthcare provision will also be witnessed. For example, increases in food productivity and supply and in transport infrastructure mean a population will be fed more reliably and is less prone to famine or diseases related to malnutrition. Similarly, developments in sanitation and public hygiene reduce the chances of water-borne infections such as typhoid and cholera. With advances in medical technology and vaccination programmes, the risk of infectious diseases has been reduced in many parts of the world.

Economic developments linked with improved health include:

- technology to improve food productivity and supply (for example, the Green Revolution)
- improved transport infrastructure to distribute food and medical supplies
- investment in drainage and sewage systems, especially in urban areas
- trading of resources or manufactured goods in exchange for a wider variety of foods and medicines.

Social developments improving health include:

- improved awareness of the importance of hygiene
- better education about sanitation and disease transmission
- advances in medical technology such as antibiotics and vaccines
- better training for doctors, nurses and midwives
- aid programmes from the UN or from NGO provision to improve healthcare facilities.

Socio-economic development is usually accompanied by a significant increase in population growth rates as life expectancy increases and death rates decrease. At the same time fertility initially remains high as social norms and prior experience of high infant mortality influence decisions to continue to have large families. Healthcare developments eventually reduce infant mortality rates and with greater survival, parents start to have fewer children, particularly in areas where development is accompanied by improved access to contraception. This leads to subsequent declines in fertility rates.

Epidemiological transition (see also the Demographic Transition Model)

The idea of **epidemiological transition** was posited by Abdel Omran (1971) who suggested that the process of socio-economic development accounts for a transition, over time, from infectious diseases to chronic and degenerative diseases as the most important causes of death (see Table 10.3 and Figure 10.29).

The transition occurs as a country undergoes the process of modernisation from developing to developed nation status. It is the element of demographic transition (examined later in this chapter) explained by advances in medical technology, disease prevention and sickness therapy and treatment.

Omran divided this transition into three distinct phases:

1. **The age of infection and famine**: Mortality is high and fluctuates; this prevents sustained population growth. The cyclical low growth patterns are associated with wars, famine and epidemic outbreaks, interspersed with periods of relative 'prosperity'.
2. **The age of receding pandemics**: The rate of mortality declines as disease epidemics occur less frequently. Population growth is sustained and begins to rise exponentially.
3. **The age of degenerative and man-made diseases**: Mortality continues to decline and eventually approaches stability at a relatively low level. Infectious agents as the major contributor to **morbidity** and mortality are overtaken by anthropogenic causes. Fertility becomes a more important contributory factor to population growth.

The theory was revisited and modified in the 1980s by other public health researchers. Olshansky and Ault added a fourth stage to the transition:

4. **The age of delayed degenerative diseases:** Declining death rates are concentrated at advanced ages.

The four recognised stages of epidemiological transition mirror the four stages of the demographic transition model (see pages 510–1). Epidemiological transition is clearly one of the key factors in explaining natural population change over time.

A fifth stage?

Since the establishment of a recognised fourth stage of epidemiological transition in the 1980s, it has been suggested that the world's population may be starting

Table 10.3 The five stages of epidemiological transition

Phase	Life expectancy	Change in socio-economic conditions	Causes of morbidity and mortality
1. Age of infection and famine	20–40	Poor sanitation and hygiene; unreliable food supply	Infections; nutritional deficiencies
2. Age of receding pandemics	30–50	Improved sanitation (e.g. clean water provided by public utilities); better diet; advances in medicine/ healthcare (e.g. discovery of penicillin in 1928 reduced deaths from infectious disease)	Reduced number of infections; increases in occurrence of strokes and heart disease
3. Age of degenerative and man-made diseases	50–60	Increased ageing; lifestyles associated with poor diet, less activity and addictions	High blood pressure, obesity, diabetes, smoking-related cancers, strokes, heart disease and pulmonary vascular disease
4. Age of delayed degenerative diseases	c. 70+	Reduced risk behaviours in the population; health promotion, disease prevention and new treatments	Heart disease, strokes and cancers remain the main causes of mortality but treatment extends life. Dementia and ageing diseases start to appear more
5a. *Age of inactivity and obesity*	*Potential reversal of life expectancy*	*Obesity increases at an alarming rate* *Lack of exercise* *Ignore risk behaviours (especially in high-income countries)*	*Chronic heart disease, stroke, congestive heart failure, diabetes, peripheral vascular disease*
5b. *Emergence of new or re-emergence of known communicable diseases*	*Reversal of life expectancy; increases in mortality*	*Globalisation – increase in disease diffusion by travel* *Increases in population density; poverty* *Concerns about vaccines* *Evolution of microbes, resistant to treatment*	*Viruses evolve and infectious diseases spread rapidly causing pandemics; e.g., MERS, SARS and Coronavirus (COVID-19)* *Re-emergence of known diseases, especially TB, measles, whooping cough, scarlet fever*

to experience a fifth stage of transition. Table 10.3 outlines two suggested ideas for this fifth stage.

Other suggestions for a stage 5 of epidemiological transition have been more positive. For example, the idea that the world will enter an 'Age of Healthy Living', based on good diet and regular exercise. In reality, these ideas are not mutually exclusive and it is likely that the world's population will experience a mixture of all these trends in the future.

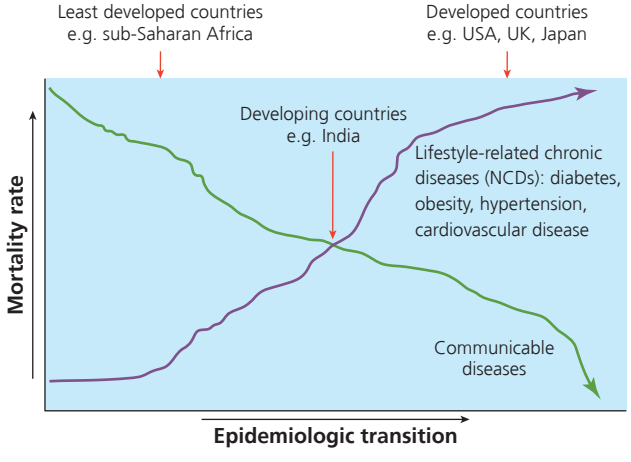


Figure 10.29 Line graph outlining epidemiological transition

The pace of transition

Omran discovered variations in the pattern, pace and the determinants of population change within the transition and so identified three types of model:

- **Classical/western model:** (for example, Western Europe) – a slow decline in death rates followed by lower fertility.
- **Accelerated model:** more rapid transition; falls in mortality take place over a shorter period of time (for example, many parts of Asia; some parts of Latin America).
- **Contemporary/delayed model:** recent decreases in mortality are not accompanied by a decline in fertility because infant and maternal mortality remain relatively high (for example, sub-Saharan Africa).

Evidence to support epidemiological transition

Epidemiological transition theory asserts that, over time, infectious diseases such as influenza, yellow fever, smallpox, malaria and tuberculosis (TB) are replaced by chronic and degenerative diseases such as heart disease, cancer and dementia as the more important causes of death.

Evidence to support the idea of transition can be seen in Figure 10.30, which shows a clear shift in the major causes of death in the USA between 1900 and 2010.

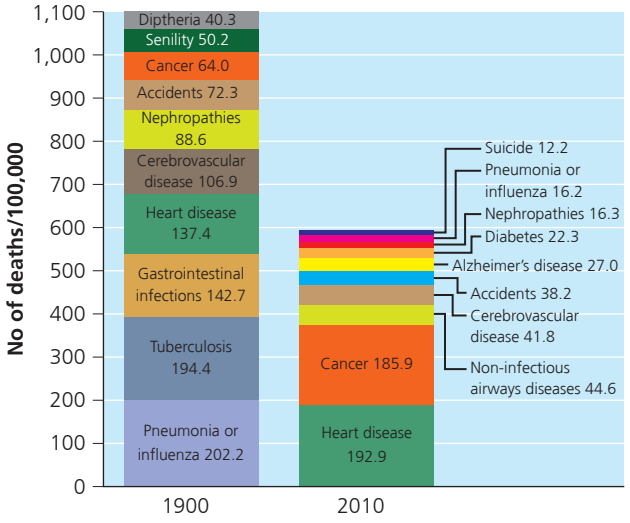


Figure 10.30 Epidemiological transition in the USA (deaths per 100,000 of population)

This shift in demographic and disease profiles is caused by factors including better healthcare, vaccination programmes, the use of antibiotics and improved public sanitation and hygiene. It should be acknowledged that some critics suggest that the degree of transition may be exaggerated; they argue that more chronic degenerative conditions existed in earlier times but were undiagnosed.

Table 10.4 shows differences in causes of death by rank between low-income countries and high-income countries and provides further evidence to support the concept of epidemiological transition.

According to WHO, in 2001 chronic, degenerative diseases accounted for 46 per cent of the global burden of disease. By 2020 that was expected to be 57 per cent. The proportion of degenerative disease is increasing rapidly in poorer countries. Epidemiological transition is well advanced, suggesting that public health policy in poor countries, with its traditional emphasis on infectious disease, will need to adapt. In a similar way, however, if the outbreak of global pandemics of infectious diseases, such as COVID-19 is to become more frequent, it means that high-income countries may have to adapt and redirect healthcare resources to tackle these outbreaks.

Environmental variables and their links to the incidence of disease

Climate

Links between the natural environment and disease tend to focus on climatic conditions. Drought leads to crop failure, reduction in food consumption and the potential for famine. At the other extreme, flooding caused by heavy rains or tropical storms can lead to water-borne disease and respiratory infections.

Seasonal affective disorder (SAD) is a type of depression that has a seasonal pattern. The episodes of depression tend to occur at the same time each year, usually during the winter, and have been linked to reduced exposure to sunlight during the shorter days of the year (Figure 10.31).

Table 10.4 Differences in the top six causes of death between low-income and high-income countries, 2016 (deaths in rank order, per 100,000 population)

Low-income countries		High-income countries	
Cause	**Deaths/100,000**	**Cause**	**Deaths/100,000**
Lower respiratory infections (e.g. pneumonia)	76	Ischaemic heart disease	146
Diarrhoeal diseases	58	Cancers (all major cancers)	96
Ischaemic heart disease	53	Stroke	63
HIV/AIDS	44	Alzheimer's disease	61
Strokes	42	Chronic obstructive pulmonary disease (COPD)	45
Malaria	38	Lower respiratory infections	36

Source: WHO

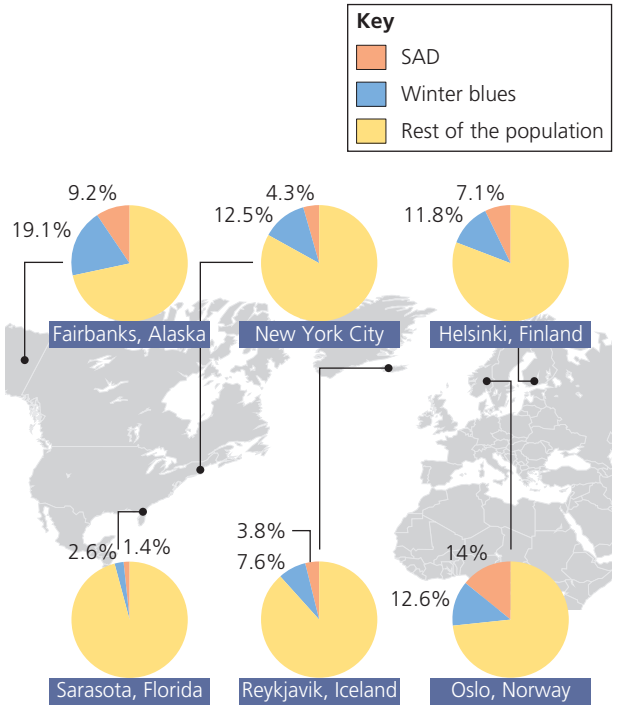


Figure 10.31 The changing incidence of seasonal affective disorder with latitude

Source: Norman Rosenthal

In the UK, common everyday ailments linked to the natural environment include allergies such as **hay fever** and **asthma**, where the body reacts to pollen, dust and other substances in the air during the spring and summer.

Climate and infectious diseases

Weather conditions, such as rainfall, flooding, heat waves and humidity have well-documented effects on infectious diseases.

- Prolonged periods of heavy rainfall increase the opportunities for vector-borne diseases to spread.
- Longer periods of mild temperatures increase the transmission potential of vector-borne diseases. Vectors become more infectious more quickly in warmer temperatures.
- Vector-borne viruses take too long to replicate if temperatures are too cold but if it is too hot, the lifespan of the vectors is decreased.
- Airborne respiratory diseases, such as influenza, thrive in temperate winters when temperatures are lower (but not too cold) and there are damper conditions.
- Some diseases are spread more readily in specific conditions. For example, coccidioidomycosis (Valley Fever – a chronic pulmonary disease) outbreaks are associated with 'grow and blow' conditions. Wet weather encourages growth of the fungus, *Coccidioides*, but subsequent dry weather increases transmission of the spores, which when inhaled cause the disease.
- Excessive rainfall can cause sewage system outflow, leading to water contamination and the outbreak of both viral (for example, norovirus) and bacterial (for example, cholera) infections.

Table 10.5 (page 492) gives three examples of infectious diseases in which climatic conditions are considered to be a contributory factor to their prevalence. Malaria is covered in more detail on pages 495–9.

Climate change is another consideration, particularly for vector-borne viruses such as Zika. As climatic conditions change, the spread of disease vectors into new regions is facilitated, placing new populations with limited immunity at risk.

Topography, drainage and health

In many low-income countries of Asia, high densities of people occupy the flood plains of major rivers (for example, the Ganges, the Yangtze and the Mekong). They are attracted by the flat land and the seasonal flooding that enables them to grow rice. Flooding, for them, is a good thing; that is, until they have exceptionally wet years. In 2018, unplanned flooding affected the greatest number of people of all environmental hazards and was responsible for 24 per cent of all deaths from natural disasters. Most floods occur in low-income countries and tropical regions where the impact on health is significant, the number of people displaced is often large, and the death toll is high. Floods contaminate freshwater supplies, increase the risk of water-borne diseases and create breeding grounds for disease-carrying insects such as mosquitoes. Increased rates of diarrhoea (including cholera and dysentery), respiratory infections, hepatitis A and E, typhoid and leptospirosis have been reported after flooding in developing areas.

Other environmental hazards

Other hazard events such as earthquakes can have a major impact on the incidence of disease, especially in low-income countries. The secondary environmental effects of earthquakes on public health will vary

Table 10.5 Examples of disease transmission associated with environmental factors

Transmission category	Example disease	Route/sources of transmission	Environmental factors	Other factors
Vector-borne	Zika virus (of the flavivirus family, similar to dengue, and West Nile virus)	Transmitted primarily through mosquitoes; sometimes through ticks or other insect bites	Tropical conditions of warm/hot humid conditions; the Pan-American Zika virus in 2015–16 spread during a long drought in Brazil (caused by an El Nino event) and spread northwards into the USA in summer 2016	Like other flaviviruses, it thrives in urbanised environments where there are more breeding sites for vectors in stagnant water in containers such as old oil drums, discarded tyres etc.
Airborne	Influenza	Respiratory – person to person via aerosol droplets	Clear seasonal cycle – relative humidity and pathogen survival Relationship with temperature less clear – appears to favour low-moderate temperature but also occurs in the tropics	An explanation for influenza's seasonal cycle is more indoor crowding in temperate climates during winter leading to greater transmission
Water-borne	Cholera	Oral – ingesting contaminated water or seafood	Excessive rainfall and water contamination, especially in marine and estuarine waters and when warm waters are brought to the coast (for example, by El Nino events)	Water treatment and management of faecal waste reduce its effectiveness to survive

depending on the magnitude of the earthquake, the nature of the built environment and the secondary physical effects (for example, tsunamis, landslides). Cholera, pneumonia and tetanus are the most critical post-earthquake infectious diseases. Risk factors lead to an increase in the incidence of disease because of:

- damage to health facilities and transport infrastructure – disrupting disease prevention and treatment services, medical supplies and availability of emergency equipment
- absence of health workers who may not be able to reach health facilities that are still functional
- aggravating weather conditions and lack of shelter
- prolonged exposure to large dust clouds generated by landslides or inhalation of tsunami water
- increased risk of communicable diseases in overcrowded emergency shelters
- increased exposure to disease vectors
- weakened immune systems of young children and the elderly
- lack of screening for blood-borne infections before blood transfusions or emergency surgery.

Air quality and health

Ambient (outdoor) air pollution is a major environmental health problem with the potential to affect everyone in both developed and developing countries. According to the WHO, air pollution is now the world's largest single environmental health threat. Terms such as 'airpocalypse' have been used to describe the increasing number of deaths and poor health associated with polluted air.

Air pollutants include **particulate matter (PM)**, photochemical smog and gases such as sulphur dioxide. Photochemical smog is a haze (rather than a dense smog) caused by the reaction of pollutants to sunlight and is comprised mainly of nitrogen oxides and ozone. Particulate matter includes solid particles and liquid droplets in the air. These are created by vehicle emissions, industrial processes (including energy generation), forest fires and burning waste. PM is measured by size at either less than 10 microns (PM_{10}) or less than 2.5 microns ($PM_{2.5}$) in diameter.

- Fine particulate matter ($PM_{2.5}$) has the greatest effect on human health. The smaller particles have a greater probability of being inhaled more deeply. If inhaled in large quantities, they are able to penetrate deep into the lungs causing infections, which can ultimately be fatal.
- Air pollution causes and exacerbates a number of diseases, ranging from asthma to cancer, pulmonary illnesses and heart disease. In 2017, 4.2 million deaths were caused by ambient air pollution.

- Worldwide, ambient air pollution is estimated to cause about 16 per cent of the lung cancer deaths, 25 per cent of chronic obstructive pulmonary disease (COPD) deaths, about 17 per cent of heart disease and stroke deaths and about 26 per cent of lower respiratory infection deaths.

People living in low- and middle-income countries are disproportionately affected. Globally, air pollution contributed to 9 per cent of deaths in 2017, ranging from a low of 2 per cent across high-income countries, to close to 15 per cent across many countries in south and east Asia. (India, Bangladesh, Nepal and China were the worst affected.) In the past decade, some rapidly growing African cities are showing very high air pollution levels as they start to industrialise, notably Luanda in Angola and Kampala in Uganda.

The level of ambient air pollution is measured by the concentration of particulate matter in micrograms per cubic metre. A Greenpeace report published in 2019 analysed air pollution readings from 3,000 cities around the world and found that 22 of the 30 most polluted were in India, with Delhi ranked as the most polluted capital (Figure 10.32). Every city in Africa and the Middle East measured in the report exceeded the WHO health guidelines for air pollution, as well as 99 per cent of cities in south Asia and 89 per cent in east Asia.

Most sources of outdoor air pollution are well beyond the control of individuals and demand action from authorities and policymakers (from city to international level). Policies and investments which support cleaner transport, power generation and industry, energy-efficient housing and better municipal waste management would reduce key sources of urban outdoor air pollution. Rural air pollution in developing regions can be tackled by reducing emissions from agricultural waste incineration, forest fires and certain agro-forestry activities (for example, charcoal production). By reducing air pollution levels, countries would reduce the burden of disease.

Figure 10.32 Children protecting their faces from the Delhi smog

In addition to outdoor air pollution, the use of fuelwood for indoor cooking and heating produces harmful pollutants which negatively affect the health of millions of people around the world. Exposure to smoke increases the risk of respiratory infections, lung cancer, cardiovascular disease and cataracts. In low-income countries, 2.5 billion rural households use solid fuels, with a high preference for fuelwood. In 2017, WHO reported that 3.8 million deaths were attributable to household air pollution. This amounts to 6.9 per cent of global mortality and almost all of it is in low- and middle-income countries. Women and children tend to be most at risk because they spend more time in fuelwood-burning environments. However, unlike ambient air pollution, the death rate caused by household air pollution is falling quite rapidly because of raised awareness of the problem. Government and NGO-led education programmes and the provision of alternative fuels have contributed to improved indoor pollution control.

Water quality and health

Water pollution and poor water quality have important impacts on health and can lead to many diseases. The WHO claims two million deaths annually are attributable to unsafe water and poor sanitation and hygiene.

Water-related diseases and morbidity include the following.

- Diarrhoeal diseases such as cholera are caused by bacteria and chemicals in water that people drink. Inadequate drinking water, sanitation and hygiene are estimated to cause 829,000 diarrhoeal disease deaths per year (WHO, 2017). Diarrhoea is a leading cause of malnutrition and the second leading cause of death in children under five years old. WHO reported a 60 per cent fall in cholera cases in 2018 as key endemic countries claimed gains in cholera control.

- *Schistosomiasis* is an acute and chronic disease caused by parasitic worms which have part of their lifecycle in water; people are infected during exposure to infested water. Over 240 million people are sufferers.
- Diseases like malaria have water-related vectors.
- *Legionellosis* (or Legionnaire's disease) is a respiratory disease caused by bacteria which infect the lungs and cause a severe form of pneumonia. The most common form of transmission is inhalation of tiny droplets (aerosols) of water that contains the bacteria that cause the infection. It is uncommon but is usually caught from the water supply of public buildings such as hotels, where the bacteria have collected in the water supply or air conditioning system. Infection can also occur by aspiration of contaminated water, particularly in susceptible hospital patients.
- Cancer and tooth/skeletal damage: millions are exposed to unsafe levels of naturally-occurring arsenic and fluoride.

Human sewage is one of the main pollutants of water. Toilets may be a simple hole in the ground (latrine) or a drop over a water source. The same water may then be used for drinking, washing and watering animals. WaterAid estimate that 785 million people live without ready access to safe drinking water and 485,000 children under five die every year from diarrhoea caused by unsafe water and poor sanitation.

Much of the burden of water-related diseases (water-related vector-borne diseases in particular) is caused by methods of water resource development and management. In parts of the world, attempts to manage water supplies through dam construction, irrigation development and flood control have led to adverse health impacts from water pollution causing significant preventable disease. For example, stagnant waters created by man-made reservoirs in sub-Saharan Africa create the perfect conditions for malaria-spreading *Anopheles* mosquitoes to breed and multiply. A recent study found that cases of malaria in high risk areas across the region are associated with people living in the vicinity of a reservoir. Even attempts at more sustainable strategies such as increasing the use of wastewater in agriculture (grey water), which is important for livelihood opportunities, are also associated with serious public health risks, as in areas without separate sewage systems the wastewater can contain faecal matter (black water).

Most of the diseases and health problems related to water quality are preventable and relatively affordable and give plentiful opportunities to improve health and reduce disease, even in the world's poorest countries. In 2017, 2.1 per cent (1.2 million) of all global deaths were caused by diseases relating to an unsafe water supply, though this was as much as 6 per cent in low-income countries. Much progress has been made because of the focus provided by the Millennium Development Goals (MDGs). The number of children dying from diarrhoeal diseases has fallen steadily and between 2000 and 2017, 1.6 billion people gained access to an improved drinking water supply. Practical measures have included:

- better tools and procedures to improve and protect drinking water quality at the community and urban level, for example, through Water Safety Plans (including better education and awareness to reduce faecal contamination and disinfecting supplies with chlorine)
- availability of simple and inexpensive approaches to treat and safely store water at the household-level (for example, boiling and covered storage).

As the world focuses its attention on the Sustainable Development Goals (SDGs), controlling human exposure to waterborne pathogens associated with faecal waste remains a priority in attaining the goal of safe drinking water and sanitation for all (SDG 6).

There are many other contaminants which are detrimental to human health. Oil spills contaminate land and water, leading to digestive problems and diarrhoea. In the Niger Delta (see the case study in Chapter 11, page 631), oil spills, waste dumping and gas flaring are endemic and this type of pollution has damaged the soil, water and air quality for decades. Hundreds of thousands of people have been affected, particularly the poorest and those who rely on traditional livelihoods such as fishing and agriculture.

Malaria: a biologically transmitted disease

Malaria is a tropical **vector-borne disease**, biologically transmitted by insects, which has a devastating impact on people's health and livelihoods around the world.

Malaria is caused by a unicellular protozoan parasite from the *Plasmodium* family. The parasite is transmitted through the bite of a female Anopheles mosquito. If the mosquito bites a person, the parasite may be injected into the person's blood stream. The parasites multiply in the liver, and then infect red blood cells. Symptoms of malaria include fever, headache and vomiting, and usually appear between 10 and 15 days after the mosquito bite. If not treated, malaria can quickly become life threatening by disrupting the blood supply to vital organs. It can either kill a person directly or weaken the immune system so that other diseases affect the sufferer.

Global prevalence and distribution

It is important to acknowledge that just a few generations ago malaria was common in many more places around the world than it is today. Over the course of the twentieth century, the disease was eliminated in many populous regions of the world, including many parts of the USA and Europe, saving the lives of millions.

According to WHO, in 2018:

- 3.6 billion people (nearly half the world's population) in 91 countries lived in areas at risk of malaria transmission
- an estimated 228 million cases occurred, killing around 405,000 people
- children under five years of age are the most vulnerable group; they accounted for 67 per cent (272,000) of all malaria deaths worldwide.

As Figure 10.33 shows, sub-Saharan Africa carries a disproportionately high share of the global malaria burden. In 2018:

- the region was home to 93 per cent of malaria cases and 94 per cent of malaria deaths
- 19 countries in sub-Saharan Africa plus India carried almost 85 per cent of the global malaria burden
- six countries accounted for more than half of all malaria cases worldwide – Nigeria (25 per cent), the Democratic Republic of the Congo (12 per cent), Uganda (5 per cent), and Côte d'Ivoire, Mozambique and Niger (4 per cent each).

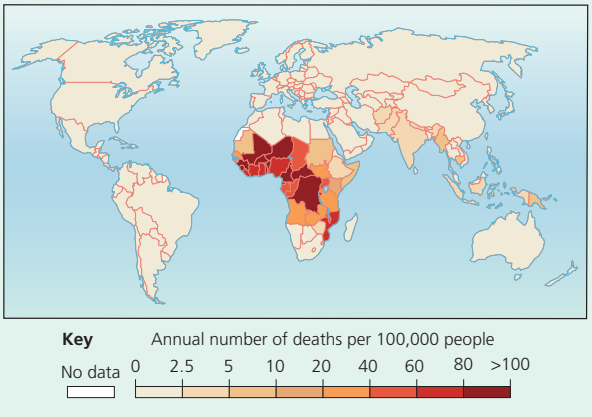


Figure 10.33 Death rate from malaria, 2017

Source: Our World in Data

Links to physical environment and seasonal incidence

Malaria is strongly linked to the natural environment. Mosquitoes breed in stagnant water and transmission is greatest in areas during and just after the rainy season. Figure 10.34 (page 496) shows how, in Africa, the length of the malaria transmission season is directly related to the length of the rainy season. At altitudes above 1,500 m and where rainfall is below 1,000 mm, malaria transmissions fall. A high temperature is also important. The parasites require temperatures of between 16°C and 32°C to develop inside the mosquito and this is one reason why the disease is largely concentrated in the tropics and subtropics. Climate change will gradually alter the pattern of the transmission season in all areas over time. Areas receiving more flood events may show higher incidence of transmission over a longer period. Those regions that become more prone to drought may have shorter transmission seasons but will, of course, have other problems to contend with.

On a smaller scale, coastal areas have much less seasonal variation in temperature, are lower in altitude and generally have higher relative humidity. Providing the temperatures lie within the transmission vector's range of tolerance, coastal areas will often show a higher prevalence of the disease.

Land use is another important factor which can determine malaria risk. Studies in India have shown that those living in proximity to forested areas (again with higher humidity) are more susceptible to infection.

Links to socio-economic environment

While environmental variables are the overriding determinant of malarial mosquito presence, socio-economic factors will also influence the level of risk of contracting the disease. Malaria is often referred to as a disease of poverty and when the burden is measured as disability adjusted life years (DALYs), nearly two-thirds of the total global burden due to malaria is concentrated among the poorest 20 per cent of the world's population.

Malaria is a common cause of morbidity in many countries but is concentrated in the world's poorest continents because of the lack of investment in projects for developing new drugs and vaccines, and the limited

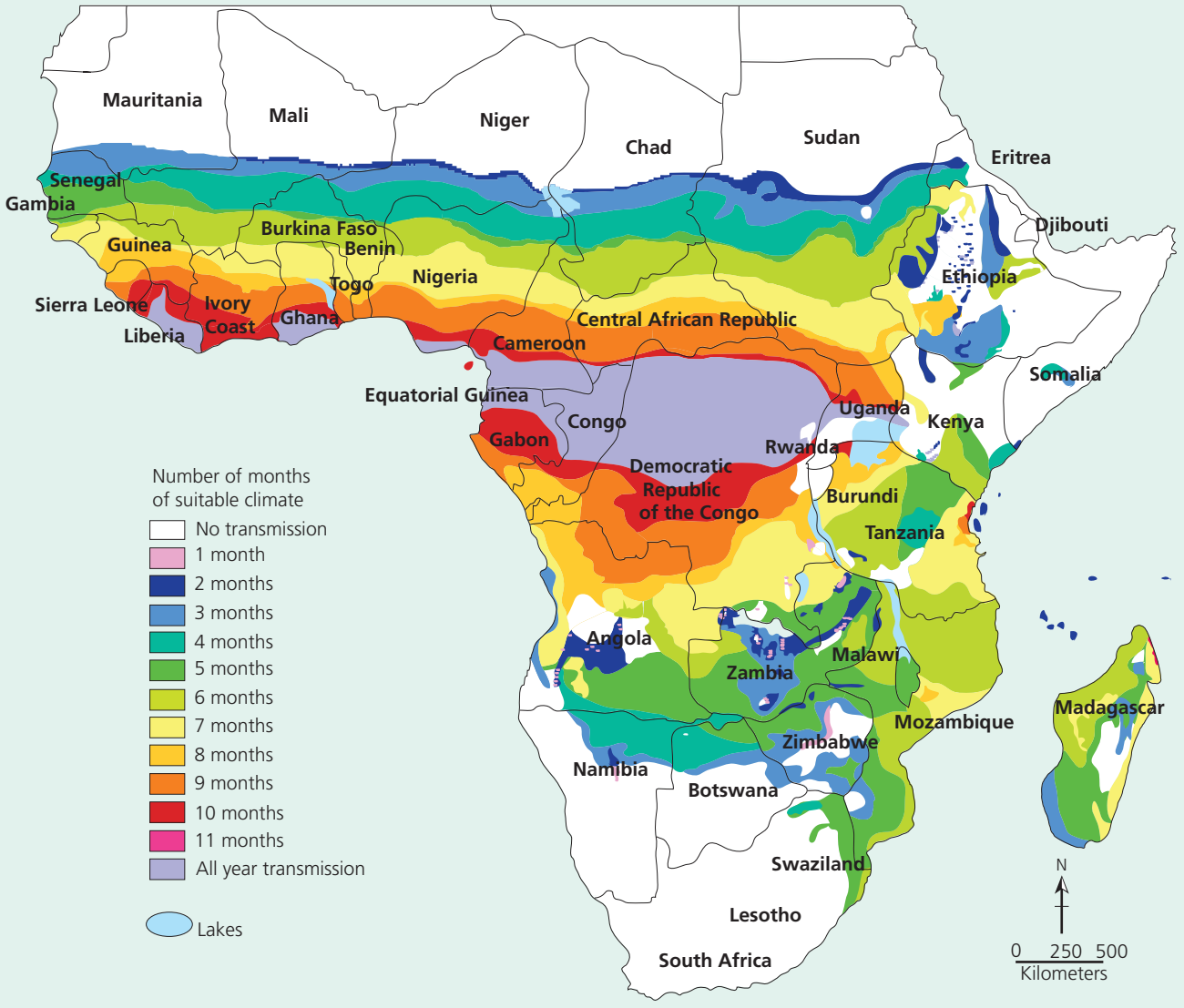


Figure 10.34 The duration of the malaria transmission season

Source: MARA/ARMA (Mapping Malaria Risk in Africa/Atlas du Risque de la Malaria en Afrique)

number of existing programmes to control the disease. Studies undertaken in Africa, Asia and South America have identified a number of prevailing socio-economic variables that increase the incidence of malarial infection:

- **Poor housing quality:** Homes with earth/sand floors using materials such as mud, bamboo cane or wooden trunks for walls and grass or palm leaves to provide a roof; those with poorly-fitted windows or doors, or windows without glass, screens, curtains or shutters.
- **High density occupancy:** Densely clustered built-up areas and overcrowded rooms (for sleeping) is another important factor in increasing risk. High malaria incidence is associated with high infant mortality and subsequent high fertility rates. A consequence of this is that children (and sometimes parents) often share a room with at least five occupants.
- **Unsanitary conditions:** Studies of malaria 'hotspots' in Chennai identified that although individual houses were clean, surrounding areas were dirty and polluted by rubbish and waste outflows. Spitting and open defecation are also commonly associated with attracting mosquitoes.
- **Occupation:** Agricultural workers are more exposed to the mosquito vector. This particularly applies to those living near to stores of irrigation water and to migrant workers who may sleep outdoors at night. For example, studies in Ethiopia have shown

that highland farmers, who seasonally migrate to find work in lowland areas, are prone to malarial infection.

- **Rural vs urban environments:** Generally, those in rural areas are seen to be more at risk, however contamination rates in urban slums and squatter settlements are high due to building density, unsanitary conditions and stagnant water being retained in rubbish.
- **Age and gender:** These are not significantly associated with the risk of infection, though children aged under 5 are more likely to suffer the extreme consequences of the disease. However, studies in both The Gambia and Tanzania indicate a shift in malaria burden from the under-fives to older children aged 5 to 14 years. This is thought to be as a result of the focus on prevention in the younger age groups through better education, awareness and provision of nets.

It is important to distinguish between the overall risk of infection, determined by the mix of physical and socio-economic environmental factors listed above, and the utilisation of prevention and treatment methods. These have a much clearer association with poverty and socio-economic status such as the level of income and educational attainment as well as accessibility to intervention methods. There is more consistent evidence that vulnerability to the consequences of malaria is higher among lower socio-economic groups.

- **Income:** There is a strong positive correlation between income and the use of prevention methods. Those with higher incomes spend more on repellents, insecticide-treated nets (ITNs) and mosquito coils, to reduce their risk of infection. Higher income is associated with better nourishment and studies show that cerebral malaria is less common in well-nourished children. Again, this disproportionately affects agricultural workers, many of whom receive seasonal incomes at harvest time and cannot afford nets or treatment during the main malarial season.
- **Education:** Those with a clearer understanding of the link between malaria and surrounding environmental conditions, including hygiene and sanitation, are more likely to use prevention strategies. For example, studies in Malawi found that net ownership was largely absent in homes where the head of the household had not completed primary education.
- **Distance and accessibility**: Greater distance to the nearest clinics or hospitals is associated with fewer seeking treatment for symptoms and less expenditure on prevention methods. For example, repellents and coils are less available and so used less in rural areas; instead there is a reliance on burning medicinal plants as a repellent.

Most of the variables listed are proxies for poverty and evidence suggests that malaria infection, and particularly death resulting from malaria, is strongly associated with low socio-economic status. It is children from the poorest income groups living in poor housing who are most at risk.

The impacts of malaria

Well-being can be defined as 'the state of being comfortable, healthy or happy'. Clearly if an individual is ill then they cannot be any of those things. If groups of people are ill then it has an effect on the economic well-being of a community.

Impacts on health and well-being

Malaria kills a child somewhere in the world every two minutes. Ninety-four per cent of malaria deaths occur in Africa, where malaria accounts for about one in six of all childhood deaths.

Malaria causes the patient to develop a high fever, which comes and goes. The pattern of fevers varies according to the species of malaria. Initially, it feels like the flu with high fever, fatigue, and body aches, with hot and cold stages. Signs and symptoms in children may be nonspecific, leading to delays in diagnosis. People also may have headache, nausea, shaking chills (rigors), sweating and weakness. Anaemia is common in patients with malaria, in part due to the effects of the *Plasmodium* parasite on the red cells.

Plasmodium falciparum causes a particularly severe form of malaria. In addition to fever, patients may experience complications such as destruction of red blood cells, yellow skin discoloration, kidney failure, fluid in the lungs, cerebral malaria, convulsions, coma or death.

People who have lived for years in areas with malaria may develop a partial immunity to new infections.

The disease also contributes greatly to anaemia among women and children. Infection during pregnancy is associated with severe anaemia and other illness in the mother and contributes to low birth weight among newborn babies – one of the leading risk factors for infant mortality and causes of sub-optimal growth and development.

Impacts on economic well-being

Malaria has serious and far-reaching impacts, slowing economic growth and development and prolonging the vicious cycle of poverty. While it is true that poverty and

lack of development may be key causes of the presence of malaria, there is a strong argument that malaria causes underdevelopment.

- **Costs to individuals and families** – travel to and treatment at clinics; purchase of drugs to treat malaria; lost income from absence from work; reduction in crop production; absence from school; purchase of preventative measures; burial expenses in cases of death.
- **Costs to government** – building, staffing and maintaining healthcare facilities; purchase of drugs and of public health intervention equipment such as insecticides and nets; loss of production and revenue from taxes during absences; loss of tourism revenue.
- It has been estimated that in many parts of west and central Africa, malaria accounts for up to 40 per cent of public health expenditures, 30–50 per cent of inpatient hospital admissions and 60 per cent of outpatient health clinic visits.
- **Direct and indirect costs** – the direct costs of malaria resulting from the provision of healthcare are thought to be around US$15–18 billion each year globally. In addition, there are many indirect costs resulting from lost economic growth and lost productivity, as well as the lack of foreign investment in industry, commerce or tourism.

Management and mitigation strategies

Malaria **elimination** is defined as the interruption of local transmission of the Plasmodium parasite species in a defined geographical area as a result of deliberate intervention (WHO). **Eradication**, however, is defined as the permanent reduction to zero of the worldwide incidence of malaria infection. Globally, the elimination net is widening, with more countries moving towards the goal of zero cases of malaria each year.

Vector control

This is the main way to prevent and reduce malaria transmission. When malaria was more prevalent and widespread globally, many countries eliminated the disease by removing mosquito habitat and breeding sites. This was done by changing land use, for example, by draining swamplands containing stagnant water, or by regular, extensive fumigation of areas using insecticides such as DDT. Fumigation using DDT is still carried out in many tropical and sub-tropical countries and can be effective in reducing mosquito numbers (Figure 10.35).

Figure 10.35 The use of DDT fumigation to prevent malaria in Sri Lanka

Both of these methods have environmental consequences; changing habitat impacts on other species and the application of insecticides kills honey bees and other pollinators, as well as affecting human health. For example, it is thought that DDT can be ingested and passed through the food chain. WHO recommends the following methods of vector control:

- **Insecticide-treated mosquito bed nets (ITN)** – sleeping under an ITN is seen as a particularly effective low-cost strategy as it provides a physical barrier and an insecticide effect (see Figure 10.36). The insecticide woven into the net kills mosquitoes so they cannot go on to bite others who may not be protected. Bed nets can reduce transmission by as much as 90 per cent in areas with high coverage. In 2018, about half of all people at risk of malaria in Africa were protected by ITNs, though this proportion has not increased.

Figure 10.36 Insecticide-treated bed nets are seen as a cost-effective strategy in the fight against malaria

- **Indoor residual spraying (IRS)** – spraying indoor surfaces with pyrethroid insecticides can also reduce malaria transmission. To gain significant community protection, IRS should be implemented at a high level of coverage, but usage has declined across most WHO regions. This is partly because there is evidence of Anopheles mosquitoes developing an increasing resistance to pyrethroid insecticides. Many countries are switching to more expensive alternatives to mitigate mosquito resistance to pyrethroids.
- **Mosquito coils** – mosquito coils are used in some parts of the world. They are burned to emit a repellent smoke that controls mosquito invasion, but studies have shown they are only between 24 and 64 per cent effective. Their use has also been discouraged by WHO because they emit other pollutants such as sulphur dioxide, nitrogen dioxide and volatile organic compounds (VOCs).

Medicine

- **Anti-malarial drugs** – some medicines can be used to prevent malaria by suppressing the blood stage of the infection. WHO has recommended intermittent preventative treatment for pregnant women and infants living in moderate to high transmission areas and additional seasonal doses of medicine in areas of sub-Saharan Africa. Unfortunately, *Plasmodium* resistance to anti-malarial drugs, such as chloroquine, is a recurring problem.
- **Vaccines** – until the last few years, there were no vaccines available to prevent malaria. The first and only one to be developed that shows that it can significantly reduce malaria in young African children is RTS,S/AS01 (RTS,S), sold under the trade name Mosquirix, which acts against *Plasmodium falciparum*. In large scale clinical trials, the vaccine prevented approximately four in ten cases of malaria over a four-year period.

Treatment

Early diagnosis and treatment of malaria reduces disease, contributes to reducing transmission and prevents deaths. The most effective treatment for malaria sufferers is artemisinin-based combination therapy (ACT).

Management strategies

To follow up the progress made by the Roll Back Malaria programme (part of the MDGs), WHO has launched a global technical strategy for malaria from 2016–2030 as part of the SDGs. The strategy sets ambitious but achievable global targets, including:

- reducing both malaria case incidence and mortality rates by 90 per cent
- eliminating malaria in at least 35 countries by 2030
- preventing a resurgence of malaria in all countries that are malaria free.

The Global Malaria Programme

The programme co-ordinates WHO's global efforts to control and eliminate malaria. The programme is supported by an Advisory Committee (MPAC) that is made up of a panel of malaria experts.

High burden, high impact approach

In 2018, WHO called for an aggressive new approach to make more progress in tackling malaria. This new approach is being driven by 11 countries that carry a high burden of the disease (including Burkina Faso, Cameroon, DRC, Ghana, India, Mali, Mozambique, Niger, Nigeria, Uganda and Tanzania). The approach builds on the principle that no one should die from a disease that can be diagnosed and prevented. Key elements include:

- political will to reduce malaria burden
- measures taken based on strategic information
- better guidance and a co-ordinated national response.

Coronary heart disease (CHD) – a non-communicable disease

Coronary heart disease CHD (also known as ischaemic heart disease) is the leading cause of death worldwide, accounting for around 9.5 million deaths annually. CHD is one part of a group of related diseases known as cardiovascular diseases (CVD), which includes cerebrovascular disease (strokes) and vascular dementia. CHD rates vary considerably among populations and the risk of heart attacks within populations from across the world can be ascribed to the varying levels of a number of risk factors among individuals. A heart attack occurs when the blood vessels supplying the heart muscle become blocked, starving it of oxygen and leading to heart muscle failure or death. A wide range of risk factors can be responsible for a heart attack, often acting in combination. Many of these factors are significant in all populations but their incidence varies, as does the occurrence of the disease globally.

Global distribution

The global distribution suggests the highest mortality rate from cardiovascular diseases occurs in Eastern Europe, northern and Central Asia, parts of north Africa and parts of South East Asia (Figure 10.37, page 500).

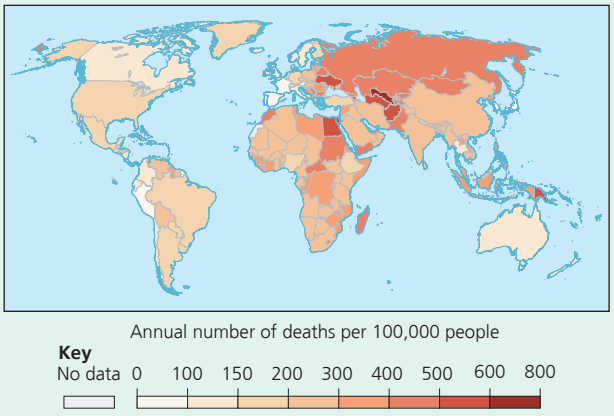


Figure 10.37 Global distribution of death rate from cardiovascular diseases (including CHD) in 2017

Links to physical environment

Air quality: In a historical context, the changing physical environment has been a contributory factor to increases in heart disease. Industrialisation has brought a number of negative effects. Studies in the US have shown that increased exposure to airborne pollutants, especially particulate matter (PM) emanating from industry and transport systems, increases the risks of CHD. This suggests that rural environments, relatively clean of air pollution, would indicate a lower risk, though there is no clear evidence that this is the case in the developed world.

Climate: Studies have shown that temperature is significantly associated with CHD mortality rates. Results suggest that there is an optimum range of temperatures at which CHD mortality is lowest. Outside of this optimum range, CHD mortality increases as temperatures fall or rise. Extremes of cold or heat (beyond what the individual body is accustomed to) puts additional strain on the cardiovascular system.

The optimum temperature range varies regionally depending on the prevalent temperature conditions that the population of the region is acclimatised to. For example, the optimum temperature range for low CHD mortality rates in northern Finland is lower (14.3 to 17.3°C) than in London (19.3 to 22.3°C). Climate is important in that cold and damp winters in temperate climates have a negative impact on the cardiorespiratory system, which increases the strain on the heart and the subsequent risk of heart attack.

Relief and topography: There are no clear links between landscape and CHD risk other than the notion that more challenging relief requires more physical effort when walking, which can be both advantageous in increasing exercise and activity (which reduces risk) but can be a threat for individuals with other underlying risks.

Links between CHD and the physical environment, with the possible exception of climate, are limited and minimal. Lifestyle choices within a variety of physical environments are more important risk factors. For example, it could be argued that the Japanese diet of fish (which reduces CHD risk) is influenced by Japan's physical environment as much as it is by culture. However, studies have shown that societal factors are influential. For example, Japanese migrants to the USA have, over generations, adopted American values, diets, lifestyle habits and ultimately higher CHD rates.

Links to socio-economic environment

In developed countries, socio-economic factors and lifestyle choices are much more important determinants of CHD prevalence. It is estimated that 75–85 per cent of people dying from CHD have one or more major risk factors that are influenced by negative lifestyles, such as poor diet, physical inactivity or smoking.

There is an increasing risk of CHD with age but gender is a less important determinant, despite it traditionally being thought of as a disease of men. In the UK, one in seven men and one in twelve women die from coronary heart disease. CHD is multifactorial and it is the interaction between some of the genetic, lifestyle and other social factors which increases the risk:

- **Social deprivation:** There is a positive correlation between deaths from circulatory diseases and deprivation. Some studies have looked at the general increasing deprivation in the northern regions of the UK compared with southern regions. Premature deaths (under 75) from CHD are most common in the north of England, central Scotland and the south of Wales, and lowest in the south of England. For example, the premature death rate caused by CHD in Glasgow (138 per 100,000) is over three times higher than that for Hart in Hampshire (39 per 100,000).
- **Tobacco use:** Up to 20,000 CHD deaths each year can be attributed to smoking as it lowers 'good' cholesterol (HDL) and makes blood more sticky and more likely to clot, which can block blood flow to the heart and brain.
- **Alcohol use:** Excessive alcohol consumption contributes to raised blood sugar, increased fats and bad cholesterol in the blood, leading to high blood pressure – one of the most important risk factors for a heart attack. Some studies have shown that moderate consumption can reduce risk by raising the level of HDL cholesterol.
- **High blood pressure:** Around 47 per cent of heart attacks worldwide are attributable to hypertension. Consistently high blood pressure means that the heart has to work harder and the extra strain causes the coronary arteries to slowly become narrowed from a build-up of fat, bad (non-HDL) cholesterol and other substances, known as plaque. This slow process is known as atherosclerosis.

- **Poor nutrition:** A WHO report stated that a diet high in saturated fat, sodium and sugar and low in complex carbohydrates, fruit and vegetables increases CHD risk.
- **Overweight and obesity:** Obesity is an independent risk factor for CHD but is also associated with other risk factors such as high blood pressure, high cholesterol and diabetes.
- **Diabetes:** Men with Type 2 diabetes have two to four times greater risk of CHD; women have three to five times greater risk. High blood glucose levels from diabetes can damage the blood vessels and the nerves that control the heart.
- **Infrequent exercise:** Physical activity reduces risk; it is estimated that approximately 35 per cent of CHD mortality in the US is due to physical inactivity. Even for people with no other risk factors, sedentary lifestyles can lead to an increase in the likelihood of developing CHD or other risk factors, such as obesity and high blood pressure.
- **Ethnicity:** Racial and ethnic background seem to influence heart disease risk, but the picture is complex and lifestyle habits play a bigger role. For example, in the UK, south Asian people are more likely to develop heart disease and have a higher premature death rate from CHD than white Europeans.

  In the US, African Americans are at greater risk. It is thought that a genetic predisposition to salt sensitivity results in African Americans being more susceptible to high blood pressure. Hispanics generally display higher rates of obesity, diabetes and other CHD risk factors but have lower rates of heart disease and fewer deaths from CHD. There is no clear reason for this but it is thought that it is a mixture of genetics and lifestyle choices.
- **Family history**: There are inherited genetic conditions which can be passed on through families, affect people of any age and may be life-threatening. First degree relatives of patients with premature CHD are at increased risk of developing the disease themselves.

In developing countries with low mortality, such as China, the same risk factors apply, with the additional risks of under-nutrition and communicable diseases. In developing countries with high mortality, such as those of sub-Saharan Africa, low vegetable and fruit intake are also important factors. Most of the risks identified above are modifiable in that they can be prevented, treated and controlled. There are considerable health benefits at all ages, for both men and women, in stopping smoking, reducing cholesterol levels and blood pressure, eating a healthy diet and increasing physical activity.

Urbanisation and cardiovascular diseases

According to the World Heart Federation, one of the major factors increasing the risk of CHD in developing societies is the rapid rate of urbanisation taking place. They cite the following negative impacts of urbanisation, which increase risk, particularly for children growing up in fast-developing cities:

- areas with insubstantial housing conditions and poor access to healthcare services, healthy foods and safe, green places that are free of environmental toxins and pollutants for outdoor activity
- crowded living environments can spread diseases such as rheumatic fever, which if left untreated can cause rheumatic heart disease
- city dwellers are more likely to be exposed to marketing and advertising for unhealthy foods, tobacco and alcohol
- higher levels of particulate matter air pollution
- urban environments may discourage physical activity and encourage sedentary habits
- individuals use more prepared and heavily processed convenience foods that are often high in sugar, salt and saturated fats
- city dwellers are more frequently tobacco users than rural dwellers
- children in cities may be particularly susceptible to second-hand smoke given the number of smokers in urban areas, along with crowded living conditions.

Impacts of CHD

Impacts on health and well-being

The most common symptom of CHD is angina; a low level but fairly constant chest pain which can spread to other parts of the upper body. The pain can increase when the heart is put under more stress, for example, from physical activity. These symptoms can be relieved with the use of nitrate tablets or sprays, beta blockers (to slow heart beat), calcium channel blockers (to relax arteries), aspirin or other clot-preventing drugs.

As well as angina, the other main symptoms of CHD are heart attacks and heart failure. Heart attacks can permanently damage the heart muscle and, if not treated straight away, can be fatal. Heart failure can also occur when the heart becomes too weak to pump blood around the body. This can cause fluid to build up in the lungs, making it increasingly difficult to breathe.

It is possible to recover from heart attacks and lead a normal life. Those who are at high risk or have suffered a heart attack can undergo heart bypass surgery to reduce the risk, but this in itself can be stressful and risky.

Rehabilitation programmes for those recovering from heart attacks or surgery focus on exercise; education about lifestyle choices; and relaxation and emotional support to build confidence. Those suffering from heart attacks can lose confidence in undertaking physical

activity. CHD sufferers at all stages of the disease may be required to take continued medication for the rest of their lives. This may include warfarin (an anticoagulant which 'thins' the blood) which may have subsequent side-effects. In some cases, electronic devices are fitted to patients to regulate heart beat and blood flow.

Impacts on economic well-being

The economic impacts of heart disease include:

- cost of healthcare to the individual and their family
- loss of income due to time off work
- cost to the government of providing healthcare, treatment and medication
- cost to the economy because of lost productivity.

These are sometimes difficult to quantify but estimates suggest:

- In the UK, healthcare costs relating to heart and circulatory diseases (all cardiovascular diseases) were £9 billion in 2018.
- The cost to the UK economy (including premature death, disability and indirect costs) is around £19 billion per year.
- Heart disease costs the US economy around $220 billion each year (including cost of healthcare, medicines and lost production).

Management and mitigation strategies

WHO have identified two types of cost-effective interventions that can be implemented to prevent and control CVDs. They recommend that these are used in combination to reduce the disease burden:

- '**Population-wide**' interventions that affect the whole population and are aimed at primary prevention.
- **Individual level** interventions targeted at those who are at high total CVD risk and treatment of symptoms for those with established disease.

At the national level, governments have a responsibility to protect the health of citizens. They use a combination of approaches to implement these measures, including:

- providing health education for the public to raise awareness of the causes of CHD and promote healthy lifestyles
- introducing policy and legislation to discourage certain activities and lifestyles that lead to poor health or that encourage activities which improve health
- providing affordable medication and healthcare to control the symptoms or treat the disease.

Health education

Interventions are only effective if they have public support and understanding. Health education is essential to promote healthy choices. Schools are an ideal starting point for many health campaigns as implanting the message in a receptive learning environment may have a more lasting effect. The World Heart Federation (supported by WHO) has initiated a number of activities to assist schools around the world and co-ordinate World Heart Day events and activities on 29 September each year (see Figure 10.38). Activities and events include:

- blood pressure testing
- encouraging physical exercise
- scientific conferences
- promoting a heart-healthy diet.

Figure 10.38 Pupils at a primary school in Lianyungang, China studying the heart in preparation for World Heart Day

In the UK, dieticians promote the benefits of eating oily fish, fruit and vegetables and less saturated fats for a heart-healthy diet.

Policy and legislation

Only governments can legislate for the prevention and control of disease and WHO recommends the following population-wide policies:

- **Tobacco** – comprehensive smoking control policies including taxation, advertising bans, designating smoke-free areas and bans in public places, enforcing health warnings on packets.
- **Foods** – taxation on foods that are high in saturated fat, sugar and salt; requiring food processing companies to label food products by levels of fat, sugar and salt and so on.
- **Alcohol** – taxation and other measures to reduce harmful use.
- **School meals** – ensuring children receive healthy balanced meals.
- **Physical activity** – providing cycle paths, walking routes and other incentives to increase activity.

Medical treatment and secondary prevention

Many different medicines are used as secondary prevention to treat CHD. Usually they aim to reduce blood pressure or widen the arteries. For example:

- **blood thinning medicines** – reduce heart attack risk by preventing the blood clotting, for example, aspirin
- **statins** – reduce cholesterol levels by blocking its formation
- **beta-blockers** – block the effects of a particular hormone in the body, which then slows down the heartbeat and improves blood flow
- **nitrates** – widen the blood vessels
- **angiotensin-converting enzymes (ACE) inhibitors** – treat high blood pressure, for example, ramipril
- **calcium channel blockers** – reduce blood pressure by relaxing the muscles that make up the artery walls, causing the arteries to become wider, for example, amlodipine
- **diuretics** – flush excess water and salt from the body.

If the blood vessels are narrow due to a build-up of fatty deposits or if symptoms cannot be controlled, then costly surgical procedures may be required such as a coronary heart by-pass, valve repair and replacement or balloon angioplasty, where a small balloon-like device is threaded through an artery to open the blockage. Medical devices, such as pacemakers or prosthetic valves, are fitted to treat some types of CHD. In the most extreme cases, a heart transplant may be necessary.

Role of international agencies and NGOs in promoting health and combating disease

International agencies

Leading the effort within the UN system to promote and protect good health worldwide is the World Health Organization (WHO), which was established in 1948.

The World Health Organization

The WHO constitution states that, 'Health is a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity.' Its primary role is to direct and co-ordinate international health within the UN system. WHO's underlying objective is the attainment of the highest possible level of health for all people. At the outset, its top priorities were to:

- strive to combat communicable diseases including influenza, malaria, tuberculosis (TB) and venereal diseases
- address mothers' and children's health to help them survive to be healthy
- improve nutritional levels and environmental sanitation.

These remain on WHO's agenda, which now also includes providing advice and support in reducing non-communicable diseases such as cancer and CVDs. More recently, the organisation has had to turn its attention to newer communicable diseases including HIV/AIDs and Ebola, and since 2020, COVID-19.

As of 2020, WHO has 194 member states and 6 regional offices. Key facts about the WHO and their main areas of work:

- responsibility for the International Classification of Diseases, which has become the worldwide standard for clinical and epidemiological purposes
- advising national ministries of health on technical issues and providing assistance on health systems and care services
- advising on the prevention and treatment of both communicable and non-communicable diseases
- working with other UN agencies, NGOs and other partners on international health issues and crises. For example, ensuring the safety of the air that people breathe, the food they eat and water they drink, as well as the medicines and vaccines they need.

The WHO has been criticised for being over-bureaucratic and lacking the practical 'front-line' application to health issues. However the organisation's ability to focus, promote and co-ordinate efforts to tackle health problems on an international scale, using international experts and collating the most recent research, has achieved undeniable progress and success. Most notably:

- In the 1970s, WHO efforts led to the eradication of smallpox – the only major infectious disease that has been completely eradicated (probably WHO's greatest achievement).

- In 1988, the WHO launched its Global Polio Eradication Initiative and by 2006 the number of cases was reduced by more than 99 per cent.
- It has given special attention to adapting global HIV/AIDS policies to fit the specific needs of different regions. For example, using a different approach in sub-Saharan Africa, where the epidemic is largely spread through heterosexual sex, compared with Eastern Europe, where injecting drug use is the primary mode of transmission.
- It worked in partnership with other agencies to help achieve some of the health-related Millennium Development Goals (MDGs) by 2015.
- From 2010–2015, WHO's Strategy for Women's and Children's Health successfully mobilised global resources under the 'Every Woman, Every Child' movement to save the lives of an estimated 16 million women and children. To build on this success, the Global Strategy for Women's, Children's and Adolescents' Health was launched in 2016 to end preventable deaths among all women, children and adolescents.

WHO interventions cover all areas of the global healthcare spectrum, including crisis intervention; response to humanitarian emergencies; officially declaring pandemics; establishing international health regulations, which countries must follow to identify disease outbreaks and stop them from spreading; and the prevention of chronic diseases.

Other international organisations

It would be misleading to suggest that the entire work of supporting global health rests with the WHO. Other UN bodies are also engaged in this critical task, for example, the Joint United Nations Programme on HIV/AIDS (UNAIDS); the United Nations Population Fund (UNFPA) in support of reproductive, adolescent and maternal health; and the health-related activities of the United Nations Children's Fund (UNICEF).

- **UNAIDs**: an innovative partnership that leads and inspires the world in achieving universal access to HIV prevention, treatment, care and support. It was formed in 1994 as the main advocate for accelerated, comprehensive and co-ordinated global action to combat the HIV/AIDS pandemic.
- **World Bank:** A vital source of financial and technical assistance to developing countries around the world to fight poverty, including making investments to improve health.

UNICEF

UNICEF works with a number of other agencies, including NGOs, to establish and develop work on the SDGs. It is partly responsible for eight of the SDGs and custodian for 19 SDG indicators. UNICEF has a number of priorities:

- reducing child mortality by aiming to reach vulnerable children everywhere
- working to end preventable maternal, newborn and child deaths by scaling up immunisation programmes
- supporting and promoting curative services for pneumonia, diarrhoea and malaria
- tackling health emergencies in places affected by conflicts, natural disaster, migration or political or economic instability, to ensure children are secure
- helping to develop resilient health systems that can withstand crises.

Non-governmental organisations (NGOs)

A non-governmental organisation (NGO) is any non-profit association that operates independently of both government and of profitable businesses. They are organised on a local, national or international level and are primarily funded by public donations. NGOs have a long history of involvement in the promotion of human well-being. They have increasingly been promoted as alternative healthcare providers to the state, especially in developing countries; governments cannot always cope with the enormity of some health issues such as HIV/AIDs, malaria or viral outbreaks and pandemics without support. NGOs are increasingly instrumental in the implementation of international health programmes, as they possess attributes enabling them to function as effective agents, creating links between the community, private sector and government. The attributes of NGOs that increase their potential effectiveness include:

- their ability to reach areas of severe need
- their promotion of local involvement
- their relatively low cost of operations
- their adaptiveness and innovation, independence, and sustainability.

Links

There are links here to Chapter 7, Global systems and governance, specifically Global governance – both International agencies, page 328, and Interactions of governance between different scales, page 330.

They have more flexibility and freedom to respond in creative ways to a range of situations and to promote the same health goals, but are less hampered by bureaucracy and resource constraints.

The role of NGOs in health promotion and provision can be divided into four main areas:

- **Service providers:** front line providers of clinical healthcare: treating illness in the absence of, or complementary to, any existing government or private provision.

Médecins Sans Frontières (MSF)

Médecins Sans Frontières (MSF) (also known as Doctors Without Borders) is an NGO that exists to save lives by providing medical aid where it is needed most. Since 1971, MSF has cared for millions of people caught up in crises such as natural disasters, conflicts, famines, refugee exoduses or people excluded from healthcare. MSF employs 35,000 personnel globally; 90 per cent of these are recruited locally by European managers. The majority are local doctors, nurses or other medical specialists but they also recruit logistics experts and water and sanitation engineers. They have projects in over 74 countries, requiring both rapid response and longer-term specialist medical help to tackle health crises and diseases including COVID-19, cholera, Ebola, malaria, meningitis and HIV/AIDs. In 2019, 96.2 per cent of MSF's funding came from 6.5 million individual donors around the world, which helps to ensure operational independence and flexibility. The remaining funds come from governments and international organisations. MSF does not accept contributions from companies in industries whose core activities may compromise their ability to provide independent care, for example, the tobacco, mining and pharmaceutical industries. Since 2016, MSF have taken a principled stance by refusing to accept funding from the EU. This is in protest at the EU's reluctance to accept more migrants fleeing from conflicts in the Middle East and the EU's deal with Turkey to hold more refugees there.

Figure 10.39 Médecins Sans Frontières mobile clinic in the Central African Republic

MSF also treats malnutrition with ready-to-use therapeutic food (RUTF) which includes all the nutrients needed by children; it is mainly a peanut butter paste. RUTF has revolutionised the treatment of severe malnutrition as it can be stored for long periods and can be used in a variety of settings. MSF healthcare teams work with locally trained midwives, and traditional birth attendants, to set up programmes so that complicated births can be identified quickly to help prevent maternal deaths.

As well as providing vital medical assistance, MSF undertakes invaluable field research to generate evidence that improves the effectiveness and quality of the clinical care they provide. Research topics include treatment of multi-drug-resistant TB, HIV/AIDS, and combating neglected tropical diseases (such as Chagas and kala-azar), as well as mental health.

MSF acts as a watchdog to protect public health needs against corporate interests. MSF initiated the international 'Campaign for Access to Essential Medicines', which has since been joined by other NGOs. Its objective is to increase the availability of essential medicines in developing countries. The campaign lobbies both governments and pharmaceutical companies, challenging the high cost of existing medicines and the absence of treatment for many of the diseases affecting patients in developing countries. The campaign pushes for price cuts to medicines, vaccines and diagnostic tests by encouraging the production of more affordable generic products. It also aims to steer the direction of research towards new drugs and vaccines.

- **Social welfare activities:** providing or ensuring that the health infrastructure is in place, for example, food supply, clean water provision, public hygiene, sanitation and shelter for the unwell.
- **Support activities:** working with and training locals as health workers, for example, in family planning advice, hygiene and disease control.
- **Research and advocacy:** involvement in research to improve efficacy of prevention campaigns for both communicable and non-communicable diseases.

Despite their good work, NGOs have to be careful that their efforts do not undermine the ability of the host country to provide and maintain strong public health services by diverting local health workers into parallel operations. By attempting to plug the gaps in healthcare left by underfunded public health systems in LDCs, NGOs often reinforce geographical inequality in healthcare provision and in the types of disease treated. As a result of WHO concerns, the Global Health Workforce Alliance introduced the NGO Code of Conduct for Health Systems. The code:

- provides guidance to NGOs on working in a way that supports the primacy of the government's responsibility for organising the delivery of healthcare
- strengthens government-supported health provision through clinics, leading to more equitable care.

Review questions

9 Outline the four stages of the epidemiological transition.

10 Explain how *either* water quality *or* air quality impact on health.

11 Assess the various mitigation strategies used to manage malaria.

12 To what extent are lifestyle choices the most important factor in determining a person's likelihood of developing coronary heart disease?

13 Evaluate the role of *either* international agencies *or* NGOs in combating disease at the global scale.

10.4 Population change

There are **two** main components of population change:

- **Natural change:** The change in population brought about by the difference between **crude birth rate** and **crude death rate**. If the birth rate exceeds the death rate, then the population size will grow – this is called a **natural increase**. If the number of deaths exceeds the number of births there is a **natural decrease**.
- **Migration change:** Migration of people into or out of an area or country will also cause the population size to change. Those people moving into an area are called **immigrants** and those moving out are **emigrants**. The difference between the number of immigrants entering and emigrants leaving an area or country is known as the **net migration change**.

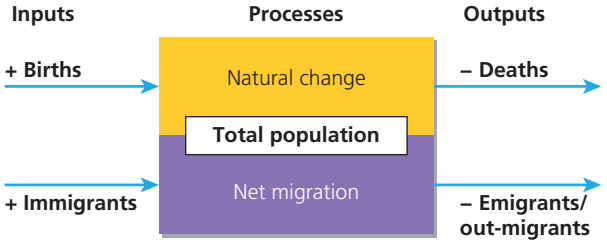


Figure 10.40 Systems diagram showing the two components of population change

Source: After Waugh D., Geography: An Integrated Approach

Figure 10.40 shows that these two components are different and distinctive but not necessarily mutually exclusive and will affect each other. For example, many international migrants are of **reproductive age**, consequently an increase of immigrants into a population will increase the child-bearing potential of that population, leading to an increase in birth rate. There will also now be a smaller proportion of elderly in the population so the movement of younger immigrants into a country will also reduce the death rate. If birth rates or death rates (or both) are changed by a migration event, then migration will have had an impact on natural change.

Factors in natural population change

Key vital rates

Natural population change will give either a natural increase or a natural decrease. Table 10.6 (page 507) shows birth rates and death rates for a selection of ten countries. Two other **demographic** indicators (**total fertility rate** and **infant mortality rate**) are also included in the table.

Table 10.6 shows there is a wide variation in birth rates and death rates in countries across the world and that both birth rates and death rates are able to rise or fall over time. The majority of countries in the world have a natural increase but some have a natural decrease (Figure 10.41).

Niger has the highest birth rate on the table, although it fell to 46 per 1,000 in 2018. Niger's seemingly low death rate (at 8 per 1,000) is explained by its youthful population structure, and the fact that healthcare has improved, largely because many NGOs operating in the country have supported the public health system to provide more welfare and medical care. The difference between the birth rate and death rate figures in 2018 was 38 people per thousand, which is the highest natural increase on the table.

Japan is the only country in the table with a natural decrease in 2018 (four people per 1,000). This comes as a result of falling birth rates and rising death rates, both associated with having an ageing population structure.

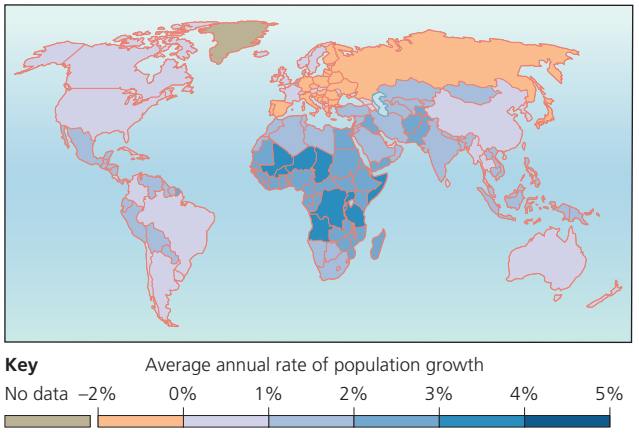


Figure 10.41 World map showing variations in natural population change in 2020

Source: Our World in Data

Growth rates and replacement fertility rates

From the natural change, demographers can calculate the growth rate of a population by multiplying by 100 to get a percentage. Niger's natural increase in 2018 is 38 per 1,000 of its population, which can be converted to a growth rate of 3.8 per cent. Niger's population in 2019 was thought to be around 23.3 million and with this growth rate, the population is expected to double by 2045.

Skills focus

Using Table 10.6:

a Calculate the percentage population change rates for Canada and The Gambia.

b Construct a Spearman's Rank Correlation table to compare the relationship between total fertility rate and infant mortality rate in 2018.

c Calculate whether there is a significant correlation between the two sets of data.

Table 10.6 shows some correlation between birth rates and total fertility rates, as both are measures of the reproductive potential within a population. Total fertility rate is a more accurate measure and a better indicator of future population change. This is because it focuses only on those who can give birth, by calculating the average number of children that each woman of reproductive age gives birth to.

The fertility rate can be used as an indicator of future population growth. A fertility rate of 2 means that, on average, each mother will have one daughter to

Table 10.6 Differences between vital rates of natural change in ten selected countries

A	B		C		D	E
	2011		2018			
Country	Birth rate – per 1,000 population	Death rate – per 1,000 population	Birth rate – per 1,000 population	Death rate – per 1,000 population	Total fertility rate 2018 (average births per woman)	Infant mortality rate 2018 per 1,000 < 1 year old
Brazil	15.3	6.4	14	6	1.7	13
Canada	11	7	10	8	1.5	4
The Gambia	43.2	10	39	8	5.2	39
India	21	7.9	18	7	2.2	30
Japan	8.3	9.9	7	11	1.4	2
Mexico	19.2	4.5	18	6	2.1	11
Niger	49.9	11.5	46	8	6.9	48
Russia	12.6	13.5	12	12	1.6	6
Sri Lanka	18.3	7	16	7	2.2	6
United Kingdom	12.8	8.7	11	9	1.7	4

Source: World Bank – World Development Indicators

replace herself reproductively in the next generation. This is known as the **replacement fertility rate**. In reality, the replacement rate is 2.1 because not all girls reach childbearing age (plus there are slightly more male births recorded globally). The global average replacement fertility rate is currently 2.3 because some populations have a higher child mortality rate. Replacement rates for different countries range from below 2.1 in richer countries to 3.4 in countries with high child mortality such as Sierra Leone.

In Table 10.6 (page 507), the sub-Saharan countries of Niger and The Gambia have the highest fertility rates. If they maintain these high fertility rates, their populations will continue to grow rapidly as each new generation will be larger than the previous one. Many of the other countries in the table have lower than replacement fertility, which is an indication that their population will shrink in the long term. It is possible to have below net replacement rate but still have a natural increase because the populations consist of more than just the youngest two generations and they have more people living longer, for example, in the UK.

Infant mortality rates

The number of children dying before their first birthday forms part of the death rate statistics but is considered to be a particularly important measure of mortality because:

- it is **age specific** as it relates to one particular group. This group is vulnerable but arguably is the group that society should be taking most care of
- it gives an **indication of the level of healthcare** available in the population, particularly maternity and post-natal care, as well as the prevalence of diseases and ability to combat them
- it gives **an indication of the wealth** of the country. Higher income countries will be able to afford the healthcare, medicines, healthy diet and clean water that will keep infants alive
- it will have an **impact on fertility rate**. Families in areas with high infant and child mortality (deaths under the age of five) will tend to have more children to ensure the survival of some of their family into adulthood.

Cultural controls

Birth rates and fertility rates vary considerably on a global scale and are influenced by a number of cultural controls as well as other socio-economic and political factors, as shown in Figure 10.42.

Key terms

Crude birth rate – The number of live births per 1,000 of a population per year – also known as 'birth rate'.

Crude death rate – The number of deaths per 1,000 of a population per year – also known as 'death rate'.

Demography (demographic) – Demography is the study of human population. A demographer is someone who studies the statistics and characteristics of populations.

Emigrant – A person leaving their home area or country in order to settle elsewhere.

Immigrant – A person moving into an area or country in order to settle there.

Infant mortality rate – The number of children who die before their first birthday per 1,000 live births per year.

Life expectancy (at birth) – The average number of years a person born in a particular year in a location is expected to live for.

Natural change – The difference between birth rates and death rates. If birth rates are higher, there is a natural increase in population; if death rates are higher, there is a natural decrease.

Net migration change – The difference between the total number or average rate of immigrants and emigrants in an area or country over a given period of time. More immigrants than emigrants will give a positive net migration and more emigrants than immigrants gives negative net migration.

Replacement fertility rate – The number of children each woman needs to have to maintain current population levels or give zero population growth by each subsequent generation. It is a measured total fertility rate (see below). If fertility is above the replacement rate, population will grow and if it is below, population will decline.

Reproductive age – The age at which women can give birth. In official demographic data it is usually considered to be between 15 and 44.

Total fertility rate – The average number of children born per woman in an area or country if all women lived to the end of their childbearing years. It is considered to be a more direct and accurate measure of fertility than birth rate as it refers to the average number of births per woman in a given population.

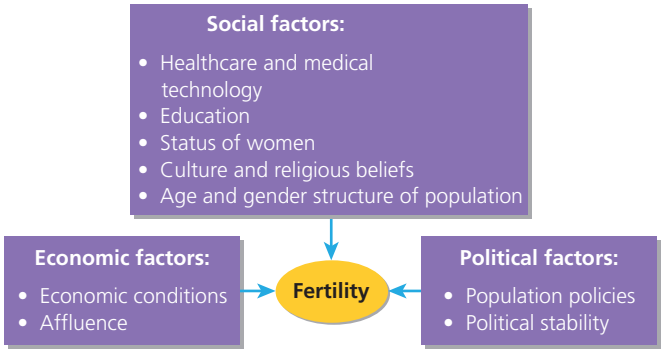


Figure 10.42 Factors affecting fertility

Decisions about family size are influenced by socio-economic and sometimes by political factors, which in turn will have long-term effects on culture. For example, larger families were once the norm in the UK but as society has become more economically developed, parents tend to have fewer children. Having more children may be regarded as an expensive option, especially when contraception is so readily available. Another key factor is that more women have careers, which they want to return to, so decide to have fewer children. Family planning is increasingly based on economic decisions that enable both parents to work to maintain personal living standards.

Religion

The doctrine of certain faiths such as Roman Catholicism and Islam forbid the use of artificial contraception and abortion as methods of population control. Most countries where Catholicism is the main faith demonstrate relatively low fertility and only slow population growth (and in some cases decline). It may be that other cultural values such as equal opportunities for women or materialism override religious doctrine. On the other hand, countries where the majority of the population are Muslim show some of the highest fertility rates and fastest growth, even though certain types of artificial contraception are permissible. However, it may be that economic factors, other cultural factors or even environmental factors are the main drivers of higher fertility rates in these countries.

Gender

- **Gender preference:** For cultural and socio-economic reasons, some societies have a distinct preference for having male children. This is particularly true in rural areas of developing countries where the male will be expected to farm the family land and look after his parents in their old age. This is the case in parts of Pakistan, India and in sub-Saharan African countries such as Niger and Mali. An obvious consequence of this will be higher fertility rates and larger families as parents will continue to have children until they have one or two boys.
- **Status of women:** Women are discriminated against in many cultures, often not receiving education. When married they are expected to leave their own family and look after in-laws. The degree of equal opportunities, in terms of education and careers, afforded to females in a society will inevitably influence the number of children each mother will have in her lifetime.

Child marriage

In the least developed countries, particularly in sub-Saharan Africa, 40 per cent of girls are married before they are aged 18 and 12 per cent before the age of 15. According to UNICEF, more than 50 per cent of girls in six countries of sub-Saharan Africa and Bangladesh were married before they reached 18. Many African countries have legal minimum marriageable ages of between 16 and 18 and in certain countries (for example, Niger) the minimum age for girls is 15 with parental or judicial consent. This will clearly have an impact on the number of children that a mother will have in these countries, as well as on the mother's health and future opportunities.

Effects of population policies

The legacy of China's enforced one-child policy is that parents now choose to have one child as the social norm, even after the policy was discontinued in 2015. From 2016, parents were encouraged to have two children but after a slight initial increase, the number of births has decreased, suggesting that fertility rate is likely to remain very low in China. Young people are faced with high house prices, labour market pressures and the expense of bringing up and educating children. For economic reasons, as much as societal norms, most people intend to have only one or two children, averaging less than the replacement fertility rate. With an ageing population, China may have to consider introducing pro-natal policies in the future.

Even in countries with less coercive population policies, government strategies advocating population control can influence culture and societal norms. For example, some Indian states have adopted a two-child policy by introducing penalties and disincentives for families with more than two children. As a result,

average Indian fertility rates have fallen from 3.2 to 2.2 in the last 20 years. Fertility rates in 23 Indian states are now lower than replacement fertility. Kerala, a state in south India, initiated the culture of smaller families being the social norm by introducing compulsory female secondary education in the 1980s. It now has the lowest fertility rate of 1.2.

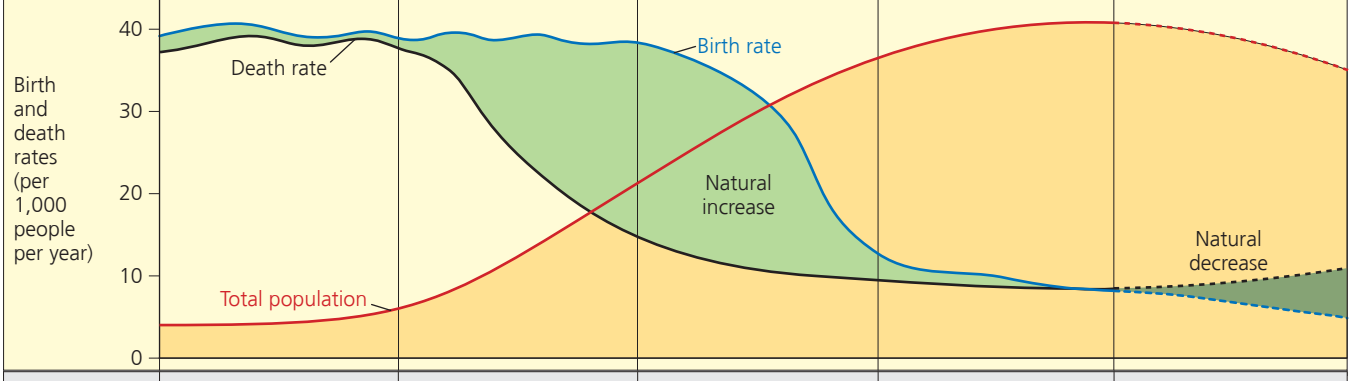
Models of natural population change

A 'model' provides a structure which is representative of the characteristics, processes and relationships that occur in the real world. Demographers have presented various conceptual ways of characterising population change and its impacts. This section focuses on two ideas that can be linked. The demographic transition model considers how population is thought to grow over time and can be associated with the epidemiological transition covered in Section 10.3. Population pyramids represent the age–sex composition of a population but should also be considered as a dynamic structure that changes through time.

Demographic transition model

The **demographic transition model** (DTM) was developed by American demographer Warren Thompson in 1929 and was based on recorded changes in birth rates and death rates in industrialised countries since the start of the Industrial Revolution. It traces a change from high birth rates and high death rates to low birth rates and low death rates, as a country develops from a pre-industrial economy to an industrialised economic system. The rate of change will vary from country to country and countries will, at any one time, be located at different stages along the transition. Look back to Section 10.3 and consider how epidemiological transition (pages 488–90) accounts for the different stages and changes in the DTM.

Figure 10.43 shows an updated version of Thompson's model. It has five stages, each identified by what is happening to birth and death rates or total population.



Stage	1 High stationary High fluctuating	2 Early expanding	3 Late expanding	4 Low stationary Low fluctuating	5 Decline
Examples	A few remote groups	Niger, Mali, Angola	Egypt, Pakistan, Philippines	USA, Canada, UK, China	Germany, Japan, Italy, Bulgaria
Birth rate	High	Remains high	Starts to fall	Low and steadies	Falls below DR
Death rate	High	Falls rapidly	Falls more slowly	Starts to level out	Remains low, may increase
Natural increase	Stable and fluctuating	Very rapid increase	Increase slows down	Slow increase, some fluctuations	Small natural decrease
Reasons for change in birth rate	Many children needed to work on the land. Many children die at an early age. Religious and social reasons. No family planning		Conditions change so fewer children needed. Lower IMR. Contraception	More family planning. Improved status of women with careers	More elderly population, smaller proportion in reproductive ages
Reasons for change in death rate	Disease, famine. Poor medical knowledge so many children die	Improvements in food and clean water supply, sanitation and hygiene. Medical care advances, more vaccines. Fewer children die		Good healthcare established. Reliable food supply	More elderly population; higher proportion reaching life expectancy.

Figure 10.43 The demographic transition model

The DTM records the rate of total population growth, which is at its most rapid when natural increase is greatest during stage two and the early part of stage three. It then starts to slow down during the latter part of stage three and stage four, when there is less difference between birth and death rates and so a smaller natural increase. In recent years, a fifth stage has been added to the original model to account for those countries with declining population numbers.

The model applied to countries with contrasting physical and human settings

The countries in Table 10.7 have been ranked from lowest to highest according to their level of economic development, measured by **gross national income per capita**. The data gives an overview of whether there is a relationship between the level of economic development and progress in terms of demographic transition. Table 10.8 outlines contrasts between Niger and Canada and considers whether physical conditions and other human factors play a part in the demographic transition experienced by these two countries.

Table 10.7 Countries at various levels of economic development and application to DTM

Country	Birth rate 2018	Death rate 2018	GNI per capita 2019 (US$)	Life expectancy 2018
Niger	46	8	550	62
Sri Lanka	16	7	3,892	77
Botswana	25	6	7,449	69
Brazil	14	6	10,826	76
Japan	7	11	41,690	84
Canada	10	8	51,196	82

Source: World Bank – World Development Indicators

Table 10.8 Contrasting physical and human settings and demographic transition in Niger and Canada

Niger	**Physical setting**	One of the Sahelian nations with a very arid, subtropical desert climate. Much of the north and east of the country is Sahara desert. The terrain is predominantly desert plains and sand dunes. The extreme south has a tropical climate near the edges of the Niger River basin. Non-desert areas in the south and west are threatened by periodic drought and desertification (Figure 10.44, page 512). Recurring drought events are a hazard and have occurred more frequently in recent years as a result of climate change
	Human setting	It is mostly sparsely populated, especially in the north. There is a higher concentration of population in a band of towns and cities in the south along the Niger basin. The economy is very reliant on its primary sector, especially agriculture and mining. Most people live in rural areas and are subsistence farmers or nomadic pastoral herders. Niger is rich in uranium deposits and has developed oil production steadily since 2010. Around 98 per cent of the population are Muslim, coming from a number of different tribal groups. Droughts, failed crops, insect plagues and internal conflicts as well as rapid population growth have led to food shortages, high food prices and hunger for many
	Application to Demographic Transition Model	High birth rates (largely as a result of religious and cultural beliefs) and relatively low and falling death rates put Niger in Stage 2 of the model but death rates are much lower than would be expected for this stage. This is partly due to their young population structure and also because the government has made great strides to reduce child mortality by reducing hunger and malnutrition and improving healthcare. They have been supported by 118 registered NGOs operating in the country, including Save the Children and WaterAid
Canada	**Physical environment**	Comprises a wide range of climatic types including arctic, temperate continental and temperate maritime. The west of Canada is mountainous and the southern central area features rolling fertile plains (Figure 10.45, page 512). The north is a rugged and mountainous wilderness of taiga and tundra. Canada is extremely rich in mineral resources
	Human setting	It is a very large country with a relatively small population of around 37 million people, so on a national scale it is sparsely populated but highly urbanised with a concentration of population in the large cities in the south east of the country, bordering the USA. Smaller concentrations appear in the southern parts of central provinces and on the west coast. The economy is based very much on its tertiary sector, particularly financial services, and manufacturing, though it also has thriving mining and oil industries. It is a multicultural society, mostly welcoming to immigrants and tolerant of different cultures, languages and traditions
	Application to Demographic Transition Model	Canada has a low birth rate and death rate with a low natural increase. This places its demographics fairly clearly in Stage 4 of the model. Canada has not progressed into Stage 5, although slight decreases in birth rate and increases in death rate have reduced the natural increase in recent years. These demographic trends are typical of a country with an increasingly ageing population. However, Canada encourages controlled migration in order to keep a balanced age structure to avoid this problem, which is encountered by some other developed countries. Having plenty of space and being rich in resources, it is a country that can accommodate population increase

Figure 10.44 Environmental conditions make reliable food production in Niger difficult

Figure 10.45 A variety of landscapes in Canada provide various economic opportunities for its population

The demographic transition model is not perfect but still provides a useful framework against which to compare real data. Its limitations can be used to explain some of the anomalies that occur between a nation's transition and what the model predicts. Countries do not all progress smoothly through the model; sometimes the transition falters or even reverses. The model might have more validity if social, cultural and political factors were given as much importance as economic development, however these are less predictable and may render the model unnecessarily complex.

Table 10.9 Summary evaluation of the Demographic Transition Model

Strengths	Limitations
The model can be: • easily and universally applied • used to compare stages of demographic development • used by demographers to make predictions about future change	The model does not account for: • migration and its impact on birth rates • government policy which may influence birth rates • wars and other conflicts which will impact on death rates • the impact of major disease pandemics on death rates, for example, COVID-19 • environmental limitations placed on future economic development, for example, meagre resources, harsh climatic conditions, natural disasters

Population structure – age and sex composition

Population structure is a term used to describe the age distribution and sex composition of a population. It is usually examined at a 'national' scale, by country, but it can be used at many different scales.

Population structure is usually depicted by the construction of a **population pyramid** as shown in an annotated version in Figure 10.46.

A population pyramid is a snapshot at any given time, but population structure is constantly changing as each age group moves up the pyramid over time. Pyramids provide insights into past trends in population, such as changes in fertility, mortality and international migration, as well as what is currently happening to population.

Interpreting population pyramids

Population pyramids can:

- show past changes in population
- be used to predict short-term and long-term future changes in population
- show the effects of wars/diseases and famine
- indicate life expectancy
- give an idea of the stage of demographic transition
- show effects of migration on a population structure (including the age and gender of migrants).

Dependency and its implications

For analytical purposes, population structures and their representative pyramids are divided broadly into three categories according to their economic productivity:

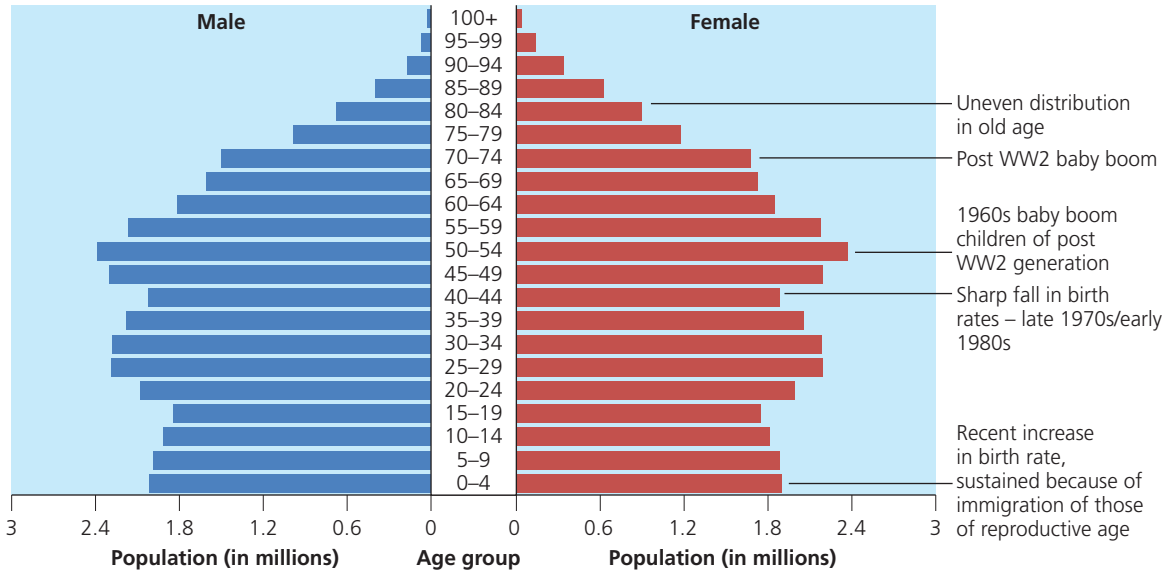


Figure 10.46 Population pyramid showing population structure for the UK, 2018

Source: Central Intelligence Agency

- **Ages 0–14** – known as the '**young dependent**' population.
- **Ages 15–64** – a group known as the '**economically active**' in a population.
- **Ages 65 plus** – usually known as the '**elderly dependent**' population.

The economically active tend to be the **working population** who earn an income and pay taxes. The two dependent groups, the young and elderly, are known as the non-economically active population. They are dependent on the economically active to make sure they have a means of living. Young dependants rely on their parents to support them while the elderly rely on pensions and/or family support, depending on the development of social services where they live.

The **dependency ratio** is a measure of the level of dependency and is expressed as:

$$\frac{\text{Young dependants (0–14)} + \text{elderly dependants (65 and over)}}{\text{Economically active (those aged 15–64)}} \times 100$$

In 2019 the UK's estimated population structure comprised of:
0–14 = 11.96 million (17.8 per cent of population)
15–65 = 42.87 million (63.8 per cent of population)
65+ = 12.37 million (18.4 per cent of population)

$$\frac{11.96 + 12.37}{42.87} \times 100 = 56.75$$

This means that for every 100 people of working age there are around 57 people dependent on their earnings.

Application of population structure pyramids to natural change

The shape of the population pyramid evolves based on changes in fertility and mortality. The demographic transition model gives a predicted sequence of change in birth rates and death rates over time. The two concepts can be linked and a 'definitive' shape of population pyramid outlined for each stage of the model (Figure 10.47, page 514).

The model determines what stage of demographic development a country is at so we can suggest an associated population structure for countries at each stage of development.

These contrasting population structures have different implications for the use of resources and services and for the economic potential of countries at different stages of development.

The composition of dependency will vary between more and less developed economies. In richer countries a greater proportion of the dependency comprises the elderly population, whereas in less developed countries the higher overall dependency is largely of a youthful population. Both types of structure have problems resulting from dependency but they also have potential benefits (see Table 10.10, page 514).

Table 10.10 Implications of contrasting population structures for resources and development

Problems	Potential benefits
Youthful populations (for example, Niger)	
Increasing demands for improving maternal and child healthcare and education Providing food, water, energy and shelter for growing population Governments need to keep up with demand for schooling Lack of attendance in schools (especially in rural areas) leads to low levels of literacy and poorly educated workforce	Young people of today are the country's human resources of tomorrow Can foster growth and development if there are favourable political and economic conditions Can provide a cheap workforce and a growing market for foreign investors
Ageing populations (for example, Japan – see page 545)	
Welfare and healthcare costs escalating as the elderly are more likely to need support Pensions will cost increasingly more in the future (the **'pensions time bomb'**): – costs have to be borne by a smaller economically active population – workers may have to pay higher taxes Smaller proportion of the population are 'economically active' – this may affect economic growth and the overall standard of living; may be skills shortages	Some pensioners are healthy and affluent: – growth in leisure/tourism industry – growth in private health and residential care businesses – companies able to target growing markets Fewer people of working age so there should be less unemployment in the population. Some elderly still 'work': – looking after grandchildren, enabling parents to work – volunteering in their local community

Population structure will also influence the growth rate and future size of the population, which has implications for the balance between population and resources. These arguments are extended in more detail in Section 10.5, Principles of population ecology applied to human populations (page 521).

The demographic dividend

The **demographic dividend** refers to a period when the population structure of a country means there is low dependency.

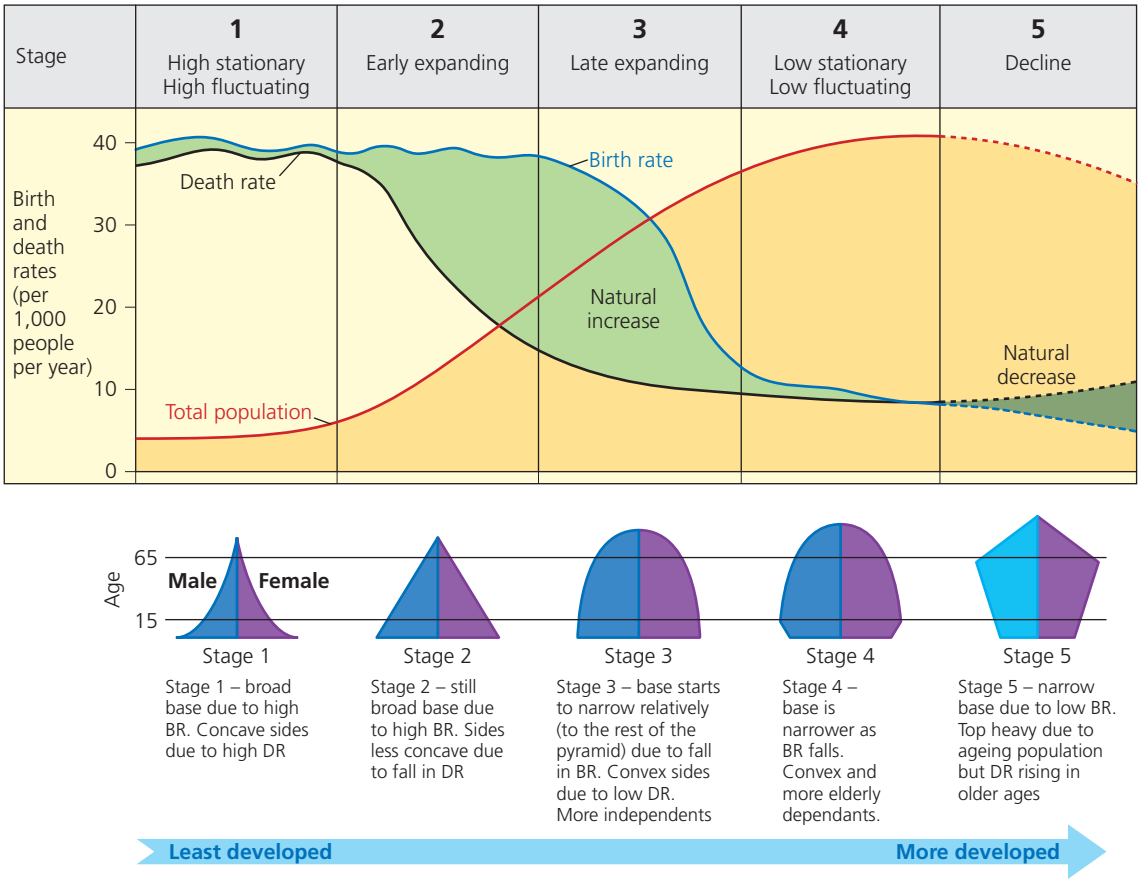


Figure 10.47 Population pyramid outlines for each stage of the demographic transition model

Key term

Demographic dividend – The benefit a country gets when its working population outgrows its dependants, such as children and the elderly. A boost in economic productivity results from growing numbers in the workforce relative to the number of dependants.

The lag between falling death rate and subsequent reduced birth rate may last for one or two generations before parents adjust to falling infant mortality by having fewer children. This lag period is one of rapid population growth and produces a 'generational bulge' that rises up the population pyramid, eventually into the economically active sector (as seen in the population pyramid for Brazil in Figure 10.48). With few elderly dependants because of previously low life expectancy, and fewer young dependants because of declining fertility, the dependency ratio will fall. For countries progressing through demographic transition relatively quickly, just as China, Brazil and India have done in the past 50 years, this 'bulge' appears in Stage 3, just as fertility rates are starting to fall.

This large group of young and aspirational people will give rise to a period of higher productivity and a boom in economic growth and social development because:

- a large, young, educated **workforce attracts investment** from footloose global companies (TNCs)
- workers with fewer children **invest more of their income**, leading to financial stability and growth
- fewer children also means **more women join the workforce**, promoting more gender equality
- a larger proportion of economically active salaried workers provides a **growing market** for consumption of goods and services.

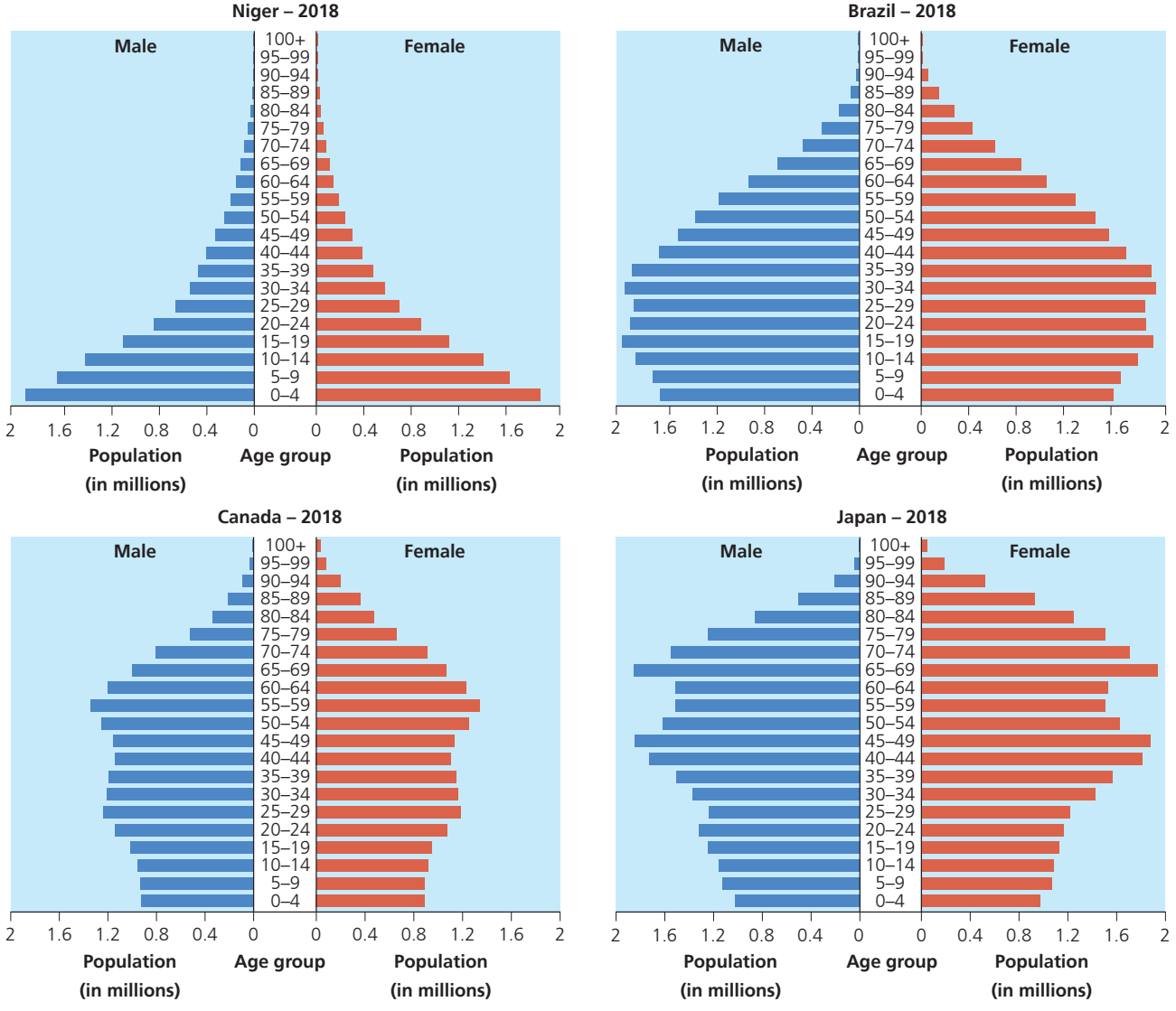


Figure 10.48 Population pyramids for countries at different stages of development

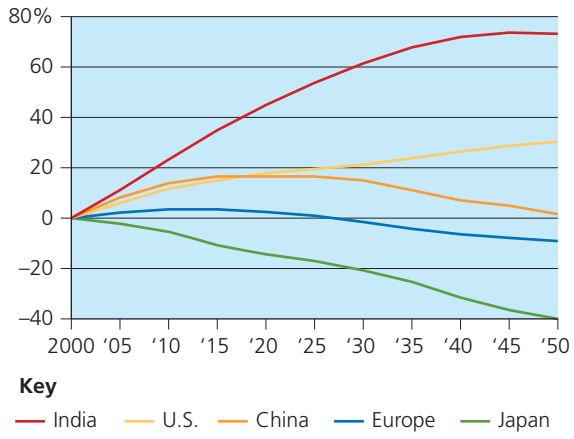


Figure 10.49 Changes in working-age populations (actual and projected)

Source: UN Population Division

Links

There are links here to Chapter 7, Global systems and governance, specifically Globalisation – Flows of capital, page 285; and Global systems – Unequal flows of people and Unequal flows of money, page 300.

Countries can only take advantage of their demographic dividend providing there is investment in education and employment and the political will to encourage foreign investment to stimulate growth. The Asian Tiger economies are good examples of where this has occurred. Countries may further maximise the dividend if there is a rich natural resource base to support growing demand, such as in Brazil and China. However, the dividend may be a wasted opportunity if there is poor educational attainment or a lack of transparency and democracy in government. Political instability will deter investment, and lack of employment opportunities leads to the emigration of young people, thus jeopardising the dividend itself. Myanmar, for example, has not benefited fully from its dividend for these very reasons.

The boost in productivity will create extra demand in the economy and this may cause environmental degradation and put additional pressure on the use of natural resources, unless they are managed sustainably. The demographic dividend itself can only be sustained by managing fertility rates at replacement level. Even then, improved life expectancy will mean, in the longer term, the bulge will move up the pyramid to form an ageing population structure. This in turn will put pressure on specific services such as healthcare. Thus it could be argued that taking advantage of the demographic dividend is a 'one-time opportunity' for most countries.

International migration

The other major component of population change is migration. The geographical definition of migration is 'the movement of people across a specified boundary in order to establish a new permanent or semi-permanent residence.' Migration can take place at various scales and is the result of a number of factors. Whatever the scale and causes, it is likely to have an impact on population change at both the **origin** and the **destination** of the migrants.

International migration is an important dimension of global systems and a high-priority issue for both developing and developed countries.

International migrants are defined as persons who are living in a country other than their country of birth or their country of citizenship. According to UN estimates in 2019, 272 million people (3.5 per cent of the world's population) are international migrants. The majority of migrants cross borders in search of better economic and social opportunities. Others are forced to flee crises such as wars, conflict, persecution or natural disasters. This suggests that the majority of migration is one directional from less developed to high-income countries. However, migration flows can be multi-directional especially if re-migration is taken into account. In some cases, immigration and emigration between two countries can be seen as interlinked rather than separate processes, for example, flows of people between Poland and the UK or between China and Canada.

The UN International Migration Report for 2019 identified a number of key findings including:

- two-thirds of all international migrants were living in just 20 countries
- one-third of all international migrants originated from only ten countries (Table 10.11 shows the top five)
- 74 per cent of all international migrants were of working age (20–64 years old)
- forced migration across borders has risen significantly in the last 10 years
- women and girls comprise around 48 per cent of all international migrants
- few governments are seeking to lower the number of immigrants providing they arrive through legal channels. Promoting orderly, safe, regular and responsible migration is a priority for most governments.

Table 10.11 Top five countries of origin and destination of migrants currently living outside their country of origin (2019)

Country of origin	Size of diaspora (millions)	Country of destination	Number of immigrants (millions)
India	17.5	USA	51
Mexico	11.8	Germany	13
China	10.7	Saudi Arabia	13
Russia	10.5	Russia	12
Syria	8.2	UK	10

Figure 7.7 (page 288) in Chapter 7 shows the main flows of international migration.

Types of migrant

During large-scale international migration events, such as the European migrant crisis, which has been ongoing since 2015, there can be confusion in the media and in public debate about the differences between **asylum seekers**, **refugees** and **economic migrants** and the complex legal distinctions of each status can be lost. The 1951 UN Refugee Convention (and subsequent 1967 Protocol on the Status of Refugees) remain the key documents in international law for the protection of refugees. The Convention provides the most comprehensive statement in defining a refugee, their rights and obligations and the legal obligations of states. There are 145 signatories and the Convention is overseen by the UNHCR (the UN's Refugee Agency).

Key terms

Asylum seeker – A person who has fled their country of origin and applies for asylum under the 1951 Convention on the grounds that they cannot return to their country of origin because of a well-founded fear of death or persecution. An asylum seeker is someone seeking international protection but whose claim for refugee status has not yet been determined.

Economic migrant – A person who has voluntarily left their country of origin to seek, by lawful or unlawful means, employment in another country.

Refugee – In its **broader context** it means a person fleeing, for example, from civil war or natural disaster but not necessarily fearing persecution as defined by the 1951 Refugee Convention. **Legally**, however, a refugee is an asylum seeker whose claim for asylum has been successful under the 1951 Convention and they are granted refugee status in a new country to live there permanently.

There is some overlap between the different categories. For example 'asylum seekers' and 'refugees' are frequently conflated, which gives rise to confusion and suggests that all asylum seekers are automatically granted refugee status. The reason for the conflation is that people fleeing from an area of conflict or disaster are initially labelled generally as 'refugees'. It is only when they reach a destination country that they seek asylum, and only when their application for asylum is granted in the new country are they given legal 'refugee status'. The position of asylum seekers is similar to those who enter on short-term visas. They are granted temporary admission while their applications are pending, but if their claim is unsuccessful they have no lawful right to remain in their chosen country of destination.

Some migrants claiming to be asylum seekers may be economic migrants who hope to secure entry into a country by claiming asylum. This is more feasible for migrants coming from areas of ongoing conflict or civil disorder.

Causes of migration – push and pull factors

Global migration has declined a little since reaching an average of 6.4 million migrants per year between 2005 and 2010. The numbers of international migrants between 2015 and 2020 are estimated to be around 4.9 million on average each year, though there are signs of it growing again.

People migrate for a number of reasons. These can be categorised as **push factors** or **pull factors.**

- 'Push' factors are reasons which are based at the origin of the migration and which initiate the migrant's desire to move.
- 'Pull' factors are based at the intended destination of the migrant and attract people.

Figure 10.50 lists examples of push and pull factors. Observers often distinguish between **'forced'** migration, when the individual or family has little choice but to move, and **'voluntary'** migration. Forced migration is always caused by extreme push factors at the origin, such as natural disasters or wars. A voluntary decision to migrate may be based on factors present at both the origin and the destination as indicated by Lee's Push–Pull Model of migration shown in Figure 10.51.

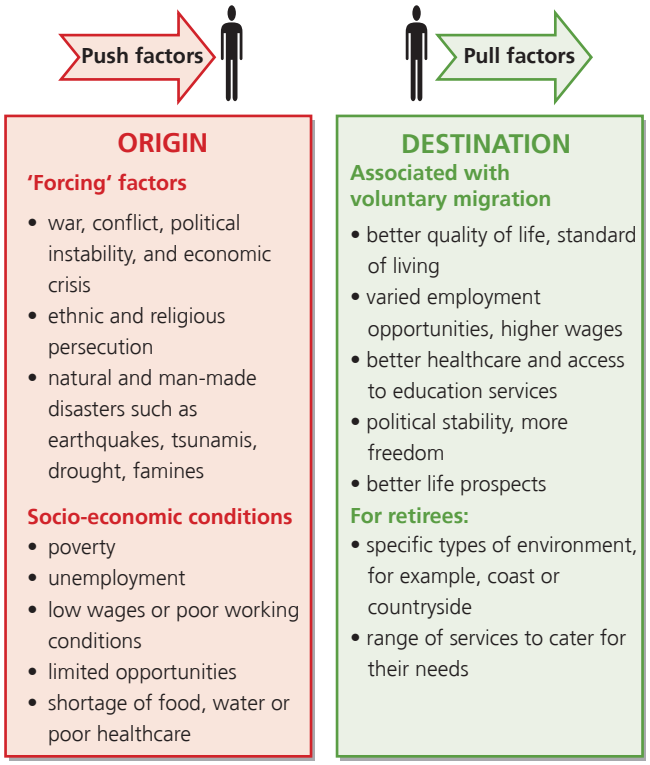


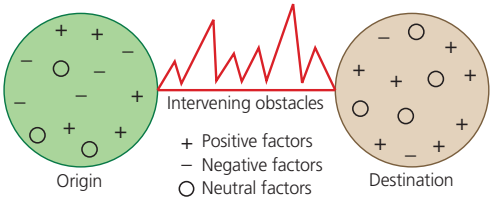
Figure 10.50 Push and pull factors

A combination of war, religious and ethnic persecution and poverty is causing the movement of refugees and economic migrants on a global scale. UN figures suggest that 1 in every 117 people globally is either a refugee, a displaced person or is seeking asylum. In 2018, UN figures stated that the total global refugee population was 25.9 million and the number of internally displaced persons due to violence and conflict reached 41.3 million.

Processes of migration

Lee's push–pull model of migration suggests that the process of migration from one area or country to another is not usually a straightforward one.

The decision to migrate is made by individuals or families based on existing conditions at the origin. For example, during the peak of Europe's 'migrant crisis' in 2015–2016, the majority of the migrants were fleeing civil war in Syria. Innocent civilians were caught up in fighting between government forces and various rebel groups including the militia of the so-called Islamic State (of Iraq and Syria) or ISIS. Migrants genuinely feared for their lives and so the negatives at the origin far outweighed any positives for them to stay in Syria. Equally the prospect of higher living standards and the promise of better opportunities in the EU made the positives at the destination outweigh any negatives. For many the decision to migrate was an obvious choice. However, both refugees and economic migrants still have to overcome the intervening obstacles, which for many involved negotiations with people traffickers in Turkey and a hazardous sea voyage, often on unstable craft (Figure 10.52). On reaching EU territory in Greece, many refugees continued their journey by foot to reach intended destinations in Austria or Germany. On these relentless marches through Balkan countries they were met with further obstacles in the form of closed borders and often had to find other routes.



- For people to move, they need to be 'pushed' from their country of origin and 'pulled' to another country.
- The majority of migrants are voluntary movers, doing so for largely economic reasons.
- The negative factors at the origin are the 'push' factors; the positive at the destination are 'pull' factors.
- Migrant has to evaluate these factors and obstacles before they move.
- Intervening obstacles might include: travel costs, family pressures, language barriers, misinformation, immigration controls, bureaucracy, border controls.

Figure 10.51 Lee's push–pull model of migration

The intervening obstacles make the migration a 'step by step process'. Initial moves by migrants may be to seek refuge in 'safer' towns, countries or refugee camps before deciding on a longer and more permanent move.

Figure 10.52 Overcrowded migrant boat in the Mediterranean

Implications of migration

Table 10.12 (page 520) outlines the demographic, environmental, social, economic, health and political implications of migration.

Links

There are links here to Chapter 7, Global systems and governance, particularly Unequal flows of people, page 300; and Chapter 8, Changing places, particularly Demographic change and cultural characteristics, page 361.

Economic implications

International migration is a feature of the interdependence that exists between different countries in global systems. Large numbers of unemployed in poor countries are seeking better opportunities abroad, usually in HDEs, and causing an unequal flow of labour.

As a consequence of globalisation, there has been a 'scaling up' of economic migration in recent years, which is taking place over longer distances for a number of reasons:

- More skilled workers are required to work abroad, especially if they are working for TNCs.
- Wealthier countries, especially those with ageing populations, may encourage immigration of skilled workers to fill skill shortages.
- Economic migrants who cross borders illegally are better informed and seem willing to take greater risks, including paying large sums of money to organised traffickers.

Economically, richer countries often gain from migration at the expense of LDEs who continue to lose younger, talented people. In such flows, the more developed countries also gain additional skills and labour and create, through migration, a partial demographic dividend. Immigrants' spending on goods and services, together with paying taxes, will help to support and possibly boost the developed economy. Migration will have the reverse effect at the origin, where dependency increases and services close due to lack of support. Remittances sent back to relatives in their country of origin may offset these negative effects to an extent but otherwise economic inequalities between HDEs and LDEs tend to be reinforced.

Demographic implications

Any notable migration event will have implications for the origin and for the destination area or country. In the first instance the population **size** will change as people move from the origin to the destination. The age and gender characteristics of the migrants will also have an impact on the **structure** of the population.

Voluntary migration is age and gender specific, so inevitably the movement of a large number of people from one population to another will have an impact on the population structure of both the origin and the destination. This can happen at all scales, from rural to urban migration occurring in developing countries, up to an international scale. Migration causes a fairly **immediate impact** on the size and structure of the population, unlike natural change, which tends to have a more gradual, long-term effect.

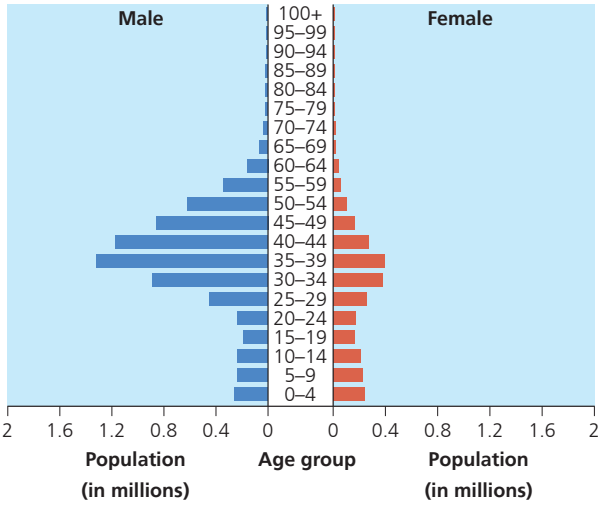


Figure 10.53 Population pyramid for United Arab Emirates, 2018

Source: Central Intelligence Agency

Table 10.12 Implications of migration for the origin and for the destination

Implications at origin (Home country)	Implications at destination (Host country)
Demographic implications	
Lower birth rates, people of childbearing age leave Population structure – ageing population remain; population unbalanced Loss of population of working age	Balances population structure, if previously ageing population Migrants in reproductive age groups means an increase in birth rates Increase in population of working age (in some migration events, this is dominated by male immigrants)
Social implications	
Advantages Reduced pressure on healthcare, though this will be dependent on population structure remaining (see healthcare) Reduced pressure on education **Disadvantages** Loss of traditional culture Break-up of family units Break-up of communities May lose qualified workers such as doctors, nurses and teachers	**Advantages** Cultural advantages of new foods, music, fashion, etc. **Disadvantages** Pressure on maternal and infant healthcare Pressure on schools (particularly primary) Can give rise to ethnic and racial tensions Potential increase in crime Migrants may settle in concentrated areas and become segregated
Economic implications	
Advantages Reduced pressure on food, energy, water, etc. Less unemployment as there are fewer in the workforce Remittances sent back home by migrants Migrants develop new skills which they can bring back home **Disadvantages** Lose better educated/most skilled from workforce May create a dependency on remittances Less agricultural and industrial production Decline in services – not enough people to support them	**Advantages** Overcomes any labour/specific skill shortages May provide a more competitive workforce, willing to work longer and harder for relatively lower pay Working migrants spend money/pay taxes Increases size of workforce – can provide economic boom and multiplier effect Reduced dependency – creates a partial 'demographic dividend' **Disadvantages** Pressure on jobs which may lead to unemployment during periods of economic downturn Resentment towards migrants in time of recession
Political implications	
Pressure to re-develop areas in decline May introduce pro-natal policies or incentives to retain skilled workers Alternatively some governments of 'over-populated' countries may encourage emigration to gain socio-economic advantages (as listed above) or as an anti-natal strategy (for example, the Philippines)	Pressures to control immigration Rise of anti-immigration political parties Growth of right-wing racist organisations
Environmental implications	
Farmland, buildings and sometimes whole villages may be abandoned Less environmental management	Pressure on land for development – roads, housing, other infrastructure Increased demand for energy, water and food puts pressure on natural resources
Health implications	
Migrants leave areas where infectious diseases are endemic or sometimes epidemic Less pressure on limited health services but… …demographics of migration mean that the most vulnerable (children, elderly and poor) remain at risk … many developing countries lose doctors and other medically trained staff to richer countries that can afford to pay higher wages and offer better prospects, this puts more pressure on health services in LDEs.	Spread of communicable diseases being transmitted may become more likely as migrants move from areas of higher disease prevalence. Increased pressure on health services because of rise in infectious diseases Increased pressure on health services to treat non-communicable/ chronic diseases

Specific impacts on structure:

- To bolster the economically active proportion of the population and therefore reduce dependency.
- Those of economically active age are also of reproductive age and there is a consequential effect where the birth rate will increase in the destination country. This has happened over the past 15 years in the UK and has increased the numbers in the younger age groups.
- Migration events can be gender imbalanced. For example, the population pyramid for the United Arab Emirates (Figure 10.53, page 519) shows a large protrusion of males of working age groups from 25 to 59.

Health implications of migration

Disease transmission

Disparities in disease prevalence and healthcare often exist between a migrant population's place of origin and its destination. These can be among the driving push and/or pull factors initially triggering migration. There are a number of implications for increased migration across disease disparate regions, including the transmission of certain diseases globally and particularly in countries receiving migrants. This has been evident during the COVID-19 pandemic when rates of disease incidence have been carefully monitored by WHO and national governments. As a result, many borders have been closed during the crisis, especially for those travelling from areas where COVID-19 incidence rates are higher than their intended destination. Most forms of permanent legal migration have been put on hold while the pandemic persists, though there is some evidence to suggest that attempts to migrate illegally have increased.

However, it is important to consider that a rise in infectious disease incidence may be wrongly attributed to immigration; it may be only one (if at all) of many contributing factors.

Impact on health services

In 2019, a Health Foundation statement concluded that evidence showed that immigration made a positive contribution to the UK health service:

- Data suggests that migrants make a disproportionately low demand on health services, mainly because they tend to be younger and healthier but also because of language barriers and a lack of knowledge of accessing the health system.
- As taxpayers, immigrants contribute to the cost of public services.
- Migrants are an essential part of the healthcare workforce in the UK. Around 13 per cent of NHS staff are non-British and 28 per cent of doctors working in the NHS were trained in other countries. Considerable staffing shortages are predicted over the coming years. It is likely that the UK will be continually dependent on recruiting both healthcare and social care staff from overseas to fill these shortages, For example, it is thought that the NHS will need to recruit 5,000 international nurses each year until 2023–24.

On the other hand, consider the reverse and the negative impacts that this will have on health services in many LDEs as they often lose many of their most talented clinicians to more developed countries.

Review questions

14 Suggest why infant mortality rate is a better measure of development than crude death rate.

15 To what extent does epidemiological transition explain the stages of the demographic transition model?

16 Outline the environmental and socio-economic causes that would initiate outward migration from an area.

17 Evaluate the impacts of migration on a destination area receiving a significant number of immigrants.

10.5 Principles of population ecology applied to human populations

Population growth dynamics

In ecology, population growth is determined by birth rates and death rates. The number of births is controlled by the natural reproductive potential of the species. This is also known as the **biotic potential**, which has evolved over time for individual species and is related to the survival rates of the young of that species. Thus fish species such as herring produce many young because their chances of survival are very low. In contrast, most mammals (including humans) have a small number of offspring as their rate of 'survivorship' is high, though they can be vulnerable to increases in death rate. The

number of deaths is controlled by environmental limiting factors that prevent survival. This group of factors is collectively known as **environmental resistance.** Some environmental limiting factors are '**density independent**' and are not influenced by the population size and density. For example, natural hazards such as drought, floods, volcanic eruptions, etc., will increase the death rate whatever the population size. On the other hand, '**density dependent**' factors such as food supply and disease will become more prevalent in limiting growth as the population size and density increase.

When the biotic potential is greater than environmental resistance, the population will grow. From these ecological concepts, model 'growth curves' can be plotted to show how the population of a species grows over time (Figure 10.54).

Key terms

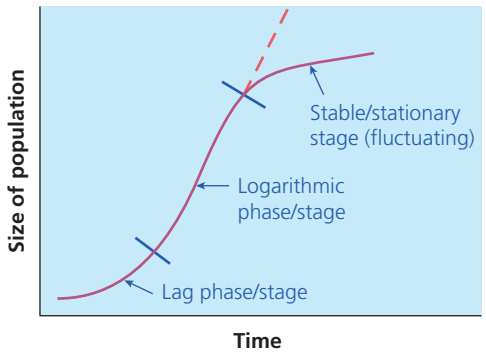
Biotic potential – In population ecology, the natural reproductive potential of the species.

Environmental resistance – A term used in population ecology to explain mortality rates controlled by environmental factors that prevent survival, for example, disease or shortage of food. Sometimes known as 'limiting factors'.

Population ecology applied to the human population

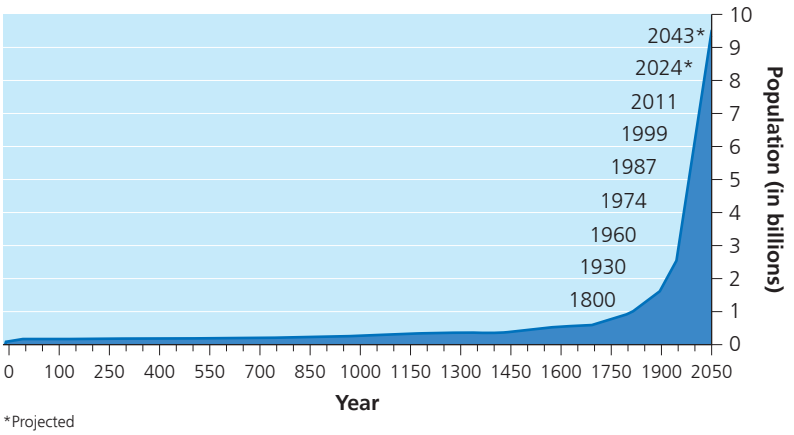
For the greater part of human history, human population size remained relatively stable. Although birth rates were high and family sizes were larger in the past, most parents were only survived into adulthood by two of their children, who effectively replaced them in the population. The reason for the stabilisation of population was therefore the high death rates and low life expectancy brought about by the many diseases and lack of food security that threatened human lives.

Developments in technology improved food supplies in many parts of the world. This, together with improvements in medical science, healthcare and sanitation, have been the main reasons behind the exponential growth in population over the past two hundred years. According to demographer Nicholas Eberstadt, 'We didn't start breeding like rabbits, we stopped dying like flies'. Medicines and vaccines have been developed that combat many of the infectious diseases, such as smallpox, measles and tuberculosis, which were killing people well into the twentieth century. Recognising that improved sanitation would also reduce the spread of water-borne diseases, such as typhoid and cholera, also helped to control the death rate and increase life expectancy. In 1800, the Earth's population was estimated to be around one billion, whereas by 2020, it was estimated to be 7.8 billion, largely due to the control of death rates. In effect, human population has overcome much of the 'environmental resistance' that was limiting growth (Figure 10.55). For the last two hundred years, human population has been through a 'log phase' of exponential growth. World population is still growing rapidly though the rate of growth is starting to slow down.



When **biotic potential** > **environmental resistance** = Rapid population growth during **log phase**

Figure 10.54 Model growth curve showing how populations grow over time



*Projected

Figure 10.55 J curve of human population growth showing exponential growth since 1800

Overpopulation, underpopulation and optimum population

Figure 10.56 A crowded Oshodi market in Lagos, Nigeria – the country's population is expected to soar from its current 210 million to 800 million by 2100. How will this impact food resources, health and well-being?

Optimum population

The relationship between human population and the resources available to support that population is a complex one, which can be addressed at all scales. An ideal balance between population and resources is known as **optimum population** (Figures 10.57 and 10.58). This is a theoretical balance that refers to a population size in an area which, working with all resources, will give the highest standard of living for the people of that area. It maximises the income per capita and is a balance which most nations will try to achieve through direct or indirect population or resource management. It is constantly changing, for example, the development of new technology may increase the resources available.

Overpopulation

Overpopulation exists when there are too many people in the area relative to the available resources, putting pressure on those resources. It means that a continued increase in population will reduce the average standard of living for all. Conversely it implies that with no advances in technology and no increase in resources, only a fall in population would increase living standards. It is a mistake to assume that the concept only relates to the idea of population density, suggesting that areas or countries which are densely populated are overpopulated. This is often incorrect because the measure applies only to the relationship between population size and available resources. Thus a densely populated country such as the Netherlands (511 people per km^2) can enjoy a relatively high average standard of living, while less densely populated Mali (16 people per km^2) struggles to provide the basic needs for its population because of a poor resource base.

Underpopulation

Underpopulation occurs when there are too few people to use the resources efficiently for a given level of technology. It suggests that an increase in population would mean a more effective use of resources and increased living standards for all. Developed countries with access to resources and advanced technology but with relatively small populations such as Canada, Australia and Norway are typical examples of countries which may be regarded as underpopulated and with potential to support larger populations.

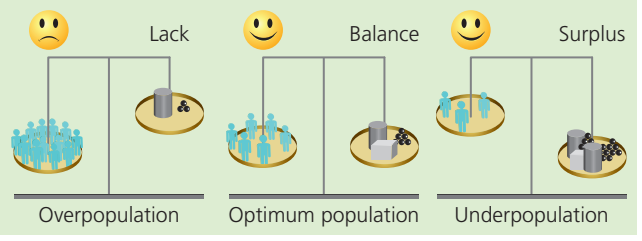


Figure 10.57 The balance between population size and available resources.

Table 10.13 Characteristics of overpopulation and underpopulation compared

Overpopulation	Underpopulation
Low per capita income/ poverty	High per capita incomes (not maximised)
Unemployment/ underemployment	Low unemployment
Outward migration	Inward migration
Insufficient food, minerals and energy resources	Good living conditions
Prone to natural disasters	High levels of technology
War, conflicts, ethnic/tribal tensions	

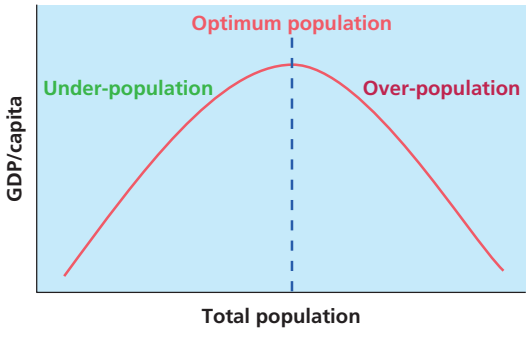


Figure 10.58 Relationship between population size and standard of living.

Implications of population size for the balance between population and resources

Continued global population growth will put pressure on the Earth's resources and its environment. This situation will affect us all in some way. It raises the questions, 'at what point will breaking-point be reached?' and 'how many people can live on the planet sustainably?'.

Carrying capacity and ecological footprint

Carrying capacity is an ecological concept usually applied to animal and plant populations. It is a calculation of how large a population any given environment can support. Population ecologist Professor William Rees has applied the concept to humans and suggests that the measure of carrying capacity can be flexible depending on the average lifestyle of the population concerned. Together with his PhD student, Mathis Wackernagel, Rees developed the concept of an **ecological footprint**. Their arguments about carrying capacity hinge on the varying levels of consumption of resources in different parts of the world. For example, the Earth could support a larger population of 10 billion at a modest level of consumption but it will have a much lower carrying capacity if we use and waste more resources in order to live the more comfortable lifestyles enjoyed in rich industrialised countries.

These ideas suggest that Earth has bio-productive limits; it can only accommodate so much consumption of its resources. The basis for Rees's calculation of carrying capacity is measuring the '**total productive bio-capacity** of the Earth' and dividing this by the total population.

Key terms

Biosphere – The biological component of Earth systems (the parts of the Earth where life exists – also known as the ecosphere).

Carrying capacity – The maximum population size that an area or environment can sustain indefinitely.

Ecological footprint – A measure in global hectares (gha) of the land/water area needed to produce the resources that humans (individually or as society/ies) demand, and to assimilate the waste generated from this production.

Overshoot – In ecological terms this refers to a point when a population's consumption of resources exceeds the long term capacity of the environment to regenerate the resources being consumed.

Calculating carrying capacity and ecological footprint

$$\frac{\text{Total productive biocapacity}}{\text{Total population}} = \text{Global hectares (gha) available per person}$$

Total productive bio-capacity = the total biologically productive area of Earth's natural systems (measured in gha) available to provide the food, water, energy and other resources that we use as humans and to absorb our waste.

One global hectare (gha) = a unit of measurement which represents the average productivity of all biologically productive areas (cropland, forests, fishing grounds etc.) on Earth in a given year.

Global hectares per person (in this calculation) = the average amount of global hectares (gha) available to each person to provide for their consumption of resources and assimilation of waste.

The calculation gives us the global hectares **available** for each person on the planet. This is the ecological footprint that we are each entitled to, with a population nearly 8 billion. From the average ecological footprint available to each individual we can calculate how many people the Earth can support (carrying capacity) at different levels of consumption. The ecological footprint is a measure of the demand each of us places on the Earth's **biosphere** and is a useful tool to assess our pressure on the planet's resources. One of the largest components of the ecological footprint is each individual's carbon footprint, which is in essence a waste product that needs to be assimilated. Rees's ideas challenge us to ask 'how much of the Earth's surface is needed to support me in the lifestyle to which I am accustomed?' There are a number of ecological footprint calculators available on the internet that will give an approximate answer to this question. For example, try answering the questions at www.footprintcalculator.org.

In 2019, there were 12.2 billion hectares of biologically productive land and water on Earth. If this total productive biocapacity is shared **evenly** by the global population of 7.7 billion, each person would have nearly **1.6 global hectares** (gha) available to them. However, ecological footprint values vary enormously and are distributed unevenly around the world. Although in some parts of the world, average footprints are at or below the 1.6 gha level, most regions are well above this in terms of consumption. Figure 10.59 gives an impression by

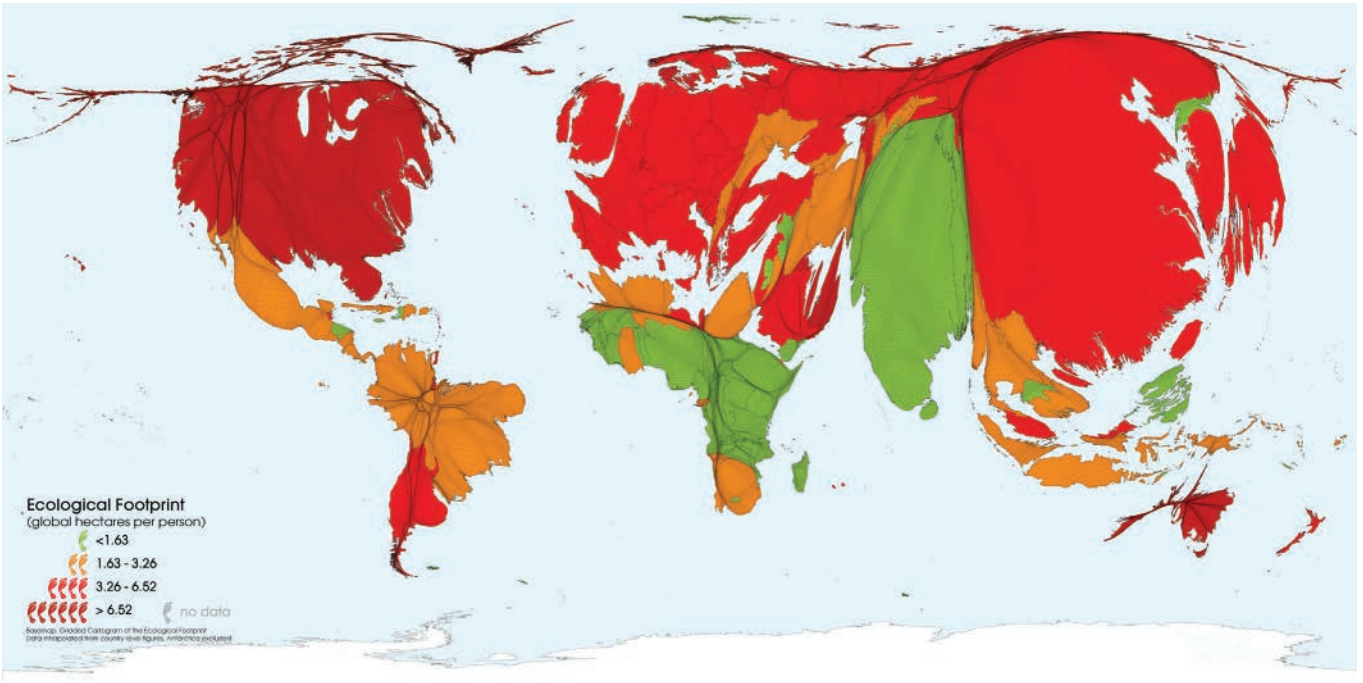


Figure 10.59 Distorted map showing global distribution of ecological footprints, 2019 (based on consumption levels)

distortion of size of the total ecological footprint of countries globally. (The colour shading shows each country's average ecological footprint per person.)

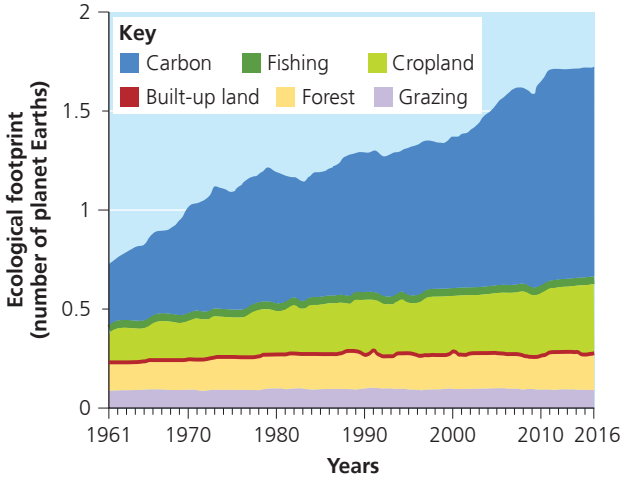


Figure 10.60 Compound graph showing components of ecological footprint from 1961–2016

Skills focus

Using the compound line graph in Figure 10.60:

a State the amount of footprint used i) by carbon in 1975 and 2011 and ii) by cropland in the same two years.

b Estimate the percentage of the total ecological footprint taken up i) by carbon in 1961 and in 2016 and ii) by cropland in the same two years.

Ecological footprint accounting represents human consumption of resources and generation of waste. Figure 10.60 gives an idea of the component parts of the ecological footprint and how it has changed over the past 60 years. Carbon waste is the largest and fastest growing component.

Consumption and waste production rates vary considerably from country to country and Figure 10.61 (page 526) shows average ecological footprint values for selected countries across the world in 2016. According to these figures, if all humans had an ecological footprint similar to the average Indian, the Earth could support 10 billion people. However, if everyone lived as we do in the UK, then the Earth could only support 2.8 billion, and only 1.5 billion if we lived the lifestyle of the average US citizen. In general, ecological footprints in many HDEs are slowly decreasing in size because of deindustrialisation, increased environmental awareness and strategies to mitigate consumption and waste (for example, recycling). However, footprints in developing countries such as China and India are rising because of industrialisation and increased consumption (see Table 10.14, page 526).

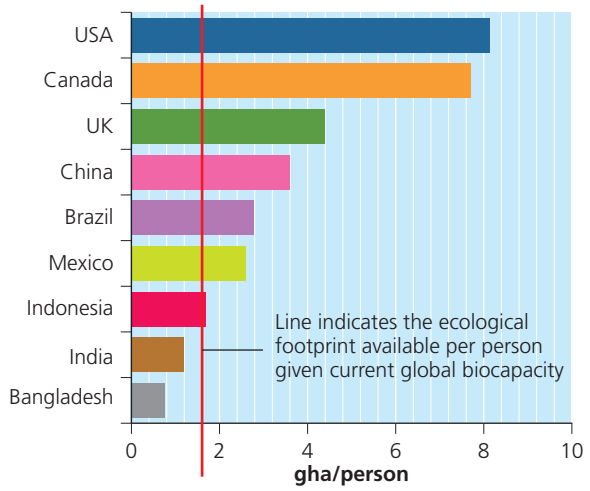


Figure 10.61 Bar graph showing per capita ecological footprints for selected countries, 2016

The productive biocapacity available also varies between countries because of the different land/sea areas they occupy and the productive potential of those areas (for example, Brazil has the highest total biocapacity of any country at 1.8 billion gha). Table 10.14 compares the ecological footprint for six selected countries from Figure 10.61 with the biocapacity available per person within that country, to discover whether the country has a biocapacity reserve or deficit.

Table 10.14 Average ecological footprint per person compared with productive biocapacity per person for six countries from Figure 10.61

Country	Biocapacity available per person (gha)	Ecological footprint per person, 2000 (gha)	Ecological footprint per person, 2016 (gha)	Biocapacity reserve or deficit per person, 2016 (gha)
Bangladesh	0.4	0.6	0.8	- 0.4
Brazil	8.7	3.0	2.8	**+ 5.9**
Canada	15.1	9.0	7.7	**+ 7.4**
China	1.0	1.9	3.6	- 2.6
India	0.4	0.9	1.2	- 0.8
UK	1.1	5.7	4.4	- 3.3

Source: Global Footprint Network

Implications of carrying capacity and ecological footprint

The concept of carrying capacity introduces a 'limit' to population growth, so the exponential 'J curve' growth of human population shown in Figure 10.55 (page 522) needs to be reviewed in terms of what will happen if population size exceeds this limit. Figures for humanity's global ecological footprint are already unsustainable and evidence suggests that we are already in a state of **overshoot**. The implications of current measures of carrying capacity and ecological footprint mean that:

- **humanity uses more resources than the Earth provides** – each year levels of consumption exceed the productive bio-capacity of the planet, not only in terms of its ability to produce but also in its ability to assimilate our pollution and waste
- **global consumption is unevenly distributed** – it favours a few and is inefficient because of excessive waste and unnecessary usage of resources.

According to the think tank the Global Footprint Network (GFN), we are living beyond the means of the planet to sustain us and it estimates that we need 1.7 planets to support our current way of life (Figure 10.60, page 525). The GFN also presents us with the notion of an 'Earth Overshoot Day' each year, designated as the day it is thought we have used the productive biocapacity of the planet for that year; after that day we go into 'ecological debt'. In 2010, the overshoot day was 21 August and by 2019 it was 29 July. However, according to the GFN, the sudden contraction of humanity's ecological footprint because of the COVID-19 pandemic in 2020, meant the overshoot day was three weeks later on 22 August.

The implications for human population will depend on different scenarios:

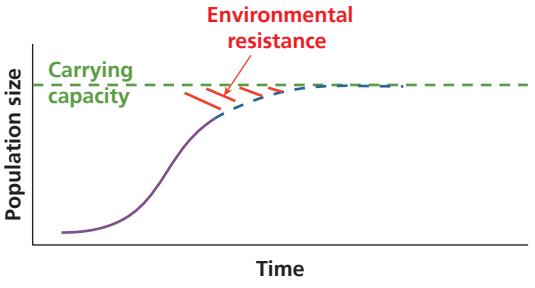
- If we continue to consume resources at existing rates:
  - a number of population 'checks' may continue to occur in the form of famines, water shortages or disease pandemics, which increase mortality rates; this is already evident in poorer countries of sub-Saharan Africa despite the efforts of the UN's SDGs
  - an increase in wars or conflicts over limited resources, which will also increase mortality and may exacerbate the problem by reducing biocapacity (positive feedback)
  - a rapid and significant decrease in consumption of resources to reduce ecological footprints would mean a substantial reduction in the standard of living for most of the world's population, which may not be palatable for those in wealthier developed economies.

These arguments are presented as Malthusian perspectives later in this section (see page 530). It is important to understand that the negative implications of breaching carrying capacity are not related to population size alone. While efforts to control population growth may alleviate some of the pressure on productive biocapacity, the consumption issue is more critical and needs to be addressed, particularly in those countries with growing per capita ecological footprints. The negative environmental implications of enlarging ecological footprints are particularly acute and include:

- climate change, exacerbation of global warming
- more land used for settlement, industry and transport
- degradation of natural ecosystems
- increased threat of species extinction
- over-cultivation and overgrazing reducing land and soil quality
- depletion of fish stocks beyond recovery.

The relationship between carrying capacity and ecological footprint is complex and more positive views of the future are suggested by:

- **evidence that population growth rates, on a global scale, are slowing down** – the view of 'transition' theorists is that the population growth curve is gradually forming as a 'sigmoidal' or 'S' shaped curve as it adjusts more gradually to carrying capacity. This adjustment suggests that as population size increases, the rate of increase declines as more environmental resistance is encountered (for example, dwindling food and water resources). This will eventually lead to an equilibrium population size fixed by carrying capacity (see Figure 10.62)
- **the notion that carrying capacity can be increased** – advances in technology in the past have enabled increases in the productivity of food and access to more resources than were thought available. However, many of the technologies that have increased the carrying capacity have also increased humanity's ecological footprint. Solutions are now needed that can increase carrying capacity but without placing an increased demand on resources; green technologies such as the transition to renewable energy have made some progress in this direction. This is explored in more detail in alternative perspectives on population growth on pages 539–43.



In population ecology this pattern is characteristic of long-lived species with relatively few predators, *e.g. primates*

Limiting factors slow down the growth rate by creating **environmental resistance**

Figure 10.62 The 'sigmoidal' or 'S' curve of population growth

The population, resources and pollution model

At the beginning of this chapter we considered the relationship between humans and their environment; how humans exploit and utilise resources provided which, in turn, has an impact on the environment. The population, resources and pollution (PRP) model (Figure 10.63, page 528) represents a fundamental ecological relationship true to all organisms including humans, and so the model:

- provides an insight into the human–environment interaction
- illustrates several important relationships between the two
- adopts a 'systems' approach – understanding that a consequence of changing one variable in the model is that others will be affected
- promotes 'systems thinking' which is vital to sustainable development
- provides an insight into sustainable solutions – for example, to control resource depletion, population control measures might be adopted; encouraging less demand for resources will reduce levels of pollution, for example, in energy generation
- uses the concepts of positive and negative feedback.

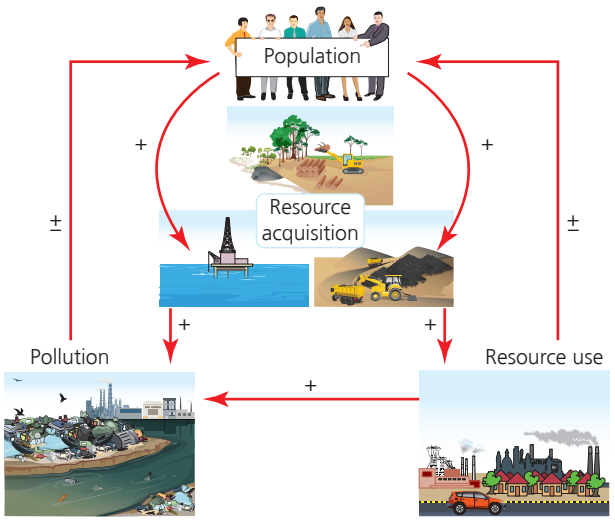


Figure 10.63 The population, resources and pollution model (after Dr D. D. Chiras). The model outlines how all organisms interact with their environment.

Links

There are links to Chapter 1, Water and carbon cycles, specifically Systems frameworks and their application, page 2.

The model shows that populations acquire resources from the environment. This acquisition of resources will alter the existing biotic and abiotic conditions of ecosystems. For example, open-cast coal mines damage and destroy wildlife habitat; they can also lead to deforestation and soil erosion that pollutes nearby streams. On the diagram this is represented by the arrow between resource acquisition and pollution.

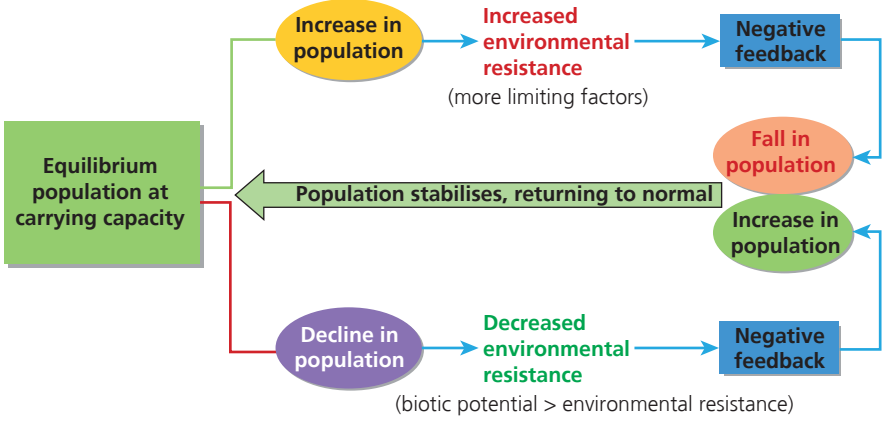
Resources that are extracted from the environment are utilised; for example, coal may be used in power stations to generate electricity, or iron ore deposits are crafted into finished consumer products such as cars. The conversion of raw materials into energy or into finished products results in pollution and this is indicated on the diagram by the arrow connecting resource use to pollution. Pollutants usually contaminate specific ecosystems, such as those on land, in water (aquatic and marine) or in the atmosphere, but evidence suggests that many pollutants cross these boundaries. For example, acid rain is an atmospheric pollutant but when it falls from the sky, acid is deposited, contaminating forests, rivers and lakes.

Positive and negative feedback mechanisms

Positive and negative feedback either enhance or counter changes that occur in a system (see Chapter 1).

- **Positive feedback** enhances or amplifies changes – moving a system away from its equilibrium state and making it more unstable.
- **Negative feedback** is an opposing force which counters any change, regulating the system in a more stable equilibrium.

Negative feedback loops are the 'default' for controlling biological systems such as population growth and regulating homeostasis, as shown in Figure 10.64.



The mortality rate changes depending on whether the population is above or below the carrying capacity – to return the population to the carrying capacity

Figure 10.64 Homeostatic regulation of population growth

These ideas are applied in population ecology and can equally be applied to human population growth. Population regulation uses negative feedback to keep plant and animal populations within the carrying capacity of the environment.

In the PRP model, as population expands, resources become depleted and pollution increases (positive feedbacks). In turn, this disrupts the natural ecosystem balance, reducing productivity and biocapacity, which will lead to food, water and energy insecurity. The negative feedback loops from overuse of resources and increased pollution are of great concern. For example, continued soil erosion from overgrazing, over-cultivation or pollution of land and water from chemical fertilisers or industrial waste will eventually lead to a decline in food production. This will lead to an increase in mortality rates and could have devastating effects on human population size (negative feedback).

The PRP model suggests that negative feedback will be the consequence of continued population growth. As population nears or exceeds the carrying capacity then negative feedback will mean that there are extremes of pollution (including climate change) and insufficient resources so that death rates overtake birth rates, causing a population decline back to a stable equilibrium. This is also known as homeostatic regulation.

The positive feedback loops coming from resource use and pollution back to population on the model indicate that resource acquisition and use, together with the pollution they create, enhance survival and enable human population growth. One way to view positive feedback is to consider agriculture and population together. They can be viewed as being in a **positive feedback spiral**, in that one drives the other with increasing intensity. Population growth has prompted an increase in agricultural yields and production

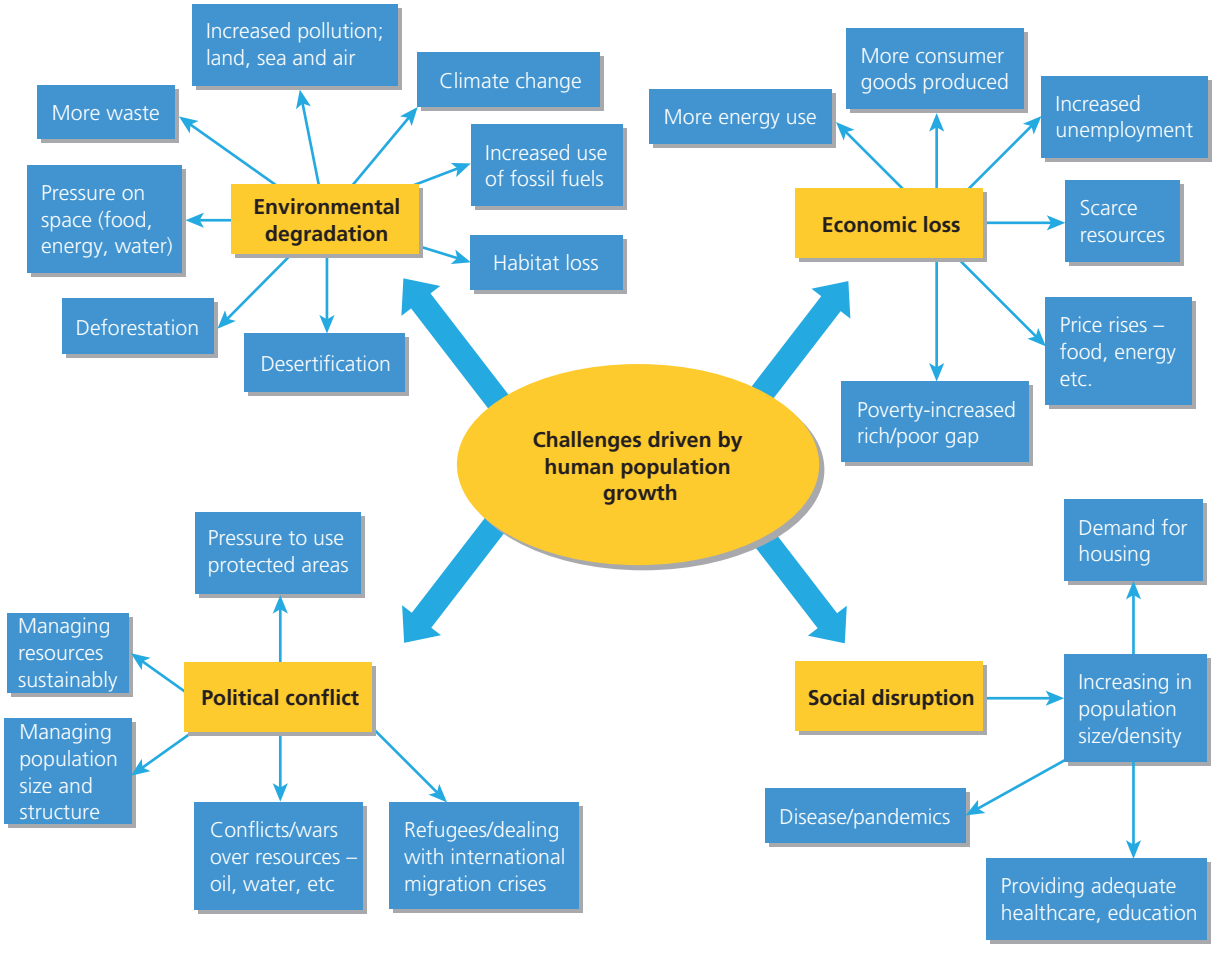


Figure 10.65 Challenges faced by the modern world driven by population growth

through technologies, such as high yield variety crops, which in turn means that more population can be supported and so population continues to grow.

Positive feedback mechanisms may cause serious problems, such as the devastating cycles of depletion and environmental destruction. Eventually, in agricultural production, despite technological advances, soils are depleted, the carbon cycle is disrupted and yields diminish, leading to instability. Another problem is the use of fossil fuel energy (for example, in producing chemical fertilisers using the Haber process), which has increased our capacity to produce and distribute food but is creating a rise in global temperature because of greenhouse gas emissions. The resulting shift in rainfall and drought patterns may devastate crop production in many areas. The negative feedback effect of climate change (resulting from pollution) on population has already been seen in regions of sub-Saharan Africa, with droughts leading to famines in the Sahel in 2010 and in East Africa in 2011. Over 100,000 people died during these events, which are likely to become more frequent occurrences.

Contrasting perspectives on population growth

The negative implications of continued population growth are all too obvious. As countries develop economically and take advantage of their demographic dividend, then it is likely that the consumption of resources will continue to grow. An increasing population with an ever-decreasing supply of resources is an unsustainable position.

In addition to the environmental pressures, Figure 10.65 (page 529) identifies a range of social, economic and political implications. There are, however, some positive implications to consider:

- Increasing education and awareness can influence and change lifestyles, thereby reducing ecological footprints – though more needs to be done.
- Biotechnologies and appropriate technologies could increase the carrying capacity in specific areas and globally.

These contrasting views concerning the balance between population growth and resources are outlined by theorists such as Thomas Malthus and Ester Boserup and their respective followers.

Malthus and neo-Malthusians

Just as human population was starting its unprecedented exponential growth curve around the beginning of the nineteenth century, Thomas Malthus, an English clergyman, published *An Essay on the Principle of Population* in 1798. Malthus was one of the first writers to make an observation about the relationship between humans and resources and he used his observations to predict the future. He pointed out that 'the power of population is infinitely greater than the power in the earth to produce subsistence for man'. In other words, food production cannot increase as rapidly as human reproduction.

The principle behind Malthus's theory is that food production can, at best, only increase arithmetically (so from an index value of 10 to 20, 30, 40 and so on) so at any one time there is a fixed 'carrying capacity' that can only support a given population. However, population growth will occur geometrically (so from a given index value of 1 to 2, then 4, 8, 16, 32 and so on). Eventually the population value will exceed the carrying capacity value and there will be a population 'crash' caused by his predicted catastrophe (Figure 10.66).

Time periods (years)	25	50	75	100	125	150	175	200
Population (index)	1	2	4	8	16	32	64	128
Food supply	1(0)	2(0)	3(0)	4(0)	5(0)	6(0)	7(0)	8(0)

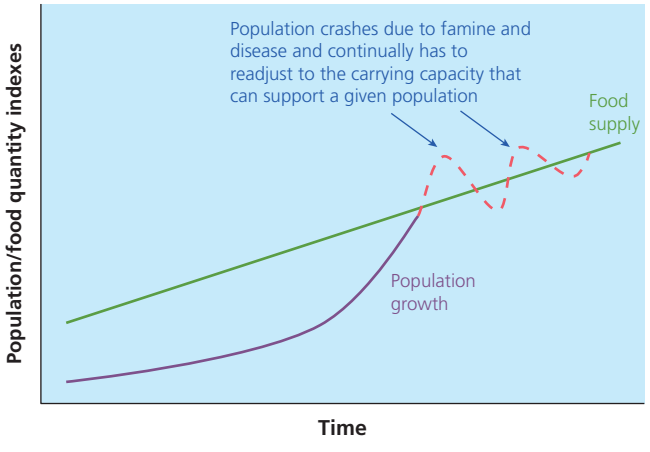


Figure 10.66 Malthusian diagram showing the relationship between population and resources

Malthus's views on the population/resource balance are **pessimistic**. He believed that there is an ultimate limit to how much human population can grow and this limit is determined by the environment and its ability

to produce food. He also believed that the balance is maintained by various checks:

- **Positive checks:** Increased deaths through war, famine and disease; increased incidence of abortion, infanticide, etc. (negative feedback).
- **Negative checks:** He advocated moral restraint, for example, celibacy and later marriages (his theory was written before the advent of modern contraception).

Malthus predicted that unless population growth was controlled, life would end in misery, leading to what became known as a 'Malthusian catastrophe'.

Although Malthus's ideas originated over two hundred years ago, they gained much support during the 1960s as global population experienced its most rapid growth phase. Demographers supporting these ideas are known as **neo-Malthusians** and include, among others, US biologist Paul Ehrlich and a global think tank known as the Club of Rome, which formed in 1968 and still see their work as relevant today.

Ehrlich

In 1968, Ehrlich authored a controversial book, *The Population Bomb*. The book predicted the mass starvation of millions by the 1970s, and also advocated coercive control of population and consumption. In response to criticisms about his predictions of doom, Ehrlich claims that humanity merely deferred the disaster by the use of more intensive agricultural techniques, such as those introduced by the Green Revolution. Ehrlich claims that many of the problems resulting from population growth have materialised, such as:

- at least 600 million would be undernourished
- increase in disease occurrence
- global warming and climate change
- increased urbanisation
- overfishing
- loss of biodiversity.

Alternative perspectives

There are alternative, more 'cornucopian' views of our future such as those expressed by economists Ester Boserup (Danish) and Julian Simon (American). Cornucopian futurists believe that the continued progress and provision of materials for mankind can be met by continued advances in technology.

Boserup

Boserup, like Malthus, recognised that environments have limits that restrict production. Unlike the 'determinism' (limits determined by the environment) claimed by Malthusians, however, she suggested a 'possibilistic' philosophy, where human ingenuity could expand the carrying capacity and enable it to extend upwards in line with population growth (Figure 10.67).

Time periods (years)	25	50	75	100	125	150	175	200
Population index value	1	2	4	8	16	32	64	128
Food supply	1(0)	2(0)	3(0)	4(0)	5(0)	6(0)	7(0)	21(0)

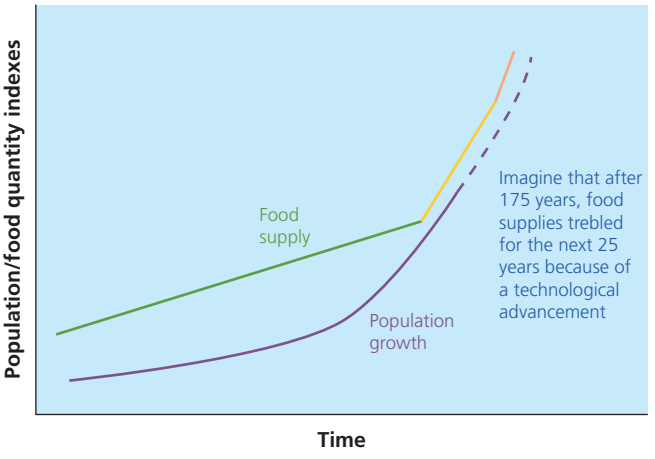


Figure 10.67 Boserup's perspective of the population–resource relationship

In 1965 Boserup suggested that population growth often stimulates innovation and limits can be increased by improvements in technology. These would increase resource availability so much that the carrying capacity could sustain a much larger population than Malthus had suggested. It seems that our history is such that when there is a crisis for survival in an area, we find a solution that eventually averts that crisis, thus proving that 'necessity is the mother of invention'. This is shown in agricultural innovations such as:

- intensification of farming in Europe in response to food shortages after the Second World War
- the 'Green Revolution' in Mexico, South and South East Asia
- recent developments in biotechnologies including genetic modification (GM).

Table 10.15 Summary of contrasting perspectives on population growth

	Neo-Malthusian perspectives	'Cornucopian' perspectives
Philosophic basis	Deterministic – the environment imposes physical limits and determines population size	Possibilistic – human ingenuity and creativity can extend the physical limits
Carrying capacity	Fixed – or at best grows at a slower rate than population growth	Can be expanded by technological innovation that increases biocapacity
Population size	Limited by the carrying capacity of the environment	Not limited – population growth is enabled by technology and ingenuity
Resource use and consumption	Increased ecological footprints, greater resource consumption leading to depletion Output per person declines, rising prices for goods	As certain resources become depleted, they are replaced by others Switching from non-renewable to renewable; more repair, re-use and recycling
Pollution and environment	Environmental degradation, increased levels of pollution and loss of biodiversity	Increased awareness of damage to environment – mitigation introduced
Consequences for population growth	Famine and disease cause severe and rapid population crashes (negative feedback) Population control measures need to be introduced	Humanity adapts to changing conditions Population growth slows down 'naturally' with economic and social development (transition theory)
Evidence to support this perspective	>800 million undernourished Frequent droughts and famines, especially in sub-Saharan Africa Wars and potential conflicts fought over access to food, water and energy resources Global warming and climate change	Increases in food production through intensification and technology (e.g. GR and GM) Transition to new technologies (e.g., renewable energy) Adoption of appropriate technologies designed to be more sustainable

Simon

Julian Simon, an academic contemporary of Ehrlich, argued that every important long-term measure of human material welfare shows improvement in all parts of the world, demonstrating that despite rapid population growth, humans are measurably better off. He used facts to support the idea that raw materials have become less scarce and that the relative costs in accessing them have also reduced. Using other facts to support his claims, Simon suggests:

- the air in rich countries is safer to breathe
- water cleanliness has improved
- the condition of cropland is improving rather than worsening
- historically, food production increases have always been at least as fast as population increases.

In Simon's book *The Ultimate Resource* (1981), he argues that the only resource that the Earth was running short of is people, even though there are more of us. There are many challenges to his arguments and more recent trends suggest flaws in his claims, but the underlying premise, that as population has grown resources have increased, is largely supported by fact.

Evidence and arguments are offered to support both perspectives on population growth. As students of population and the environment, you should research more on this and formulate your own view.

Review questions

18 Explain why overpopulation is not necessarily related to population density.

19 Compare the concepts of carrying capacity and ecological footprint.

20 Outline both positive and negative feedback loops from resource use to population size, as indicated by the population, resources and pollution model.

21 To what extent do you agree with the neo-Malthusian perspective on population growth?

10.6 Global population futures

There has been much speculation about what will happen to global population size during the twenty-first century following its rapid increase in the twentieth century. With more people on the planet the demand for resources will undoubtedly increase, especially if people in growing populations become more affluent and desire a higher standard of living.

At the same time, the Earth's physical environment is undergoing change, particularly with regard to its atmosphere and climate, which will impact on human health.

Health impacts of global environmental change

Ozone depletion

Depletion of ozone (otherwise known as triatomic oxygen or O_3) is a relatively recent phenomenon in human history and is known to be caused by halogenated chemicals such as chlorofluorocarbons (CFCs) used in refrigeration, insulation and spray-can propellants. These chemicals react with and destroy the ozone molecule in the extreme cold of the polar stratosphere. This destruction of ozone occurs mainly in late winter and early spring.

The stratospheric ozone layer utilises and filters out much of the incoming solar ultraviolet radiation (UVR), especially the biologically more damaging, shorter wavelength UVR. If the ozone becomes depleted, then more UV rays will reach the Earth, which can have serious impacts on humans, flora and fauna. Extended exposure to UV rays has harmful impacts on human health including increased prevalence of skin cancers, cataract formation, increases in other eye diseases and an increase in infectious diseases due to excessive UV radiation weakening the human immune system.

Skin cancer

Most skin cancers are caused by exposure to the Sun. This may be long-term exposure, or short periods of intense sun exposure and burning. The ultraviolet light in sunlight damages the DNA in the skin cells. There are two main types of skin cancer: melanoma and non-melanoma. Non-melanoma cancer (the most common of which are basal cell carcinomas) is the fifth most commonly occurring cancer globally, with over one million diagnoses in 2018. Malignant melanoma is less common but a more dangerous cancer because it is more likely to spread to other organs if not treated early. Incidence rates of malignant melanoma have risen faster than most other types of cancer over the past 50 years to become the nineteenth most common, with over 300,000 cases worldwide in 2018. Melanomas are more likely to affect:

- fair-skinned populations
- the elderly
- those living in lower latitudes
- males – incidence rates are slight higher among men; this is thought to be related to relatively more men being employed in outdoor occupations.

Australia and New Zealand have the highest rates of skin cancer in the world at 33.6 and 33.3 per 100,000 respectively (see Figure 10.68). The combination of the predominantly light-skinned population, lower latitudes with high levels of ultraviolet radiation and cultural emphasis on outdoor activities has contributed to this problem.

- At least two out of three Australians will be diagnosed with one form of skin cancer by the age of 70.
- Around 2,000 Australians die from skin cancer each year.
- Of all the new cancers diagnosed in Australia each year, approximately 80 per cent are skin cancers, including 11 per cent that are malignant melanomas. The melanomas contribute to 3 per cent of all deaths from cancer.
- It has been calculated that treating all skin cancers cost the Australian health system around A U$ 900 million in 2017 (diagnosis, treatment and pathology).

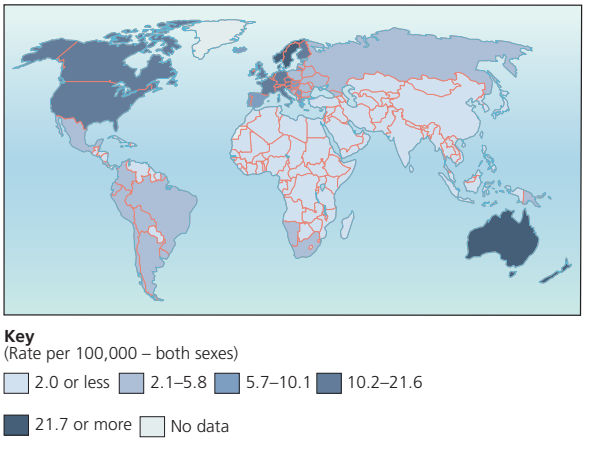


Figure 10.68 Global incidence of melanoma skin cancer, 2018
Source: American Cancer Society

Cataracts

Cataracts are a form of eye damage which causes a loss of transparency in the lens of the eye, leading

to cloudiness of vision and can eventually result in blindness. There are different types of cataracts and a number of factors can lead to their formation, including their being a natural part of the ageing process or as a result of diabetes. Smoke from burning fuelwood and cigarettes as well as poor nutrition (particularly a lack of vitamin A) can also cause cataracts.

Exposure to UV radiation can damage different parts of the eye, including the lens, cornea, retina and conjunctiva, and research shows it appears to be a major risk factor for cataract development. Cataracts are the leading cause of blindness in the world, especially in LDEs where there is often lack of access to surgery to remove cataracts. Proteins in the eye's lens unravel, tangle and accumulate pigments that cloud the lens and eventually lead to blindness. Increased ozone depletion means anyone who spends a lot of their time outdoors is at risk of eye problems from UV radiation, but the extent of the risk depends on a number of factors including:

- **Geographic location:** UV levels are greater in lower latitudes within or near the tropics.
- **Altitude:** UV levels are greater at higher altitudes.
- **Time of day:** UV levels are greater when the Sun is high in the sky, typically from 10 a.m. to 2 p.m.
- **Setting:** UV levels are greater in wide open spaces, especially when highly reflective surfaces are present, like snow and sand. Exposure is less likely in urban settings, where tall buildings shade the streets.

UV levels are not significantly affected by cloud cover and risk of UV exposure can be quite high even on hazy or overcast days. This is because UV is invisible radiation, not visible light, and can penetrate clouds. Wearing good quality sunglasses blocks harmful UV radiation and close fitting 'wrap around' styles are better as they prevent sunlight entering the peripheries of the glasses.

The only effective treatment for cataracts is surgery to remove the cloudy lens and replace it with a clear plastic intraocular lens implant. Although curable, cataracts severely diminish the eyesight of millions of people in both the developing and developed world. A combination of variables means that those living in rural areas in developing countries are disproportionately affected by UV-induced cataracts and they are less able to afford prevention or treatment for the problem. A sustained 10 per cent depletion of the ozone layer is expected to result in nearly 2 million new cases of cataracts globally each year.

Climate change

A change in climatic conditions can have three kinds of health impacts:

- Relatively direct impacts, usually caused by increases in the frequency or severity of extreme weather events such as:
  - storms which increase the risk of dangerous flooding, high winds, and other threats
  - warmer average temperatures leading to hotter days and more frequent and longer heat waves.
- Consequences of environmental change and ecological disruption that occur:
  - increased concentrations of unhealthy air and water pollutants
  - changes in temperature and precipitation patterns could exacerbate the spread of some vector-borne diseases.
- Consequences associated with demoralised populations in the wake of climate-induced economic dislocation, environmental decline, and conflict situations. For example:
  - greater frequency of infectious disease epidemics following floods and storms
  - substantial traumatic, nutritional or psychological health effects following population displacement from sea level rise or increased storm activity.

Although global warming may bring some localised benefits, such as fewer winter deaths in temperate climates and increased food production in certain areas, the WHO suggests that the overall health effects of a changing climate are likely to be overwhelmingly negative (Figure 10.69). Between 2030 and 2050, they estimate that there will be approximately 250,000 additional deaths per year from malnutrition, malaria, diarrhoea and heat stress.

For each potential impact of climate change, certain groups will be particularly vulnerable. Vulnerability depends on factors such as population density, level of economic development, food availability, income level

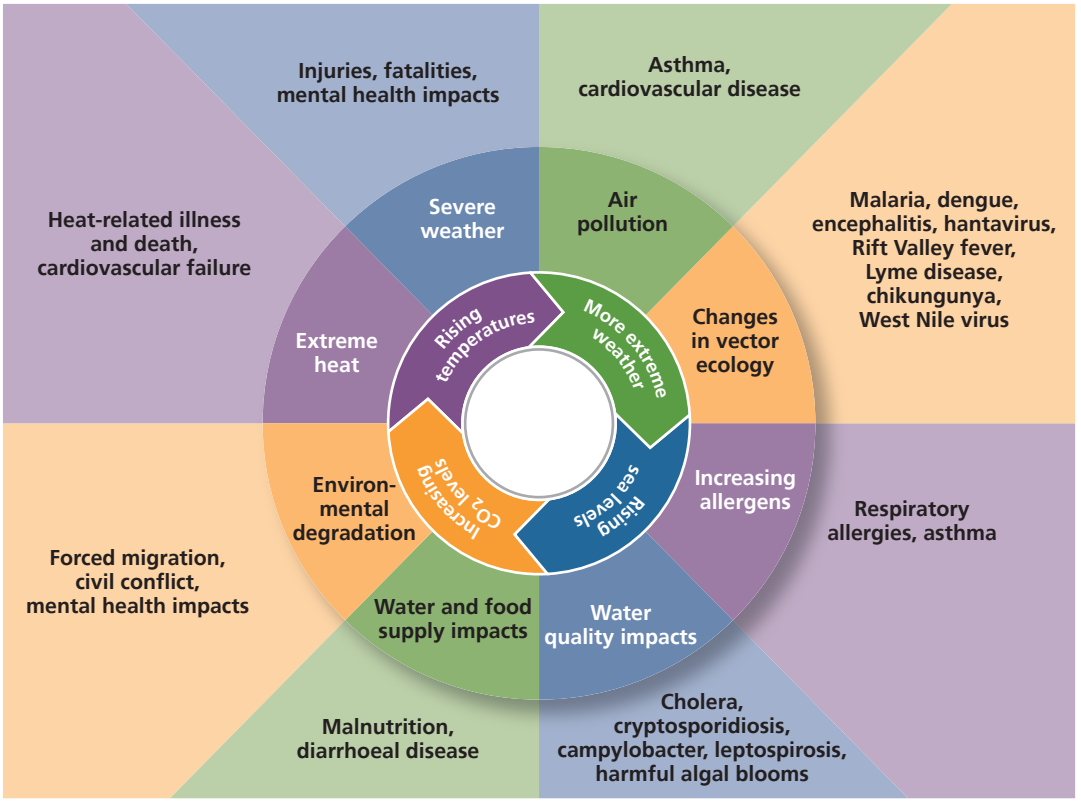


Figure 10.69 Impact of climate change on human health

Source: US Department of Health and Human Services

and distribution, local environmental conditions, pre-existing health status and the quality and availability of public healthcare. For instance, those most at risk of being harmed by thermal extremes include the elderly and the poor.

Thermal stress

One of the potential effects of human-induced climate change is a change in mortality related to thermal stress. Research by the Intergovernmental Panel on Climate Change (IPCC) concludes that climate change will cause increased heat-related mortality and morbidity but decreased cold-related mortality in temperate countries. A change in temperature at both extremes affects both cardiovascular and respiratory mortality.

Heatwaves

The main risks posed by a heatwave are dehydration, overheating, heat exhaustion and heatstroke. The very young, the elderly and the seriously ill are the groups who are particularly at risk of health problems when the weather is very hot. Northern and western Europe are experiencing more frequent heatwaves. During August 2003 in France and Spain, unprecedented high temperatures lasted for three weeks and resulted in 15,000 excess deaths. The WHO reported that there were 70,000 excess deaths across the whole of Europe during the same period. Nearly 900 excess deaths were caused by summer heatwaves in the UK in 2019, 320 of those in a six-day heatwave in late August. In both cases the elderly were disproportionately affected.

Heatwaves can lead directly or indirectly to other health risks. These include:

- photochemical smogs, caused by UV light reacting with nitrogen dioxide in the atmosphere to form ground level ozone, nitric acids and other pollutants. These can cause eye irritation and respiratory problems
- the growth of blue-green algae in water courses, which can cause problems for aquatic life, including fish, as well as toxic algal blooms, causing problems for public recreational water activities
- health and environmental problems including odour, dust and vermin infestation. Measures may be necessary to mitigate these problems, including more frequent waste collections and extra pollution controls at landfills and other waste treatment facilities
- wildfires
- water shortages.

Heatwaves have a much bigger health impact in cities than in surrounding suburban and rural areas. The impact on mortality from heat stress may be more significant in developing-country cities (such as Mexico City or New Delhi) where populations are especially vulnerable as they lack the resources to deal with heatwaves.

Cold spells

In many temperate countries, there is clear seasonal variation in mortality. Death rates in winter can be as much as 25 per cent higher than in summer. Extremes of cold impact severely on those suffering with both cardiovascular and respiratory diseases and lead to an increase in mortality. Over 34 per cent of excess winter deaths in the UK were caused by respiratory diseases in 2017–18. However, annual outbreaks of winter diseases such as influenza, which have a large influence on winter mortality, are not strongly associated with colder temperatures.

Climate change is likely to bring milder winters in temperate regions. Results from research indicate that, for most of the cities studied, global climate change is likely to lead to a reduction in mortality rates due to decreasing winter mortality. This effect is most pronounced for cardiovascular mortality in elderly people in cities which experience temperate or cold climates at present. There is conflicting evidence as to whether the decreases in winter mortality are greater or less than the increase in summer deaths due to heatwaves. Severe cold is thought to have a larger impact on mortality than extreme heat but other factors affecting winter mortality suggest that the negative impacts of more frequent heatwaves outweigh the benefits of fewer cold days. The net impact on mortality is likely to vary between populations.

Emergence and changing distribution of vector-borne diseases

Important determinants of vector-borne disease transmission include:

- vector survival and reproduction
- the vector's biting rate
- the pathogen's incubation rate within the vector organism.

Vectors, pathogens and hosts each survive and reproduce within a range of optimal climatic conditions. Temperature and precipitation are the most important, while altitude, wind and daylight duration are also important. Any changes in temperature and rainfall regimes resulting from global climate change will alter the geographical distribution of optimal conditions for most vectors.

The once limited geographic ranges of many vector-borne diseases are expanding, spurred largely by anthropogenic factors. By 2100 it is estimated that average global temperatures will have risen by 1.5–5.0°C, increasing the likelihood of many vector-borne diseases in new areas. The greatest effect on transmission is likely to be observed at the extremes of the range of temperatures at which transmission occurs. For many diseases these lie in the range 14–18°C at the lower end and about 35–40°C at the upper end.

Malaria and dengue

Malaria and dengue fever are the most important vector-borne diseases in the tropics and subtropics. Populations living at the present margins of malaria and dengue, without effective primary healthcare, will be the most susceptible if these diseases expand their geographic range in a warmer world. For example, the *Aedes* mosquito vector of dengue (which carries other viruses including yellow fever and Zika) is sensitive to climate conditions and studies suggest that climate change could expose an additional one billion people to dengue transmission by the end of the century.

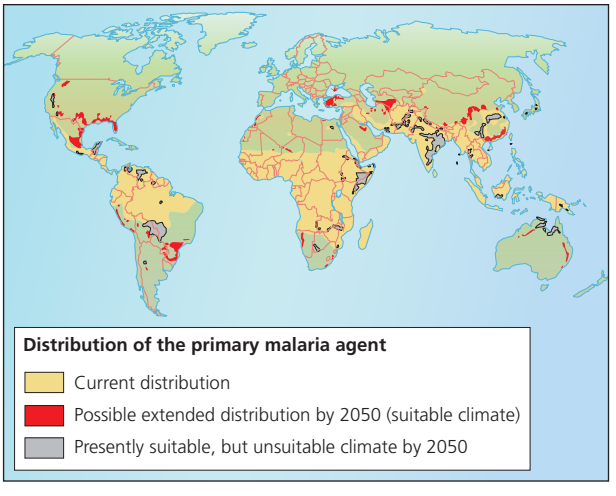


Figure 10.70 Climate change and malaria

Figure 10.70 shows that malaria is likely to spread by 2050, based on the Hadley Centre climate change model's 'high' scenario. Areas shown in yellow indicate the current distribution of malaria. Areas shown in red indicate areas where climate will be conducive to malaria by 2050. Other areas may become free of malaria as climate changes. Climatic anomalies

associated with the El Nino–Southern Oscillation phenomenon, resulting in drought and floods, are also expected to increase in frequency and intensity and have been linked to outbreaks of malaria in Africa, Asia and South America.

Lyme disease

Lyme disease (*Lyme borreliosis*) is the most common vector-borne disease in temperate climates of the northern hemisphere, including the USA and Europe. The *Borellia* bacteria is transmitted to humans by the bite of infected deer ticks of the *Ixodes* genus. It is an emerging vector-borne disease thought to be associated with warmer and more humid conditions.

West Nile Virus

West Nile Virus (WNV) is an emergent disease transmitted by the *Culex* species of mosquito, which has a different climatic tolerance range and can survive in more temperate regions. Human infections attributable to WNV have been reported in many countries in the world for over 50 years. Since 1997 it has spread widely and in 1999 the virus reached New York, resulting in a large and dramatic outbreak that spread throughout the USA in the following years. Since its introduction into the USA, the virus has spread and is now widely established from Canada to Venezuela. The WNV outbreak in the USA (1999–2010) highlighted the fact that establishment of vector-borne pathogens outside their current habitat represents a serious danger to the world.

Zika virus

Zika virus infection is caused by the bite of an infected *Aedes* mosquito, usually causing rash, mild fever, conjunctivitis, and muscle pain. Another emergent disease, the virus was isolated in 1947 in the Zika forest in Uganda. It remained mainly in Africa (with only sporadic outbreaks in Asia) until 2014 when Chile notified the WHO of the virus on Easter Island. In 2015, Zika spread into South and Central America and the Caribbean. After peaking in the Americas in 2016 it has since declined there. In 2019, the WHO reported 87 countries with evidence of mosquito-borne transmission. These countries were located in Africa, the Americas, South East Asia and the Western Pacific.

Agricultural productivity

The current impacts of climate change on agriculture (and adaptations to these impacts) are discussed earlier in this chapter (pages 473–5), but it is important to examine how climate change is predicted to affect productivity into the future (see Figure 10.71 on page 539).

Direct impacts

Higher growing season temperatures will significantly affect agricultural yields, farm incomes and food security. There will be gains and losses depending on the location of the growing region. Some gains in productivity include:

- In mid and high latitudes, the crop yields are projected to increase and extend northwards, especially for cereals and cool season seed crops, such as oil seed rape.
- Crops grown in lower latitudes, such as maize, sunflower and soya beans, could become viable further north and at higher altitudes; yields could increase by as much as 30 per cent by the 2050s.
- There are potentially large gains in agricultural land for regions such as Russia, owing to longer planting seasons and more favourable growing conditions, amounting to a 64 per cent increase over 245 million hectares by the 2080s. There are also concerns that these gains will be tempered by an increase in drought conditions, rendering extensive irrigation programmes necessary to benefit from the warmer temperatures.

Without adaptations by farmers, climate change may not necessarily confer benefits to productivity. For example, an increase in the mean growing season temperature will bring forward the harvest time of current varieties and possibly reduce final yield without adaptation to a longer growing season. There are also likely to be losses in productivity:

- In areas where temperatures are already close to the physiological limits for crops, such as seasonally arid regions, higher temperatures will be detrimental, increasing the heat stress on crops and water loss by evaporation. Significant modifications to crop strains and/or farming methods may be necessary if agricultural production is to be feasible in such conditions.
- Varying rainfall patterns will have a significant impact on productivity. The impact of climate change on regional precipitation is difficult to forecast but there is confidence in projections of a general increase in high-latitude rainfall, especially

in winter, and an overall decrease in many parts of the tropics and subtropics.

In conclusion, because different crops show different sensitivities to temperature change there are still uncertainties about yields for given levels of global warming. As an example, it is thought that a 2°C warming in the mid-latitudes could increase wheat production there by 10 per cent, whereas at low latitudes the same amount of warming may decrease yields by around the same amount.

Indirect impacts

Climate change may also impact indirectly and adversely on crops through effects such as:

- **Pests and diseases:** Indications are that pests such as aphids and weevil larvae respond positively to higher levels of CO_2. Increased temperatures also reduce the overwintering mortality of many pests enabling earlier and more widespread dispersion.
- **Changes in water availability owing to distant climate changes:** Water for irrigation is often extracted from rivers which depend upon distant climatic conditions. For example, agriculture along the Nile in Egypt depends on rainfall in the upper reaches of the river in the Ethiopian Highlands or Rwandan mountains.
- **Sea level rise:** Vulnerability of crop productivity is greatest where a large sea-level rise occurs in conjunction with low-lying coastal agriculture. It may result in saltwater incursion into aquifers used for irrigation and soil salinisation, damaging fertile farmland near to river deltas.

Nutritional standards

The impact of climate change on global food production will have uncertain and varying consequences for human health and nutrition. It can be viewed from different perspectives depending on levels of development.

Developed regions

Increasing food prices may lower the nutritional quality of dietary intakes, exacerbate obesity, and amplify health inequalities. Changes in agricultural production resulting from new crop and livestock species may lead to use of different pesticides and veterinary medicines, and in turn affect the transfer mechanisms through which contaminants move from the environment into food. This will have implications for food safety and the nutritional content of food.

The complementary promotion of healthier diets and climate change mitigation may increase consumption of foods whose production reduces greenhouse gas emissions. For example, reducing red meat consumption will have positive health effects, including reducing levels of saturated fat in the diet and lowering risk of cancers in the digestive system (though it reduces iron and zinc intake). It has the additional benefits of freeing more arable land to grow crops for human consumption.

Developed countries may be able to adapt to the food safety consequences of climate change, although the ability to respond to nutritional challenges is less certain.

Developing regions

Nations experiencing rapid economic growth often go through a 'nutrition transition' where increased affluence guides the population to a more 'westernised' diet. China, in particular, has seen a significant shift from primarily cereal consumption (mainly rice) to more meat consumption. This will have mixed nutritional results for populations and though it has not come as a result of climate change, it will exacerbate the effects through 'positive feedback'. Increased livestock production will increase methane emissions, require more land usage for fodder production and so more clearance of forests.

Least developed regions

These regions are the most vulnerable to loss of production due to climate change. A study by UNICEF found clear and alarming links between climate change and the nutritional status and migration patterns of populations. The predictability of rainfall is critical for rain-fed agriculture and climate change has given an unpredictable combination of extended periods of rainfall (leading to floods) and droughts, both of which have resulted in crop failure. When crops fail, farmers may sell livestock as an economic default, which can lead to iron and zinc deficiencies in the diet. Another important aspect of climate-induced crop failure is an over-reliance on fewer crops which are less likely to fail. Exclusive diets of particular crops lead to specific forms of malnutrition.

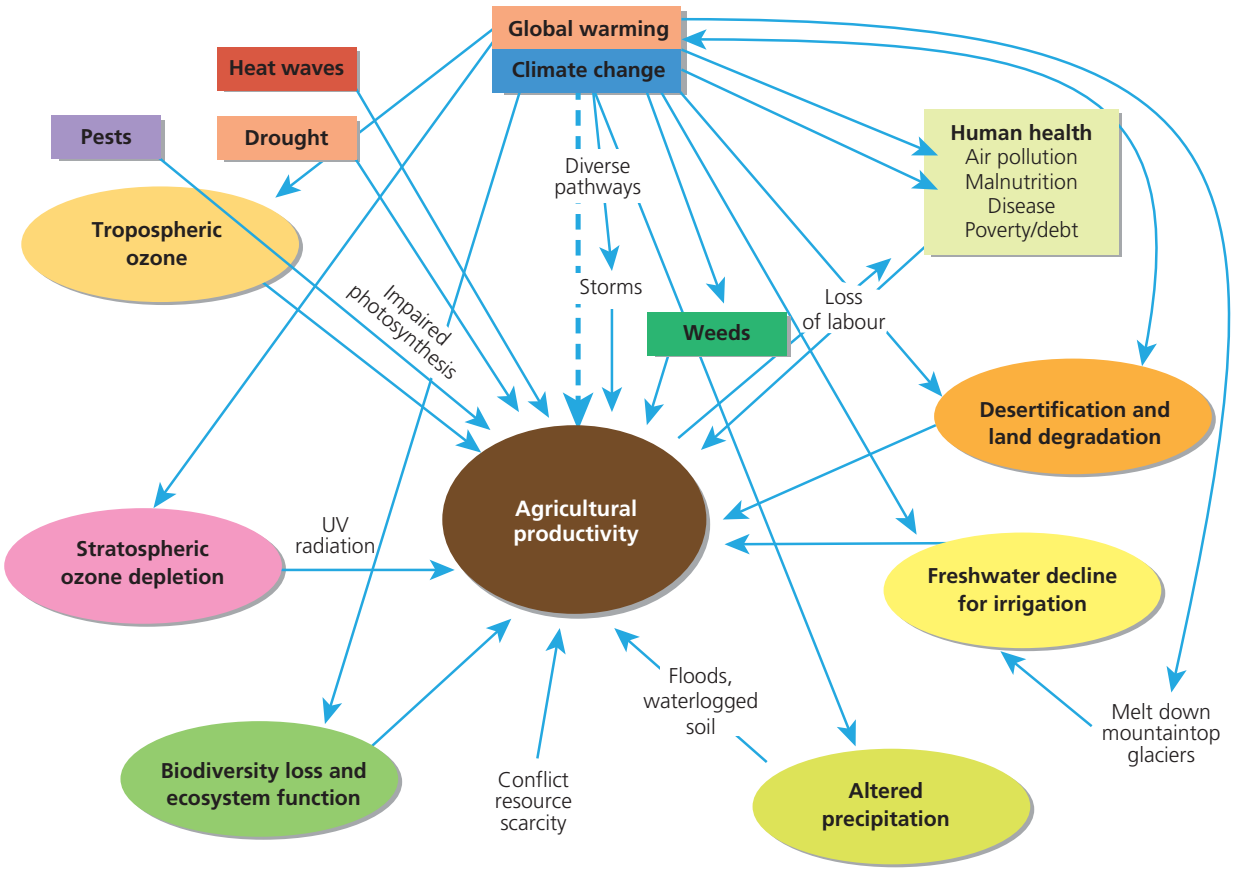


Figure 10.71 Multiple impacts of global warming and climate disruption on agriculture

Prospects for global population change

Over the course of the twenty-first century, world population is likely to rise by around 75 per cent to a UN estimate of 10.9 billion in 2100. This compares with the nearly 400 per cent rise of the last century when population grew from 1.6 billion in 1900 to 6.1 billion in 2000. The reason for this is much slower growth rates. The peak of world population growth rate was reached in the late 1960s and it has continued to fall from 2.09 per cent to approximately 1 per cent (Figure 10.72).

Drivers of world population growth

The two drivers of world population growth are the world **fertility rate** and the world **life expectancy**.

Fertility rates

A country's total fertility is the key factor in the demographic transition that each country goes through as it develops. The global average fertility rate in the 1960s was five children per woman but this has now halved. The UN forecasts that this downward trend will continue during the twenty-first century from 2.5 to just over two children (Figure 10.73 page 540). The current UN projection for global fertility is that it will be 1.9 by 2100.

As global fertility falls to the replacement level after 2050, it will become less important as a driver for future population growth.

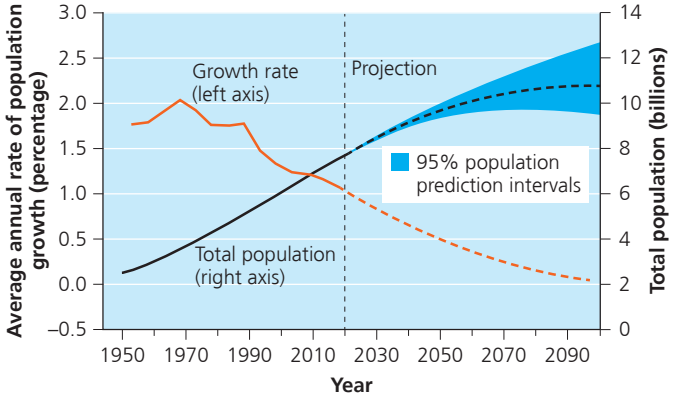


Figure 10.72 Changing and projected growth rates and total population size from 1950–2100

Source: United Nations

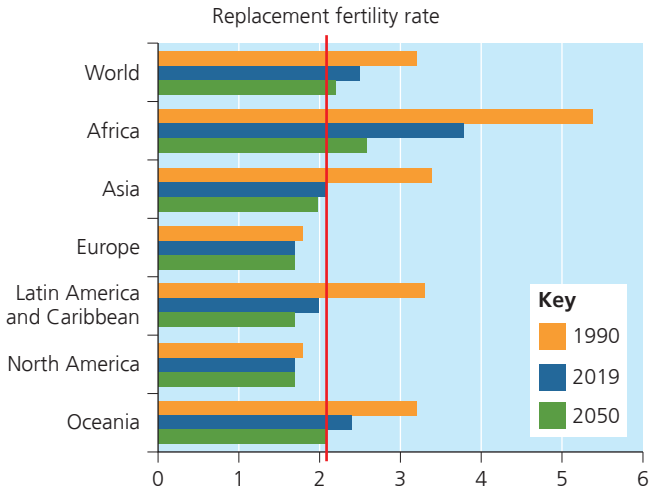


Figure 10.73 Global decline in fertility rates from 1990–2050

Life expectancy

As living standards rise and healthcare and medical science improve around the world, more and more children survive. The transfer of medical knowledge, often through NGOs, has improved the chances of survival in even some of the poorest countries. This means that average life expectancy around the world is also increasing rapidly (Figure 10.74). This has become the main reason for population growth; more people are surviving into older age.

Predicting future population size

Future population growth is highly dependent on the path that future fertility takes. The UN's Population Division makes forecasts for future population size based on different modelled fertility rates. In 2017, the UN presented four variants of population projections:

- **Constant variant** – fertility based on current fertility rates remaining constant. This is thought to be unrealistic.
- **Medium variant** – the most likely scenario and the basis for UN projections.
- **High and low variants** – these assume that fertility rates in each country are 0.5 higher or 0.5 lower than the medium variant.

The projections are shown in Figure 10.75 and Table 10.16.

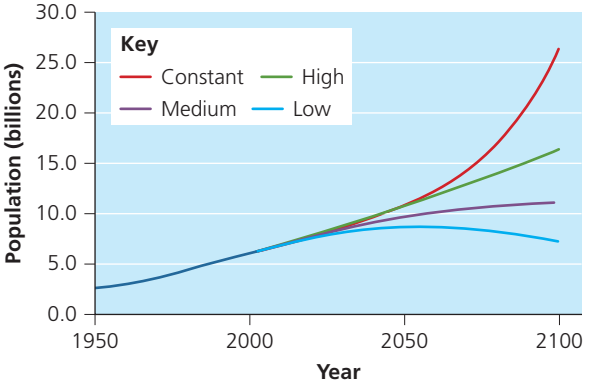


Figure 10.75 UN estimates and forecasts of world population by different fertility scenarios

Source: United Nations

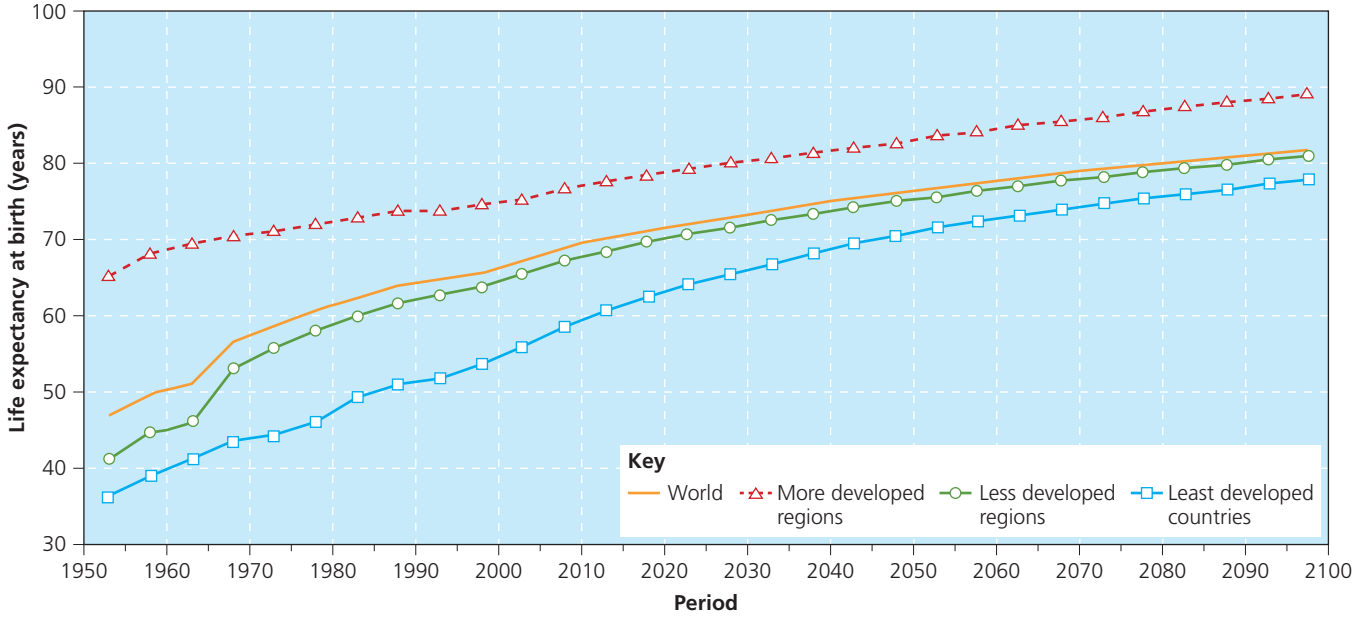


Figure 10.74 Life expectancy at birth for the world and major development regions, 1950–2100

Source: Mercedes Ayuso, Jorge Miguel Bravo and Robert Holzmann, Population Projections Revisited

Table 10.16 UN 2017 forecasts of world population size under different scenarios

Fertility scenario	2050 population in billions	2100 population in billions
High	10.85	16.52
Medium	9.77	11.18
Low	8.75	7.28*

Table 10.16 demonstrates the difficulty of forecasting population size and the importance of fertility in determining this. The 2050 forecasts show a 2.1 billion range between the different scenarios. A more dramatic divergence of 9.24 billion in forecasts appears for the end of the century. The extent of the spreads in both sets of forecasts points to the opportunity for fertility reduction, but also to the threat of more unsustainable growth. *Note that the prediction for 2100 with continued reduced fertility could give a population lower than the Earth's current population. The UN constantly review predictions and publish their findings every two years. Their 2019 report predicted the 2100 population would be 10.90 billion, with a 95 per cent chance of being between 9.40 and 12.70 billion (see Figure 10.72 on page 539).

The 2019 report indicates more confidence that the medium variant level of fertility is likely. There are signs that high-fertility populations are experiencing development-related fertility decline, partially supported by SDGs. What is less certain is whether low fertility will be sustained in countries with low fertility levels.

It seems that the future pace of fertility decline in sub-Saharan Africa is the main determinant of future world population growth over the next 30 years. It will also have implications for Africa and the rest of the world because of international migration pressure and difficulties in meeting the Sustainable Development Goals.

Fertility decline has slowed in the region relative to other areas that went through this transition. This has largely been attributed to the slow pace of introducing compulsory female education in older age groups.

Alternative projections

Many groups research population growth and in 2014 the Wittgenstein Centre for Demography, part of the International Institute for Applied Systems Analysis (IIASA), produced a set of world population projections. They used different methods and their model incorporated a crucial variable not included in the UN projections – the level of education in a given population. As a result, their projections for population size were lower and claimed that the world's population will stabilise before the end of the century. They argue that:

- the population will reach 8.3 billion by 2030, 9.2 billion by 2050 and peak at 9.4 billion in 2070 before falling back to 9 billion by 2100
- UN projections for rapid population growth in Nigeria were overestimated because improvements in education there will reduce fertility (the UN have since revised their estimates)
- educating girls is one of the most successful methods of reducing fertility in the long term – there is a strong negative correlation between female education and levels of fertility and child mortality (Figure 10.76).

Figure 10.76 Girls in Kerala, southern India, receive compulsory education, which has contributed to lower fertility rates there

Summary of prospects

Despite the variety of forecasts that have been made about the future of global population change this century, there are some common themes summarised in the UN's World Population Prospects report for 2019:

- World population growth rate will continue to slow down. The rate of deceleration is difficult to predict.
- The most rapid population increases will be in countries of sub-Saharan Africa.
- Two-thirds of the projected population growth until 2050 will be driven by people currently living and

would occur even if child bearing in high-fertility countries were to fall immediately to fertility replacement rates.

- Between 2019–2050, the populations of 55 countries or areas are projected to decrease by one per cent or more because of sustained low levels of fertility, and in some places due to high rates of emigration. These will mostly be in parts of Europe, Japan, China and other parts of Asia and the Americas.
- Countries with declining populations will face new challenges with ageing populations.
- The world's population is growing older, with the age group of 65 and over growing the fastest. Between 2020 and 2050 the proportion in this age group will nearly double from 9 per cent to 16 per cent. Globally in 2018, for the first time in history, those aged 65 or over outnumbered children under five years of age.
- Average global life expectancy at birth will continue to rise in the foreseeable future, reaching 77.1 years in 2050. Despite progress made in closing the longevity differential between countries, large gaps remain. In 2019, life expectancy in LDCs is over ten years below the global average.
- Global total fertility rates will continue to fall towards or below replacement rate in most global regions.
- India is projected to surpass China as the world's most populous country in 2027.

Projected distributions

Global population distribution will change significantly during the next century. According to the UN's 2019 World Population Prospects report, changes in the projected global population distribution among major regions are as shown in Table 10.17.

Table 10.17 Projected changes in the world's population distribution by region, 2020–2100

	Population (millions)			
Region	**2020**	**2030**	**2050**	**2100**
Africa	1,340	1,701	2,498	4,310
Asia	4,650	5,010	5,267	4,710
Europe	748	741	700	630
Latin America and the Caribbean	654	706	762	680
North America	369	391	425	491
Oceania	43	48	57	75
World	7,804	8,597	9,709	10,896

Source: UN Population Division (2019)

As with any type of projection, there is a degree of uncertainty and the results presented in the table are based on the medium projection variant. Figure 10.77 shows how the population distribution is projected to change (by UN SDG region) over the century.

The key findings from the UN report suggest that:

- **Africa**: More than half of global population growth from 2020 to 2050 is expected to occur in Africa.

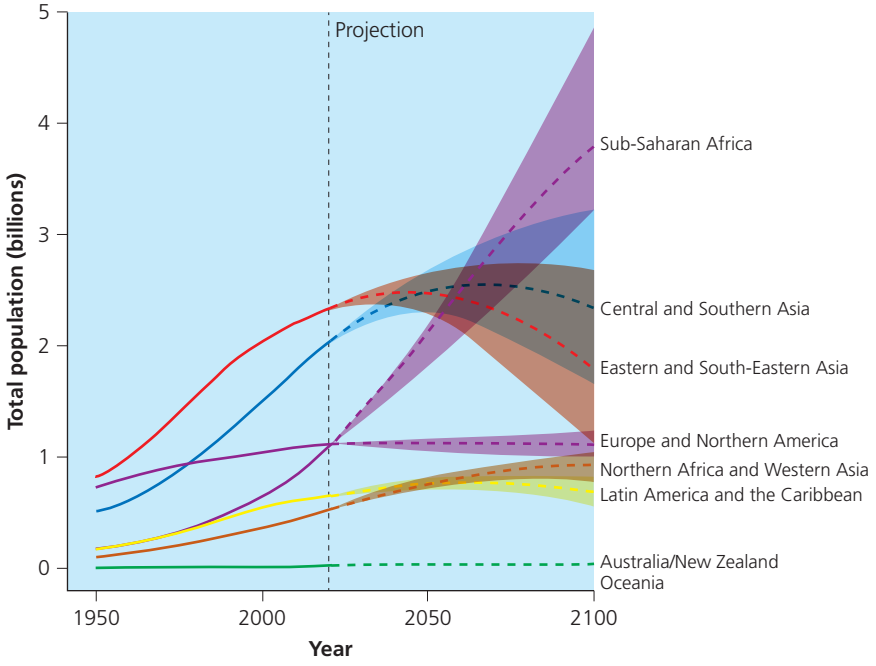


Figure 10.77 Projected changes in population distribution by UN SDG regions

Source: United Nations

Africa has the highest rate of population growth among major regions, growing at a pace of 2.5 per cent annually. The population of sub-Saharan Africa is projected to double by 2050 (99 per cent increase).

- Lower rates of population growth will be experienced between 2019 and 2050 in some regions including Australia/New Zealand (28 per cent), Central and Southern Asia (25 per cent), Latin America and the Caribbean (18 per cent), Eastern and South East Asia (3 per cent), and Europe and Northern America (2 per cent).
- **Least Developed Countries (LDCs)**: Population growth remains especially high in the group of 47 countries designated as the least developed countries (LDCs), of which 33 are in Africa and 13 in Asia.
- **Asia**: This is the second largest contributor to future population growth, but the size of increase is relatively small in comparison.
- **Short list of countries**: Just nine countries will contribute more than half the world's population growth between now and 2050: India, Nigeria, Pakistan, DR Congo, Ethiopia, Tanzania, USA, Indonesia and Egypt.
- Disparate population growth rates among the world's largest countries will re-order their ranking by size (see Figure 10.78).

In addition to the projected changes in distribution of population size, there will be significant changes in the distribution of population structure:

- In most of sub-Saharan Africa, as well as in parts of Asia and Latin America, recent reductions in fertility mean that the working age population is growing faster than in other age groups. This provides an opportunity to take advantage of demographic dividend to achieve accelerated economic growth.
- Around 30 countries, including Japan and many others across Europe, have a falling proportion of working-age populations and increased ageing dependency, which will put pressure on social, welfare and pension systems and may impede future economic performance.

Critical appraisal of future population – environment relationships

There is an ongoing debate as to whether '**population**' or '**consumption**' is a more significant threat to the environmental limits set by planet Earth. They are both challenges to sustainability and should not be set in opposition to each other. A growing and more affluent global population poses particular challenges for ecological footprints and carbon budgets. The main challenges are:

- to continue to supply the key resources needed for survival: food, water, energy and minerals
- to do this using **sustainable** methods.

The challenges and opportunities presented by the future population–environment relationship are summarised in Table 10.18 (page 544).

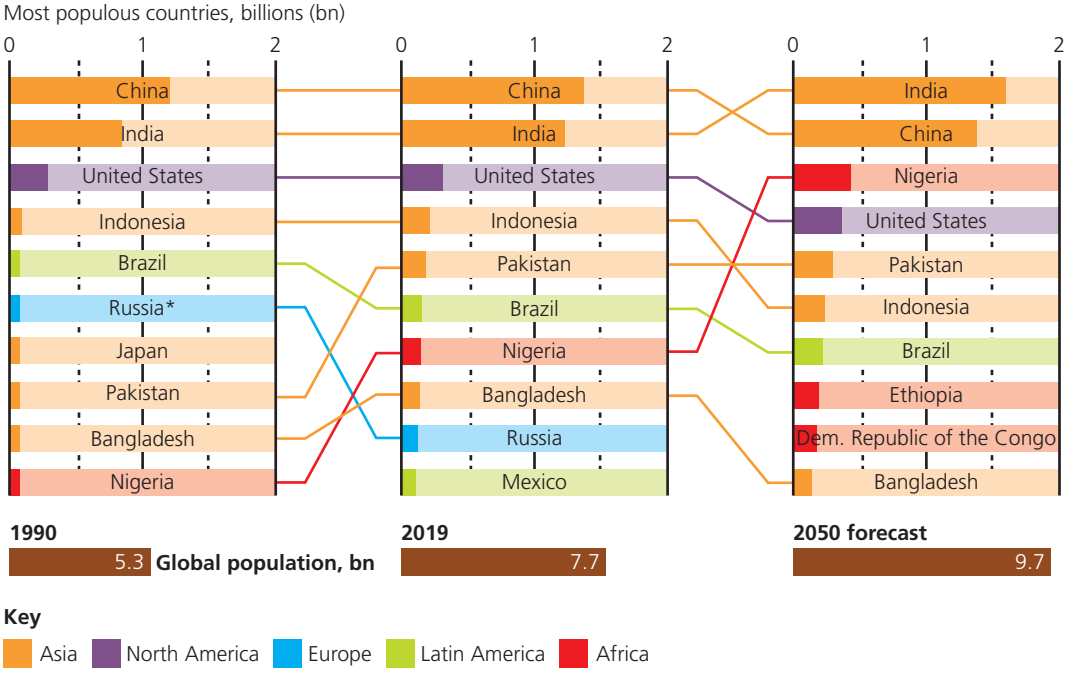


Figure 10.78 Changing rankings of the world's ten most populous countries, 1990–2050

Table 10.18 Summary of challenges and opportunities presented by future population–environment relationships

	Challenges/threats	Opportunities/possibilities
Water	• Strain on water resources locally and globally (40 per cent shortfall between demand and supply by 2040) • Chronic water scarcity • Critical for food production (70 per cent used for irrigation) • Increased water pollution • Climate change – extreme weather events leads to hydrological uncertainty	• Integrated water resource management (IWRM) – better management to increase supply and storage – control environmental stores, reduce demand and wastage • Economic water scarcity can be reduced by provision of better supply infrastructure; physical scarcity by IWRM • Use of better technologies, e.g. drip-feed irrigation • Desalination • Cloud seeding
Food	• Increased consumption puts pressure on land and farming systems – more intensification, more marginal land used • Pressure on land impacts on habitat, ecosystems and ecological interactions; could lead to soil erosion or desertification • Dietary transition – increased demand for meat from those in emerging economies exacerbates the strain on land and on GHG emissions • Intrinsically linked with water supply for irrigation – so puts further strain on water supply • Climate change will impact negatively on food production in some regions • Competes with non-food agricultural produce including cotton, rubber, palm oil, biofuels	• Sufficient food is currently produced and there is scope for increased yields and for less wastage at source and at the point of consumption • Shift to healthier more sustainable diets; more fruit and vegetables • More efficient and equitable food distribution systems are needed • Technology has enabled humans to keep food supply in line with population growth – GM and other agricultural technologies including hydroponics may increase yields in controlled environments • Growth of aquaculture may help to provide a more sustainable supply of fish protein without overfishing; it may impact further on water supply though • Effects of climate change may increase production in some regions • Substitute foods and sources of protein are being developed (e.g. insects for human consumption)
Energy	• Demand expected to increase by around 35 per cent by 2035 because of population growth and economic development • Depletion of non-renewable fossil fuels • Increased consumption of fossil fuels will exacerbate GHG emissions and climate change • Pressure on fuelwood supplies in developing areas	• Energy transition – a shift to more renewable and sustainable sources of energy • Energy conservation – energy demand can be reduced by technologies in the home, workplace and in transport systems • New technologies developed to access new supplies of fossil fuels (e.g. hydraulic fracturing) and/or to reduce emissions (e.g. carbon capture and storage) • Investment in hydrogen technologies to provide clean, renewable energy • Appropriate energy technologies in developing regions, e.g. biogas
Minerals	• Depletion of mineral resources, especially scarce base and precious metals and rare earth elements – especially from emerging economies • Destruction of landscapes and ecosystems caused by open pit mining • Pollution of water courses from tailing ponds and ore processing • Increased demand for energy to process and manufacture goods	• Technologies to exploit low-grade ores can increase supplies • Reduce consumption • Use more abundant substitute materials to replace minerals/metals • Re-use or find other uses for products that work; repair products when possible and recycle materials – all to reduce demand for new materials

The future of water, energy and mineral resources are covered in more detail in the next chapter – Chapter 11, Resources. This is a separate option topic and will not be linked to the population–environment relationship in the same way but may provide more information on challenges and opportunities.

Review questions

22 Describe the geographical spread of impacts on health caused by ozone depletion.

23 Explain the major impacts that climate change is likely to have on health around the world.

24 Outline the key population changes that are projected until 2050.

25 Evaluate the challenges and opportunities presented by population growth and sustainable management of resources.

Case study of a country experiencing specific patterns of overall population change: Japan – decline and ageing

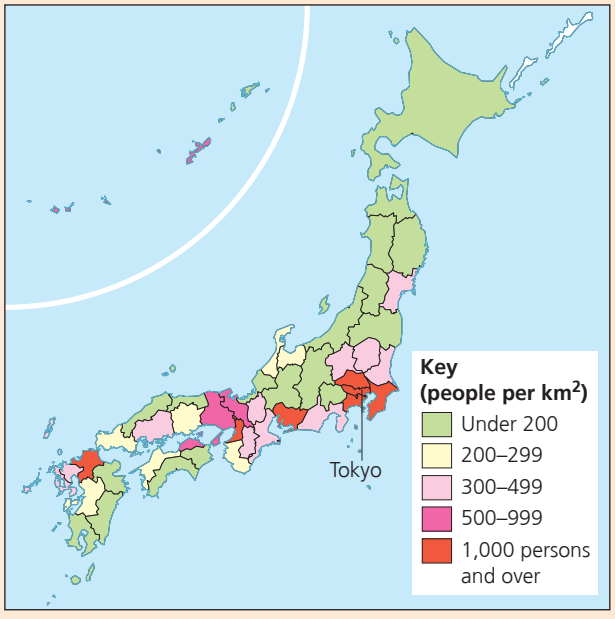


Figure 10.79 Population density map of Japan by prefecture

Japan is one of the world's oldest populations with one of the lowest birth and fertility rates. The total population of Japan in 2019 was 126.3 million (eleventh largest globally) and population density was 347 per km². This is an average figure and does not reflect the stark contrast in density between urban areas and rural mountainous areas.

Environmental factors

Japan is a mountainous archipelago consisting of thousands of islands of which Honshu, Hokkaido, Kyushu and Shikoku are the largest. The islands contain fertile valleys and narrow coastal plains where much of the population is concentrated. Around 70 per cent of the land is mountainous with relatively sparse population density. Around 92 per cent of the population live in very densely populated urban metropolises, such as Tokyo and Osaka (see Figure 10.79).

Due to the large north–south extension of the country, the climate varies strongly in different regions. It enjoys a mainly temperate climate but ranges from long, cold winters in Hokkaido and the Sea of Japan coast in the north, to subtropical weather in the southernmost islands such as Okinawa, where the mean temperature in January is 17°C. The climate in most of the major cities, including Tokyo, is temperate to subtropic and consists of four seasons. Winters are mild and the summer is hot and humid. There is a rainy season in early summer, and typhoons hit parts of the country every year during late summer.

Located in an area where several continental plates meet, Japan experiences frequent hazards from typhoons, earthquakes and volcanic eruptions. The volcanic activity in the islands has provided the population with fertile soils. These soils, together with the temperate and seasonal warmth, have fostered intensive agricultural production. Only 20 per cent of land is suitable for farming but the population has partially overcome the limitations presented by the landscape by intensively farming alluvial plains and by terracing slopes to increase farmland availability. Agriculture is highly subsidised (rice is the main crop) and food, especially meat products, is imported to ensure food security.

The archipelagic nature of Japan means it is a seafaring nation with a strong fishing industry. This includes a well developed marine aquaculture industry that also provides some food security, though it is still has to import large amounts of seafood to meet the large domestic demand.

Japan has relatively small amounts of fossil fuel and mineral deposits. In the industrial age, it relied on expansion into neighbouring countries such as Korea, Manchuria in northeast China and Inner Mongolia to access resources to support its ever-growing population and its desire to become an industrial superpower. This, together with further expansion of 'Empire' into the Pacific, was ultimately one of the triggers for the Second World War.

Character, scale and patterns of population change

Japan has a distinctive culture which is very traditional and based on order and harmony. Japanese are taught to be responsible members of their families and their country, and are expected to serve others' needs before their own. Consequently, they have a reputation for being hard-working and intensely loyal to family, employer and nation. It is the third largest economy in the world and the gross national income per person is US$41,690 per annum. Traditionally males have been the main breadwinners in Japanese society but as the economy has moved more towards a post-industrial service economy, well-educated females have become an increasingly important part of the workforce. This may have had an impact on birth rates and fertility rates, which are 7/1,000 and 1.4 respectively. Death rates are 11/1,000. This means that Japan has a natural decrease and population is steadily declining. The country is in Stage 5 of the DTM. It is thought that Japan's population will decline by 20 per cent to around 100 million by 2050 and by a third of its current size to 85 million by 2100.

The average life expectancy of 85 is one of the highest in the world and over 65s comprised 28 per cent of the population in 2019. There are more than 71,000 centenarians. Infant mortality is one of the lowest in the

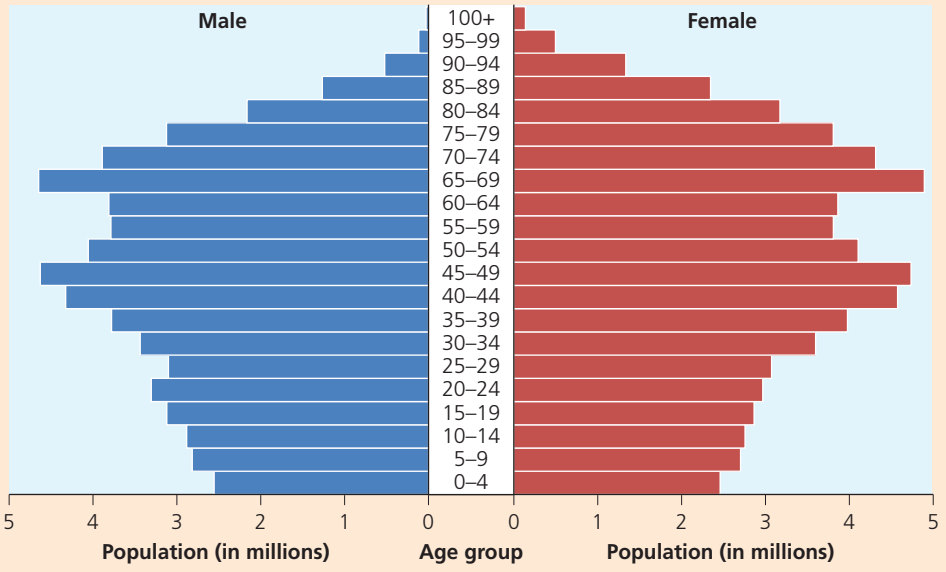


Figure 10.80 Japan's population structure, 2018

Source: Central Intelligence Agency

world at just 2/1,000 of live births. Young dependants only make up 12.5 per cent of the population. The percentage of economically active people is therefore 59.50 per cent and the dependency ratio is 68. It is estimated that nearly 40 per cent of Japan's population will be over 65 by 2050 so they face a 'demographic time bomb' where the population decline means that there are not enough young people to replace the older population who are dying out. The burden borne by the working and younger population in supporting the elderly may become unsustainable. The shape of the Japan's ageing population structure is shown in Figure 10.80.

Socio-economic factors

The two main underlying causes of Japan's ageing population are higher life expectancy and a low fertility rate. A number of social and economic factors are thought to be causes for this:

- Excellent **healthcare, sanitation and high living standards** mean high life expectancy.
- **Low fat diets** – unlike many western MEDCs, Japan has fewer obesity and heart disease problems.
- Although non-communicable diseases are still the leading cause of deaths, the probability of premature death from NCDs is low in comparison to other countries.
- Reputedly the air is clean and it is true that Japan uses fewer fossil fuels to generate energy – relying more on nuclear and HEP. (They also have a tendency to use more public transport.)
- **Women** receive good education so as a result many are more **focused on careers**, so they have **fewer children**.
- **Later marriage** reduces the child-bearing window.
- **Contraception is accessible and widely used** – there are few religious or social taboos.
- **Cost of education** is rising at a time when wages are staying relatively still. These are mostly incidental costs but funding one child through high school and university in Japan costs as much as ¥10 million (nearly US$ 100,000). These high costs are thought to be a factor in low fertility and population decline.

Figure 10.81 Japan – an ageing society. Japan has the highest percentage of over 65s of any country, at over 28 per cent

Implications for the country and Japanese society

The negative consequences arising from Japan's ageing population range from the serious – for example, the domestic economy is shrinking, to the more trivial – Japan's ageing problems means that more adult incontinence nappies are sold than baby nappies. All generations are affected by the population structure.

- Around 10 per cent of GDP is used to pay pensions and a further 21 per cent on health and social security for the elderly.
- The '**2025 problem**' is cited with widespread concern as in that year, all of the postwar baby boomer generation will have turned 75 or older, further pushing up the cost of social security benefits.
- Since the 1960s retirees pay only a small part of their medical care through a universal healthcare system, partly covered by a national insurance scheme.
- The working population is shrinking fast, which will have a negative impact on domestic economic growth – affecting both production (fewer workers) and consumption (fewer people earning means less spending). It also means fewer workers paying taxes, which mainly supports the pensions and social security system.
- A rise in home care by families, who may have to give up full-time jobs to become sole carers, takes more people out of the economic workforce. Around 65 per cent of the elderly live with their children and three generation households are typical.
- Pressure on care services for the elderly means more medical and care workers are needed. A shortfall of 370,000 medical staff, nurses and care workers is predicted by 2025.
- There is increasing pressure on NGOs (charitable organisations) to provide this care. The number of residential homes is restricted and there is a stronger emphasis on community care, which is justified on fiscal grounds but is also considered to be more supportive of well-being. An emerging voluntary sector with more reliance on unpaid volunteers (many of whom are over 65) is providing this care.
- It is estimated that nearly 6 million elderly Japanese live alone with no family support.
- There are increasing numbers of elderly living on the streets, in poor quality housing or ending up in prison (especially males). Reoffending rates are high as prison offers better care for some.
- There is a rising number of cases of dementia. It is predicted that one in five of the elderly will suffer with dementia by 2025.
- More elderly are having to return to work to make up their income – around 25 per cent of senior citizens (over 65) have jobs.

There are many positives as well:

- Due to globalisation, the Japanese manufacturing industry is still very successful and income from large TNCs such as Toyota, Sony, etc., brings in repatriated profits to support the domestic economy.
- There is a growth in the number of private hospitals specialising in healthcare for the elderly and so creating jobs and income.
- There is an increasing demand for leisure activities and tourism among the affluent elderly – also generating income and creating jobs.
- One of the positive features specific to Japan is the embracing of new artificial intelligence (AI) and robotic technology – to build 'robot carers' to assist the elderly including companion pets, bionic limbs that aid mobility, fetching, carrying and even robots that help when going to the toilet. Japan leads the world in the field of AI and the exponential rate at which robotic technology is being applied to the care industry is helping to fill some of the shortfall in healthcare workers. It is expensive and so only available to the more affluent, though it is likely that the relative costs will reduce as it is more widely used. This type of automation is already thought to be worth US $2 billion to the economy.

Responses by the government:

The Japanese government has a duty to try to tackle some of the issues for a sustainable future but has to be cautious with its policies so that it does not alienate the elderly, an important lobby group (the 'grey vote'). Responses are as follows:

- **Raise taxes** – the government raised the consumption tax rate to 10 per cent in 2019 in an effort to cover the growing cost of supporting the country's ageing population.
- **Raising retirement age** – the standard retirement age is to be raised to 65 by 2025 (compared to 62 since 2016) but workers are encouraged to work until they are 70 and employers are given incentives to retain or recruit those older than 65. This is to be accompanied by pension and social security reforms.
- **Pro-natal population policies** – the Plus One Policy introduced in 2009 is modestly pro-natalist, encouraging parents to have more children by increasing paid maternity leave, child benefit payments and cheaper healthcare/childcare support to offer a better work–life balance for mothers. Shops offer discounts for larger families and local economic incentives have also been attempted. For example, the town of Yamatsuri near Tokyo offered parents US $4,600 on the birth of a child and a further US $460 for the first 10 years. These initiatives have enjoyed moderate success and fertility has risen slightly in the past 10 years but overall the incentives have been relatively ineffective and are seen as symbolic gestures rather than being of serious practical help.
- Encouraging the **robot technology** and automation revolution in industry to offset the loss of production due to a shrinking workforce.

- **Encouraging immigration of skilled foreign workers to increase the workforce:** This is the most obvious short-term solution to labour shortage and dependency problems.

Immigration has traditionally been a problem for Japan for a number of reasons:

- Japanese is a difficult language and immigrants are given a language test which has a high failure rate.
- Foreigners have difficulty integrating with Japanese culture, particularly matching the hard work ethic that employers expect.
- Accepting immigration as a solution has traditionally been seen as a threat to social stability, cultural heritage and public safety by wider Japanese society.

There has been a softening of anti-immigration attitudes, possibly as a result of the rise in inbound tourism in recent years. Additionally, there is an increased acceptance that Japan needs foreigners to keep functioning. Facing critical blue-collar labour shortages and pressure from industry groups, the government took action. In 2019, the Immigration Control and Refugee Recognition Act officially opened Japan's doors to lower-skilled foreign workers. Fearing an anti-immigration backlash, the government's policy remains cautious and only targets certain industries where there are critical skills shortages. Skilled workers are invited on a temporary ('guest worker') basis (unaccompanied by family) for an initial five-year stay.

Japan already has a permanent immigrant population of around three million (about 2.4 per cent of the population), mainly from neighbouring East Asian countries such as China and Vietnam, who integrate more easily. It also accepts much needed skilled nurses from Indonesia and the Philippines. However, to prevent a fall in the proportion of the working age population, it is thought that 600,000 immigrants will be needed each year until 2050, numbers that it seems unlikely it will receive or accept.

Case study of a specified local area: Place, health and well-being – Blakenhall, Wolverhampton, West Midlands

Place

Wolverhampton is an industrial city in the Black Country conurbation to the north and west of Birmingham (Figure 10.82). In 2019, the population of the city was estimated to be 263,000. Wolverhampton was granted city status in December 2000 and has since received an economic boost from the kudos. The council has invested in regenerating transport, retail, hospitality and cultural facilities, especially in deprived areas such as Blakenhall, a ward to the south of the city centre where this case study focuses.

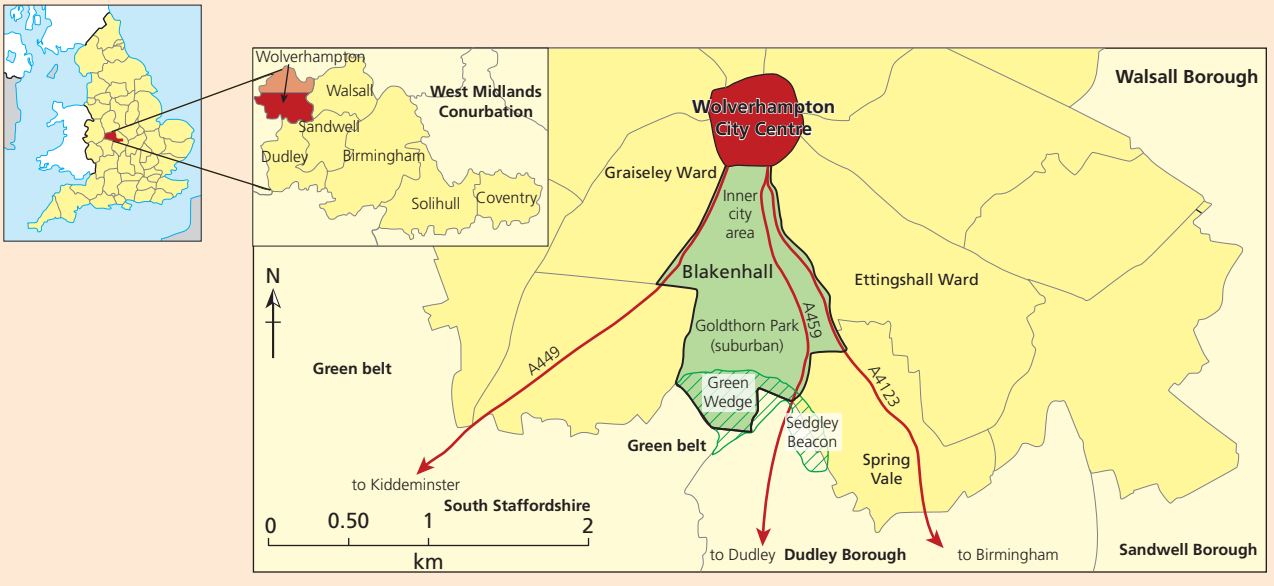


Figure 10.82 Location of Blakenhall in the context of Wolverhampton city area (the area shaded dark red on the inset map represents the area of south Wolverhampton shown on the main map)

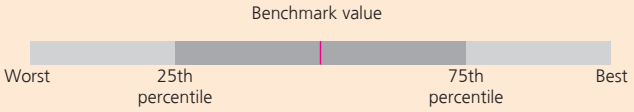
Key

Compared with benchmark:

Better Similar Worse

Quintiles:

Best Worst



	Indicator	Wolverhampton Count	Wolverhampton Value	England Value	England Worst	England Range	England Best
Life expectancy and causes of death	Life expectancy at birth (Male)	–	77.2	79.6	74.5		83.9
	Life expectancy at birth (Female)	–	81.4	83.2	79.5		87.0
	Under 75 mortality rate from all causes	2,639	432	330	544		220
	Under 75 mortality rate from all cardiovascular diseases	649	108.5	71.7	124.6		39.1
	Under 75 mortality rate from cancer	939	156.8	132.3	190.3		94.8
	Suicide rate	60	9.0	9.6	20.7		4.2
Behavioural risk factors	Admission episodes for alcohol related conditions – under 18s	40	22.2	31.6	112.7		7.8
	Admission episodes for alcohol related conditions (narrow)	2,080	876	664	1,127		389
	Smoking prevalence in adults (18+) – current smokers (APS)	30,314	15.1%	13.9%	27.5%		3.4%
	Percentage of physically active adults	–	58.0%	67.2%	46.7%		80.0%
	Percentage of adults (18+) classed as overweight or obese	–	67.3%	62.3%	75.9%		41.7%
Child health	Under 18 conception rate / 1,000	116	27.7	16.7	39.4		3.6
	Smoking status at time of delivery	542	16.7%	10.6%	25.7%		1.6%
	Breastfeeding initiation	2,301	66.8%	74.5%	37.9%		96.7%
	Infant mortality rate	62	6.0	3.9	8.2		0.8
	Year 6: prevalence of obesity (including severe obesity)	976	29.3%	20.2%	29.6%		9.5%
Inequalities	Deprivation score (IMD 2015)	–	33.2	21.8	42.0		5.0
	Smoking prevalence in adults in routine and manual occupations (18+–64) – current smokers (APS)	–	26.8%	25.4%	46.5%		4.5%
	Inequality in life expectancy at birth (male)	–	7.8	9.5	15.2		–1.3
	Inequality in life expectancy at birth (female)	–	6.3	7.5	13.6		–1.5

Figure 10.83 Part of the area health profile for Wolverhampton City

Source: Public Health England

Wolverhampton is often represented in a negative light by outsiders. The city had an overall Index of Multiple Deprivation (IMD) value of 33.20 in 2015 (Figure 10.83) – significantly above the average value of 21.80 for England. At the time it was ranked the 17th most deprived local authority area in the country. However, the rebranding that has accompanied its city designation, together with ongoing regeneration, improved its deprivation ranking by seven places by 2019.

Physical environment

Wolverhampton

- Wolverhampton is located on the Midlands Plateau and next to the main Severn–Trent watershed of England.
- The city centre is a densely built environment with a number of high-rise buildings. Beyond the inner ring road, building density is lower and interspersed with a number of public open spaces.
- To the north and west of the city boundaries lies the open countryside of Staffordshire and Shropshire.
- The city receives less rainfall (681 mm) than the national average (879 mm) but it also receives less weekly sunshine hours (26.2 compared to 29.50).

Blakenhall

Blakenhall occupies a southern sector of the city running from the inner ring road to the southern boundary with Dudley (see Figure 10.82, page 548). The ward contains three main roads radiating from the centre – all three are major thoroughfares and are often congested, which contributes to increased air pollution in the area. Blakenhall exhibits a contrast of different environments from near the city centre to its outer southern boundary:

- Nearest the city centre is a typical inner-city environment of grid-iron pattern streets with terraced housing (see Figure 10.84, page 551). There is also some older semi-detached private and council housing. The area also contains a number of older

factories. A group of six 1960s tower blocks were demolished between 2002 and 2011 and have been replaced by a regenerated area of modern town houses and retail premises. There are a few small public open spaces with children's playgrounds and a larger recreational area with playing fields.

- Further from the city centre, the building density is lower in the more suburban area of Goldthorn Park, with mainly semi-detached and some detached housing.
- The outermost part of the ward is a semi-rural fringe rising steadily towards Dudley Borough. Here a green wedge of land is formed by fields and by Sedgley Beacon – a limestone ridge which separates the two boroughs.

Socio-economic character

Wolverhampton

Many of Wolverhampton's larger engineering and manufacturing industries declined in the 1980s during a period of deindustrialisation but the area now has a diverse economy, which includes brewing and civil engineering. It also has a number of large employers in the finance industry and public sector, including education and health. More recently, in 2014, Jaguar Land Rover (JLR) opened a £500 million engine assembly plant to the north-west of the city at the i54 Business Park. Despite the variety of employment opportunities, the unemployment rate for 2019–2020 was 6.7 per cent, almost twice the national average.

Blakenhall

The inner-city part of Blakenhall is predominantly retail in function with an 'out-of-town' retail park (St John's) near the ring road, four supermarkets and a variety of small and medium-sized specialist independent shops, convenience stores and fast food outlets, many of which cater for the local Asian population. Old factories in the area have either been converted into community centres or divided into smaller units occupied by auto services businesses. There are also many personal or professional service businesses run by and aimed at the local community, for example, immigration legal services, Indian bridal wear.

Blakenhall's population was estimated at 13,000 in 2018, around five per cent of the city's total population. The city council describes Blakenhall as having the following population and housing features:

- a multicultural city suburb:
  - 74 per cent of the population are from a 'black or minority ethnic' background
  - nearly 50 per cent of the population are of Indian heritage and Sikhism is the predominant religion (36.5 per cent)
  - three Sikh Gurdwaras, two Sikh community centres, a Hindu temple and a mosque
- settled extended families – many households in the area are three generational; in smaller, older inner-city properties this can mean some overcrowding
- 21 per cent of houses have one or more category one hazards (a risk to health and safety)
- educational attainment is above the city average at both GCSE and graduate level; it is slightly below the national average at GCSE level
- a younger than average population structure, with a mean age of 39 and over 60 per cent of the population of working age
- around 32 per cent of the population work in semi-skilled or unskilled work but 13 per cent describe themselves as professional
- 16 per cent of the over-65 population are affected by fuel poverty
- with an IMD value of 32.90, Blakenhall is around the average deprivation level for the city.

Relationship with health – experience and attitudes

Figure 10.83 (page 549) shows part of the area health profile for Wolverhampton with data collected from 2015 to 2018–19. It shows that the city has below average health values on most indicators, with a particularly low score for 'under 75 mortality from all cardiovascular diseases' and 'prevalence of child obesity'. Deprivation may be an underlying factor in both of these.

From the city's Healthy Lifestyles survey in 2016:

- Blakenhall's predominant group was identified as 'healthy weight but poor lifestyle'. This group is associated with deprivation and a younger population.
- The second predominant group in the area were 'overweights', who generally eat unhealthily but are neither smokers nor high risk drinkers.

Table 10.19 shows the significant health indicators for Blakenhall compared to city and national averages.

Reasons for overall balance of poor health in Blakenhall and the surrounding areas

- Blakenhall has relatively high unemployment and below average incomes. Deprivation alone contributes to many of the indicators of poor health, such as the higher rates of infant mortality and alcohol related mortality. It may also be responsible for the poor diet that contributes to childhood obesity and 'overweights'. There may be a reliance on convenience food or fast food.
- Population density is high in the inner-city area, with poor housing quality in older properties, which increase the risk of contracting and transmitting illnesses.

Table 10.19 Comparison of health indicators with the city of Wolverhampton as a whole and the national average

	Comparison with the whole city and national average
Good health indicators	
Smoking prevalence	13 per cent – below city and national average
Over 65s with an illness that limits daily activities	Below average for the city but above the national average
Deaths and serious injuries from road traffic accidents	Below the national average Despite high traffic density, speed restrictions and regular congestion slow speeds down
Poor health indicators	
Infant mortality rate (IMR)	9.9 (2007–2016) Second highest rate in the city and well above average rates for England of 5.9
Proportion of obese children entering Reception class	16 per cent, which is the highest in the city and well over the national average 9.4 per cent Obesity at Year 6 remains high at 28 per cent compared to 19.4 per cent national average
Adult alcohol admissions to hospital and alcohol-related mortality rates	High adult alcohol admissions to hospital Alcohol-related mortality rates were one of the three highest in the city (2010–2014)
Life expectancy	Male – 76.2; Female – 79.7 Below the city average Both male and female are more than three years below the national average
Dementia prevalence	Has the second highest rate of dementia in the city The highest number of permanent placements in residential or nursing care

Figure 10.84 Terraced housing in inner-city section of Blakenhall

- Relatively high levels of fuel poverty among the over 65s may contribute to illness, especially in winter, which may also be related to poor housing quality.
- High traffic density and congestion in the area causes air pollution. This can cause respiratory infections and disease, increased morbidity and lower life expectancy.
- The high proportion of minority ethnic residents in Blakenhall may provide some explanation for increased morbidity, especially among the elderly. South Asian populations are four times more likely to develop cardiovascular diseases and are also more likely to suffer from respiratory diseases.
- For around 20 per cent of the population, English is a second language. Although multilingual information is provided, language barriers may still deter the use of healthcare services in some cases; this may impact on antenatal and postnatal care.

Obesity

Obesity, especially among children, is a particular concern in Wolverhampton. The city's Clinical Commissioning Group (CCG) has made a special plea to the authorities to adopt campaigning and other strategies to mitigate the problem. There are opportunities for exercise in Blakenhall's parks, at the local school and in the open spaces at the edge of the ward. However, there are limited opportunities for organised activities and many cannot afford to join formal groups or clubs. This problem is partly being addressed by the Bob Jones Community Hub where there is provision for a range of indoor fitness and recreational activities.

Conclusion

Despite the overall balance towards poor health indicators, the local population generally respond positively to surveys about well-being. Recent regeneration of Blakenhall with new housing, stores and support funding for local community business has given the area a positive feel. The community is strong, well connected and integrated and support groups such as dementia care operate in Blakenhall to give respite to carers and ensure well-being.

Fieldwork opportunities

Census data from the Office for National Statistics (ONS) website www.ons.gov.uk can be used to support fieldwork investigations of population change at different scales throughout the UK. For example, in order to investigate the health of your local area, start by looking at local authority health profiles available at https://fingertips.phe.org.uk/profile/health-profiles and then focus in on a smaller area utilising parish profiles and local government websites. Neighbourhood statistics from the ONS provide analysis and summaries of census data and you could then conduct your own environmental quality survey or questionnaire looking at different aspects of health. Remember that health is affected by a range of social, economic and environmental factors. Look for and survey/map provision of green spaces and recreational opportunities, look at employment data and crime figures and then try and gauge the level of community engagement and opportunities, which can all lead to improvements in health and well-being.

Further reading

General:

www.prb.org – Population Reference Bureau – a US based non-profit organisation that specialises in demographic analysis and research focused on the environment, health and structure of populations

Environment and population:

www.fao.org/home/en/ – Food and Agriculture Organisation of the UN

www.worldbank.org/en/topic/climate-smart-agriculture – Climate smart agriculture, the World Bank

www.wri.org – World Resources Institute – a global research organisation that focuses on six main areas including climate change, food security, energy and water sustainability

www.gafspfund.org – Global Agriculture and Food Security Program

Environment, health and well-being:

Davies, O. (2020), 'Communicable diseases: Twenty-first century threats', *Geography Review*, Volume 33, Number 4, April 2020, pp 22–25

Polgreen, Philip M. and Polgreen, Evelyn L. (2018), 'Infectious diseases, weather and climate', *Clinical Infectious Diseases*, Volume 66, Issue 6, March 2018, pp 815–817

www.who.int/ – World Health Organization

www.who.int/malaria/publications/world-malaria-report-2019/en/ – World Malaria Report 2019 – published by WHO

Population change:

Ley, David (2019), 'How elite migration works', *Geography Review*, Volume 33, Number 1, September 2019, pp 13–15

Ho, Elaine Lynn-E. (2020), 'Emigration, immigration and re-migration – A case study of Canadian-Chinese migrants', *Geography Review*, Volume 33, Number 3, February 2020, pp 34–37

https://ourworldindata.org/grapher/population-by-country – Our World in Data online offers a wide range of demographic data by country

http://datatopics.worldbank.org/world-development-indicators/ – The World Bank Group – World Development Indicators provide a variety of information on key vital rates

Principles of population ecology and application:

Chiras, Daniel D. (2014) *Environmental Science*, Tenth edition (Jones and Bartlett Learning)

www.populationmatters.org – Population matters – a UK-based NGO providing education, advocacy and research on population issues, particularly as they impact on the environment

Global population futures:

www.population.un.org/wpp/Publications/Files/WPP2019_Highlights.pdf – UN World Population Prospects, 2019 – Highlights

Roser, M. (2019) 'Future Population Growth', available at www.ourworldindata.org/future-population-growth

Question practice

A-level questions

1 Explain the concept of the demographic dividend. [4 marks]

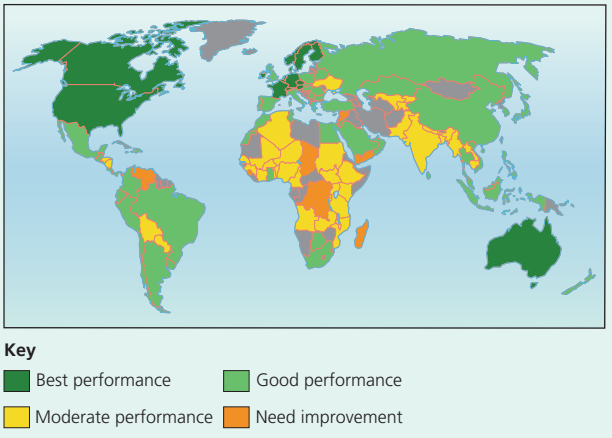


Figure 10.85a The state of global food security based on affordability, availability, safety and quality of food, 2019

Source: Statista

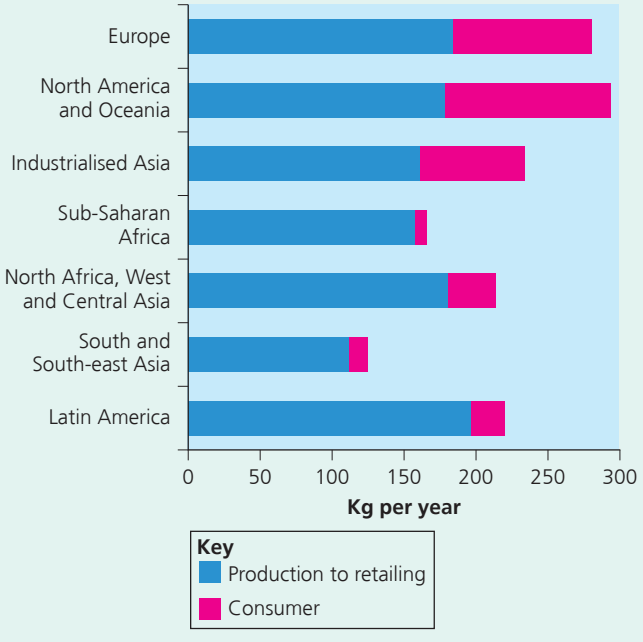


Figure 10.85b Per capita food losses and food waste in kilograms per year by region

Source: FAO

2 Analyse the data on global food security and food losses shown in Figures 10.85a and 10.85b. [6 marks]

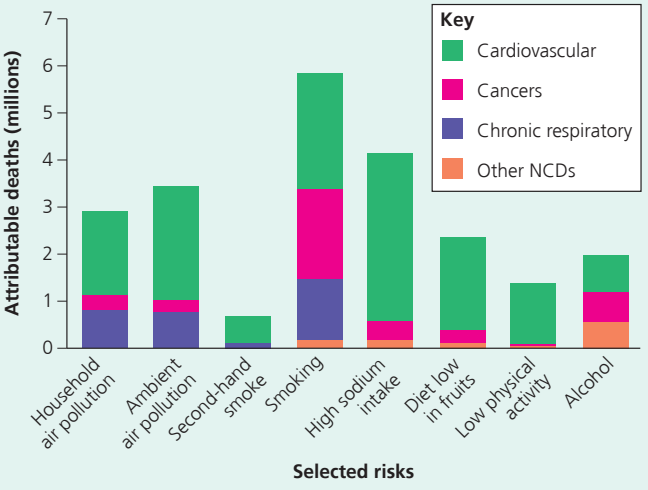


Figure 10.86 Deaths from non-communicable diseases (NCD) attributable to selected risks (all ages)

Source: WHO

Table 10.20 All premature adult deaths (30–69-year olds) caused by NCDs and the proportion of these attributable to air pollution

	Premature deaths (ages 30–69) from NCDs		Subset of which is attributable to air pollution	
	Total (millions)	Percentage	Total	Percentage
Low income countries	0.9	6%	0.16	6.5%
Lower middle-income countries	6.1	40%	1.35	53.6%
Upper middle-income countries	5.8	39%	0.88	35%
High income countries	2.2	15%	0.12	4.9%
Total	15	100	2.51	100

Source: WHO

3 Using Figure 10.86, Table 10.20 and your own knowledge, evaluate the importance of air pollution in causing deaths resulting from non-communicable diseases. [9 marks]

4 Assess the extent to which the character of a place you have studied has been affected by international migration. [9 marks]

5 'The effects of climate change on agriculture will have a greater influence on the carrying capacity for human population than growing ecological footprints.' To what extent do you agree with this statement? [20 marks]

Chapter

11 Resource security

One of the major challenges resulting from continued population growth in the twenty-first century is the large-scale exploitation of natural resources. This chapter looks at the increasing demand for **energy**, **water** and **minerals** as they are all critical for human survival and development. The uneven distribution of resources is partly resolved by geographical transfers from areas of surplus to areas of shortage or high demand. For water this is generally on a local or regional level but for energy and mineral resources it takes place on a global scale.

Understanding the nature of resource **exploitation** and demand requires an appreciation of the relationship between the human and physical environment, how the resources are used and why they are so important to people in different regions. The increasing scarcity of all three resources can lead to conflict. This means that ensuring a secure and sustainable supply of each has become a major objective for most governments.

In this chapter you will examine the following ideas:

- concepts of resource development
- global patterns of availability, production and consumption of resources
- geopolitical issues concerning energy, water resources and mineral distribution and trade
- sources of energy, water and minerals and their relationship with the physical environment
- environmental impacts of resource development
- strategically managing the supply of and demand for resources
- sustainability issues associated with the exploitation, trade and consumption of energy, water and mineral resources.

Links

There are links between Chapter 11, Resource security and Chapter 1, Water and carbon cycles; Chapter 7, Global systems and global governance and Chapter 8, Changing places.

11.1 Resource development

A **resource** is defined as any aspect of the natural environment that can be used to meet human needs. As they are useful or necessary, natural resources such as fossil fuels or fresh water are also a material source of wealth. They have economic value and can be used to improve a country's wealth and economic development. Geographically, they are unevenly distributed around the Earth. Some nations have an abundance of certain resources but experience a scarcity of others that are equally important. For example, Saudi Arabia is an 'oil-rich' nation with around 25 per cent of the world's conventional oil **reserves**, but at the same time it has a **scarcity of water** supplies. It therefore has a greater problem ensuring water security than energy security.

Resource security is usually determined at a national level though it can be applied as a global concept. It refers to the ability of a country to safeguard a reliable and sustainable flow of resources to maintain the living standards of its population, while ensuring ongoing economic and social development. Resource security relies on a country being able to develop necessary resources from its own physical environment or otherwise being able to secure supplies from a trading partner. To provide their populations with all vital resources, most nations need to employ both strategies.

Key terms

Exploitation – The action of using natural resources to the fullest or for the most profitable use.

Exploration – The process of searching an area with the intention of finding and mapping natural resources.

Flow resources – Resources that are renewable and can be replaced, examples include fresh water and timber. They are commonly expressed in terms of the annual rates at which they are regenerated.

Reserves – That part of a resource that is available for use under existing economic and political conditions and with available technology.

Resource frontier – A newly colonised region where resources have been discovered and are brought into production for the first time.

Resource peak – This marks the point in time when the maximum production rate of a resource occurs (at all scales from individual reserves to global), with production declining in subsequent years.

Stock resources – Non-renewable resources which can be permanently expended. Their quantity is expressed in absolute amounts rather than rates.

Classification of resources

Natural resources include a wide range of materials and organising them into groups helps us to understand the differences in how they are used.

Stock and flow resources

Stock resources are compound deposits of materials which have taken millions of years to form, usually on or in the Earth's crust. From an economic perspective, stock resources have a 'fixed' and finite supply as once they are used they cannot be replenished. Their quantity can be measured in absolute amounts (for example, in tonnes).

Flow resources are those which can be naturally renewed within a sufficiently short time span to be relevant to decision makers. It is useful to make a further distinction between flows which are not dependent on human activity and those which are only renewable if human use remains at or below their capacity to reproduce or regenerate. The latter are known as **critical flow resources** and can be exhausted if not carefully managed.

This distinction between these different types of resources is best exemplified with energy resources:

Energy resources are classed as either renewable or non-renewable.

- **Non-renewable energy resources** (also known as finite or stock resources) are those that have been built up, or have evolved, over time. They cannot be used without depleting the stock because their rate of formation is so slow that it is meaningless in terms of human lifespan. Non-renewable energy resources are primarily fossil fuels (oil, natural gas and coal) but they also include uranium used in nuclear energy.

Figure 11.1 A wind farm near Wellingborough in Northamptonshire. Wind power is a flow resource and a renewable source of energy

- **Renewable energy resources** (also known as flow resources) yield a continuous flow that can be consumed in any given period of time without endangering future consumption, as long as current use does not exceed net renewal during the same period. Renewable energy resources include solar power, hydroelectric power, geothermal energy, wave and tidal power, wind power (Figure 11.1) and biomass sources. Renewable resources can be subdivided into:
  - **critical** – sustainable energy resources from forests, plants and other biomass; critical resources may be depleted by overuse, for example, if they

are exploited at a faster rate than they are replaced, so they require prudent management

– **non-critical** – 'everlasting' resources such as tides, waves, running water, wind and solar power.

Resource management involves controlling the exploitation and use of resources in relation to the associated economic and environmental costs. A key element of this is the concept of **sustainable development,** which requires a controlled system of resource management to ensure that the current level of exploitation does not compromise the ability of future generations to meet their needs.

Water is a renewable resource within the confines of the closed global hydrological system. However, water is effectively **finite** since only about one per cent of freshwater is easily available for human use, so it is also important to consider what limitless exploitation of this resource might lead to. Water can only truly be considered renewable at any location when there is carefully controlled usage, treatment and release.

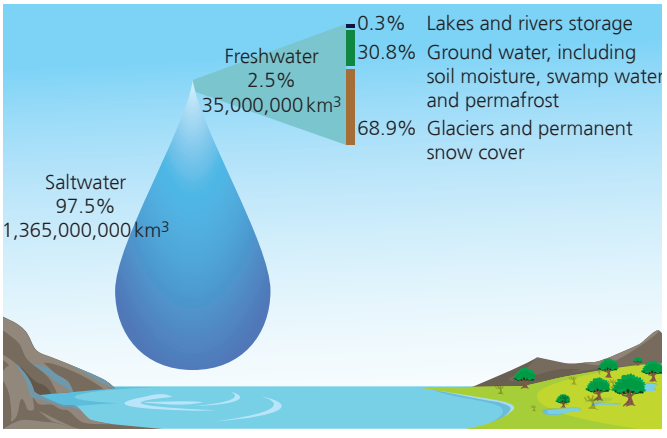


Figure 11.2 Distribution of Earth's water by store

The majority of human uses require fresh water. Figure 11.2 indicates that of the 1.4 billion km³ of water on Earth, only about 2.5 per cent of this is fresh water (97.5 per cent is salt water) and nearly 70 per cent of this is locked up in ice caps, glaciers and permanent snow. **Groundwater** represents about 90 per cent of the world's readily available freshwater resource so only a small fraction of the Earth's fresh water is available to us on the surface or in the atmosphere. Around 1.5 billion people depend upon groundwater for their drinking water but the supply is steadily decreasing. Depletion is most prominent in Asia and North America where rates of **abstraction** exceed the rates of **recharge**.

Key terms

Abstraction (water) – The removal of water available in the environment, either permanently or temporarily, from rivers, lakes, canals, reservoirs or from underground aquifers.

Groundwater – The water stored underground in the cracks and spaces in soil, sand and rock. It moves slowly through geologic formations called aquifers.

Recharge (groundwater) – A hydrological process involving the downward movement of water by infiltration and percolation causing the replenishment of natural groundwater. It occurs naturally but can be done artificially through human intervention and management.

Stock resource evaluation

Stock resources are mainly mineral deposits found in or on the Earth's crust. Mineral resources are a concentration of naturally occurring, solid, liquid or gaseous, inorganic or fossilised organic material, including precious metals and gemstones, mineral ores (such as iron and copper) and fossil fuels such as oil, gas and coal. Mineral deposits are potentially valuable, especially when discovered in sufficient quantity and of such grade or quality that reasonable prospects exist for their economic extraction.

There is a distinction between 'resources' and 'reserves'. The term 'resources' includes all deposits of mineral resources, whether known, undiscovered or discovered but unviable economically. Alternatively, 'reserves' are the parts of the resource base available for use because of being economically viable and technically feasible to extract. The distinction between reserves and resources can easily become blurred because commodity prices are so volatile and technological advances are relatively rapid. Thus a combination of continued increases in price of a particular mineral and improved technology (which reduces costs) to access and exploit it, will convert more 'resources' into actual 'reserves'.

A McKelvey box or diagram (named after US geologist Vincent McKelvey) is a visual representation that helps to distinguish between resources and reserves and the differences that exist within these two broad categories. In Figure 11.3a the larger box represents the entire 'resource base' of a given mineral such as copper ore. The smaller cube within the resource base is the actual copper reserve which is determined by economic recoverability, based on access and cost against

market price and the degree of geological assurance or certainty of its existence.

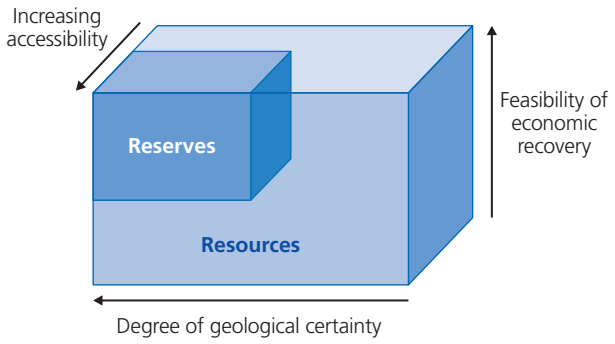


Figure 11.3a McKelvey box showing the distinction between resources and reserves based on geological assurance and economic feasibility

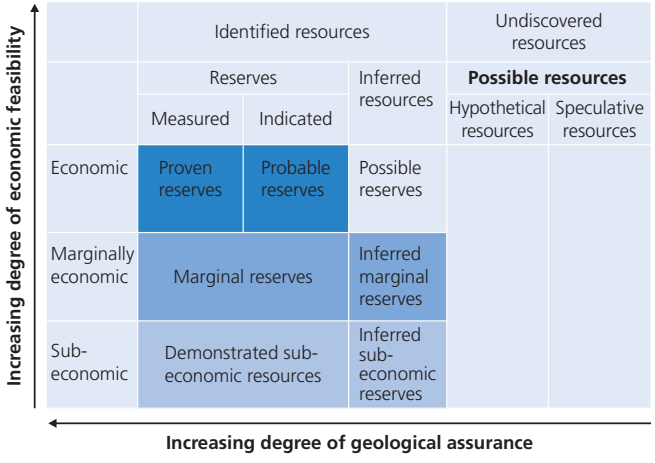


Figure 11.3b A McKelvey rectangular diagram is a useful visualisation for classifying resources by their degrees of geological assurance and economic feasibility

Figure 11.3b shows mineral resource classifications in relation to two distinct parameters. The first of these is the degree of **geological certainty** that a deposit actually exists. This is based on initial survey information about its characteristics including the quality of the deposit, the amount of material and its location and depth. The second parameter is the **economic viability** of the deposit, which is based on estimates of the amount of money received for the material balanced against the costs of extraction. Profitability can vary widely with upgrades in technology, shifts in world prices and levels of competition existing for that specific resource. Transport costs will also have a bearing on profitability.

Measured reserves, indicated reserves and inferred resources

Mineral resources and reserves are sub-divided, in order of increasing geological confidence, into **inferred**, **indicated** and **measured** categories. These categories are used as international standards to evaluate stock resources (Table 11.1). Although there are slight variations in terminology between US, Canadian and Australian standards, these three large mining nations use the same categories based on the McKelvey concept. 'Inferred' mineral resources rate a lower level of confidence than that applied to 'indicated' mineral resources. An indicated mineral resource has a lower level of confidence than a 'measured' mineral resource.

Table 11.1 Stock resource evaluation

← Increasing level of confidence for viable exploitation			
Criteria for evaluation	**Measured reserves**	**Indicated reserves**	**Inferred resources**
Assurance of quantity, grade or quality, densities and physical characteristics	Well established – estimated with confidence	Estimated with sufficient confidence to allow evaluation of economic viability	Estimates based on assumptions that existing deposits are continuous
How reserves are assessed for grade, quality and quantity	Computed from dimensions revealed in outcrops, workings and drill holes	Similar information as for measured reserves	Estimated from assumed continuity of deposits near areas already measured and tested
Degree of geological surveying and sampling	Detailed sampling; inspection sites closely examined to allow declaration of tonnages and grades	Inspection sites are spaced further apart, reducing level of confidence that the reserve is continuously viable	Limited geological sampling. Deposits are not verified by continuous samples or measurements
Summary	If economically viable – measured reserves are known as **proven reserves** and constitute mining assets with the greatest degree of confidence	Degree of assurance is lower than for measured reserves but high enough to establish an estimated tonnage sufficient to support feasibility studies and conversion to **probable reserves**	Due to uncertainty, there is insufficient information to declare tonnages or grades with any confidence. Viability for mining is not validated so they remain only as **possible reserves**

Possible resources

The diagram in Figure 11.3b (page 557) distinguishes between possible reserves and **possible resources.** Estimates of possible resources are based on broad geological knowledge of the existence of other, mostly undiscovered, deposits of the resource base and an expectation that these may become economically viable in the long term. There is less confidence of these resources becoming available and viable than for inferred resources and so the timelines for possible future production are much longer. They consist of:

- **Hypothetical resources**: undiscovered materials that are reasonably expected to exist in known mining regions under known geological conditions.
- **Speculative resources**: undiscovered materials that may occur in known types of deposits in geological settings where no previous discoveries have yet been made.

Resource development over time

Security of supply of resources depends upon a combination of **physical risks** and **geopolitical risks**. Physical risks relate to the accessibility of the resources available in an area and are evaluated against:

- the quantity of the resource that has been found
- the quality of the resource in that reserve
- its physical location and accessibility
- the technology available to access the resource economically.

Geopolitical risks include:

- the concentration of production in a relatively small number of countries
- the confidence any individual country has in trading with producers who will seek to exert their market power
- the possibility of conflict, war or other major political tensions in producing regions.

Links

See Chapter 7, Global systems and governance for further information on International trade and access to markets, pages 306–14, and Unequal power relations in global systems, page 303.

All of these factors need to be considered before deciding whether to exploit a resource in an area. Only if the economic benefits outweigh these risks will production of the reserve go ahead. Advancing technology has a key role to play in the future availability of and access to more remote resources. Equally importantly, the price or value of the resource may increase, prompting the **exploration** for reserves in more remote locations, as they become more viable for production. Exploration for mineral deposits that have the potential to become an economic resource is time consuming and expensive, and unlocking the economic value of embedded mineral deposits is even more so. For example, the cost of building a modern ore mine starts at a minimum of $500 million and can easily reach $1 billion or more. Thus reserves can be classified as **recoverable** or **possible**.

- **Recoverable reserves** are the amounts of a resource likely to be extracted for commercial use within a certain time period and at a certain level of technology. (**Proven reserves** are those claimed to have at least a 90 per cent chance of being recoverable under existing conditions.)
- **Possible reserves** are deposits that are thought to exist because the geological terrain is similar to areas that have yielded comparable deposits but no exploration has yet taken place (Table 11.1, page 557).

Cycle of natural resource development

The development of each resource will vary according to its type and location. Figure 11.4 gives the chronology in which resource development takes place. Organisations such as transnational corporations (TNCs) have to apply for licences from the appropriate authorities (for example, in the UK, from the government's Department of Energy and Climate Change) to give them the right to explore for resources in any given area.

- **Exploration** of potential sites can take a number of years. For underground reserves, geologists are employed to use satellite imagery, conduct geophysical surveys and gather field data before they can assess the extent and potential quality of the reserve.
- Discoveries have to be fully **evaluated** to determine their viability for production given existing market conditions and costs of production. In developed countries a further assessment of the **environmental impacts** of the operations is usually required before extraction can take place.

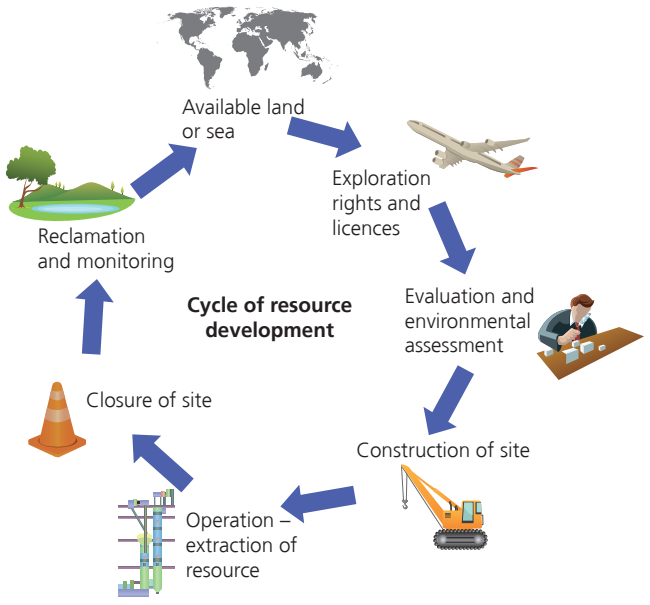


Figure 11.4 Cycle of resource development

- Once extraction licences have been awarded, the necessary infrastructure needs to be **constructed** to access the resource and start operations. This includes road building, laying pipelines (if transporting oil or gas is involved) and improving port facilities.
- **Operation – extraction** of the resource will commence and the reserve will continue to be **exploited** until all the viable reserves are recovered. The life span of the operation will vary depending on the type and quantity of the resource. (Exploration programmes will continue, to find further viable reserves in adjacent areas during operations; this will extend the life of the operation.)
- **Closure** of an operation does not necessarily mean that all the reserve has been exhausted. Any residual stock may be uneconomical to recover. In the case of valuable resources such as oil, secondary and tertiary recovery operations are put in place, for example, using steam or detergent solutions under pressure to force oil to the surface.
- **Reclamation** – in developed countries, where environmental laws are more stringent, environmental management has become an essential part of resource extraction. Operating organisations are responsible for rectifying any environmental damage resulting from the resource development. This is not always the case for resource exploitation in developing countries.

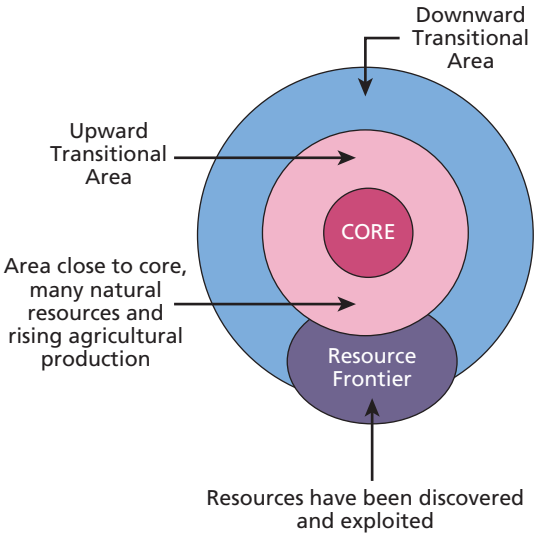


Figure 11.5 Friedman's core–periphery development model including the idea of a 'resource frontier'

Resource frontier

A **resource frontier** refers to an area where resources are brought into production for the first time.

In Friedman's core–periphery model (1963), he suggested a differential pattern of development based on a core region of a country, which attracts all the wealth and investment because of its advantageous location and existing resources. As development spreads from the core area, surrounding areas also start to benefit. In contrast, the periphery remains largely undeveloped, failing to attract investment and growth and lagging behind in standards of living and other development indicators. Within the periphery, however, is the possibility that resources such as minerals or fossil fuels will be discovered, prompting a rush of investment to the area, creating jobs directly or indirectly related to the development of the resource (Figure 11.5). Social conditions in resource frontiers may still be poorly developed but economically the region experiences rapid growth, with an influx of workers.

Resource frontiers can be identified at a variety of scales. For example, on a national level in the UK, Scotland could be identified as a 'peripheral' region far away from the core in south east England. However, the development of North Sea oil reserves from the 1980s saw the north east of Scotland become a resource frontier. The local economy boomed as it became a base for the oil industry and its associated technology and ancillary industries. Similarly, overcoming the transportation problems of crude oil in the Arctic region of North Alaska with the Trans-Alaskan oil pipeline caused Alaska to become a 'resource frontier' in the USA (Figure 11.6, page 560).

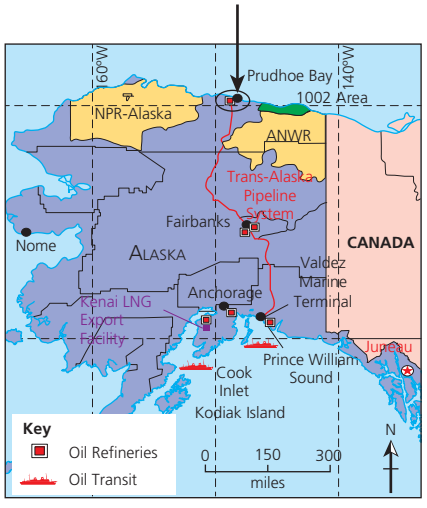


Figure 11.6 A resource frontier – the industrial petroleum extraction centre at Prudhoe Bay on north Alaska's Arctic Ocean. Its location is shown on the map

On a global scale, the depletion of resources in 'core' areas of high-income countries in North America and Europe prompted a search for mineral and fossil fuel reserves elsewhere. The increasing influence of energy and mining TNCs on the world stage, together with the process of globalisation, have led to the discovery of many new reserves of oil and gas in lower-income countries, particularly in Asia, Africa and Central America. Countries such as Azerbaijan, Chad and South Sudan (all with considerable oil reserves) have become 'resource frontiers' on a global scale.

Links

There are links here with Chapter 7, Global systems and global governance – International trade and access to markets, pages 306–14, and Nature and role of TNCs, pages 314–7.

Resource peak

The **resource peak** refers to the time of the maximum rate of production of a resource, either from a given reserve or for the resource as a whole. The idea developed from the concept of 'peak oil' originally suggested by Shell geophysicist M.K. Hubbert in 1956, and has since been applied to other finite energy resources and minerals. Hubbert accurately predicted that **conventional oil** production from the lower 48 mainland states (excluding Alaska) of the USA would peak between 1965 and 1970 and then enter a decline. His theory suggests that at any scale, from an individual oil-producing reserve to the planet as a whole, the **rate** of oil production tends to follow a bell-shaped curve (Figure 11.7).

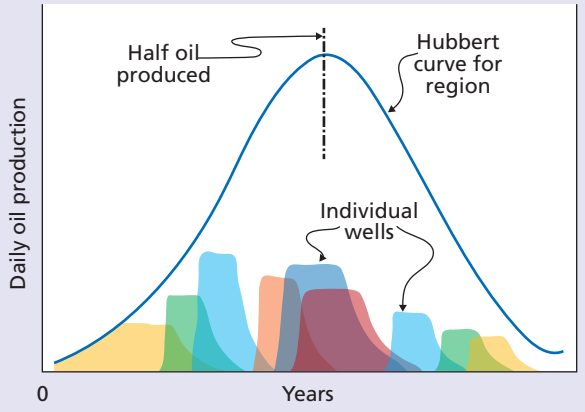


Figure 11.7 Hubbert's bell-shaped peak production curve for oil

Peak production curves for oil are also known as 'Hubbert curves' and can be plotted for all reserves as they are discovered. The concept of peak oil is sometimes confused with oil depletion; it is important to understand that peak oil is the point of maximum production, while depletion refers to a period of falling

Key terms

Conventional oil and gas – Refers to petroleum, or crude oil, and raw natural gas extracted from the ground by **conventional** means and methods.

Cost–benefit analysis – A systematic analysis of the advantages and disadvantages likely to result from a development project, where an objective value is allocated to all economic, social and environmental aspects affected.

Unconventional (oil and gas) reserves – Hydrocarbon reservoirs that have low permeability and porosity and so are difficult to produce. They require enhanced recovery techniques such as fracture stimulation. For example, shale deposits, tar sands and heavy oils.

reserves. Logically, if the resource peak represents the top of a predicted bell-shaped production curve, then only half of the total reserve has been used at that point (see Figure 11.7, page 560) and viable production can continue for many years after peak production.

Resource peaks – a moving target?

Predicting fossil fuel and other stock resource peaks at national and global scales is difficult because they are dependent on so many factors. Reports on fossil fuel peaks vary considerably because they are influenced by:

- levels of consumption
- world prices of the resource
- exploration efforts
- advances in technology.

These factors are all interdependent at national and global scales. In the global economic system, an increase in price may affect demand and consumption adversely (though, in the case of oil, relatively minimally). However, it will also prompt advances in technology and increased exploration efforts to enable exploitation of previously uneconomic reserves.

US peak oil production – a moving target?

Oil production in the USA's 48 lower mainland states (excluding Alaska) followed Hubbert's predictions quite accurately (see Figure 11.8), peaking at 9.64 million barrels per day (mbpd) in 1970 and then falling back to around 4 to 5 mbpd by the mid-2000s, as conventional reserves (with the exception of Alaska) became depleted. However, Hubbert underestimated the advances in technology, efficiency gains and effects of increasing oil prices, which enabled development and production of previously uneconomic oil reserves. The following developments have all extended peak production in the US:

- Global oil price rises in the 1970s stimulated development of the known oil reserves in north Alaska, which became economically viable, even accounting for the cost of the Trans-Alaskan Pipeline taking oil to the ice-free port of Valdez (Figure 11.6, page 560). This caused a secondary peak of US oil production by the mid to late 1980s (Figure 11.8).
- Advances in technology supported this development and enabled the exploitation of other previously unviable deep-water reserves in the Gulf of Mexico.
- Since 2008, the successful application of hydraulic fracturing ('**fracking**') to exploit unconventional reserves in shale rock on the mainland has enabled US total oil production to rebound back to its 1970 peak at 9.7 mpbd by 2015. The 'fracking revolution' (exploiting what is also known as 'tight oil') continued to contribute to increases in US oil production, which reached a new peak of 12.8 mbpd by 2020, before the global COVID-19 pandemic caused a dramatic fall in production.

While Hubbert's peak concept can be applied to individual known reserves, on a national and global scale predictions can be meaningless as economic factors and technological advances can extend or create new peaks. Whereas significantly increased levels of consumption from newly emerging economies, such as China and India, can narrow the resource peak.

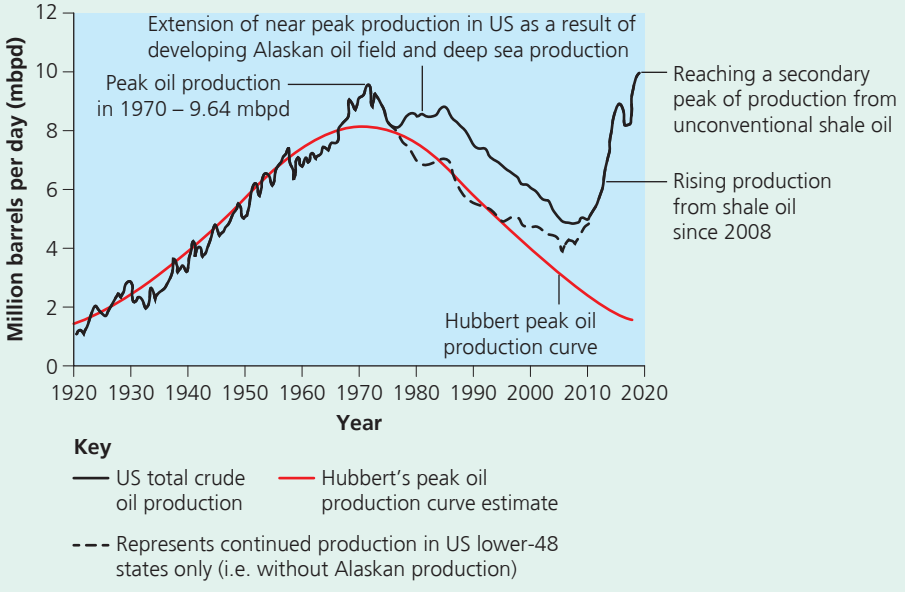


Figure 11.8 Total US production versus Hubbert's 1956 peak oil prediction, 1920–2018

Source: Energy Information Administration (1956), CEA calculations

In 2019, the **International Energy Agency's (IEA)** World Energy Outlook report stated that global oil and gas liquid production was at an all-time high and would continue to rise until at least 2040. However, oil production is expected to stabilise and decline, not because of limited oil reserves but because of:

- reduced demand resulting from emission reduction policies
- cost reductions in renewable energy generation and other alternative technologies.

With transport systems gradually moving away from carbon-emitting fossil fuels to using energy from electrical power sources (e-mobility), oil producers are now more concerned about peak oil demand than peak oil production. This change in thinking about resource peaks can be applied to all fossil fuels and has led some economists to suggest that, despite its finite condition, we will not run out of oil because there will still be sufficient reserves available at the point where there is no longer any demand.

Sustainable resource development

Resources are the keystone of every economy; when resources are used and processed they add to the wealth of present and future generations. However, the dimensions of our current resource use present major problems:

- the chance of future generations, particularly in low-income and developing countries, to have access to their fair share of scarce resources is endangered
- the consequences of our resource use in terms of environmental impacts will cause damage that goes beyond the carrying capacity of the planet
- these effects will be exacerbated once the developing world has attained levels of growth and resource use in line with the industrialised countries.

Sustainability can be achieved by different methods of managing resources, which can be broadly classified into two categories – supply side and demand side management. The two ideas are different approaches to the problems outlined above, but they are not in conflict. They are more likely to achieve success if they are used as complementary strategies.

Reducing depletion

Resource depletion is the consumption of a resource faster than it can be replenished and is mostly applied to our use of non-renewable fossil fuels and to critical renewable resources such as biomass. It equally applies to water consumption, especially in areas which rely on supplies from aquifers, when the rate of abstraction from these large underground stores of water exceeds the rate of recharge. The ability to reduce depletion depends on the nature of the resource but can generally be achieved by using demand management strategies as outlined in Table 11.2.

Table 11.2 Supply and demand management strategies for sustainable resource development

Supply side management	Demand side management
Involves seeking methods of increasing the supply of resources: • increasing exploration efforts for existing non-renewable resources • increasing research efforts to develop: – more sustainable alternative or substitute resources to replace unsustainable ones – new technologies that are more sustainable and cause less environmental impact	Involves reducing consumption of resources, individually and at all other geographical scales: • changing individual behaviour and lifestyle to discourage wasteful and/or extravagant use of resources • developing technology to enable more efficient use of resources • recycling after use • reducing population growth with population control methods so there is less pressure on resources • regulatory controls and frameworks as part of global governance, for example: – Agenda 21 – Paris Climate Agreement

Minimising environmental impacts

Another key aspect of sustainable development is minimising the environmental impact of resource use. Technological advances, such as catalytic converters on vehicles, have mitigated our polluting habits. Similarly, flue-gas desulphurisation (FGD) plant and carbon capture and storage (CCS) technology have reduced the emissions of sulphur and carbon into the atmosphere from coal-fired power stations.

Seeking alternatives/substitutes

A whole array of renewable energy alternatives has evolved to take the place of non-renewable resources as they start to run out. However, with renewable energy still only contributing around 18 per cent of the global energy mix, commentators argue that the rate at which these are being introduced has not been sufficient to significantly reduce the use of fossil fuels. While non-renewable energy is still the main source of energy the negative impacts on the environment will continue.

Environmental Impact Assessment (EIA)

EIA is a process used to identify and understand the **environmental consequences** of a proposed project, such as resource development. It informs decision makers of the environmental impacts before a decision is made on whether to proceed with the project. Assessments are often considered in conjunction with related socio-economic and human health impacts.

EIAs are an integral part of development projects around the world but the assessment process varies. With international environmental laws being complex, specialist environmental consultants are engaged by government authorities and developers to undertake an EIA and to prepare the final **environmental statement**. For example, Atkins, a London-based firm, have prepared EIAs for water and energy infrastructure projects all around the world. The EIA process in the UK derives from an EU Directive (Table 11.3) but is likely to remain similar after Brexit. In essence, the process:

- assesses the impact of changing land use on the environment
- quantifies potential environmental impacts and attempts to place a value on them so that they are considered on an equal footing with economic factors
- enables decisions to be taken with full knowledge of environmental consequences
- suggests modifications or alternatives to the proposal that would mitigate its impact.

EIA in relation to resource development projects

A typical application of EIA in relation to resources would be the development of an open-pit (also known as open-cast) mine for a mineral ore, such as copper. The assessment would consider a range of impacts caused by such a project including:

- **aesthetic problems:** degradation of landscape
- **air pollution:** dust and particulate matter (PM) from mining and vehicle movement
- **noise pollution:** from blasting and from heavy vehicles
- **water pollution and turbidity:** increased sediments in drainage water
- **toxic waste:** the prevention of **tailings** (see pages 623–4) escaping into water courses or groundwater
- **dereliction:** abandonment of site when mine exhausted and closed.

The EIA would identify the magnitude and importance of each of these problems and suggest mitigation measures so that the project could still go ahead.

Review questions

1 What is meant by 'resource security'?

2 Outline the stages necessary in the development and extraction of a resource.

3 Identify reasons why predicting resource peaks of stock resources is difficult.

4 Explain what is meant by 'the sustainable development of resources'.

Table 11.3 The main stages of the EIA process typically required by EU environmental policy

Stage	Description
1. Outline of the proposed development	Description of the site Construction, operational and decommissioning plans Identifies all sources of disturbance (e.g. noise, visual or air pollution)
2. Existing environmental conditions surveyed	Identifies all features of the environment that will be affected by the project, including water, soil, flora, fauna and social and cultural aspects
3. Assessment of all environmental impacts	Identifies all the significant impacts in the context of the project. For example, with an open-cast mine the impacts may be land degradation, loss of flora and fauna etc.
4. Modifications that mitigate the impacts	Consideration of all the measures that might be taken to reduce the impacts or that offer remediation Might include identification of alternative developments
5. Environmental Statement published	A non-technical summary that outlines the environmental impacts and may evaluate these against social and economic costs and benefits in a cost–benefit analysis
6. Decision made by the relevant authority	This is usually done after public consultation after consideration of all the costs and benefits There is often the right to appeal against the decision

11.2 Natural resource issues – Energy

Energy is one of the main drivers of economic and social development. There remains a marked **energy gap** between the rich and poor nations of the world.

- Nearly one billion of the world's population, living in low-income countries, have no electricity or other modern energy supplies and depend almost entirely on fuelwood or other biomass for their energy needs.
- Developed countries consume around 70 per cent of the total supply of the three major fossil fuels.

Global patterns of production, consumption and trade

Links

See Chapter 7, Global systems and global governance for further information on Trade agreements, pages 295–8, and International trade and access to markets, pages 306–14.

Globally, we are still largely dependent on fossil fuels, which make up around 80 per cent of the global energy mix; a proportion that will decrease only gradually in the near future. However, a number of significant changes in the global patterns of energy production, consumption and trade are taking place, such as the:

- depletion of fossil fuel reserves in many traditional producing regions
- exploration and development of fossil fuel reserves in developing countries, particularly in Asia, and in parts of Latin America and Africa
- development of unconventional oil and gas reserves from shale gas and oil, oil sands and heavy oil
- global policies encouraging a transition away from the use of fossil fuels towards sustainable energy sources to reduce carbon emissions
- rapid growth and industrialisation of emerging economies in Asia and South America, particularly the larger economies of China, India and Brazil, have led to massive increases in energy consumption there.

The main energy resources produced, consumed and traded are fossil fuels – coal, oil and gas:

- **Coal** is relatively abundant, though it is unlikely to grow in popularity as a future resource because it is so polluting. Carbon capture and storage and cleaner coal-burning technologies may reduce these concerns.
- **Oil** and **gas** are less evenly distributed globally and less abundant than coal. They are highly valued because of their flexibility of use in transport, heating and power generation. Accessing unconventional reserves will extend the longevity of both gas and oil supplies.

As technology advances, demand will most likely determine oil, gas and coal production in the future (see Resource peak, pages 560–2). The geographical distribution of resources will also determine their future use as primary energy sources.

Patterns of energy production

Global energy production is unevenly distributed. The areas of highest production are concentrated in North America, the Middle East and throughout Asia.

Some countries are '**energy rich**' because they have large reserves of fossil fuels (Table 11.4). Energy-rich nations have the wealth and technology to exploit these reserves and many are also large consumers of their own energy resources, for example, China and the USA. As reserves of fossil fuels in richer countries reach peak production, two major trends have emerged that have influenced the pattern of production:

- Exploration and production of gas and oil reserves in less developed countries has increased. Governments have given exploration rights to TNCs, whose investment and expertise have helped resource development, for example, in Indonesia, Chad, Angola and South Sudan.
- Increases in oil and gas prices, together with advances in technology, have prompted the development of **unconventional resources**, including reserves from shale rock in North America, oil sands in Canada and heavy oils in Venezuela.

Table 11.4 Top producing countries of the three main fossil fuels in 2017 (measured in Millions of tonnes of oil equivalent – Mtoes)

Top coal producers 2017	Mtoes	Top oil producers 2017	Mtoes	Top natural gas producers 2017	Mtoes
China	1,800	USA	591	USA	642
USA	373	Saudi Arabia	569	Russia	581
Australia	293	Russia	549	Iran	183
India	270	Canada	249	Canada	153
Indonesia	263	Iran	236	Qatar	150
Russia	222	Iraq	232	China	124
South Africa	145	China	192	Norway	109
Colombia	59	UAE	179	Australia	88
Poland	50	Kuwait	148	Algeria	82
Kazakhstan	50	Brazil	140	Saudi Arabia	78

Source: International Energy Agency

Some countries have limited fossil fuel reserves (even developed nations such as Japan and France) and rely on imports or other sources, such as nuclear or renewables, for their energy supply. Others are unable to secure the investment needed to develop the available resources. In some cases, conflict or political instability restricts resource development and production.

Skills focus

Using Table 11.4, calculate proportions and draw three pie charts for coal, oil and gas to show the top ten producing countries of each.

Coal production

- Coal production is distributed across the globe due to its relatively common availability (Figure 11.9).
- It is sometimes called the 'fuel of the past' because it is the most polluting of the fossil fuels. Despite this, it still accounts for around 25 per cent of the global energy mix.
- There has been a movement away from deep mining; the bulk of coal production is now from open cast mining.
- Geographically, global production of coal is dominated by China, the USA, Australia, India and Indonesia. These five countries together account for nearly 80 per cent of world production.
- Thermal electricity generation is the primary use of coal so production is often determined by a country's need for a coal based energy strategy.
- The largest producers consume much of their own coal for power generation, especially the USA, China and India.

Oil production

- Oil production is concentrated in three main areas globally: the Middle East, North America and Russia. Over 70 per cent of the world's oil production originates from these regions.
- Due to large reserves, Middle East countries contribute over 25 per cent of the world's production.

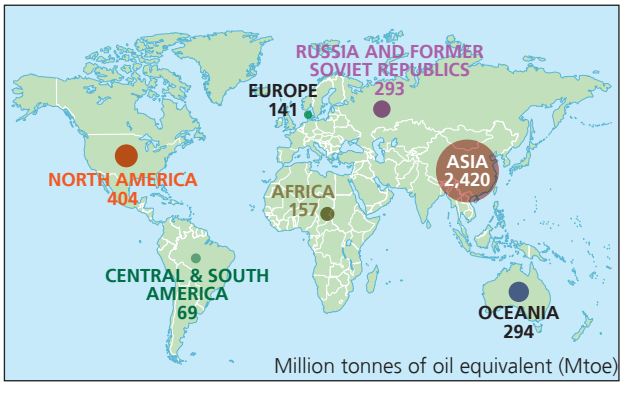


Figure 11.9 Pattern of world coal production, 2017.
Source: IEA

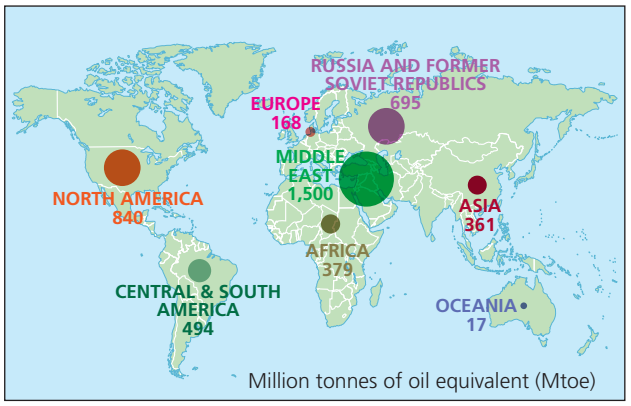


Figure 11.10a Pattern of world oil production, 2017
Source: IEA

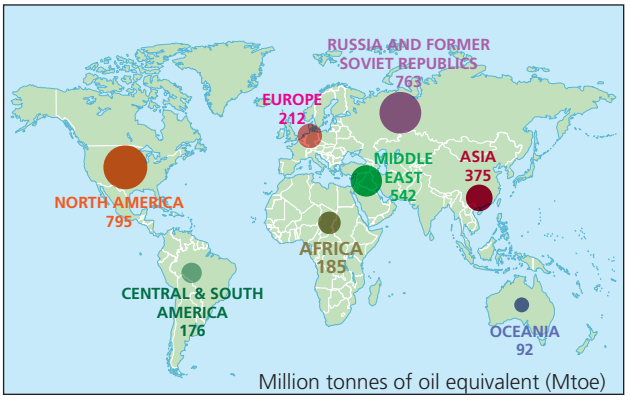


Figure 11.10b Pattern of natural gas production, 2017
Source: IEA

- Exploiting unconventional shale (tight) oil reserves using fracking has enabled the US to become the top oil producer. Similarly, less conventional reserves, such as oil sands in Canada and heavy oil in Venezuela, have enabled these countries to become important producers.
- Geological factors influence the global pattern of oil supply. Extraction tends to be more viable in Middle Eastern countries which have large oilfields, such as Saudi Arabia, compared to relatively smaller reserves found in the North Sea.
- Other factors affect production, for example, oil extraction in the physical environment of the Middle East is cheaper when compared to the harsh cold climates found in Alaska, where oil is deep below preserved permafrost layers. Conversely, the Middle East has a more politically volatile environment than other oil producing regions.

Natural gas production

- Natural gas has become an increasingly important energy resource within the last 50 years. It contributes about 22 per cent of the global energy mix.
- In 2019, world natural gas production hit a new record of almost 4,000 billion cubic metres (bcm).
- Natural gas deposits are generally found in the same locations as oil reserves. Global distribution of reserves are similar but gas production is usually determined by the availability of infrastructure to store and transport it. The USA is the largest producer and the largest natural gas reserves are found in Russia.
- Production is dominated by the USA (increasingly producing shale gas) and Russia. The top five gas producers (Table 11.4 page 565) represent over half of the world's natural gas production.

Nuclear energy

The primary energy source for nuclear fission is the chemical element uranium. Mining uranium is dependent on the quality and quantity of discovered ores. Larger concentrations of higher grade ores tend to occur in continental shield areas of the Earth. The five largest producers of uranium in 2017 were Kazakhstan (38 per cent), Canada (16 per cent), Australia (10 per cent), Namibia (7 per cent) and Niger (6 per cent).

Electricity is produced from uranium using nuclear fission. Wealthier countries are the largest producers of nuclear energy. They can afford to invest in it to diversify their energy mix and secure a domestic energy supply. The USA produces the most nuclear energy and France, lacking in fossil fuels, relies on nuclear energy to produce 75 per cent of its electricity.

Renewable energy

Renewable energy sources include solar, wind, hydroelectric power (HEP), bio fuels and geothermal energy. Together they account for around 18 per cent of world energy (and more than 26 per cent of electricity production) and are the fastest growing sources of energy production.

- Renewable energy sources can be established in many countries so the distribution of production is scattered globally.
- Table 11.5 shows the five largest producers of renewable energy.
- Around 40 per cent of the world's renewable energy is produced in low-income countries from solid biofuels (mainly fuelwood) and other forms of biomass.
- Significant developments in solar and wind energy, due to supportive policies by national governments in developed countries, have raised the importance of renewable energy production, especially for electricity generation.

Patterns of energy consumption

Global energy use is increasing by approximately 2.5 per cent per year while population growth is about 1.1 per cent. The increase in energy consumption is therefore not due to population growth alone. The amount of energy used per person globally is increasing 'on average' by about 10 per cent but varies between different countries.

Global energy consumption grew by 3 per cent in 2018 to 14.5 Gtoes (gigatonnes of oil equivalent), stimulated by the sustained economic growth and rising demand in China, the world's largest energy consumer since 2009. Energy consumption in China is mainly driven by power generation, strong industrial demand and increasing fuel consumption by private transport.

Energy consumption in the United States reached a record high of 2.3 Gtoes in 2018, an increase of 3.5 per cent from 2017, partially driven by weather conditions (hot summer, cold winter).

In contrast, energy consumption decreased in the European Union (-1 per cent) and in particular in Germany (-3.5 per cent) due to decreasing consumption in the power sector, a milder winter and energy efficiency improvements.

The US Energy Information Administration (EIA) projects that global energy consumption will grow by nearly 50 per cent between 2018 and 2050. Most of this growth will come from regions such as Asia, where strong economic growth is driving demand. The majority of that demand will continue to be met by non-renewable energy sources. However, with electricity generation expected to increase by 79 per cent over that period, renewable energy will be the fastest growing energy source and will nearly double its contribution to the global energy mix by 2050.

Table 11.5 Top producers of renewable energy, 2016

Country	Total renewable energy capacity in gigawatt (GW)	Main sources of renewable energy
China	545	Mainly HEP (319 GW); wind (149 GW) and solar (77 GW) Three Gorges Dam is the largest HEP plant in the world (22.5 GW)
USA	215	Mainly wind and solar (new wind farms under construction (26 GW) and two major solar projects in California, each 550 MW) Some HEP and geothermal
Brazil	123	Mostly HEP (98 GW) and wind (11 GW)
Germany	106	Mainly wind (50 GW) and solar (41 GW)
Canada	97	Mostly HEP (81 GW), some wind (11 GW) and solar (4 GW) Developing tidal energy in Bay of Fundy
India	91	Mostly wind (60 GW), remainder is solar, some HEP and biomass

Source: International Renewable Energy Agency (IRENA) and NSEnergy

Table 11.6 Reasons for changing patterns in energy consumption

High-income developed countries	Developing and emerging economies
High standards of living demanding more electrical and electronic appliances	**Rapid industrialisation** – as manufacturing moves to lower-cost economies – industrial development relies on energy supplies
High rates of **vehicle use**	**Urbanisation** – as more people live in urban areas, they have greater access to energy supplies
Growth has levelled because appliances and vehicles are becoming more **energy efficient**	**Rising incomes** – as populations become more affluent they aspire to the same living standards enjoyed in richer countries, so they use more appliances and vehicles
Heavy **energy-intensive industries are in decline** because of global shifts in manufacturing to developing countries	**Energy producing countries** are using their energy wealth as a trigger for development and as more reserves are exploited, some production is used by the home market
Population growth is more stable (this does not affect per capita figures)	
Increased **awareness** of environmental impacts of energy use, especially campaigns and policies to reduce carbon emissions	
Better technologies and incentives for **energy conservation**	

Consumption of fossil fuels

Coal: As well as being the leading producer, China dominates consumption of coal followed by India and the USA. These three countries account for over 70 per cent of global coal consumption (Figure 11.11).

Oil: The USA is the largest oil consumer, followed by the EU, China and the Middle East (Figure 11.11). Oil's share of global primary energy consumption has declined as it has been displaced by natural gas and nuclear for electricity production. However, it still accounts for over 90 per cent of energy consumed in the transport sector.

Natural gas: The consumption of gas has tripled over the past 40 years and it now accounts for over 20 per cent of global energy consumption. This dramatic increase in gas consumption is a result of:

- large discoveries in many countries
- development of long gas pipeline connections
- development of liquefied natural gas (LNG) plants and carriers to transport it
- environmental issues (gas is less polluting; it emits only half of the carbon dioxide of coal)
- development of gas **combined cycle technology** for more efficient electricity generation.

The largest consumers in rank order are the USA, the Middle East, the EU and Russia (Figure 11.11), and together they account for 60 per cent of world consumption.

Global trade in energy

Countries initially try to achieve energy security with their own sources, but many countries consume much more energy than they can produce so they have to import (for example, Japan and Spain). Conversely, some countries produce more energy than they need and for many low-income countries this provides a much needed source of foreign exchange.

Links

There are links here with Chapter 7, Global systems and global governance – International trade and access to markets, pages 306–14.

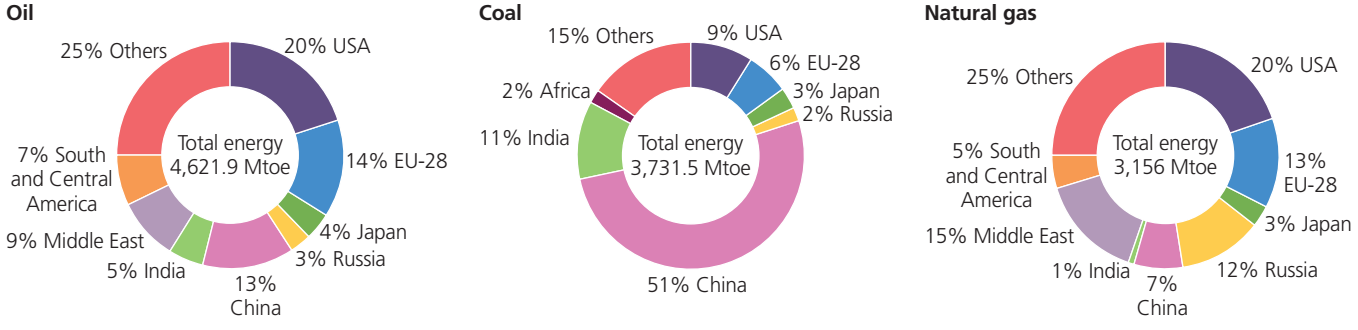


Figure 11.11 Pattern of global energy consumption of fossil fuels, 2017

Source: BP Statistical Review of World Energy

By far the most traded energy resources are fossil fuels because they are needed by many countries for electricity generation and transport. Oil is the most traded because of the mismatch between major areas of production and areas of most consumption, so the fuel will travel vast distances to reach its consumers. These distances create many challenges:

- the environmental risks of long-distance pipelines
- environmental hazards such as oil spills at sea
- political instability in producing regions disrupts supplies.
- The largest net exporters of oil are Saudi Arabia, Russia, Iraq, Canada and the United Arab Emirates.
- The largest net importers are China, USA, India, South Korea and Japan.
- Even though they are major producers, China and the USA still rely on oil imports.
- Until recently the USA was the largest net importer of oil but became a net exporter in 2019 due to increases in shale oil production.
- If shale production continues to expand, it is expected that the USA will be a major net exporter by the late 2020s.

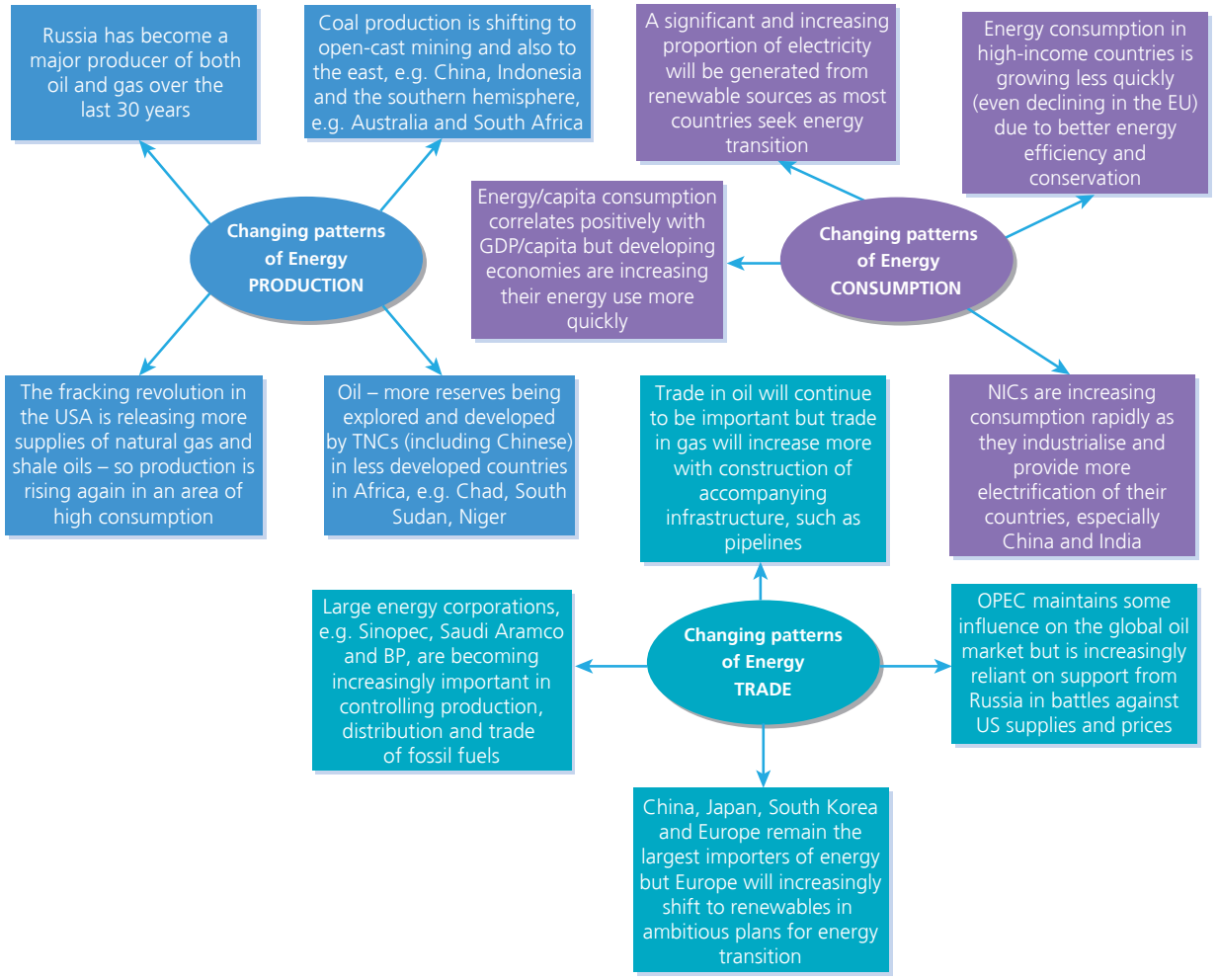


Figure 11.12 Changing patterns of energy production, consumption and trade

Transport and storage difficulties mean that natural gas is traded less than oil. International trade is hindered by the need to build pipelines, requiring huge investment, though conversion into liquefied natural gas (LNG) enables it to be carried by sea.

- Proximity to large consuming economies makes pipelines more viable.
- Russia is the largest exporter of natural gas, able to supply both Europe to the west and China, Japan and South Korea to the east by pipeline.
- Norway and Algeria are also major exporters to Europe, as is Canada to the USA.
- China, Japan, Germany, Italy and South Korea are the main importers of gas.

As coal is a low value, bulky fuel, it is less economically viable for transportation so most coal is used in the country in which it is produced. About one-fifth enters the world market and the major exporters are Indonesia and Australia as they have excess beyond their domestic needs. China, India, Japan, South Korea and Europe are the main importers.

Renewable energy resources are not usually traded as they are provided *in situ* within the country of production, though the electricity produced from renewable sources can be exported.

Geopolitical issues – Energy

Links

There are links here to Chapter 7, Global systems and global governance – Trade agreements, pages 295–8, Unequal power relations in global systems, pages 303–5, and International trade and access to markets, pages 306–14.

Key term

Geopolitics – The study of international relations, as influenced by geographical factors.

Geopolitics involves understanding the relationships between a country (or group of countries) and the rest of the world. Each nation (or group of nations, for example, the EU) has a sphere of influence that it exerts over other nations in terms of trade, economic aid and military intervention.

From the 1990s the political alliances of countries that were based on the Cold War divisions started to dissolve, though some are still rooted in the old East/West divide. The geopolitics of energy created its own alliances, for example, OPEC is an economic alliance of countries with a surplus of oil that they are able to export to developed countries which need it more. OPEC still has a major influence on the trade in oil as the organisation is responsible for about 50 per cent of crude oil exports.

Trade and management (co-operation and conflict)

Good political relationships concerning energy will lead to **co-operation** between countries in order to secure energy supplies for the future. There are also a number of potential threats and **conflicts** regarding the future security of energy supplies:

- Saudi Arabia at present has strong links with the USA and other Western countries, but this could change with leadership transition, economic reform and the perceived threat of terrorism.
- The Arab Spring uprisings that have occurred in the Middle East and North Africa (MENA) region since 2010 have led to differing outcomes for individual oil producing states, such as Libya, which are likely to affect future supplies.
- Iran is a major oil and gas supplier. Its relationship with the Western world, particularly with the USA, has been extremely difficult for many years. Sanctions were imposed on Iran in 2006 because of concerns that it was developing a military nuclear capability. Although Iran continued to trade with China and Russia, its economy was hit hard by the sanctions. An international agreement, signed in 2015, lifted the sanctions enabling Iran to trade oil with Europe and North America. However in 2019, Iran's growing military influence, particularly against US allies in the region, caused President Trump to re-impose sanctions on trade between the US and Iran.
- Trade relations between Russia and Europe are threatened because of Russia's support for separatists in eastern Ukraine and its annexation of Crimea from Ukraine. Ukraine's government has allied itself with Europe so further escalation of the conflict in the Ukraine (as well as Russia's involvement in the Syrian conflict) has led to the EU placing more sanctions on Russia, even though this puts Europe's gas supply at risk.

Organization of Petroleum Exporting Countries (OPEC) – an example of geopolitical co-operation

OPEC is an example of a geopolitical relationship between countries. The organisation's 13 member countries (Algeria, Angola, Congo, Equatorial Guinea, Gabon, Iran, Iraq, Kuwait, Libya, Nigeria, Saudi Arabia, United Arab Emirates and Venezuela) are all developing countries located in the Middle East, Africa and South America.

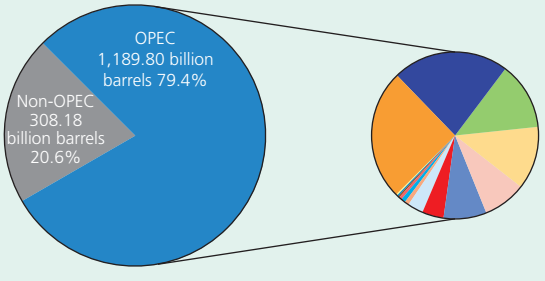
OPEC operates as a 'cartel' of co-operative producers attempting to control supply to the world market in order to influence the price of oil. Twice in the 1970s OPEC restricted the supply of oil on the world market to cause huge price hikes in oil and its derivatives. This strategy reduced depletion of OPEC oil reserves but prompted wealthy importing countries in North America and Europe to develop their own reserves (in Alaska and the North Sea respectively) which became more viable because of the price rises.

The development of oil resources in many non-OPEC countries means that the cartel has less control of world markets than it used to but its influence is still strong (Figure 11.13). For example, the fall in oil prices in 2014–15 has been attributed to an increase in production of US shale oil and a fall in demand as economic growth in China slowed down. However, OPEC also released more oil supplies onto the world market in an attempt to undercut US unconventional oil production (which is more costly), therefore making it less viable to produce.

OPEC has not always been a united organisation and is influenced by external geopolitics, with members having different political allegiances outside of the group. For example, Saudi Arabia and Kuwait have good trade relationships with the West (for example, USA and Western Europe) while Venezuela, Algeria and Iran look instead towards former communist states of the East for support and trade.

Membership frequently changes as countries leave or join the cartel. For example, Ecuador withdrew in 2020, feeling that OPEC-related constraints on production deterred investors from developing their oil reserves.

In recent years, OPEC has increasingly relied on other major oil producers, such as Russia and Mexico, known as OPEC+, to support their strategies to influence global prices. However, in early 2020, Saudi Arabia and Russia entered a price war with each other by flooding the market with supplies. This coincided with the COVID-19 pandemic, when demand for oil collapsed. A deal, brokered by the US, was eventually reached between OPEC and Russia to reduce production by a massive 9.7 million barrels per day to stabilise the market. The deal, which helped the US oil industry, was still unable to prevent oil prices temporarily becoming negative as producers paid customers to take oil from them due to overproduction and a lack of storage.



OPEC proven crude oil reserves, at end 2018 (billion barrels, OPEC share)

Venezuela	302.81	25.5%	Algeria	12.0	1.0%
Saudi Arabia	267.03	22.4%	Ecuador*	8.27	0.7%
IR Iran	155.60	13.1%	Angola	8.16	0.7%
Iraq	145.02	12.2%	Congo	2.98	0.3%
Kuwait	101.50	8.5%	Gabon	2.00	0.2%
UAE	97.80	8.2%	Equatorial Guinea	1.10	0.1%
Libya	48.36	4.1%			
Nigeria	36.87	3.1%			

Figure 11.13 OPEC share of world proven crude oil reserves, 2018

Source: Organization of the Petroleum Exporting Countries

Political developments in the Middle East and in the Ukraine are of great importance to the energy security of the rest of the world and the West has an obvious interest in supporting stability in both regions.

Review questions

5 Describe how the pattern of global energy production is changing.

6 Explain why the pattern of global energy consumption is changing.

11.3 Natural resource issues – Mineral ores

Key terms

Joint venture – A commercial enterprise undertaken by two or more parties who agree to pool their resources and share ownership, risks and returns. In mining, the agreement is often between two or more large TNCs or between a large corporation and the government of the host country.

Resource curse – Also known as the 'paradox of plenty' or 'resource trap'; refers to the notion that (some) countries endowed with an abundance of natural resources struggle to benefit from this resource wealth and instead experience minimal economic growth or social development.

An ore is a type of rock containing sufficiently high concentrations of minerals, including metals, which can be economically extracted for use in industrial processes. For example, cassiterite is an ore of tin and sphalerite is an ore of zinc. The ores are extracted from the earth by either underground or open-pit mining; they are then refined, often via smelting at high temperatures, to extract the valuable metal. Metals can be subdivided into a number of categories including:

- **ferrous and ferroalloy metals:** those with high iron content or those which improve the properties of steel, such as chromium, cobalt, manganese and tungsten
- **base metals:** mainly copper, lead and zinc
- **non-ferrous metals:** aluminium, magnesium, tin, nickel and also base metals of copper, lead and zinc
- **precious metals:** gold, silver and platinum
- **rare earth metals:** such as scandium and yttrium (extraction is complex and difficult).

A number of key issues affect the geography and economics of mineral ores. These are interrelated and based on their value as stock resources. The key issues are:

- **Finite nature:** As we use more, so stocks will decrease.
- **Price volatility:** Scarcity influences price but so does fluctuating supply and demand for specific minerals; this can be harmful to development in emerging economies reliant on ore exports.
- **Geopolitics:** Ores can be so valuable that they can influence relationships between countries.
- **Globalisation:** Its influence is seen in the increased movement and trade of minerals and particularly the role of large transnational mining corporations.
- **Environmental impacts:** The exploration and exploitation of more mineral deposits will inevitably lead to greater impacts on the global environment.

Global patterns of production, consumption and trade

Links

There are links with Chapter 7, Global systems and global governance – International trade and access to markets, pages 306–14, and Nature and role of TNCs, pages 314–7.

The global distribution of mineral deposits and the patterns of production and consumption of metal commodities are very uneven. For example, the USA has around 5 per cent of the world's population but consumes around 25 per cent of the world's mineral ore. This demand is partly met by domestic production but most of their minerals are now imported.

Industrial countries are very dependent on mineral resources to produce finished goods such as appliances and vehicles, and to provide services such as gas and electricity supply.

Global patterns of mineral ore production

Production of many mineral ores has shifted significantly in the past 50 years as most of the developed nations such as the UK and USA have depleted their own reserves. This shift in production of ores has been to developing parts of the world, where it has enabled some nations to become important emerging economies and middle-income countries.

The gap between consumption and production of iron, steel and non-ferrous metals (aluminium, copper, lead and zinc) has widened in the developed countries of North America, Europe, Japan and South Korea. More metals are consumed than extracted in these regions. Despite being a major producer of many metals, China has exhibited a similar trend during its industrialisation of the last 20 years, with rapidly growing consumption overtaking its extraction and production of metals.

The largest reserves of many minerals are based in some of the larger countries and, of the developed countries, Australia and Canada continue to be large producers of metal ores (Figure 11.14). They have extended production overseas to developing countries, so many of the large mining corporations have Canadian and Australian bases. The gap between increasing consumption and production is largely being filled by increased production in emerging economies:

- **South America**: Brazil, Chile, Bolivia and Peru are all leading producers of different mineral ores (see Table 11.7).
- **South and South-east Asian countries**: Countries including India, Indonesia, Myanmar and the Philippines are also major producers of iron, chromium, nickel and tin.
- **Africa**: Guinea (bauxite), South Africa (chromium and iron ore), Democratic Republic of Congo (cobalt) and Zambia (copper) are all listed in the top six producers of these metal ores.

These developing regions produce more mineral resources than they consume. Much of this is achieved from foreign investment in their mineral resource base but this can lead to 'mineral booms' and the so-called **'resource curse'** which do not always serve these economies well in the long term.

Countries	Production (in million metric tonnes)	Countries	Production (in million metric tonnes)
1. Australia	901	8. Canada	52
2. Brazil	460	9. USA	49
3. China	335	10. Kazakhstan	36
4. India	205	11.Iran	36
5. Russia	96	12. Sweden	36
6. South Africa	74	13.Mexico	22
7. Ukraine	60	14. Chile	14

Figure 11.14 World major producers of iron ore, 2018 (million metric tonnes)

Table 11.7 List of top producing countries of metal mineral ores, 2018

Mineral ore	Uses	World total production (thousand metric tonnes)	Top four producing nations
Bauxite	ore of aluminium	326,000	Australia, China, Brazil, Guinea
Chromium	alloys, stainless steel, electroplating	40,800	South Africa, Kazakhstan, Turkey, India
Copper	alloys, pipes, wire	20,600	Chile, Peru, China, USA
Iron ore	steel	2,923,000	Australia, Brazil, India, China
Lead	alloys, batteries	4,800	China, Australia, Peru, USA
Nickel	alloys, stainless steel	2,233	Indonesia, Philippines, Russia, Norway
Tin	alloys, containers	352	China, Indonesia, Myanmar, Bolivia
Zinc	steel alloys, galvanised metals	12,400	China, Peru, Australia, USA

Smelter and refinery production of metals remains located mainly in developed countries, although the balance is changing as China and other emerging economies have developed their own capacity to produce refined metals such as steel, copper and aluminium.

Consideration should also be given to the contribution made by secondary resources, mainly recycled metals, to the overall production of new materials.

Global patterns of mineral ore consumption

Historically, more developed industrial nations, such as the USA, Japan and those in Europe, dominated consumption of mineral ores. However, as with production, patterns of consumption are also shifting to newly industrialising countries, especially in Asia, as they increase their use of mineral ores in their growing manufacturing sectors. Consumption in more developed countries has slowed down, largely due to de-industrialisation in many regions. China, by becoming a global manufacturing powerhouse, has become the world's largest consumer of most mineral ores. Ores have played a crucial role in the rapid growth and industrialisation of the Chinese economy:

- Iron and steel have been used in major construction programmes in urban areas and infrastructure projects such as the Three Gorges Dam.
- Industrialisation demanded the development of reliable energy supplies. Plans to have a national grid infrastructure for electricity increased its consumption of copper to 11 million metric tonnes, nearly half of the global demand in 2019.
- China has reduced its imports of refined copper, purchasing copper concentrate instead or partially processed ore as it grows its own smelting capacity.
- Consumption of aluminium is also dominated by China, as is the use of rare earth metals (REEs), which are essential to many hi-tech products.
- As China has become the dominant consumer of mineral ores globally, recent slowing of its economic growth has curbed demand for some metals (aluminium and steel in particular), causing repercussions for world trade.

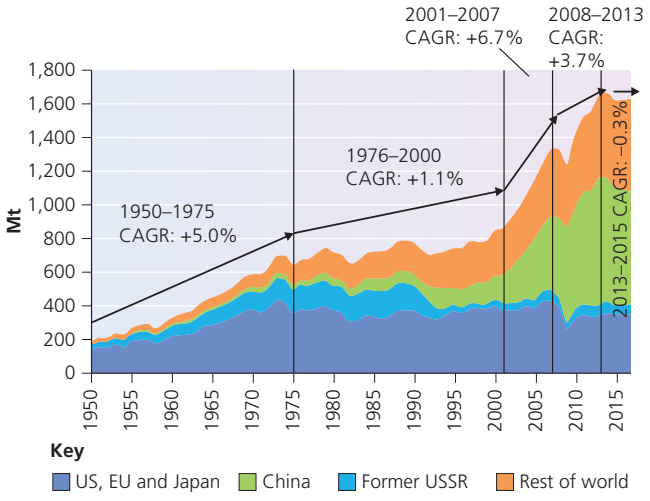


Figure 11.15 Evolution of world steel demand by region, 1950–2017

Source: Worldsteel Association

Growth in mineral consumption is also high in South Korea and India. Interestingly, on a per capita rather than total measure of consumption, South Korea surpasses all other countries in consumption of steel and copper.

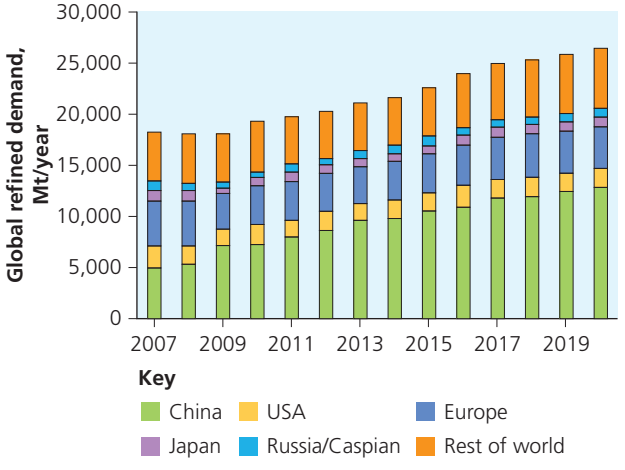


Figure 11.16 Copper demand by region 2007–2019

Global patterns of mineral ore trade

Links

See Chapter 7, Global systems and global governance for further information on International trade and access to markets, pages 306–14.

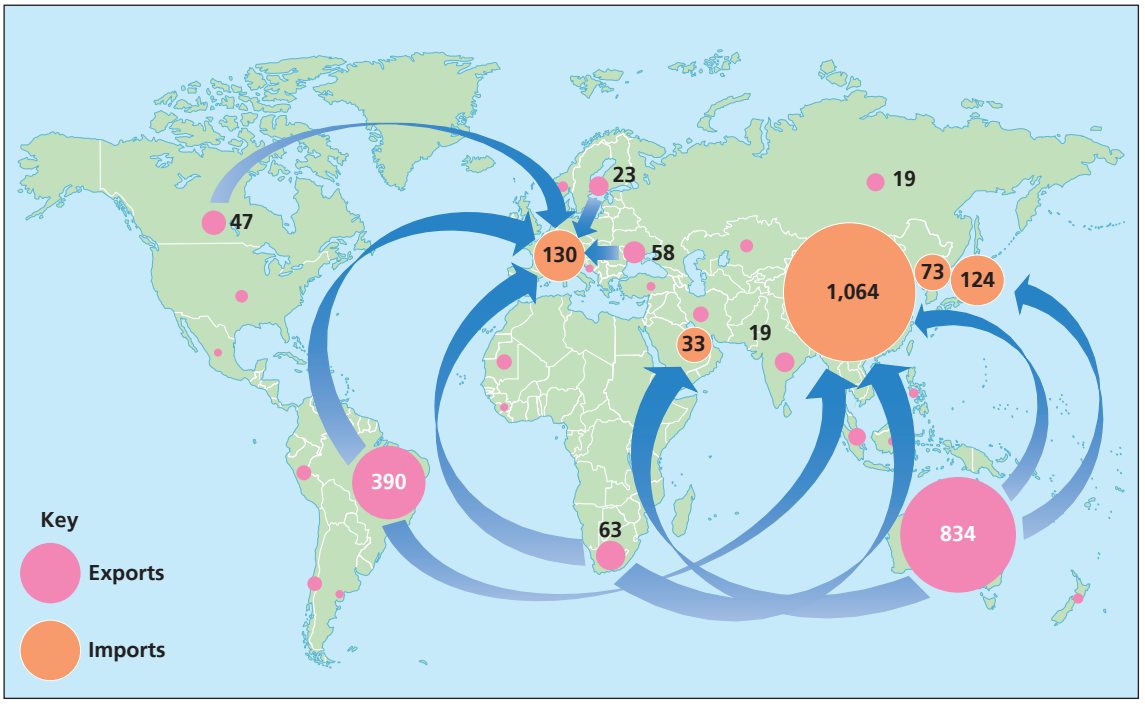


Figure 11.17 Global iron ore trade 2018 (million metric tonnes) (Iron ore exports are a function of steel production and China by far remains the largest importer)

Source: Engineering and Mining Journal

Markets in mineral commodities are more volatile than those of fossil fuels as energy demand is reasonably consistent, whereas manufacturing cycles vary. Mineral ores are usually semi-processed in the country of production in order to discard waste material and to concentrate the mineral before it is transported by bulk ore-carriers to ports near to manufacturing centres. The ore is often subject to final processing into the finished metal once it has reached its market.

With patterns of both production and consumption changing, the trade in mineral ores has also changed over the past 40 years. The geographical focus has moved away from the Atlantic towards Asia and the Pacific, particularly China and South Korea. The combination of China's resource wealth, geographical position, large population and manufacturing growth has led to its emergence as the global leader of the mineral ore trade:

- Balance between supply and demand for minerals in China is extremely vulnerable and shortages of some ores were met by importing large quantities from elsewhere in the world.
- China imports iron ore from Australia and Brazil and other ores, such as copper, from Africa. (In 2009, during a global recession, China increased imports of iron ore by 33 per cent, of copper by 14 per cent and of aluminium by eight times in order to fill the gap in resource shortages.)
- In 2014–15 the Chinese economy started to slow down, leading to a decline in the demand for mineral ores and some metals stockpiled in China.
- China became a net exporter of iron and steel. Despite still being the largest consumer of steel, they were still producing more than they needed domestically.

With the exception of China, in the recession following the global financial crisis in 2008, there was a general weakening of the global mineral ore market leading to less trade. Increased recycling of metals, such as steel and aluminium, together with the use of alternative materials (such as silicon optical fibre instead of copper in telecommunications) also reduced the demand for ores. Lower demand has led to falling prices for many mineral ores, which has caused production cutbacks by the mining TNCs that dominate global trade. Recoveries in the production of some minerals (particularly copper) by 2018/19 are likely to be countered by a further global economic recession resulting from the COVID-19 pandemic of 2020.

Trade in mineral ores is increasingly affected by a number of factors relating to the demand or supply side of the global market, such as:

- the inelasticity of demand for specific minerals needed for certain uses (determined largely by the availability of substitute materials)
- global recession and declines in manufacturing lead to falling demand and prices
- mining companies are reluctant to invest in new production with falling prices
- technological change has brought about cost reductions which have increased supply
- environmental concerns discourage new exploration and extraction but encourage recycling and alternative supplies.

Geopolitics of mineral ores

Links

There are links here to Chapter 7, Global systems and global governance – Trade agreements, pages 295–8, Unequal power relations in global systems, pages 303–6, International trade and access to markets pages 306–14 and Nature and role of TNCs, pages 314–17.

The geopolitics of minerals arises from a number of issues including the following:

- the inter-dependency existing in the international trade of raw materials
- the potential for trade wars resulting from fluctuating prices of different minerals
- the potential future conflict concerning access to 'common' resources
- the environmental and socio-economic issues arising from mine developments
- the role and dominance of large transnational mining corporations
- the role of China in the world's mining and metals industry.

Trade and management of development

Growth in exploration and mining in Africa, Latin America and parts of Asia has been spurred by:

- technological advances giving increasing viability to mining previously inaccessible deposits in remote, less developed regions
- the development of large ocean-going vessels able to carry large quantities of bulk mineral commodities such as iron ore and bauxite.

Trade in mineral ores is largely one-directional, from developing to developed countries. As with any kind of international trade, there are questions about 'who benefits' from it. It seems simplistic to assume that it will bring economic and social benefits to the developing country as a result of foreign investment. If resources are managed well by the developing country's government, it can create greater prosperity. However, many 'resource-rich' low-income countries are also subject to a phenomenon known as the **resource curse** (sometimes referred to as the 'paradox of plenty').

The resource curse means that many poorer countries, particularly in Africa, are not maximising their mineral wealth. However, the problem is not universal and some developing countries have responded by introducing measures to expand the income from their mineral deposits. For example, the Indonesian government prohibited the export of unprocessed bauxite and nickel. They negotiate licences with mining TNCs requiring them to process the mineral ores in Indonesia. This encourages processing industries, ensures a fairer return on profits and a greater share of the 'higher value' exports. The ban also resulted in falling supplies of both nickel and bauxite onto the world market since it took effect, leading to a price rise for these minerals.

China's impact on mineral ores

The role of China in the global mineral trade is geopolitically significant for countries around the world. China's actions in securing access to reserves has affected both the lower-income countries producing mineral ores and the TNCs from countries competing with China, in attempting to gain access to similar reserves. This especially applies to the iron/steel and copper industries. China is well endowed with its own mineral resources including iron ore and copper, however rapid industrial growth meant that demand for metals outstripped supply. In order to secure sufficient mineral ores for its needs, China has continued to increase domestic production as well as importing huge amounts from overseas. This has had three main geopolitical effects globally:

1. **Chinese investment in resource mining in developing countries**: China's mining industry is dominated by a number of large state-controlled corporations who operate as TNCs. These companies, such as Chinalco and Jiangxi Copper, invest heavily in establishing mining operations in developing countries, especially in Africa. The exploitation of resources has brought infrastructural benefits, such as restoration of the Benguela railway in Angola, as well as much needed investment to countries such as Zambia in the African copper belt. In Zambia, with new technologies, they have restored old copper mines and the China Nonferrous Mining Corporation (CNMC) purchased the Baluba mine in 2018 to maintain control of copper and the supply of REEs.
2. **Trade with other developing countries**: Such is China's demand for steel that they have negotiated deals with other emerging economies, such as Brazil, to import iron ore. This has been worth millions of dollars to the Brazilian economy, though a slowing-down of China's economic growth has caused problems for Brazil because of an over-reliance on exports to China.
3. **Impact on developed countries**: China has been able to manufacture large amounts of steel economically with cheap labour and state support. The slowdown in the Chinese economy led to reduced demand and from 2015, with excess production, China started to supply the world market with finished steel produced more cheaply than elsewhere. Steel manufacturers in many developed countries could not compete, which caused major impacts on the global steel industry:
   - Major consumers of steel, such as vehicle manufacturers, in the USA, South Korea and Europe could import Chinese steel at a lower price than they could buy it from steel producers in their own country.
   - In the US, the steel belt of Pennsylvania and Ohio (where the industry was already in decline), cheap steel imports resulted in closure of remaining steelworks, meaning further economic decline and rising unemployment.
   - In the UK, many large steelworks, such as Redcar, were closed or threatened with closure (Port Talbot), causing increases in unemployment and devastation of communities.
   - China's flooding of the world market with cheap steel prompted some countries to take 'protectionist' action such as increased tariffs and led to an ongoing trade war with the USA.

Resource curse

The resource curse refers to when countries with an abundance of non-renewable resources, such as minerals, tend to have lower economic growth and development than resource-poor countries. If resources are poorly managed, it can lead to economic instability, social conflict and lasting environmental damage. This is thought to be due to a number of possible causes:

- the over-dependence on the production and export of a narrow range of resources, which bring a low financial return and are subject to falling global prices
- other sectors of the economy, particularly manufacturing, do not receive the investment required for growth
- dependence on mining TNCs to provide the investment and technical expertise needed to develop the resource and the infrastructure required for extraction, processing and transport
- control of the resources by large TNCs results in leakages; only a limited share of profits remain in the country
- mining is increasingly mechanised, so job creation is limited and often low-skilled; many of the higher-income jobs go to foreign workers
- higher wages drive up the value of the local currency, impacting negatively on the competitiveness of exports in world markets (also known as '*Dutch disease*')
- authoritarian rule and, additionally, government corruption result in unfair regulation of the industry and unfair distribution of the wealth gained from it.

Role of transnational corporations in mineral ore distribution

Mining became 'global' before most other branches of industry and the distribution of mineral ores is still largely controlled by transnational mining corporations. They are competitors but will often work together in **joint ventures** to combine exploration efforts and develop mine operations. Most are multivariant in the materials they mine (in some cases this includes coal and oil). Table 11.8 shows a list of the major metal mining corporations and their operations.

Table 11.8 Examples of transnational mining corporations

Corporation	HQ/country of origin	Main mining operations
BHP Billiton	Australia	Operates some of the largest mines in the world, mainly in South America and Australia: mines copper, iron ore, uranium, aluminium, manganese, and nickel
Rio Tinto	UK/Australia	Mainly in Africa, South America and Australia, produces a wide range: iron ore, bauxite, copper, uranium
Zijin Mining Group Co. Ltd	China	State owned operating predominantly in China and Africa – gold mining and smelting, also copper, lead, zinc, silver and iron
Glencore	UK/ Switzerland	Operates globally producing copper, nickel, zinc/lead, aluminium and iron ore
Vale (CVRD)	Brazil	Primarily in Brazil – largest iron ore and nickel producer; nickel mines in New Caledonia, Indonesia and Canada; copper production in Chile and Zambia

These TNCs set criteria for the reserves that interest them, usually targeting projects that offer a lifespan of at least 20 years. They provide the expertise, technology and capital investment needed to explore and extract ores in developing countries. They are usually responsible for:

- negotiating exploration rights and leasing land from national governments
- establishing subsidiaries or entering into joint ventures with mining companies from the host nation
- developing infrastructure such as access roads and nearby accommodation to support projects.

TNCs can bring economic benefits in terms of jobs and skills to local communities but they have also been criticised for negligence of the environmental and the social impacts of their operations. Local disasters such as tailings dam failures have devastated local communities as well as the environment. Corporate social responsibility has become a more important consideration for many of these companies, so mining operations should be managed more sensitively by these operators in the future.

Review questions

7 Analyse the data presented in Figure 11.17, Global iron ore trade in 2018 (page 575).

8 Evaluate the role of large mining TNCs in the development of resources in low-income countries.

11.4 Natural resource issues – Water

In the global context, water is a renewable resource but, like energy, it will come under increasing pressure because of population growth. It is a vital resource for many aspects of life. In addition to providing drinking water, it is needed for irrigation to grow food, and in maintaining health and preventing disease through its use for hygiene and sanitation purposes. Water is also used in a variety of production processes, all of which contribute to its consumption (Figure 11.18).

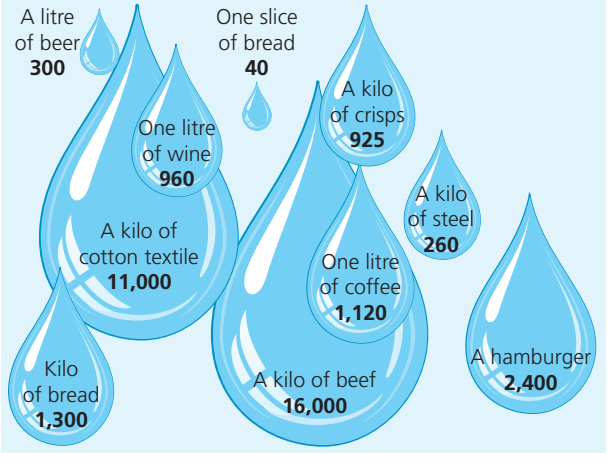


Figure 11.18 Number of litres of water used in different production processes – so called '**virtual water**'

A population increase of another two to three billion by 2050, combined with more industrialisation and urbanisation, will result in an increasing demand for water and will have serious consequences for the environment. Already there is more waste water

generated and dispersed today than at any other time in the history of our planet.

The key issues concerning water security are:

- the variable access to clean, **safe**, **potable** or **improved drinking water**
- the need to balance and conserve water supplies
- improving the supply and quality of drinking water
- reducing tensions and conflict caused by access to water supplies
- improving trans-boundary co-operation of shared water resources
- ensuring sufficient supply for food production
- reducing the environmental impacts of water use
- **water stress**.

Key terms

Improved drinking water – A source of water that by using intervening treatment is protected from outside contamination, particularly from faecal matter.

Potable (water) – Water that is suitable for drinking (it does not have to be completely pure but must not contain unacceptable levels of hazardous materials, nor look, taste or smell unpleasant).

Safe drinking water – Water that is safe for human consumption. The water must be free from harmful pollutants and bacteria that could make people ill.

Water stress – Water stress occurs when the demand for water exceeds the amount of water available during a certain period, or when poor quality of water restricts its availability for human use.

Water surplus (areas) – Areas where there is more than sufficient water available to meet human needs.

UN Sustainable Development Goals

- The UN's Sustainable Development Goal 6 is to achieve 'clean water and sanitation'. This global goal states 'clean, accessible water for all is an essential part of the world we want to live in and there is sufficient fresh water on the planet to achieve this'.
- The Millennium Development Goals (MDGs) from 2000–2015 made significant progress towards this so 90 per cent of the world's population now has access to improved sources of drinking water.
- To improve sanitation and access to drinking water further by 2030, there needs to be investment in management of freshwater ecosystems and sanitation facilities on a local level in several developing countries within sub-Saharan Africa and in parts of Central, South and East Asia.

Source: UN SDGs

Links

There are links to Chapter 7, Global systems and global governance – Global governance, pages 327–33.

Key facts about water availability

- 71 per cent of the global population (5.3 billion people) use a safely managed drinking-water service (one located on premises, available when needed, and free from contamination).
- 10 per cent of the global population (785 million) lack access to even a basic drinking-water service (an improved drinking-water source within a round trip of 30 minutes to collect water).
- Two billion people use a drinking water source contaminated with faeces.
- Contaminated water can transmit diseases such as diarrhoea, cholera, dysentery and typhoid; contaminated drinking water is estimated to cause 485,000 diarrhoeal deaths each year.
- By 2025, half of the world's population will be living in water-stressed areas.

Source: WHO

In July 2010, the Human Rights Council adopted a UN General Assembly resolution to recognise that having access to sufficient water and sanitation is a human right. Supplies must be safe, acceptable, affordable and physically accessible (the water source has to be within 1,000 metres of the home and collection time should not exceed 30 minutes) According to the World Health Organization (WHO), between **50 and 100 litres of water per person per day** are needed to ensure that most basic needs are met (including irrigation for food supply).

Most of the people categorised as lacking access to clean water use about five litres a day, one-tenth of the average daily amount used in rich countries to flush toilets. Most people need at least two litres of safe water per day for drinking and for food preparation.

Even wealthier countries, which are more able to build reservoirs to store water, or afford borehole drills and pumps to tap into aquifers underground, will put more pressure on their sources of water as population increases. For example, in the USA a major concern is the depletion of the giant Ogallala aquifer. The aquifer lies under the High Plains stretching from north to south across

eight states (see Figure 11.19). It is the most important underground source of water for domestic, industrial and most critically for agricultural needs in the mid-West states. Water is being abstracted from the aquifer at rates thousands of times greater than the recharge rate, which is not sustainable. It is estimated that it will be fully depleted within the next 25 to 30 years.

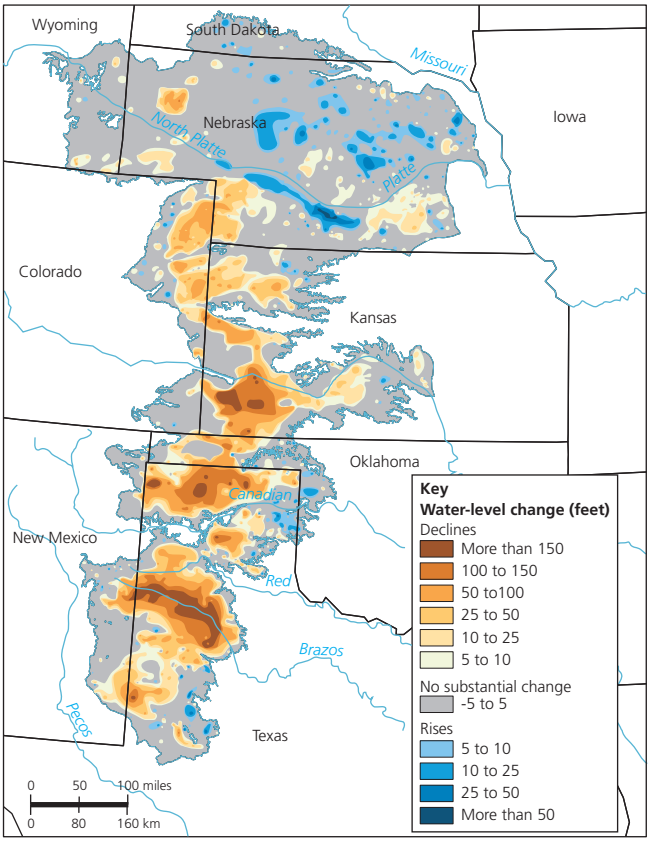


Figure 11.19 Depletion of water reserves in the Ogallala aquifer
Source: NOAA

Water resources – global patterns of availability and demand

There is enough fresh water on the planet for seven billion people but it is distributed unevenly and too much of it is wasted, polluted or unsustainably managed.

Table 11.9 Populations without access to safe water

Region	Population size without access to safe water
Developed countries in North America, Europe, Russia, Japan, Australia and New Zealand	8 million
South East and East Asia (including China) and Oceania	187 million
West, Central and South Asia (including India)	180 million
Africa	376 million
Latin America and the Caribbean	34 million
	Total = 785 million

Source: Our World in Data

Areas of the world most likely to suffer water shortages are:

- those countries experiencing rapid population growth such as poorer countries in sub-Saharan Africa. This growth will result in an increase in demand for drinking water, for both humans and livestock, for irrigation to maintain food supplies and for hygiene to prevent the spread of disease
- those regions suffering from physical shortages of water because of prevailing climatic conditions, in many cases worsened because of climate change. For example, countries in the Middle East and North Africa (MENA) are particularly prone to these problems.

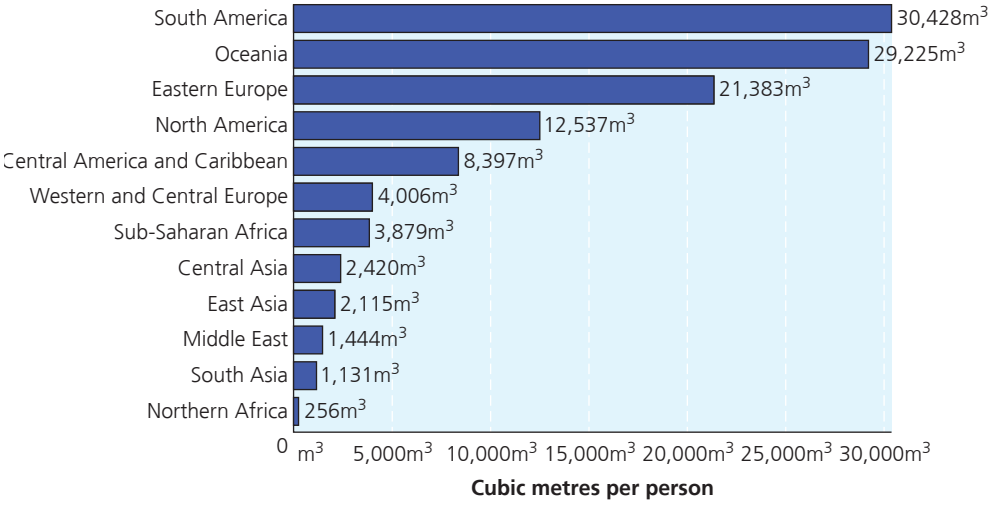


Figure 11.20 Average renewable freshwater resources per person by region, 2015
Source: Our World in Data – FAO AQUASTAT

Shortages will result in:

- droughts which kill crops, livestock and eventually people
- increases in water-borne diseases such as cholera because of using contaminated water
- barriers to industrial development
- over-exploitation of water resources.

The availability of water is affected by both physical and human factors:

- **Physical factors:** Supply is largely determined by factors such as latitude, wind direction and proximity to coasts. These factors will influence rainfall and thus whether areas receive plentiful water supplies. Those regions receiving more than adequate rainfall to meet the needs of supporting the population are areas of **water surplus**.
- **Human factors:** Availability of water for the population is largely dependent on human intervention, such as treatment facilities to provide **improved drinking water** and the infrastructure to supply safe drinking water to a population through pipelines or wells (Figure 11.21).
  - Another key feature of water availability is the general discrepancy between urban and rural areas, especially in the poorest countries. Around 82 per cent of those who lack access to improved water live in rural areas, while just 18 per cent live in urban areas.
  - Availability may be adversely affected by human factors such as overuse, even in areas where nature provides a surplus.

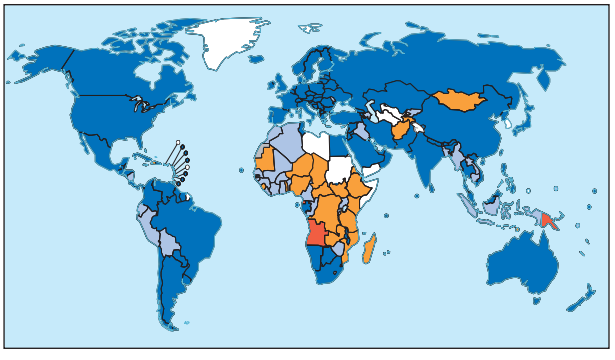


Figure 11.21 Proportion of population using improved drinking water, 2015

Source: WHO

Areas of water surplus

Water surplus countries have certain geographical, water management and water usage characteristics that determine water availability:

- **Climates giving regular, plentiful rainfall** resulting in good run-off and stores such as aquifers, large bodies of surface freshwater sources (lakes, rivers, etc.) and low evaporation rates.
- **Effective water treatment and management** protects against water pollution and ensures premium water quality and supply. (Agreements for water resources shared with other countries are also an effective management strategy.)
- **Low water-usage characteristics**, for example, low population or efficient usage of available water supplies.

Countries with water surpluses are located in temperate or tropical areas, including much of South America, North America, Northern Europe (including Russia), South East Asia and Australasia, but there is not always a correlation with rainfall levels. For example, both Russia and the USA have areas with low rainfall but both have an overall water surplus. They have a surplus because of their ability to manage their water resources, but on a smaller scale both countries have regions with a water deficit.

Areas of water scarcity

Water scarcity is both a natural and a human-made phenomenon and already affects every continent. In 2019, the World Resources Institute (WRI) identified that around 1.8 billion people (25 per cent of the world's population), living in areas of **physical water scarcity**, are facing water shortages. Hydrologists assess scarcity by looking at the population–water equation. They offer a range of parameters for areas suffering from water deficit, as the degree of deficit varies.

Most countries with water scarcity are located in the Middle East and North Africa (MENA region), Central Asia and parts of the Indian sub-continent. Sub-Saharan Africa has the largest number of water-stressed countries of any region.

Arid regions are most associated with physical water scarcity and where this occurs, the water resources cannot meet the demands of the population. There are, however, an increasing number of regions in the world where water scarcity is a situation brought about

Table 11.10 Measures of water scarcity

Degree of scarcity	Characteristics	Example areas
Areas of water stress	Annual water supplies drop **below 1,700 m³/person** Demand for water exceeds supply over a period of time causing shortages	India, parts of China, sub-Saharan African countries
Water scarcity	When annual water supplies drop below **1,000 m³/person**	South Africa, parts of Western Asia (Iraq)
'Absolute' water scarcity	When annual water supplies drop **below 500 m³/person**	North Africa – Algeria, Libya and Egypt Middle East – Saudi Arabia
Physical water scarcity	Prevalent in **arid areas** Also known as areas of water deficit Where nature's provision of water is insufficient to meet population needs	MENA region, South-west USA, parts of Central America, parts of Australia, Central and South Asia and Northern China
Economic water scarcity	Exists when a population does not have the monetary means to utilise an adequate supply of water Characterised by unequal distribution and poor infrastructure	sub-Saharan Africa; parts of South America and South and South-East Asia

by human factors. In the USA, as well as the depletion of the Ogallala aquifer, further problems occur in the Colorado River basin. Here, in south-west US states, a seemingly abundant source of water is also being over abstracted and unsustainably managed, leading to serious physical water scarcity downstream.

There is an obvious correlation between areas of scarcity and rainfall levels but there are anomalies, with some areas suffering physical water scarcity still being able to provide water to their population, for example, Australia. Australia was able to effectively manage water supplies during the worst drought in their history lasting from 1997 to 2009.

Conversely, some areas with an adequate natural supply of water are, for economic reasons, unable to fully utilise their sources because they cannot afford to abstract it, transport it or treat it if it is polluted. This is a situation known as **economic water scarcity** and it affects the poorest countries in the world, particularly much of sub-Saharan Africa, parts of Central America and South Asia (Figure 11.22). Around 20 per cent of the world's population faces water shortages because their countries lack the infrastructure to supply them.

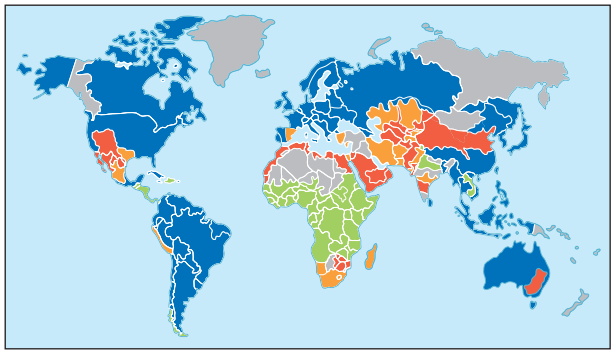


Figure 11.22 Areas of physical and economic water scarcity for 2025
Source: International Water Management Institute

Economic water scarcity is almost entirely due to a lack of investment and good governance. However, Millennium Development Goal (MDG) 7, to halve the proportion of people without access to safe drinking water, was met five years ahead of schedule in 2010. The UN Sustainable Development Goal 6 aims to build on that progress by improving levels of sanitation and by providing access to improved water supplies for the remaining 10 per cent of the world's population by 2030.

Pattern of demand for water

North America and West and Central Asia have the highest per capita water withdrawal globally (Figure 11.23). Both areas have vast plains dedicated to agricultural production and are much drier in summer, needing more irrigation. Water for cotton and cereal production is particularly important in western central Asian countries such as Turkmenistan and Uzbekistan. North America, being a more developed region, also places great demand on water withdrawals for industrial processes and for domestic uses such as in dishwashers and washing machines. Personal hygiene levels are high and water is also abstracted for recreational purposes such as swimming pools and maintenance of golf courses.

Sub-Saharan Africa has the lowest per capita usage, simply because rainfall is unreliable and they lack the infrastructure to supply water on tap. They have to be careful in using the resources they can access and this is used primarily for irrigation.

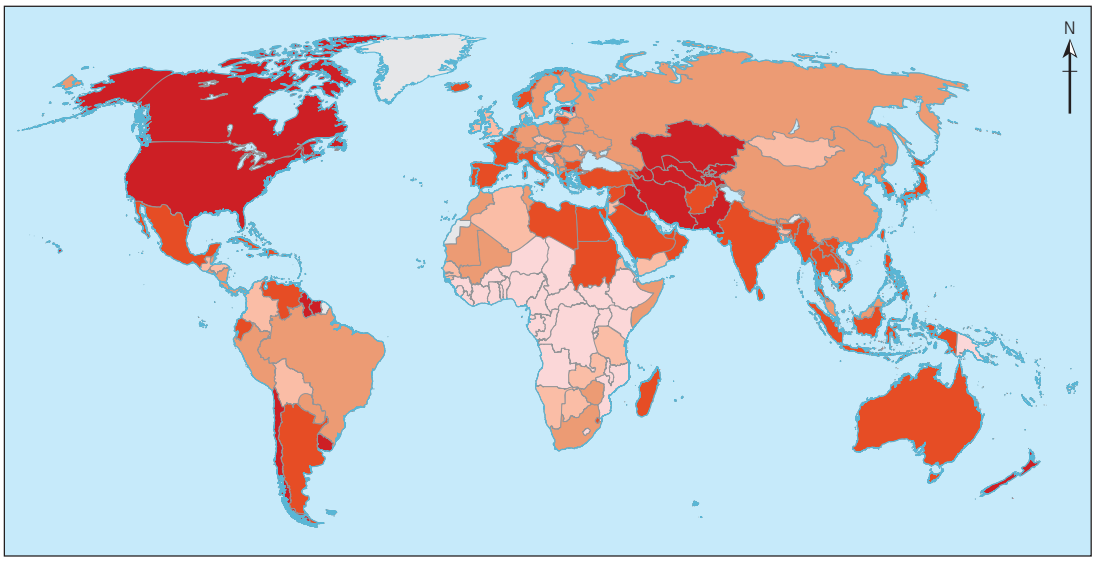


Figure 11.23 Water withdrawal per capita 2015
Source: FAO - Aquastat

Geopolitics of water resource distributions

Links

There are links with Chapter 7, Global systems and global governance – Unequal power relations in global systems, pages 303–5 and Global governance, pages 327–33.

Access to and control over sources of water gives a strategic benefit. Water therefore has an underlying geopolitical importance. Negotiations that focus on resolving conflicts caused by the shared use of water resources are called '**hydropolitics**'. As water becomes scarce, tensions among different users will intensify, at local, national and international scales. It is a less traded commodity than energy but securing access to water is potentially a greater cause of conflict. This is because it is a resource shared by many countries and for which there is no alternative. For example:

- **276 river basins are transboundary** (shared by two or more countries). Without formal agreements, any changes made within a river basin can lead to transboundary tensions
- **200 transboundary aquifers** have also been identified and over-abstraction by one country can lead to tensions and conflict, particularly in areas prone to water scarcity
- **148 countries** have territory within at least one transboundary river basin (39 of these have more than 90 per cent of their territory within transboundary basins).

When water supplies are polluted, diverted, blocked or over-abstracted by one country, it can affect the life chances of those in neighbouring countries. If major projects such as dams or abstraction proceed without regional collaboration, they can heighten regional instability.

Potential for conflict

As the demand for water reaches the limits of supply in certain areas, conflicts may arise between nations that share freshwater reserves. The UN recognises the potential for up to 60 such conflict 'hotspots' globally; more than 50 countries on five continents might be caught up in disputes unless agreements on how to share rivers, reservoirs and aquifers are made (Figure 11.24, page 584).

Problems arise when one country unilaterally makes development plans without consulting those with whom it shares water resources. For example, China controls the Tibetan Plateau, known as the 'Water Tower of Asia' because it is the source of ten major Asian

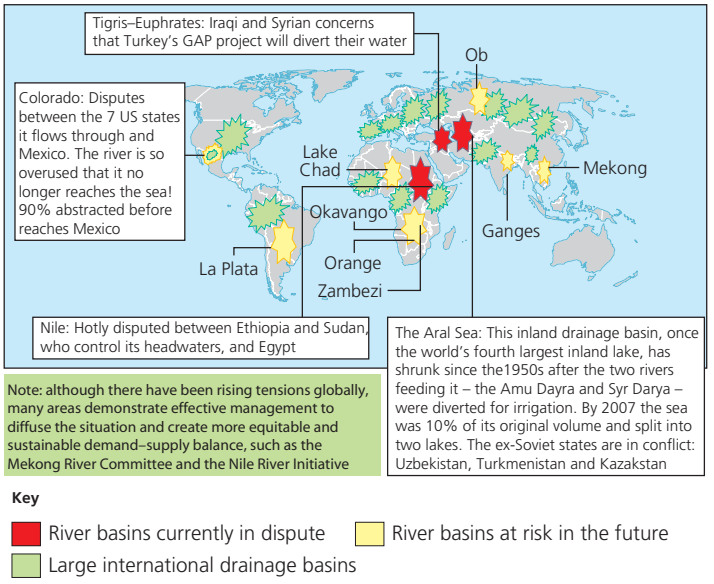


Figure 11.24 Present and potential water conflict hotspots

river systems. Such is China's demand for energy to fuel industrialisation that it has built HEP dams in the headwaters of major rivers, including the Indus, Ganges and Mekong, without agreements with downstream **riparians** (those with access rights to the river banks).

Improving transboundary cooperation

The Berlin Rules on Water Resources (2004) summarise international law customarily applied to sharing freshwater resources. The Rules require nations to take appropriate steps to sustain and manage water resources, minimising environmental harm. Countries are encouraged to see water management as an opportunity for regional economic development to be strengthened, rather than an inevitable drift towards conflict. There are around 450 transboundary agreements that exist regarding international waters. Legal agreements on water sharing have been negotiated and maintained even as conflicts have persisted over other issues. For example:

- Cambodia, Laos, Thailand and Vietnam signed the Mekong Agreement in 1995 to form the Mekong River Commission, agreeing to jointly manage the shared water resources and sustainable development of the lower Mekong Basin. In 2016, Laos started construction of the Don Sahong Dam, a HEP project, just two miles north of the Cambodian border. The dam threatens the fishing industries downstream in Cambodia, Thailand and Vietnam, so these countries demanded a say in the plan. The dam started operating in 2020 and connected its power grid to Cambodia to export a much-needed reliable supply of electricity there.
- The Indus River Commission has been a successful agreement between India and Pakistan to manage the Indus whereby they divided six tributaries equally. The agreement has survived two wars between India and Pakistan but there have been no conflicts about the Indus waters, though Pakistan has disputed technical features of dams being built by India.
- The Nile Basin Initiative (NBI) was formed in 1999 to manage the Nile River Basin in order to achieve sustainable socio-economic development in the region by promoting equitable use of common water resources. The NBI was threatened by disagreement in 2010 between the five upstream nations who wanted a greater share, opposed by Sudan and Egypt. In particular, Egypt objected to the construction of Ethiopia's Great Ethiopian Renaissance Dam in the headwaters of the Blue Nile. The disagreement was eventually resolved in 2015.

These cases demonstrate the need for institutions such as the UN to develop a process of engagement with countries to seek international water resources co-operation.

Review questions

9 Identify the factors that cause an area to suffer from water stress.

10 Explain the difference between physical and economic water scarcity.

11.5 Water security

Water security is defined as the capacity of a population to safeguard sustainable access to adequate quantities of acceptable quality water for:

- sustaining livelihoods
- human well-being and socio-economic development
- ensuring protection against water-borne pollution and water-related disasters
- preserving ecosystems in a climate of peace and political stability.

Source: UN Water

Providing stable freshwater supplies is a priority for every country in the world. Yet stable supplies are increasingly hard to come by in many countries, as water-related risks increase. By 2025, 1.8 billion people will be living in countries or regions with absolute water scarcity, and two-thirds of the world's population could be living under water-stressed conditions due to overuse.

Sources of water

The water used to satisfy these demands can be obtained from the following sources:

- **Surface water supplies** – rivers and reservoirs
- **Underground stores** – groundwater aquifers
- **Seawater** (after desalination).

Components of water demand

The major uses and demand for water can be categorised into the following three main sectors:

- **Agricultural use:** Mainly irrigation of crops; also used in cleaning/washing and watering livestock.
- **Industrial and commercial use:** Used as a coolant in electricity generation; also used in heating, steam turbines, transport and various industrial processes especially textiles, paper and fibre manufacture, metal processing, food processing and the construction industry.
- **Domestic/household use:** Known as **public or municipal water** supply and used for drinking and food preparation, personal hygiene, sanitation and washing/cleaning.

Globally the world uses fresh water in the following proportions:

- **Agriculture = 69 per cent**
- **Industry = 19 per cent**
- **Domestic (municipal) = 12 per cent**

The composition of demand will vary in different parts of the world (Figure 11.25). For example, less developed countries in arid parts of the world use a higher proportion of their water for agriculture than wealthier temperate regions, where plentiful rainfall and lower evapo-transpiration mean that less irrigation is needed. People here are more affluent and industry is more advanced, so a higher proportion is used in the domestic and industrial sectors.

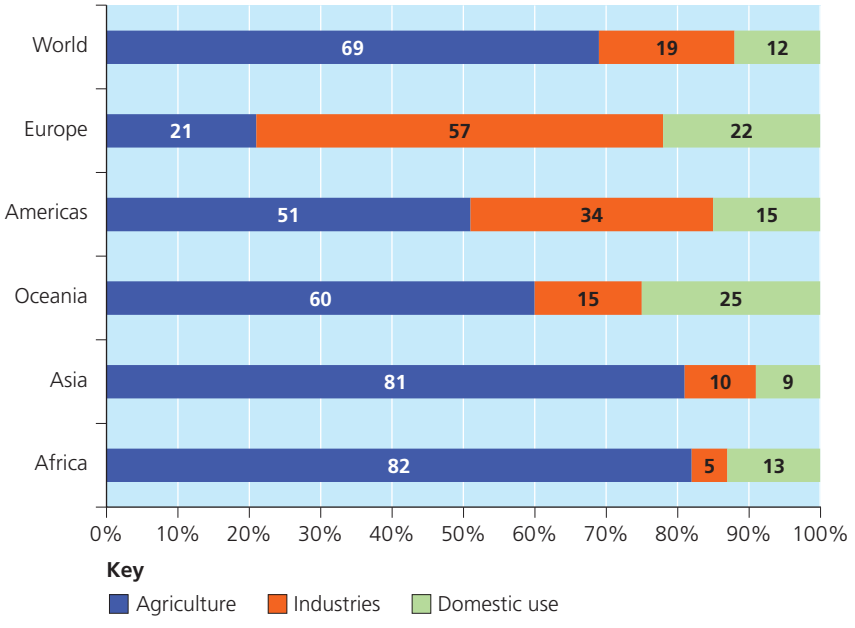


Figure 11.25 Variations in end use of water in different parts of the world

Source: FAO

The concept of water stress

Water stress results from an imbalance between water use and water resources available and leads to a situation where the lack of water is a major constraint on human activity. International organisations have assigned a measure to water stress which is, water availability of less than 1,700 m^3 per person per year (Table 11.10, page 582).

Another water stress indicator is the amount of water withdrawal as a proportion of the total water resource available. Figure 11.26 shows exposure to water stress in different parts of the world based on the ratio of withdrawals to supply. Areas with a higher percentage of withdrawal mean that water users are competing for limited resources.

Water stress causes a deterioration of freshwater supplies in terms of:

- **Quantity** – for example, over-abstraction from rivers, reservoirs or aquifers
- **Quality** – for example, organic pollution or eutrophication of surface water or saltwater incursion into aquifers.

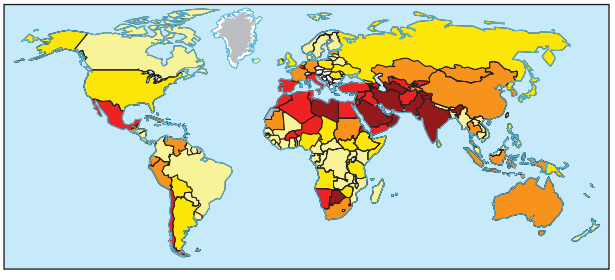


Figure 11.26 Level of water stress by country, 2019
Source: World Resources Institute

Relationship with aspects of physical geography

Links

There are links with Chapter 1, Water and carbon cycle – Drainage basin systems, pages 10–14, Water balance, Run-off variation and Water abstraction, pages 14–24; Water, carbon, climate and life on Earth, pages 41–2, and the case studies on pages 46–51.

Physical factors are important in ensuring reliable sources of surface water and aquifers. The usual requirements include:

- reliable annual rainfall (this may be seasonal but storage can help with temporal discrepancies)
- large bodies of surface fresh water – lakes and rivers provide the most accessible source
- low evaporation rates to reduce surface losses
- suitable geological structures that provide for groundwater aquifers.

Rivers and lakes

In deciding to abstract river water for human use, the following factors about the river's regime need to be considered:

- river discharge – the annual flow must be adequate
- variations in flow over time should be included when determining abstraction rates
- water quality – the level of contamination and turbidity before treatment
- other uses made of the river (for example, uses such as transport or wildlife conservation).

Over-abstraction from rivers and lakes will have a number of negative impacts including:

- reduced volume of water following abstraction concentrates pollutants
- decreased velocity and discharge will increase sedimentation downstream
- reduced downstream flooding will adversely affect soil quality
- reduced flow downstream will impact on ecosystems, habitats and may kill aquatic life.

Reservoirs

Reservoirs are created by flooding valley floors; a dam is built across the river valley to allow water to be stored behind it in a reservoir (Figure 11.28, page 589). Creation of reservoirs for storing and abstraction of water is a controversial decision for two reasons:

- conflict with other uses in the valley – such as settlement (people are displaced), food production, etc.
- disruption to the hydrological cycle and impacts on the whole river system.

In addition, good sites for reservoirs are often those that have the highest value for scenery and wildlife. This makes decisions on where they should be located difficult.

Table 11.11 Physical factors affecting the most suitable location of reservoirs

Physical characteristic	Preferred site and reason
Topography	A long narrow valley basin with steep sides that will provide a large volume relative to the surface area (minimises evaporation loss)
Geology	Impermeable rock so that water is not lost to seepage. Geology must be stable as earthquakes or faulting may damage or destroy the dam. Avoiding areas where the rocks or minerals may have an economic value that exceeds the value of the reservoir
Catchment area	Large area to increase water volume for storage
Climate: Water supply	Adequate and sustainable supply of water from upstream requires regular and frequent precipitation. Greater volume gives more flexibility and reliability
Climate: Flow fluctuation	Rainfall needs to be reliable. Cooler and less windy to reduce losses due to evaporation. Avoiding areas prone to freezing for long periods

Table 11.12 Evaluation of the potential benefits and environmental impacts of reservoirs

Potential benefits gained from reservoirs	Environmental effects of reservoirs
Flood control: Prevents flooding downstream **Stored water supply:** For agricultural, industrial or public use **Multiple purposes including:** HEP generation, Recreational use, Fishing	**Sedimentation/siltation:** Silt/sediment is deposited in the reservoir: reduces capacity and increases the erosive power of the river downstream **Creates microclimate:** Smaller temperature fluctuations: higher wind speed, higher humidity downwind **River regime downstream of dams:** Changes in flow, temperature, turbidity, erosion and fluvial landforms such as deltas **Habitat change:** Flooding removes habitat, creates a barrier to migration and dispersal of seeds

Aquifers

An aquifer is a layer of rock that holds groundwater. They occupy specific geological formations and need porous or permeable rocks such as chalk, limestone, sandstone or gravel to hold a 'reservoir' of underground water. Porosity is a measure of the proportion of rock which is 'space' and can hold fluids, while permeability is the rate at which fluids flow through a rock because of the connected pores or fissures. Permeable rocks enable aquifers to 'recharge' from natural sources.

Aquifers usually form in sedimentary basins and the different layers of permeable and impermeable rock will form confined and unconfined aquifers (Figure 11.27). The layer below the aquifer must be impermeable to prevent the escape of the water. Confined aquifers will have impermeable layers of rock both above and below them. They are called artesian aquifers and can be recharged where permeable layers meet the surface.

Water occurring below the ground is brought to the land surface by wells or springs. If the head of the well is below the natural water table level then it will flow naturally under atmospheric pressure, otherwise the water is pumped to the surface.

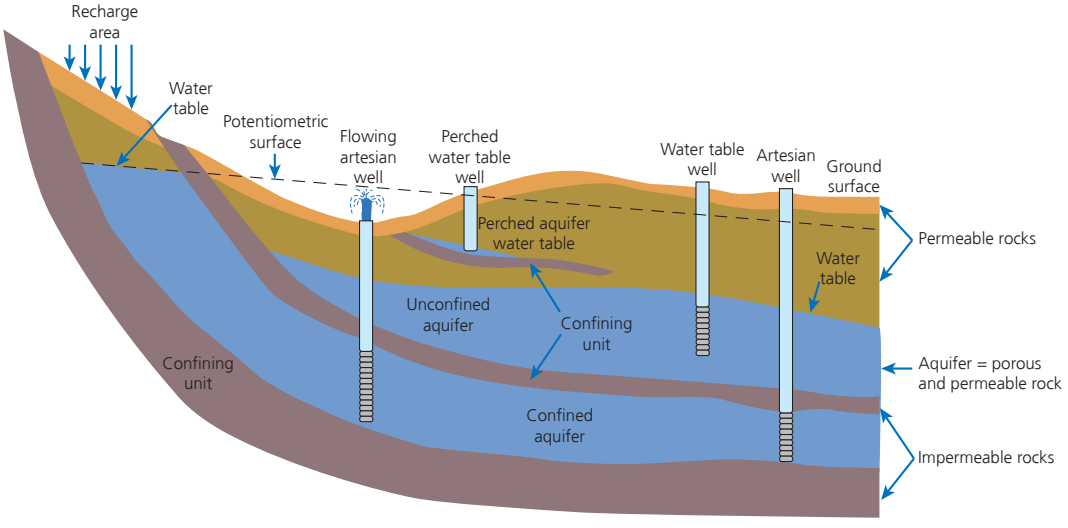


Figure 11.27 Geological structure of confined and unconfined aquifers

Source: National Ground Water Association

Groundwater is vulnerable to overuse and the improper disposal of chemicals at the land surface. The trend of falling groundwater levels may indicate climate change or prolonged over-abstraction.

Over-abstraction occurs when the rate of extraction of water from the aquifer is greater than the recharge rate. It can lead to a number of serious environmental impacts including:

- subsidence
- loss of wetlands/vegetation changes
- drying of soils leading to 'osmotic dehydration' of crop roots
- lower water table
- rivers drying up, leading to loss of habitat/loss of recreation
- saltwater incursion into groundwater.

Aquifers are naturally recharged by rainfall but it can take many years for water to percolate down and to top up aquifers, so they are also recharged artificially.

Strategies to increase water supply

Water supply and demand are prone to spatial and temporal mismatches. For example, in the UK the greatest demand for water is in south-east England, because of the higher population density and greater agricultural needs in arable farming. The south-east receives less rainfall and has less surface water than other areas, so is more reliant on groundwater sources. It has the highest potential water stress of any part of the UK.

Catchment management

Freshwater resources are limited in certain areas and at certain times. Measures can be taken to protect quantity and quality of water supplies. These measures will involve:

- increasing the total amount of water available for use, for example, by increasing abstraction
- using water more efficiently, for example, by reducing waste use
- making better use of existing supplies by:
  - reducing pollution of existing sources
  - distributing it more effectively.

In the UK, the Catchment Abstraction Management Strategy (CAMS) is an integrated approach to manage the supply of water. It is used to assess the amount of water available for abstraction licensing, taking into account what the environment needs. Environmental flow indicators (EFIs) are used to assess whether river flows are sufficient to support a healthy ecology. Surface water management can best be achieved by storage to even out temporal differences in supply or by diversion and inter-basin transfers to even out spatial inequalities.

Storage by reservoir

Reservoirs are a way of increasing and managing water supply on a regional or national scale.

- Dams are built to trap and store surplus winter rainfall or regulate river flow, to maintain habitats and prevent flooding downstream.
- Direct supply reservoirs, such as Lake Vyrnwy (Figure 11.28) store water for release via pipeline to the municipal supply nearby or further afield.
- Reservoirs may fill from rainfall, from natural river flow or via pumping from nearby rivers.
- Large scale multi-purpose reservoirs such as those created by the Three Gorges Dam (Yangtze River) and Aswan Dam (River Nile) can also provide additional benefits including flood control, hydro-electric power (HEP) as well as fishing and tourism. They are not without their environmental problems, as indicated by Table 11.12 on page 587 (see also Environmental Impacts of the Aswan High Dam and Lake Nasser as a case study on page 591).

Diversion and inter-basin transfer

River diversion is used to transfer water from one river catchment to another when there is a surplus of water in one area and too little in another. Water is drawn from rivers, reservoirs and aquifers, and pumped across the region in pipelines and aqueducts, enabling localised droughts to be tackled quickly. Transfer can also take place by river diversion or canal networks.

In the UK there is a huge variation in average rainfall between the different regions; Snowdonia receives over 400 cm per year, whereas parts of East Anglia receive only 55 cm per year. Demand also varies due to population distribution. For example, heavy rainfall in sparsely populated mid-Wales is stored in reservoirs and then transported via rivers and pipelines to the West Midlands conurbation.

Figure 11.28 Aerial view of the dam forming Lake Vyrnwy reservoir in mid-Wales and its location on the map. The reservoir was constructed in the late nineteenth century to supply water to the growing city of Liverpool, 70 miles away.

Water transfer schemes operate on a regional scale in Britain and larger national scale schemes have also been proposed to move surplus water from the north and west to the drier south-east. Proposals include taking water from the Kielder reservoir in Northumberland via Newark to London and a River Severn to Thames inter-basin transfer has also been considered. The notion of taking water from areas of water surplus to an area of water stress is attractive. However, the prohibitive construction costs, ecological concerns and public resistance to more reservoirs being built to store the transferred water have meant that these projects are still on hold.

In water-stressed developing countries which are still heavily reliant on water for agriculture, irrigation canals transfer water from major river basins to farming areas. The Sindh province of Pakistan has a network of irrigation canals abstracting water from the Indus basin. High evaporation rates and poor maintenance of the linings of the canals cause large losses of water and leakages. This leads to the dual problem of waterlogging and salinity, which destroys crops and renders the soil infertile.

Table 11.13 Evaluation of inter-basin water transfer

Advantages of inter-basin transfer	Disadvantages of inter-basin transfer
Localised drought can be remedied quickly Habitats can be protected Addresses spatial mismatch	Capital and running costs are very high – water is expensive to pump There are regional differences in water quality parameters (for example, chemical composition) Requires development of storage reservoirs in receiving basin

China has built the world's largest transfer of water between river basins. The South to North Water Diversion Project has three routes taking water from the Yangtze basin northwards by canal and pipeline to treatment plants near Beijing, an area experiencing severe water shortages:

- The Eastern route uses the ancient Grand Canal.
- The Central route, opened in 2014, takes water 1,432 kilometres from the Danjiangkou reservoir on a tributary of the Yangtze river and now supplies two-thirds of Beijing's tap water.
- The proposed Western route transferring water from the Tibetan plateau has been postponed indefinitely as the area is prone to earthquakes and landslides.

The two working stretches can transfer up to 25 billion tonnes of water annually from south to north.

Figure 11.29 The South to North Water Diversion Project, China

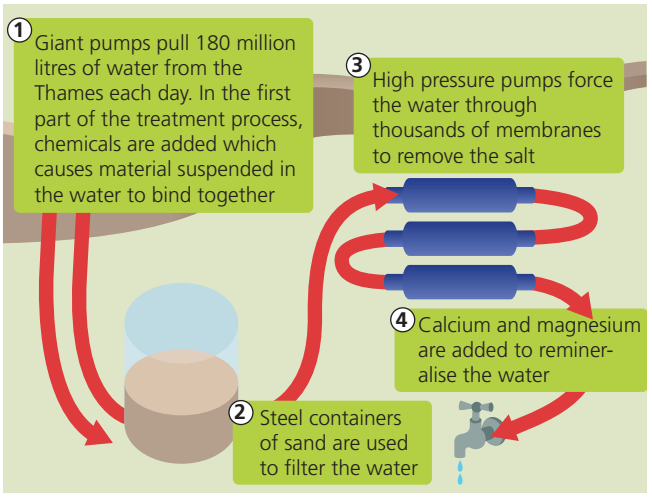


Figure 11.30 The desalination process at Thames Water's Beckton operation using reverse osmosis.

Source: Daily Mail

Desalination

The removal of salt from seawater would provide a huge supply of water but desalination is an expensive process which is not considered sustainable because it is very energy intensive and a significant source of greenhouse gas emissions. There are two main methods used in desalination:

- **Reverse osmosis**: Involves filtration of seawater at high pressure (600–1,000 PSI) through a partially permeable membrane, using small polyamide tubes, to convert it to potable water. This process collects fresh water and rejects very saline water, which is pumped back into the sea.
- **Distillation:** Water is boiled by heating under reduced pressure to save energy (water boils at a lower temperature in a vacuum). The steam produced is condensed and collected separately and the salt is left behind in the boiler.

Desalination plants have generally been developed in wealthier countries with water stress or scarcity issues, especially those in the Middle East such as Saudi Arabia, Oman and the UAE, but also in the Mediterranean, the USA and Australia. In the UK, Thames Water's desalination plant at Beckton in East London provides water for 400,000 households by reverse osmosis membrane filtration (Figure 11.30).

Strategies to manage water consumption

Where people are not facing scarcity, water is often taken for granted and wasted. Whatever the use of fresh water, huge savings and improvement of water management are possible.

Domestic use

Probably the most effective way of avoiding wastage is to ensure households pay for what they use. The introduction of water meters in the UK ensures that households with a meter fitted pay for the amount of water used, rather than a standard rate. This gives householders an incentive not to waste water. Because the south of England has been declared an 'Area of Serious Water Stress', water companies such as Southern Water, Thames Water and South East Water are installing meters to encourage customers to use less water.

A number of simple methods are available to reduce water consumption:

- toilets – around 30 per cent of the water in homes is used to flush toilets. This can be reduced by:
  - installing a displacement device (for example, a 'hippo bag') in the cistern to reduce the volume of water used per flush
  - installing low flush or dual flush toilets
- personal hygiene – having showers instead of baths saves water use
- machinery – using newer technologies such as water-efficient washing machines and dishwashers
- filling kettles with just enough water for your needs; this will also reduce energy bills
- collecting rainwater to use on gardens.

Industrial and commercial use

Similar savings can be made in commercial and industrial settings. Newer buildings and complexes are designed to reduce water use by collecting rainwater, which is filtered and then sent into toilet and washbasin areas. This not only reduces abstraction from traditional sources but also saves the energy required in treating water from those sources.

Agricultural use

It is possible to reduce water consumption when irrigating by using water efficient methods. Micro-irrigation is the most efficient method. This employs low-flow technology to deliver water directly to plant roots at rates that prevent deep percolation, evaporation and run-off losses. Typical devices include:

- drip feed irrigation: small irrigation heads that deliver a drip or trickle directly at the plant base
- soaker hoses: hoses with small openings that allow water to seep out; they are placed at the plant base and are suitable for using with rows of crops (Figure 11.31)

- micro-sprayers: overhead sprinklers which deliver a fine mist that can be controlled.

Additional measures that growers can use to reduce water use include:

- using mulches wherever possible to retain water content
- collecting storm water and irrigation run-off in ditches or drains
- watering early in the morning to reduce evaporation losses
- minimal or no ploughing (leaving at least 30 per cent of the previous season's crop residue in the field)
- contour ploughing reduces run-off
- using organic fertilisers because organic matter retains soil moisture

Figure 11.31 Micro-irrigation devices such as soaker hoses reduce water losses

- using cover crops in winter and fallow times to improve the soil and reduce water loss and soil erosion.

Environmental impacts of a major water supply scheme – The Aswan High Dam and Lake Nasser

Background to the Aswan scheme

The Aswan Dam is situated across the Nile River in Aswan, southern Egypt. The High Dam, which is larger and further upstream than the earlier Low Dam, was constructed between 1960 and 1970. It was funded with assistance from the Soviet Union (aiming to gain geopolitical support in the region). It is nearly 4 km wide across the valley. The reservoir of Nile waters stored behind the dam is called Lake Nasser and stretches back 550 km southwards into northern Sudan (Figure 11.32). It is 35 km at its widest point and holds around 132 km^3 of water. The scheme was designed to be multi-purpose, regulating the annual river flooding and providing storage of water for agriculture, and later, generating hydroelectric power (HEP). The water in Lake Nasser represents the national freshwater bank of Egypt and has had a significant impact on the economy.

Figure 11.32 Panoramic view looking south towards Lake Nasser from the Aswan High Dam

Before the dams were built, the Nile River flooded each year during late summer. These floods brought high water and nutrients and minerals that enriched the fertile soil along the floodplain and delta; this made the Nile valley ideal for farming. However the floods varied: in high-water years the whole crop might be wiped out, while in low-water years widespread drought and famine occurred. There was a desire to control the floods in order to protect farmland and to provide irrigation for the economically important cotton crop. With the capability of reservoir storage, the floods were reduced and the water could be stored for later release.

Socio-economic impacts

Lake Nasser flooded much of lower Nubia and around 100,000 Nubian people had to be resettled in purpose-built villages in Sudan and Egypt. However, the lake created a lot more land for people to farm on because the stored water was used to irrigate land around the lake. The lake also created a large fishing industry producing over 25,000 tonnes of fish per year.

Unfortunately, the Aswan Dam had a negative impact on the lives of farmers downstream. Since the dam was built, the annual flood has been controlled and fertile alluvial sediments are held back in the reservoir. Farmers downstream now have to use fertilisers to grow their crops, which makes it more expensive and creates indirect environmental impacts.

The HEP capacity of the Aswan High Dam is 10 billion KWh per year and in its early operation was providing half the country's electricity. This proportion has fallen in recent years as energy demands have increased and it now only provides around 10 per cent.

Environmental impacts

Critics of the project in the 1970s pointed out potential failings of the dam (Figure 11.33).

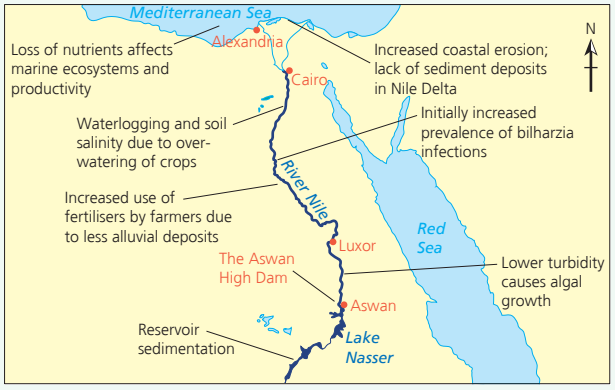


Figure 11.33 Outline map of the Lower Nile valley in Egypt labelled with some of the key environmental impacts

Coastal erosion

The erosion of the Mediterranean coastline has accelerated due to lack of sediment, once brought down to the Nile Delta by the river. Coastal erosion is severe in some areas, especially at the Rosetta and Damietta promontories. The increased erosion is the result of the trapping of sediments behind the dam. It has made the construction of expensive coastal protection works in the Nile Delta necessary. Efforts to prevent overall coastline retreat have been largely unsuccessful.

Reservoir sedimentation

Sediment deposited in the reservoir is lowering the water storage capacity of Lake Nasser. The reservoir storage capacity is 162 km^3, including 31 km^3 dead storage at the bottom of the lake. Sediment accumulates much faster in the upper reaches of the Lake where sedimentation has already affected the live storage zone. It will take about another 900 years until the storage zone is completely filled with sediment but operation of the dam will become impossible some time before that. Efforts to dredge the reservoir mitigate this problem to some extent but it is easier to dredge parts of the river before it enters the reservoir.

Marine ecosystems

Marine productivity in the Mediterranean has decreased because nutrients that used to flow down the Nile to the Mediterranean are trapped behind the dam. For example, the sardine catch off the Egyptian coast initially declined massively, though this now seems to have recovered but only to half the levels prior to the building of the dam.

Less freshwater entering the sea because of the dam, indirectly increases the salinity of the Mediterranean. This affects the Mediterranean's outflow current into the Atlantic Ocean, which can be traced thousands of kilometres into the Atlantic.

Health and disease impacts

Schistosomiasis, also known as **bilharzia**, is an infection caused by a parasitic worm that lives in fresh water in subtropical and tropical regions. The standing water in irrigation canals is a breeding ground for snails carrying the bilharzia parasite. The incidence of bilharzia increased due to the water supply to the channels becoming more constant thus inhibiting natural fluctuations in water height. However, other important factors contributed to the prevalence of schistosomiasis including poor sanitation and limited awareness of how the disease was transmitted. The provision of clean water, sanitation, health education and rural clinics has reduced the overall prevalence of the disease.

Waterlogging and increased soil salinity

Before the construction of the High Dam, groundwater levels in the Nile Valley fluctuated by as much as 8–9 m/year as a result of changing water levels. During summer when evaporation was highest, the groundwater was too deep to allow salts dissolved in the water to be pulled to the surface through capillary action. With the continuous irrigation enabled by the scheme, fields in the Nile Valley tend to get waterlogged as groundwater levels remain high, with little fluctuation. Soil salinity has also increased because the distance between the surface and the groundwater table is small enough to allow water to be drawn up by evaporation. As a result, relatively small concentrations of salt in the groundwater gradually accumulate on the soil surface over the years. Since the water is not flushed out annually by flooding, the soils are saltier and less fertile.

Most of the farmland did not have proper subsurface drainage to lower the groundwater table so salinisation gradually affected crop yields. It took twenty years after the completion of the Aswan scheme to address this problem by providing two million hectares of farmland with subsurface drainage (costing more than the construction of the High Dam). Overall, agricultural production in Egypt has increased as a result of the dam. The more reliable and constant water supply allows multi-cropping during the year and Egyptian farmers have also adopted more successful seeds and harvesting methods.

Algal growth and water quality

The lower turbidity of the water downstream from the dam means that sunlight penetrates deeper in the Nile water. As a result of this and the increased presence of nutrients from fertilisers in the water, more algae now grow in the Nile. This has increased the costs of water treatment to supply potable water. The decrease in water quality was an unforeseen consequence of the High Dam scheme.

Sustainability issues

Sustainable water management aims to achieve a greater long term balance between supply and demand while minimising environmental impact. More sustainable management will help to maintain water supplies for future generations, particularly in areas of water stress.

Virtual water trade

The water used in the production process of an agricultural or industrial product is called the 'virtual water'. The virtual water trade is the idea that when commodities are exchanged, so is virtual water. It requires water to be considered as an 'economic good' with a value. Professor Tony Allan introduced the concept as a way of understanding how water-scarce countries could provide food, clothing and other water-intensive goods to their inhabitants without depleting their own water resources. A water-scarce country might wish to import products that require a lot of water in their production, and export products or services that require less water (water-extensive). This means that water-scarce countries, such as those in the Middle East, become net importers of 'virtual water', thus relieving pressure on their own water resources. For example, on average, it takes 1,340m^3 of water to produce one tonne of wheat. When a country imports a tonne of wheat it is saving this amount from its own water supplies. Equally, water-scarce countries like Israel discourage the export of water-intensive products such as oranges to prevent large quantities of water being exported.

Conservation measures

Land use, ecological management and afforestation

One of the strategies to conserve water is to control land use in a river basin. This will affect run-off, infiltration rates and groundwater storage. Impermeable surfaces in and around urban areas increase run-off, taking water to one of its shorter-term stores in the river channel. An increase in permeable land uses increases infiltration and percolation rates, taking water into longer-term groundwater storage, thus increasing supply for future use.

Links

There are links here with Chapter 1, Water and carbon cycles – The water cycle, pages 4–24.

Vegetation cover such as forests and tree plantations also regulate the water cycle by reducing run-off and improving the replenishment of the water table. Providing that there is not excessive evapotranspiration from leaves, they will help to maintain groundwater stores at a more sustainable level. Afforestation can be a good tool for water conservation and forested catchments can supply a high proportion of water for domestic and agricultural use. On a micro-climatic to regional scale, tree planting has often been proposed as a way to increase rainfall. However, trees also consume water. The more the aerial system of trees is developed, the more water they transpire. The desirability of tree planting in arid lands is debated because trees may consume more water than they provide to the water cycle.

Recycling and grey water

Other conservation methods include:

- **Recycling of water:** This is done through extensive sewage treatment. In London, 90 per cent of used water is recycled by being treated immediately after being returned to the drains. In the Middle East and North Africa (or MENA countries), 82% of wastewater is not recycled; this presents an opportunity to relieve water stress.
- **Greywater use:** Greywater is water that has been used for washing or cleaning but has not been in contact with faecal matter (black water). It can be re-used for purposes where non-potable water is adequate (for example, flushing toilets, gardening or irrigation). Greywater recycling is ideally suited to hotels, leisure centres and large office and residential blocks as it recycles water used in baths and showers to flush toilets.
- **Leakage control:** Better maintenance of supply pipes, taps and appliances will prevent losses.

Groundwater management – aquifer recharge

Aquifers take a long time to recharge naturally but they can be recharged artificially by:

- artificial infiltration – pumping water underground
- diverting rivers to permeable surfaces
- diverting storm water into recharge wells. This is an aspect of **rainwater harvesting** to replenish subsurface water supplies. It is practised extensively in India and in Australia (where they are known as 'leaky wells').

This is a good way of storing water that is surplus to requirements so that it is available when supplies are less plentiful. Maintaining aquifers may help maintain rivers and ensure surface stability.

Water conflicts

Water conflict is a term describing a conflict between countries, states or groups with opposing interests over access to water resources.

Local scale

Where demand for water exceeds supply and no effective management operates, conflicts between the various users of that water source may ensue. Conflict at a local level is likely to be caused by different users wanting a scarce resource; it can lead to sabotage, or escalate into violent protest.

For example:

- In Chile, there are competing demands for limited water supplies. Under Chile's Water Code, water is treated as a commodity, allowing companies to request water rights allocated by the government. In the north, there are conflicts between agriculture and copper mining for access to the scarce resource. As mining companies can afford to pay more, water has been diverted away from farms, leading to conflict and protests from locals.
- More recently, the central town and province of Petorca has become the epicentre of conflicts over water. The area is a key avocado growing region. Rising global demand for avocados (known locally as 'green gold') has led to worsening tensions over water rights between large commercial producers who have moved into the area and local citizens and small-scale farmers. Campaigners have accused the big producers of water theft, claiming that they are using larger amounts of water than their allocations allow. They point to evidence showing illegal underground pipes and systems delivering water from rivers to orchards on the hillsides. Meanwhile, rivers are drying up, farmers are unable to irrigate their crops and citizens of Petorca have to rely on water delivery by truck.

National scale

Disputes at this scale are often the result of water scarcity and an uneven distribution of available resources. Although Israel and the Palestinian Authority are two separate states, they can be examined at a national scale for water management as much of the Palestinian territory is controlled by Israel. The climate of the region is very arid and with the growing populations that share resources, there is increased tension concerning the availability and use of water.

Figure 11.34 Pressure on limited groundwater resources in Israel and Palestine

Because of its importance for irrigation and settlement, many argue that access to water is used as a weapon in the ongoing political and religious conflict in the region. The Jordan River runs dry most years as Israel commands sole use of the Sea of Galilee near the river's source. Apart from the river, the mountain aquifer, mostly in Palestinian territory on the West Bank, is the main source of water. The West Bank is controlled by Israel, and they have built settlements on Palestinian territory there. The mountain aquifer in the occupied West Bank is largely used by Israel (80 per cent), leaving only 20 per cent for the Palestinians. Israel claims they have the right to use the aquifer, because some groundwater flows into Israeli territory (Figure 11.34, page 594). Meanwhile, Palestinian farmers have insufficient water. Israel relies on desalination to top up their water supply but in Palestinian Gaza this is not affordable. Israel sells some water to Gaza but the national water carrier from the north stops short of the border.

International scale

International disputes result from sharing transboundary river basins. They usually arise because an upstream nation's use of the water affects the quantity or quality of the water for nations using it downstream. In particular, dam construction will have hydrological consequences and industrial development may increase pollution levels and cause negative ecological impacts. Countries sharing the same section of an international boundary river will have concerns about each other's use of the water. International agreements on the management of transboundary basins help to prevent conflict but disputes still occur (see pages 583–4, The geopolitics of water).

Links

There are links with Chapter 7, Global systems and global governance – Unequal power relations in global systems, page 303, and Global governance, pages 327–33.

There are 13 basins worldwide that are shared between 5 and 8 riparian nations. Five basins, the Congo, Niger, Nile, Rhine and Zambezi, are shared between 9 and 11 countries.

Examples of conflict regions:

- The Aral Sea (Kazakhstan and Uzbekistan) – due to the over-abstraction (for cotton irrigation) from the two rivers (Syr Darya and Amu Darya) that flow into the Aral Sea, half of its surface has disappeared, representing two-thirds of its volume – 36,000 km^2 of land is now covered by salt.
- Turkey, from its dominant position upstream, has been diverting the Tigris and Euphrates rivers, thus increasing water stress in the already volatile states of Iraq and Syria.

These 'basins at risk' are just some examples of where tensions at an international scale could escalate into so-called 'water wars'.

For more detailed information on water conflicts visit the following website: www.worldwater.org/water-conflict/.

Review questions

11 How do the main uses of water vary proportionately in different parts of the world?

12 Outline the geological conditions necessary for the formation of an underground water supply.

13 Explain how managing parts of the water cycle ecologically will conserve water.

14 With reference to examples, suggest why shared access to water resources might cause conflict at a local scale.

11.6 Energy security

In 2015 Energy Ministers of the G7 group signed a joint statement in Hamburg on an initiative for sustainable energy security. The fundamental principle they subscribed to was that sustainable energy security is a common responsibility. The annual meeting of this group to review progress of the initiative illustrates the geopolitical importance of global energy security.

For the remainder of the twenty-first century, the global energy system faces the following strategic challenges:

1 The need for **energy transition** from fossil fuels to cleaner renewable energy resources, driven by:
   - the need to reduce energy-related CO_2 emissions to limit climate change
   - the threats of other forms of environmental damage caused by extraction, transport and burning of fossil fuels
   - the continued depletion of non-renewable energy supplies.

2 The growing risk of disruptions to energy supply caused by political tensions, conflict and wars.

3 The persistence of energy poverty in developing countries.

Key terms

Energy mix – The composition of different sources of energy used in a country or globally. Primary energy mix refers to those used to produce all secondary energy. Some energy mix data refers only to sources used to produce electricity.

Energy transition – a pathway toward transformation of the global energy sector from fossil based to zero carbon by the second half of this century. At its heart is the need to reduce energy-related CO_2 emissions to limit climate change.

Primary energy (resources) – Energy sources obtained in their raw material or natural form, such as oil, natural gas, wind or running water; they are not usable until they are converted into heat or mechanical action to produce secondary energy.

Secondary energy (resources) – Energy that is transformed or converted from primary energy sources into manufactured sources of power which are usable, such as heat, electricity or petroleum.

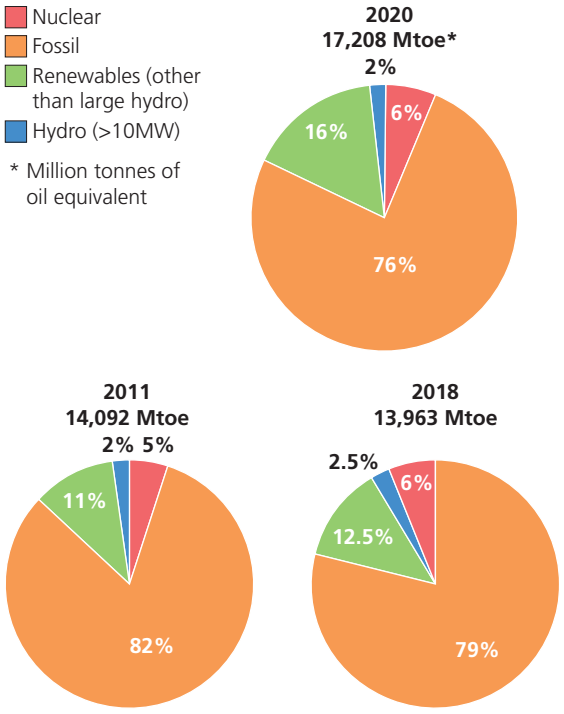


Figure 11.35 Global energy mix and changing composition of total primary energy supply (TPES)

Source: World Energy Council

Primary and secondary sources of energy

Primary energy resources are potential sources of energy found in their natural form (Table 11.14), such as wind or running water or raw materials such as coal, oil, wood and uranium. Primary energy sources need to be converted or processed into **secondary energy** sources such as petrol or diesel used to run vehicles, or into electricity which is then used to power domestic, industrial and commercial premises. In more developed countries, energy sources such as oil, natural gas, coal, hydroelectric power and uranium provide the main sources of power. In contrast, many less developed countries, rely on biomass sources such as fuelwood and animal waste.

Fossil fuels

Oil: Its share of total primary energy supply has fallen to 30 per cent, but it is still the most widely used energy source.

Natural gas: Accounts for 22 per cent of the global energy budget. It is the fastest growing fossil fuel source as it is cleaner and has lower environmental impact.

Coal: Accounts for around 24 per cent of the world's energy supply; in terms of electricity generation it is around 37 per cent. Used widely in China and India to power industrialisation, but concerns about greenhouse gas emissions and acid rain mean a gradual decline in use.

Nuclear energy

Nuclear power generates around six per cent of the world's energy. Nuclear production requires advanced technology and huge investment, so is mostly found in developed, high-income countries. Developing countries, such as China and India, are starting to invest more of their wealth in nuclear power plants. Political factors have had negative impacts on nuclear energy development. Nuclear accidents at Chernobyl, Ukraine (1986), and more recently at Fukushima in Japan (2011) forced the evacuation of the local populations. Large amounts of radioactive material escaped into the surrounding areas causing loss of lives and wide-scale environmental damage. These incidents, together with the inherent problems of disposing of

Table 11.14 Main primary sources of energy and how they are converted into secondary energy

Primary energy resources			Converted by/energy systems	Secondary energy form
Non-renewable sources	**Fossil fuels**	**Crude oil**	Oil refinery	Petrol, diesel, fuel oils, etc.
			Thermal power station	Electricity (limited use)
		Natural gas	Thermal power station (combined cycle)	Electricity
			Combustion after transfer by National Grid	Heat for cooking, central heating, etc.
		Coal	Thermal power station	Electricity
			Direct burning	Heat
	Mineral fuels	**Uranium**	Nuclear fission (nuclear thermal power station)	Electricity
Renewable sources		**Solar energy**	Photovoltaic cells (unit, array, farm)	Electricity
			Solar furnaces/tower and parabolic reflectors	Electricity
			Harnessing heat/light passive solar architecture	Heat
		Wind energy	Wind turbines (farm)	Electricity
		Biomass sources	Burning fuelwood directly	Heat (for cooking and heating)
			Fermentation/distillation and refining into liquid	Bioethanol or other alcohol-based fuels
			Anaerobic digestion to produce gas	Methane gas for burning – heat for cooking/heating or electricity generation
		Hydroelectric power (HEP)	Potential energy released and converted to kinetic energy to turn turbines	Electricity
		Tidal energy	Tidal turbines or tidal barrage (kinetic energy)	Electricity
		Wave energy	To drive hydraulic turbines or compressed air chambers	Electricity
		Geothermal energy	Thermal power stations	Electricity
			Directly harnessing underground heat	Heat

nuclear waste, have dampened the appetite of some advanced economies to seek security in nuclear energy and have, until recently, resulted in a levelling off in new reactor construction.

Figure 11.36 EDF's Sizewell B nuclear power station in Suffolk

Advantages of nuclear energy:

- highly energy efficient fuel with newer reactors even more efficient
- relatively large reserves of uranium
- moderate and predictable cost of electricity over the life of a power station
- excellent replacement for fossil fuels in generating thermally sourced electricity
- atmospheric pollution is much less than with fossil fuels.

Disadvantages of nuclear energy:

- high investment and compliance costs in constructing nuclear power stations
- public concerns about operation and radioactive waste disposal
- potential dangers and impacts of nuclear accidents
- difficulty in finding suitable sites
- plutonium is a by-product which can be used in nuclear weapons – so can cause geo-political tensions or threats of terrorism.

Table 11.15 Advantages and disadvantages of each of the fossil fuels

	Advantages	Disadvantages
Oil	Important in road transport and petrochemical industries Leading tradable commodity Flexible, easy to transport fuel	High price volatility Geopolitical tensions in areas with the largest reserves Market is dominated by leading oil-producing countries and large TNCs
Natural gas	Cleanest of fossil fuels Flexible and efficient fuel for power and heat generation Increasing proved reserves from unconventional sources (reassessments of speculative reserves and shale gas)	Reserves increasingly offshore or in more remote areas Large investment requirement for gas transport and distribution Increasingly long supply routes with high infrastructure costs
Coal	Wider geographical distribution and more plentiful reserves New technologies to improve environmental performance Stable prices	High emissions of carbon, particulates and other pollutants Environmental mitigation such as carbon capture (use) and storage (CCS/CCUS) have a negative impact on energy efficiency

Source: World Energy Council, World Energy Report 2018

Renewable energy sources

The contribution of all renewables to the world's TPES has increased to around 16 per cent and this growth will continue. Renewable energy is the third largest contributor to global **electricity production**. For some countries the proportion of electricity generation from renewable energy sources can be very high, for example, 100 per cent of Iceland's electricity is generated from geothermal and HEP sources. In Paraguay and Norway, respectively 100 per cent and 98 per cent of their electricity is produced by HEP.

Solar energy

The Sun contributes significantly to heating buildings through windows and walls. Features to maximise this energy are incorporated into new buildings using **'passive' solar architecture**, for example, larger south-facing windows. The potential for solar energy in less developed tropical countries is excellent because of the reliability and intensity of the sunlight. It is already used in 'appropriate technologies' such as solar cookers, but more electricity could be generated from photovoltaic (PV) units/arrays. The principle of using **heliostats** (parabolic mirrors) to direct the Sun's energy onto a focal point on solar furnaces and solar towers is utilised to produce extremely hot temperatures for industrial uses or alternatively to generate steam for turbines to produce electricity (Figure 11.37).

Figure 11.37 The PS10 solar power tower, near Seville, Spain

Wind energy

The potential for wind energy varies enormously because of variations in wind speed and reliability. Turbines are used to generate electricity and are increasingly found in large numbers on wind farms connected to an electricity grid network. Conflicts with other land uses may restrict wind energy use in many parts of the world. Countries with frequent, prevailing winds and large coastlines are more likely to develop offshore wind farms.

Biomass sources

Biomass sources currently provide around 10 per cent of TPES. 'Bioenergy' is a broad category which includes the use of any plant or animal matter converted into solid, liquid or gaseous biofuels using one of a range of conversion technologies. The term 'traditional biomass' mainly refers to fuelwood, charcoal and agricultural residues which are used for household cooking, lighting and space-heating in developing countries. These are the main fuel sources for over a third of the world's population. Biomass can also be burnt in thermal power stations or used in the anaerobic digestion process producing a biogas that can be used for cooking or heat generation or burnt for thermal electricity generation.

Hydroelectric power (HEP)

Falling water generates around two per cent of total energy supply. It is thought to be under-exploited; however large dam projects (such as the Aswan High Dam in Egypt, pages 591–2) are controversial and carry heavy environmental costs.

Tidal power

The energy potential of the oceans is massive but only a fraction is likely to be harnessed. Tidal power is harnessed either through tidal current water turbines (underwater versions of wind turbines) or by tidal barrages. Barrages operate in a similar way to HEP dams and turbines: the incoming tide is captured behind a barrage at high tide, giving a great deal of potential energy. At low tide, sluice gates are opened and the captured water is released through turbines, thus maximising its kinetic energy (see Figure 11.38). Electricity can be generated through turbines, on both the incoming ebb tide and the outgoing flow tide. Viable sites for tidal barrages are limited as specific physical conditions are necessary, such as deep coastal bays, inlets or estuaries with a high tidal range. Tidal barrage construction is controversial because of the enormous cost and because of the environmental disruption they may cause by changing natural tidal flows.

Wave power

Various technologies are being developed to make wave power commercially viable, especially as energy is lost in transmitting the electricity to shore from offshore wave farms. The most common offshore technologies are based on harnessing the kinetic motion of surface waves to move floats up and down so that hydraulic rams drive an electrical generator (see Figure 11.39(a) – the Pelamis Wave Energy Converter). Onshore technologies are an alternative and the LIMPET device (on the Isle of Islay), for example, harnesses onshore wave action inside a concrete chamber to compress air which is forced through a pair of turbines (Figure 11.39(b)).

Geothermal energy

Earth's temperature rises by approximately 1°C for every 30 m of depth within the crust, and this temperature gradient is even steeper in geologically active areas such as plate margins. Geothermal energy makes use of this heat, either directly by producing hot water or indirectly to produce electricity. The most common way of capturing geothermal energy is to tap into naturally occurring 'hydrothermal convection' systems where cooler water is pumped into the Earth's crust, where it heats up and then rises to the surface. Once this

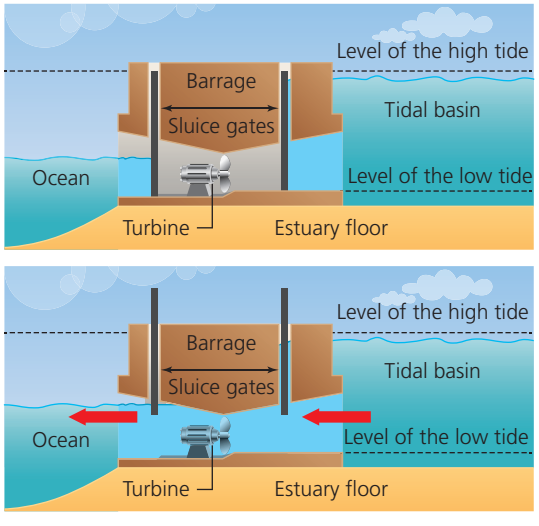
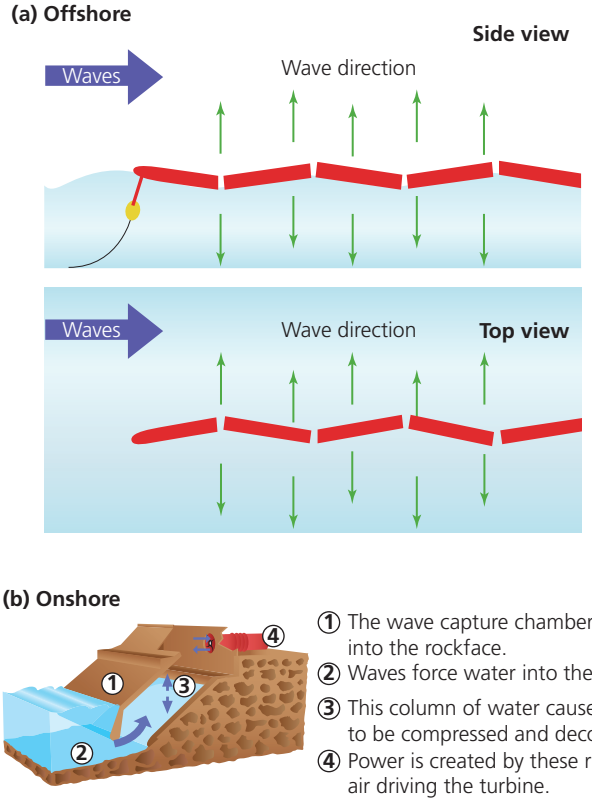


Figure 11.38 How tidal barrages work – this is a one-flow generation system



① The wave capture chamber is built into the rockface.
② Waves force water into the chamber.
③ This column of water causes the air to be compressed and decompressed.
④ Power is created by these rushes of air driving the turbine.

Figure 11.39 Offshore and onshore wave energy technologies

heated water is forced to the surface, the steam is captured and used to drive electric generators (Figures 11.40a and b).

The main and overriding advantage of all renewables is that they are sustainable in the long term and reduce our dependency on fossil fuels. Some renewable energy forms have greater environmental impacts than others, either in their manufacture, construction or operation. With the exception of biomass sources, all renewables require certain physical factors to be present to make them viable, thus rendering them **geographically specific**, which could be a limitation for their further development.

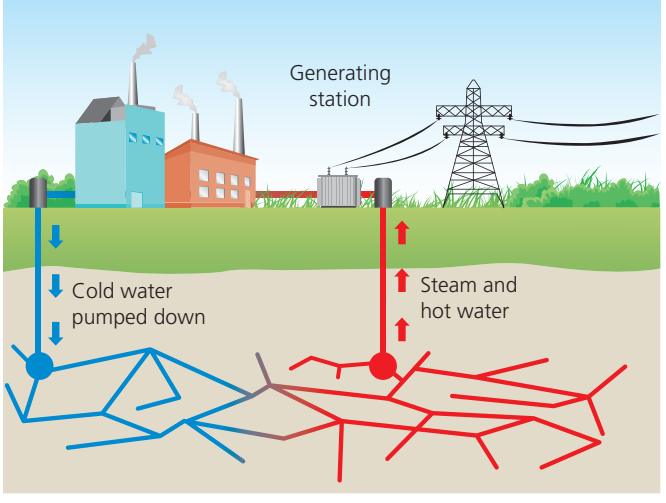


Figure 11.40a How geothermal electricity generation works

Figure 11.40b Iceland's Nesjavellir geothermal power station

Energy mix and components of energy demand

The global TPES is in effect the global **energy mix** but this will vary considerably from country to country. The utilisation of this energy will also vary from a global average, especially between industrialised and non-industrialised countries.

Energy mix

Energy mix is the composition of different primary energy sources from which households and industries in a named area (usually at national level) get their energy. Some data refers only to the combination of energy sources used to generate electricity. This is not the primary energy mix of a country as it only includes one form of secondary energy and ignores transport or other uses such as gas or biofuels used in heating or cooking, etc.

Factors influencing a country's energy mix

- **Availability of energy sources within the country:** Nations seek to use available resources first.
- **Inertia:** Retaining a mix that already exists because of economic or technical difficulties in changing (even if the reasons are outdated).
- **Government energy policy:** Making decisions to strive for energy security (for example, France) or to abide by international treaties (for example, the 2015 Paris Climate Agreement).
- **Geopolitics:** Trading partners who are friendly or the country can co-operate with; some suppliers cannot be relied on or present potential conflict.
- **Level of development (economic/technological):** Has to be 'appropriate', for example, less developed countries cannot afford to develop nuclear energy or become over-reliant on fossil fuel imports.
- **Physical/locational conditions:** Suitable for certain types of renewable energy (for example, UK – wind; Iceland and Norway – HEP; Spain – solar).
- **Diversity:** Governments may decide to diversify their energy mix so that countries are not too reliant on only one or two sources.

Table 11.16 Advantages and disadvantages of renewable energy sources

	Advantages	Disadvantages
Solar	Pollution free (except for production and disposal of pv units) Plentiful silica resources (to make pv cells) High reliability, no moving parts Relatively cheap and efficient Good potential in less developed countries	Intermittency of solar intensity Ineffectiveness of storing the energy generated Use of some toxic materials in pv cell production Grid connection challenges
Wind	Quick installation and dismantling (onshore) Relatively simple technology No atmospheric pollution (except in production of turbines) Land use does not always conflict with other operations (for example, farming) Option for offshore combined with tidal energy	Some land use conflicts Conflicts such as visual intrusion, effects on wildlife (birds) Local objections to installation: visual, noise and radio interference Intermittent nature of wind thus relatively inefficient and unreliable Offshore technology and potential problems for sea navigation
Biomass	Proven simple combustion technology Alternative to oil for transport Energy crops are an efficient way of using 'marginal' land Mostly carbon neutral Relatively high energy density Flexibility of form – solid, liquid, gas	Growing energy crops conflicts with food production Some emissions (NOx and SOx) Fuelwood use can cause deforestation, soil erosion and desertification Carbon neutrality is questionable as CO_2 sink is lost
HEP	Low operating costs No waste or CO_2 emissions Proven technology Can top-up supply in periods of peak demand	Dams have high construction costs Requires large amounts of land so causes conflict, for example, relocation of populations Environmental impacts on drainage basin and micro-climate effects
Tidal	Generation periods are predictable Potential power output is large	Power generation is intermittent Ecological impacts (for example, on marine life)
Wave	Produces electricity regularly Less environmental impact (than tidal barrages, for example) Virtually pollution free Onshore sites relatively cheap	Difficulty in transmission from offshore Offshore 'farms' require large areas of sea Onshore technology is relatively inefficient Environmental impacts – reduces turbulence, increases sedimentation
Geothermal	Clean and sustainable Relatively cheap in suitable areas	Reliant on specific geology to be viable so restricted to certain locations

Energy mixes in contrasting settings: Niger, Iceland and France

Niger

- **Energy mix:** The majority (77 per cent) of Niger's primary energy supply of 3.100 mtoe in 2018 was provided by biomass sources, mainly in the form of fuelwood, charcoal and animal waste. Around 84 per cent of the population live in rural areas and rely on these fuels for heating and cooking. Only 20 per cent of the population has access to electricity. The country is reliant on electricity imports for a significant share of its supply. Ironically, Niger is one of the ten largest uranium producers in the world, with half a million tonnes of proven reserves, but it does not have the wealth or technical expertise to develop nuclear energy. It also has significant underdeveloped oil reserves and is able to use these to supplement energy supply to the transport and power generation sectors.
- **Energy consumption:** Households consume 82 per cent of Niger's energy, (road) transport 14 per cent and industry (mainly mining) and commercial/public services use the remainder.
- **Potential:** There is potential for harnessing HEP from the Niger river. Construction started on the Kandaji HEP station in 2019, which is expected to open in 2025. There is also considerable potential for developing more solar and wind energy. Niger is reliant on a US government support organisation, called Power Africa, to support the energy sector, particularly in encouraging the development of electricity in rural areas.

Why this energy mix? There is little use for electricity except in urban areas. The main energy need is for cooking and heating, for which firewood is the cheapest and most readily available source. Coal and oil reserves

are not fully developed. Niger is a poor country and cannot afford to develop much of its energy potential without help from the World Bank.

Iceland

- Iceland has one of the highest energy consumption/capita of any country.
- It is one of the lowest carbon emitters as it only uses fossil fuels for transport.
- It straddles a plate margin which is geologically active so has abundant geothermal energy.
- It is a mountainous country with plentiful water supply and potential energy in falling water for HEP.
- As an advanced country with a relatively small population, Iceland can afford the technology to harness its geothermal potential and to construct dams to develop HEP for its electricity supply (Figure 11.41).

Table 11.17 Iceland's renewable energy, 2018

Source	Primary energy	Electricity only
HEP	21%	71%
Geothermal	67%	29%
Fossil fuels	12%	0

France

- France gets 40 per cent of its primary energy from 59 nuclear power stations (Figure 11.42).

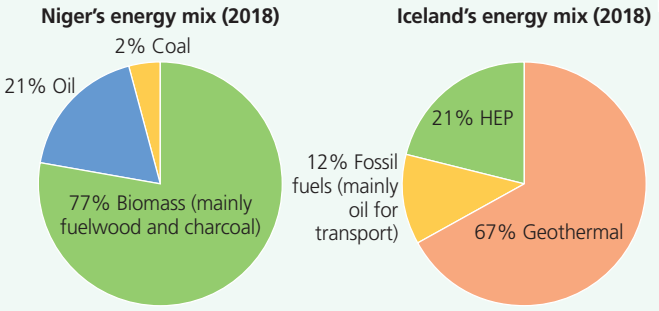


Figure 11.41 Pie diagrams showing contrasts between energy mix in Niger and Iceland, 2018

- 72 per cent of France's electricity is supplied by nuclear energy and it is able to export electricity.
- The French government decided to develop nuclear capacity following oil price rises in the 1970s. It was a less controversial choice for determining energy security at that time.
- France has a shortage of fossil fuel reserves (especially oil/gas).
- Physical conditions were suitable for developing nuclear energy with many large rivers for cooling.
- Many of the nuclear reactors are reaching the end of their lifetime, so France has started an ambitious energy transition under the Energy and Climate Change Law 2019, aiming to gradually replace lost nuclear capacity with non-hydro renewable technologies.

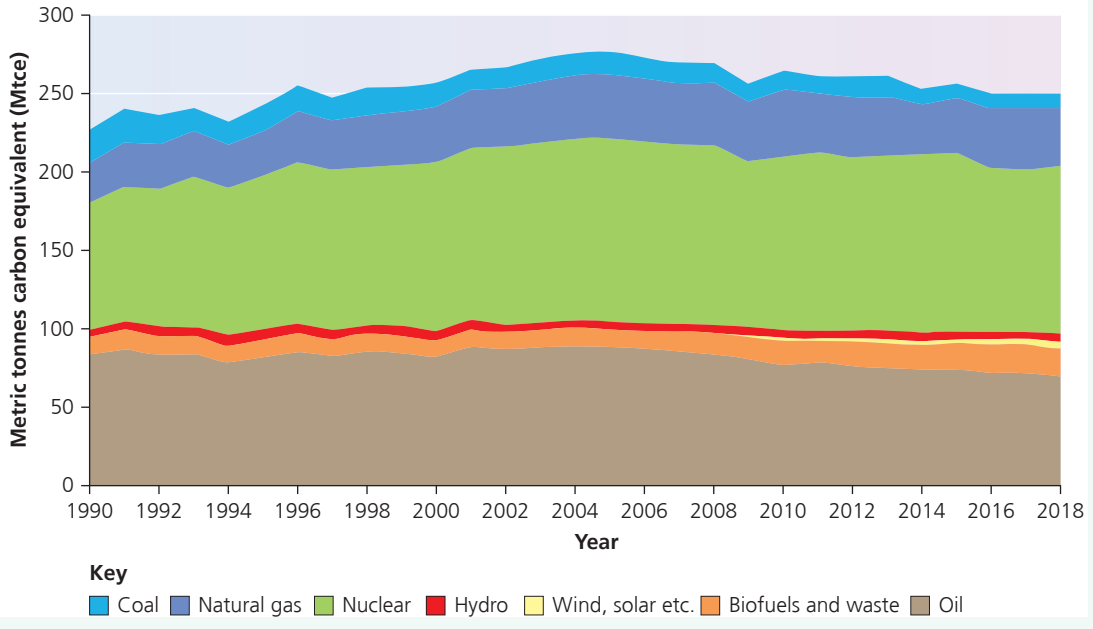


Figure 11.42 Compound line graph showing France's primary energy mix, 1990–2018

Source: International Energy Agency

Components of energy demand

Energy that is generated by primary sources is used in a number of ways, but most of the data divides the components into four categories: industrial, commercial, residential (household use) and transport. Figure 11.43 shows the world energy consumption by these four sectors in 2015 and at this scale the majority of energy is used for industrial purposes.

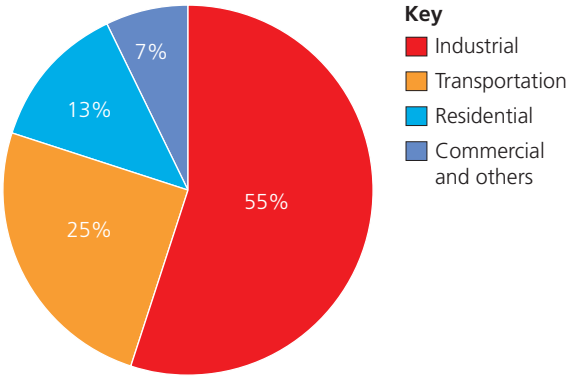


Figure 11.43 World energy consumption by sector, 2015

Consumption by sector will vary from region to region and Figure 11.44 shows a slightly different composition for the EU in 2015. Here, in some of the most developed countries of the world, energy generated for transport, services (commercial) and household use (residential) is significantly higher and the proportion devoted to industrial use is lower.

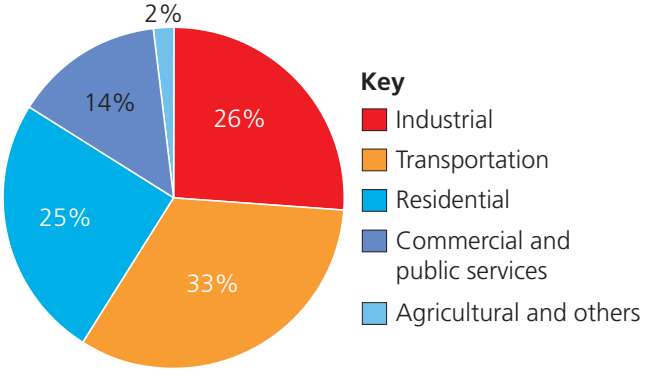


Figure 11.44 EU energy consumption by sector, 2015

Table 11.18 shows contrasts in the components of demand for energy in China, India and the USA. It demonstrates the importance of energy for industrial use in China. In India, residential, commercial and agricultural use is more important, contrasting with the USA where transportation is the dominant use of energy. Components of energy demand generally reflect the importance of each sector to a country's overall economy.

Table 11.18 Components of energy demand by sector in China, India and USA

Sector	China	India	USA
Industrial	54%	36%	20%
Transportation	17%	18%	42%
Residential	20%	31%	19%
Commercial and public services	6%	8%	18%
Agriculture and others	3%	7%	1%

Skills focus

Using the data in Table 11.18, construct a compound bar graph for the different energy uses in the three countries.

Energy supply and physical geography

The availability of different energy resources is dependent on a number of physical limiting factors; for fossil fuels and geothermal energy, geological conditions are a major determining factor, while for many renewable sources of energy, such as wind and solar energy, climate is a more important constraint.

Geology

Coal

Coal is a sedimentary rock which started to form 360 million years ago in swamp conditions on the edge of sedimentary basins such as lagoons or lakes. Plant debris accumulated and was buried under layers of mud and sand where the decomposition process was slowed down due to the anaerobic conditions. The basin gradually sank under the weight of more sediment and the layers of dead vegetation were subjected to increasing amounts of heat and pressure in a process known as 'coalification' (Figure 11.45, page 604). Increased heat and pressure over time reduces the levels of moisture and volatile materials present in the sediment. At the same time the carbon content and quality of coal that is produced will increase.

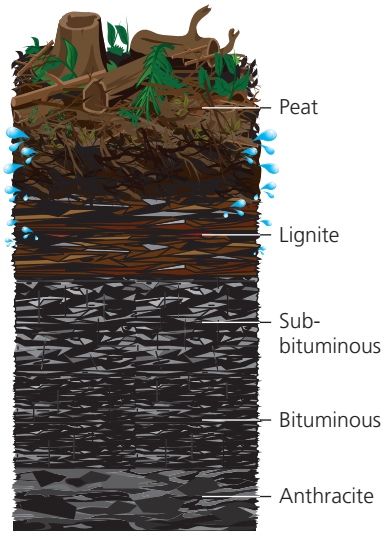


Figure 11.45 Stages in the process of coalification

The different types of coal formed are ranked according to their quality, which is dependent on carbon content:

- **anthracite:** (86 per cent to 98 per cent pure carbon) is an excellent fuel that is still used to heat homes
- **bituminous coal:** (containing 70 per cent to 86 per cent carbon) is used to make coke for furnaces
- **sub-bituminous coal:** (70 per cent to 76 per cent carbon) is burned in industrial boilers
- **lignite** (also known as '**brown coal**'): (65 per cent to 70 per cent carbon) is a low-grade fuel with a high moisture content that is also used in industrial boilers.

Thermal power stations generating electricity use a grade which is between anthracite and bituminous known as **steam coal**. The coal is first milled to a fine powder, which increases the surface area and allows it to burn more quickly.

Oil and natural gas

Oil and natural gas are hydrocarbons of organic origin (plants, algae and animals) which settled to the bottom of the sea as they died, were buried and fossilised in sedimentary 'source' rocks such as finely grained shale. The organic matter in the sediment changed into a substance called kerogen and when underground temperatures rose above 110°C, the kerogen gradually changed into oil. Under hotter conditions it formed into natural gas.

After oil and gas have formed in the shale, pressure continues to rise, squeezing the oil and gas out, upwards or horizontally into more porous and permeable rocks (for example, sandstone) that have more spaces within them for the oil and gas to flow into. Some eventually reaches the surface and escapes naturally, but most of it comes up against a layer of impermeable rock that it cannot pass through. This is known as a 'cap' rock because it forms a seal or trap, and slowly the oil and gas build up in the porous rock beneath it, forming a reservoir.

Reservoirs are sedimentary rock formations that hold oil, natural gas or both within their pores, like a fossilised sponge. All pore spaces in the rocks are filled with water, gas or oil. Crude oil is a complex mixture of hydrogen, carbon and traces of other substances. Its texture varies, but it is generally liquid and in a geological trap it will accumulate above layers of trapped water. Natural gas is mainly made up of methane; as it is the lightest substance it moves to the top of the trap.

There are two types of traps – **structural traps** and **stratigraphic traps**. Structural traps such as the one shown in Figure 11.46 hold gas and oil because the Earth's crust has been deformed in some way, such as the dome-shaped anticline shown.

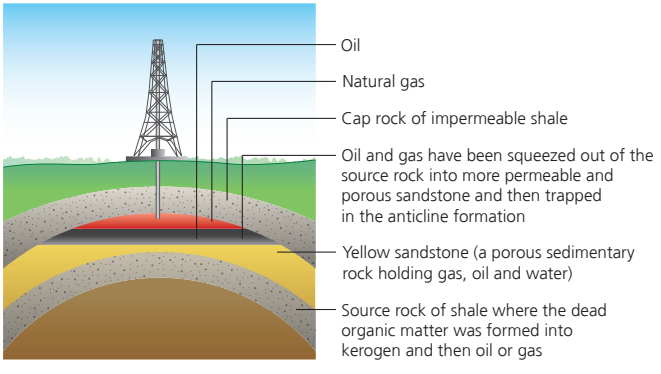


Figure 11.46 A conventional oil and natural gas reservoir: this illustrates a structural 'anticlinal trap' that contains oil and natural gas

Unconventional availability of oil and gas

Hydraulic fracturing ('fracking'): Economic pressures on fuel prices have prompted technological advances enabling the **resource development** of other known reserves of oil and gas, such as those found in shale. The pore spaces in shale are so tiny that the oil and gas have difficulty moving through it into a conventional reservoir.

New technology has enabled engineers to penetrate down to the level of the shale and then drill horizontally through the shale rock (Figure 11.47). The permeability of the shale is increased by pumping water down the well under sufficiently high pressure to fracture the rock. These fractures liberate some of

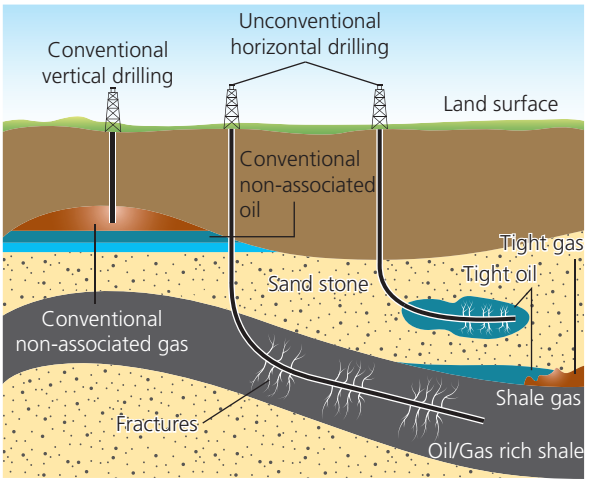


Figure 11.47 Unconventional fracking of 'tight' gas and oil deposits compared to conventional drilling

the gas and oil from the pore spaces and allow them to flow to the conventional well or directly to the surface. This technology means that the previously '**possible resources**' of oil and gas are now accessible as recoverable '**measured reserves**'.

Oil and tar sands: These are bituminous sands and are a type of unconventional petroleum deposit. The largest discoveries and development of these reserves have been on the Athabasca River in Alberta, Canada and in the Orinoco basin in Venezuela (though technically the latter fall into the category of 'heavy oils'). These areas have thus become **resource frontiers**.

The bitumen is too thick to be pumped as a liquid from the ground so has to be mined by strip mining or open pit mining. The relatively low recovery rate of oil of around 10 per cent is compensated for by the massive scale of the reserves and the relatively low cost of extraction. A number of techniques have been developed to separate the bitumen from other constituents but it still takes around four tonnes of tar sands to produce one barrel of oil (see the text box on the Athabasca tar sand oil development in Alberta, Canada, pages 607–10).

Climate

Solar energy

Solar energy can be harnessed anywhere that receives sunlight, so it is abundantly available but it is also variable and intermittent. Different locations and weather conditions will affect the amount of energy that can be harnessed for heating or electricity production. Solar power is less effective in overcast or cloudy conditions and needs storage to release energy at night when it cannot be generated.

Climatic considerations:

- Sunlight:
  - tropical and sub-tropical areas receive the most reliable and regular sunlight
  - the Sun is higher in the sky so the days are longer
  - cloudy regions make solar systems harder to justify, though not necessarily uneconomical.
- Frequent fog, smog or air pollution: Will negatively affect the amount of sunlight received.
- Air density: Solar exposure is better in the mountains than near sea level because the air is thinner and scatters less sunlight.
- Snowfall: PV systems will not operate effectively if covered with heavy layers of snow.
- Rainfall: Electrical and metallic components are susceptible to corrosion in humid conditions.
- Wind: Frequent strong winds can damage solar installations and cool down surfaces on solar water heating panels quickly.

Wind energy

Wind energy can also be variable and intermittent. For economic reasons wind farms are most viably productive in areas where average wind speed is greater than 5.5 metres per second.

- **Wind speed**:
  - The minimum wind speed at which the wind turbine will generate usable power is between 7 and 10 mph for most turbines.
  - The minimum wind speed at which the wind turbine will generate its designated power (known as 'rated power') is between 25 and 35 mph for most turbines.
  - At very high wind speeds of between 50 and 80 mph, most turbines will shut down. This is the cut-out speed and is a safety feature which protects the turbine from damage.
  - Turbines have to be spaced well apart to prevent eddying or a reduction in wind speeds in the middle of wind farms, which means large areas of land are required.
- **Air density**: The greater the density of the air, the more energy is received by the turbine. Air is less dense at higher elevations than at sea level, and warm air is less dense than cold air.
- **Prevailing winds**: Optimum conditions for wind energy generation exist when there is a prevailing wind direction.

The UK has the best wind energy potential in Europe because of its position in temperate latitudes on the northeast rim of the Atlantic. It has the long 'fetch' of uninterrupted air flow across the ocean from which prevailing south-westerlies blow.

Drainage systems

The size and shape of drainage basins, together with the volume of water in the channels, are important considerations in the development of large- or small-scale conventional HEP which involves dam construction. Two vital factors to consider are the **flow** and the **head** of the stream or river. The flow is the volume of water which can be captured and re-directed to turn the turbine generator, and the head is the height the water will fall (potential energy). The larger the flow and the higher the head, the greater will be the energy available for conversion to electricity.

The key equation is:

Power = Head × Flow × Gravity

The topography of the drainage basin is another important consideration as dam construction is expensive. A long, steep-sided valley basin with a relatively narrow exit is ideal as this maximises the volume of the reservoir and makes dam construction more manageable and less costly. The geology of the rocks should be stable and impermeable, to secure the dam and to prevent water seepage (Table 11.11, page 587).

Energy supply and globalisation

Links

There are links with Chapter 7, Global systems and global governance – Trade agreements, pages 295–9, International trade and access to markets, pages 306–14, and Nature and role of TNCs, pages 314–17.

By 2025 energy consumption in newly industrialising countries is set to surpass that of more developed countries. As countries such as China and India develop, so will their appetite for natural resources, leading to fierce demand for ever-diminishing supplies of non-renewable energy.

Competing national interests

The importance of oil and gas, combined with their uneven distribution, make these the most traded energy commodities. Both are vulnerable to events that impact on their supply and demand. Countries that depend on oil imports are defenceless against external events affecting their supply.

As reserves of non-renewable fossil fuels continue to diminish, the following scenarios are likely:

- European dependence on oil and gas from various suppliers will remain significant.
- Asian dependence on Gulf oil will increase.
- China will start to develop oil reserves in Africa (for example, Chad and Sudan).
- As the USA develops its 'tight' reserves, it will become a net oil exporter but will still rely on imports because of different grades of oil used in refining.
- Natural gas will continue to be the fastest-growing tradeable primary energy source, requiring more investment in transborder pipelines; though more will also be shipped in liquefied natural gas (LNG) tankers.
- Russia will expand its exports of oil to China, Japan and South Korea through the East Siberian Pacific Ocean (ESPO) pipeline and will want to maintain gas and oil supplies to Europe.
- The use of sustainable renewable sources of energy will increase and will gradually reduce reliance on trade in fossil fuels.

Except for the growth of trade between Russia and its eastern Asian neighbours, it is unlikely that the pattern of oil production and trade will change significantly in the near future.

Role of transnational corporations (TNCs)

Of the top ten largest TNCs by revenue in 2018, six were oil companies, including Royal Dutch Shell, BP, and Exxon-Mobil. Wealthy energy TNCs dominate the international oil and gas trade, exercising considerable control in today's globalised world. They fund development projects and even support presidential campaigns in some countries. As a result of their size and wealth, they can control and coordinate economic activities and influence governments, for example, by being granted

exploration rights in an area to commence the cycle of resource development there. Typically, energy TNCs are:

- involved in all stages of exploration, production and distribution, categorising these activities as:
  - **upstream** – exploration and drilling
  - **midstream** – transportation by pipeline or tanker
  - **downstream** – refining of crude oil into different products for distribution and sale
- often involved in **joint ventures** with other energy companies or with the host government where they operate
- increasingly positioning themselves for energy transition by putting considerable investment into renewable energy development.

Energy companies based in developing countries are also starting to exert influence on a global scale. They largely control production in their own country; some are totally or partially state-owned such as Sinopec and CNPC in China, Gazprom (51 per cent owned by the Russian state) and Petrobras (a semi-public Brazilian company).

Royal Dutch Shell

Shell is an Anglo-Dutch transnational oil and gas company with headquarters in the Netherlands.

Shell is **vertically integrated** and is active in every area of the oil and gas industry, including:

- exploration and production – using geological expertise
- refining – it owns 35 refineries
- distribution – it has its own fleet of oil tankers and lorries
- petrochemicals
- power generation and trading
- investment in renewable energy sources.

Shell operates in over 70 countries and produces around 3.7 million barrels of oil equivalent per day. Its main production activities are in South East Asia and in the Niger Delta in Nigeria (pages 631–4) where it has been accused of damaging local communities and the environment by its activities. Shell provides jobs and income, as well as improving the skills, education and training of the local workforce. Shell works in consortiums or joint ventures with other companies (for example, with Petronas, a Malaysian state company) to gain exploration rights, or with other TNCs to invest in mutually advantageous infrastructure such as oil pipelines. Shell has invested in alternative, renewable sources of energy and has ambitions to become a large electricity supplier.

Figure 11.48 Shell's Pernis oil refinery at the Port of Rotterdam, Netherlands

Environmental impacts of the Athabasca tar sand development in Alberta, Canada

Background to tar sand development in Alberta

As conventional sources of crude oil are depleted, unconventional sources, such as the bitumen found in tar sands (also known as oil sands) play a role in offsetting declining conventional production.

The bulk of Canada's tar sands reserves are found in three major deposits in the province of Alberta, where they underlie 140,000 km^2 of pristine boreal forest, an area that is approximately the size of the state of Florida (Figure 11.49, page 608). The largest and most viable of the deposits is the Athabasca River deposit. Only 20 per cent of the oil sands lie near the surface where they can be open-pit strip mined. The remainder are more than 75 metres below ground and are extracted using thermal 'in-situ' methods. The sands are estimated to contain 175 billion barrels of recoverable reserves of crude bitumen.

The oil sands have been mined since the 1960s but advances in extraction technology combined with

continued rising oil prices led to more extensive development from the start of the twenty-first century. In recent years, tar sand production has outpaced conventional oil reserves and output has risen to 2.5 million barrels per day with planned production to reach five million barrels per day. Most of the oil is exported to the US where it provides around 7 per cent of its daily consumption. There are around 20 mine projects in operation, owned and operated by a number of different partner companies, including many large transnational oil corporations such as Chevron-Texaco, Shell and Total.

Figure 11.49 Map of the tar sand deposits in northern Alberta

Environmental impacts

Development of this unconventional resource requires unconventional extraction and production processes with associated risks and consequences. To justify the investment, the scale of the project is enormous, arguably the largest industrial project in human history and possibly the most destructive.

Extraction

Most production takes place in vast open-pit mines where the sands are strip-mined layer by layer from the surface (Figure 11.50). Some of these open-pit mines can be 150 km² and 90 m deep. Before strip-mining begins, the boreal forest is clear felled, rivers and streams have to be diverted and wetlands drained. The overburden of soil, rocks and clay are stripped away (and stockpiled to be used in reclamation) to access the bitumen deposits. It is estimated that around 4 tonnes of material has to be removed to produce each barrel of bitumen. This leaves massive scars on the landscape as well as destroying

Figure 11.50 Open pit mining of Athabascan oil sands in northern Alberta, Canada

habitats and natural ecosystems, leading to Canadian Geographic terming the mining projects as 'scar sands'.

Thermal 'in-situ' methods involve injecting high-pressure steam into a well to soften and liquify the bitumen so that it can be pumped out, leaving most of the solids in the ground. In-situ development could occur in an area 30 times greater than the current mining area. These methods have a much smaller footprint and require less land disturbance, but they impose another set of negative environmental impacts (see Table 11.19).

Upgrading and refining

Producing oil from the tar sands is complex. They consist of 85 per cent sand, clay and silt, 5 per cent water and 10 per cent crude bitumen (a tar-like substance which is converted to oil). Bitumen does not flow like oil and

Table 11.19 Advantages and disadvantages of 'in-situ' extraction compared to open-pit mining

Advantages of in-situ bitumen extraction	Disadvantages of in-situ extraction
Smaller surface footprint – surface area for well heads is relatively small	**Linear disturbances** – from 3D seismic and core hole exploration, roads and pipelines; these negatively impact on wildlife
More efficient water usage – water used for steam production is mostly recovered and recycled	**More production well-pads** – smaller footprint but more of them and more scattered
No tailings pond – much of the sand is left in the ground	**Lower bitumen recovery rates** – only about 40–60 per cent; much lower than surface mining
Lower costs – with no mine and no tailings pond, projects require less capital	**Greater greenhouse gas emissions** – steam is produced by burning natural gas

extracting it from the sands uses a hot-water based separation process requiring large quantities of both water and energy. Once separated, the bitumen has to undergo further energy-intensive processes to upgrade it into a synthetic crude oil. This occurs before it is sent to refineries. In some cases, raw bitumen is diluted (using napthia) and sent by extensive pipeline networks to refineries (mainly in Midwest USA) for both upgrading and refining.

Climate change

The mining and upgrading procedure releases at least three times the CO_2 emissions of conventional oil production and is said to be the single largest industrial contributor in North America to climate change. The transition from conventional oil to tar sands actually worsens greenhouse gas emissions, as these processing emissions are added to the emissions from the end product.

The carbon emissions contribute to a rise in global temperatures, which have impacts on a global scale. Ironically, the effects of this can be seen visibly within the Athabasca basin with the rapid melting and retreat of the Athabasca Glacier upstream in Jasper National Park. The outlet stream from the glacial lake is the source of the Athabasca River. Water is abstracted from the river to be superheated for the in-situ extraction of bitumen and also for separating the bitumen extracted by open-pit mining. The emissions from these processes contribute to the greenhouse gases that are melting the glacier.

Loss of forests and wildlife habitat

The tar sands are already considered to be the cause of the second fastest rate of deforestation on the planet after the Amazon Rainforest. Canada's boreal forest is globally significant as it is a complex ecosystem of trees, wetlands and lakes representing 25 per cent of the world's intact forests. It provides 1.3 billion acres of wild habitat for a vast array of species including wolves, grizzly bears, lynx, nesting migratory birds and thousands of plant varieties.

Planned tar sand development projects are expected to see at least 5,000 km^2 of forest cleared, drained and stripped for open-pit mining. The remaining 135,000 km^2 could well be fragmented into a web of seismic lines, roads, pipelines and well pads for in-situ drilling. These will negatively affect the populations of wildlife, such as the woodland caribou, which avoid linear features.

Water abstraction and pollution

For each barrel of oil produced, as much as five barrels of water are used in extraction, separation and upgrading processes. The Athabasca River serves as the primary source of water used in these processes. Current operations are permitted to withdraw 350 million m^3/year (equivalent to the amount used by a city with 2 million people). Initially, it was thought that the river had sufficient flows to meet the demands of tar sand operations but it is becoming clear that this might not be the case, particularly in winter when flows are naturally lower. Even in-situ operations, despite recycling, leave some water underground, so a continuous supply of new water is needed. Some operations are a distance from rivers, so rely on using groundwater aquifers, which lowers the water table in the region and threatens surface freshwater.

Over-abstraction poses threats to the sustainability of fish populations in the Athabasca River and also to the Peace-Athabasca Delta, the largest boreal delta on Earth. It is a World Heritage Site and one of the most important waterfowl nesting and staging areas in North America.

Tar sand mining operations produce large volumes of waste in the form of **tailings** (around six barrels of tailings per barrel of bitumen extracted). Liquid tailings are a slurry mix of water, sand, clay and unrecovered bitumen but also contain naphthenic acids and trace metals, making them toxic to aquatic organisms and mammals. Operators are required to store tailings waste in vast wastewater reservoirs on site in large containment dykes, because the water is too toxic to be returned to the river. These reservoirs are misleadingly referred to as tailings 'ponds', as they cover an area of around 100 km^2 and are thought to contain 720 billion litres of toxic tailings (Figure 11.51). There is growing evidence that toxic chemicals from tailings ponds are leaching into groundwater and into the nearby Athabasca River.

Air quality

Air pollution from tar sands operations is increasing at both local and regional scales. Criteria air contaminants (CACs) are the most common pollutants released by heavy industry. They include lead, particulate matter (PM), carbon monoxide, nitrogen oxide and sulphur dioxide, all of which are emitted in large volumes by tar sand operations. These pollutants affect human health and contribute to acid rain.

Figure 11.51 Polluted tailings ponds near the Suncor tar sands mining operation in northern Alberta

Cumulative impacts and reclamation

In addition to the environmental devastation in northern Alberta, pipeline infrastructure to refineries and to supertanker ports crosses the continent to all three major oceans and the Gulf of Mexico.

Very little of the area directly affected by mining operations has been reclaimed and tailings ponds are expected to cover even greater areas over the next 20 years. UNEP has identified Alberta's tar sand mines as one of 100 global 'hotspots' of environmental degradation. According to Environment Canada, development of the tar sands presents 'staggering challenges for forest conservation and reclamation' and yet oil sands projects continue to be approved.

Supply management strategies

In order to ensure security of energy supply, countries have to develop strategies which will increase the reliability of future supplies and simultaneously reduce consumption by maximising energy efficiency.

Decisions about energy security will often lead to countries, such as the UK, diversifying their energy mix, so they are not too dependent on one source of energy, and thus vulnerable to economic change.

Oil and gas exploration

As the price of fossil fuels increases, extra exploration effort is invested in finding new reserves or reassessing speculative reserves to analyse whether they have become viable to develop. Countries are more willing to grant exploration rights to TNCs in these circumstances. For example, although North Sea oil and gas are past their peak, exploration continues and new discoveries are still being made. In 2019, gas production commenced from the large new Culzean field (230 km off the Aberdeenshire coast). A special three platform rig needed to access the gas cost over £3 billion. It is operated by Total and supplies five per cent of the UK's demand for gas.

Oil and gas price rises prompt large TNCs to develop new technologies to access previously unviable reserves. Such has been the case with hydraulic fracturing in the USA, but in England, fracking operations by companies such as Cuadrilla were stopped by the UK government in November 2019. Fracking operations had attracted protests from environmentalists. In August 2019, an earth tremor registering 2.9 in magnitude was recorded near Cuadrilla's site in Lancashire. The government finally cited fears of damage from earthquake tremors as the reason for placing a moratorium on operations; though they stopped short of an outright ban.

Nuclear power

Nuclear power provides a long-term, efficient and viable energy security alternative for many countries which have become over-reliant on fossil fuels. It currently provides around 20 per cent of the UK's electricity but all operating energy plants (with the exception of Sizewell B in Suffolk) will be closed by 2030. No new reactors have been opened since the Pressurised Water Reactor (PWR) at Sizewell B was commissioned in 1995.

The UK's Department of Business, Energy and Industrial Strategy (BEIS) intends to increase new nuclear capacity and sees it as an important part of increasing energy supply. The UK government relies on the private sector to build new reactors and funding remains an issue. Construction for EDF's twin reactors at Hinckley Point in Somerset (C1 and C2) started in 2012 and they are expected to be operational by 2025. EDF's investment in building the £25 billion project is based on being able to charge a guaranteed rate for the electricity generated. However, projects to build new reactors on existing sites at Oldbury (Gloucestershire) and Wylfa (Anglesey) have been shelved as the developers, Hitachi-GE, failed to attract investors. Plans for new reactors to be built at Sizewell C and Bradwell B (Essex) in the mid-2030s using Chinese technology means that generating nuclear power remains part of the UK energy supply strategy for the rest of the century.

There are important reasons why nuclear power remains part of a country's strategy to supply energy:

- It is a low carbon choice that supports climate change goals: it generates electricity without emitting greenhouse gases.
- It produces electricity 24 hours a day, whatever the weather.
- Reactors do not require a daily supply of new fuel to operate.

Development of renewable resources

Since the Paris Climate Agreement was signed in 2016, the development of renewable energy has been a key

strategy to increase energy supply in many countries. High installation costs often require government intervention in the form of subsidies. More recently, the drive for energy transition and decarbonisation has increased market demand and investment from the private sector.

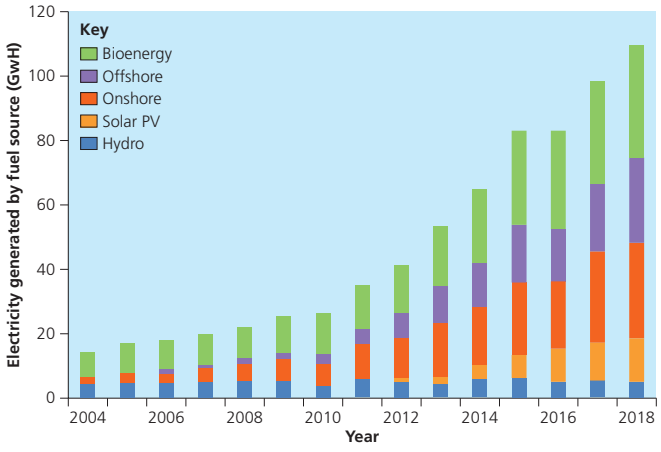


Figure 11.52 Increases and composition of renewable energy generation in the UK, 2004–2018

Source: Reddie & Grose LLP

In the UK, a variety of government subsidies were offered to provide financial support for renewable electricity suppliers. These were stopped in 2015 on the basis that falling installation costs meant they were no longer necessary. The government's Review of Energy Policy in 2019 put climate change higher on the agenda than ever before and committed the UK to net zero greenhouse gas emissions by 2050. The BEIS's main mechanism for supporting low carbon electricity generation now is the **Contracts for Difference (CfD)** scheme which incentivises investment in renewable energy but also supports nuclear and carbon capture development.

Incentives have caused the rapid growth of renewable energy in the UK, particularly from wind, solar and biomass sources (Figure 11.52). The UK is the world leader in offshore wind energy: it generates around 10 per cent of the nation's energy supply. In 2019, renewable energy competed with gas to become the largest source of power in the UK; both contribute around 39 per cent of the electricity mix. With nuclear energy still providing nearly 20 per cent, renewables have largely replaced coal in the country's electricity mix.

Skills focus

Using the compound bar graph in Figure 11.52, calculate the percentages of each type of renewable energy resource used in 2018 and then construct a pie chart.

Notable renewable energy developments in the UK include:

- **solar – Blackfriars Bridge:** the world's largest solar bridge with 4,400 solar PV panels installed on its roof spanning the Thames; it provides over 50 per cent of the energy for Blackfriars Station
- **wind – Walney Extension** (659 MW) off the Cumbrian coast and **London Array** (630 MW) are the UK's largest offshore wind farms and **Whitelee** (539 MW) on Eaglesham Moor in Scotland is the largest onshore wind farm
- **biomass – Drax** thermal power station in North Yorkshire generates two-thirds of its 4,000 MW capacity by burning wood pellets imported from Canada and the USA (the remainder is generated from coal).

Demand management strategies – managing energy consumption

The tone is set to reduce energy consumption by international agreements such as the Paris Agreement. Energy conservation efforts are mostly driven by:

- government incentives (for example, subsidising householders' loft and cavity wall insulation)
- media and other campaigns to raise awareness – these may be government led or run by energy-saving NGOs such as the Carbon Trust
- pricing strategies – the cost of fuel gives an incentive for energy saving but off-peak electricity pricing can also reduce pressure on generation.

Household energy saving

The EU Directive on the Energy Performance of Buildings is part of government strategy for tackling climate change. The directive requires an energy performance certificate (EPC) to be issued whenever a property is constructed, rented out or sold. The EPC shows the energy efficiency rating of a dwelling on an A–G rating scale similar to those

used for refrigerators and other electrical appliances. The UK government has suggested it will continue to operate an EPC style scheme after Brexit is completed. To achieve good ratings homeowners will have to consider:

- improving thermal efficiency of walls, windows and roofs
- draught-proofing of floors, doors and windows
- installing a high efficiency condensing boiler
- being part of a **heat network** (also known as a district heating system)
- using new building materials to reduce heat loss (thermal blocks)
- using low carbon technologies: solar panels, biomass boilers or wind turbines
- installing energy efficient appliances and lighting
- improved daylighting – fitting larger windows (passive solar heating)
- using environmentally friendly/recycled building materials.

Heat networks

Heat networks (also known as district heating systems) are an important part of the UK government's plan to reduce carbon emissions by providing a supply of low carbon heat to homes, offices and public buildings. They exploit larger scale, often lower cost, renewable and recovered heat sources that otherwise would be wasted. Heat energy is released from a central source through a distribution system of insulated pipes to a number of domestic and/or non-domestic buildings. The heat source might be a facility that provides a dedicated supply to the heat network, such as a combined heat and power plant; or heat recovered from industry or 'energy from waste' plants.

There are currently about 17,000 heating networks in operation in the UK, covering half a million premises. It is estimated that around 18 per cent of UK heat will need to come from heat networks by 2050 if the UK is to meet its carbon targets cost effectively.

Industrial and commercial energy saving

The Energy Saving Trust and the Carbon Trust are non-profit organisations whose respective aims are to help the business community save energy and move toward a low carbon economy. They advise companies about creating climate change strategies and provide smaller businesses with free energy audits and no-interest loans for energy efficient equipment. The energy savings strategies offered are similar to those in residential premises.

Many industrial processes generate heat, which is a source of energy that can be wasted unless it is captured. The Carbon Trust encourages industrial energy consumers to install **heat recovery systems** which collect and re-use heat arising from any process. Heat recovery can help to reduce the overall energy consumption of the process itself, or provide useful heat for other purposes. This should not be confused with **combined heat and power systems**, which are also beneficial in reducing energy consumption in industrial, commercial and domestic settings.

Combined heat and power (CHP) systems

CHP generates electricity while also capturing usable heat that is produced in the process. Conventional electricity generation in coal- and gas-fired power stations generates heat but, because of their remote location, up to two-thirds of the overall energy is lost. Energy is also 'lost' in transmission costs from generation to consumption.

CHP enables consumers to generate their own power locally using renewable or non-renewable fuels. It captures the waste heat from generation so it can be used for space or water heating, thus reducing energy consumption and carbon emissions (Figure 11.53). It can be used in homes, businesses or public settings. For example, Center Parcs' most recent site at Woburn Forest uses two CHP generators, which saves the company £200,000 and 900 tonnes of carbon emissions per year.

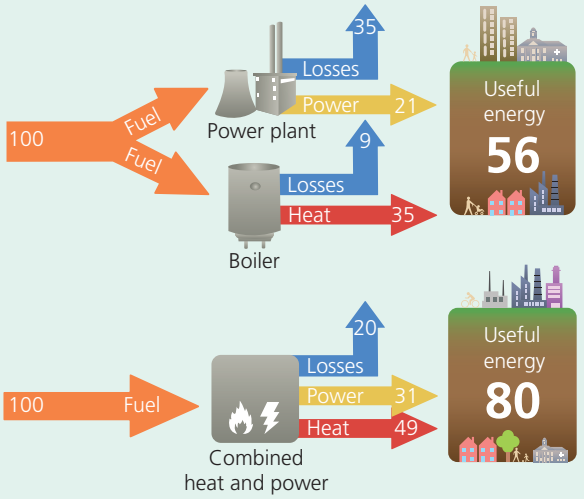


Figure 11.53 The principle of combined heat and power systems. Generating electricity locally can increase efficiency and reduce consumption

Source: The Association for Decentralised Energy

Sustainable transport strategies

Energy conservation can be achieved in transport use through technology, design and incentives to alter lifestyle and the type of transport used.

- **Technology:**
  - engine design: more efficient fuel usage (for example, fuel injection) with lower emissions
  - stop/start ignition control: reduces fuel wastage and cuts emissions
  - electric vehicles (EVs) – various types are now available including fully electric battery powered (BEVs); plug-in hybrids (PHEVs) and hybrids (HEVs). These become more sustainable options as the proportion of energy generated by renewable sources increases.
- **Methods and design:**
  - aerodynamic designs reduce fuel consumption
  - using more renewable fuels (for example, bioethanol)
  - using cleaner, more efficient fuels (for example, LPG, low sulphur alternatives).
- **Schemes and campaigns to discourage car travel or reduce congestion:**
  - government policy/legislation (for example, road tax based on emissions)
  - congestion charges (for example, London Congestion Charge Zone)
  - toll roads (for example, M6 Toll)
  - park and ride schemes in cities combined with high car park charges in city centres
  - campaigns and incentives to encourage car sharing
  - campaigns to encourage taking public transport/cycling/walking
  - planning homes, services and workplaces in close proximity to reduce travelling.

Energy sustainability issues

One of the main differences between renewable and non-renewable sources of energy is their impact on the environment. Although they have disadvantages, renewable sources are generally cleaner and less harmful to the atmosphere than non-renewables. Non-renewable sources of energy release harmful pollutants, such as carbon and sulphur compounds, into the atmosphere when they are burnt.

Environmental impacts

The environmental impacts of using energy resources take place during production, transportation (trade) and consumption of the fuels to produce energy.

In developing countries, particularly in rural areas, people largely rely on biomass such as fuelwood to meet their energy needs for cooking. The scarcity of fuel supplies has led to the widespread removal of woodland for firewood. This results in less interception of rainfall, reduced infiltration, faster run-off and greater soil erosion by water and wind, often leading to desertification.

Table 11.20 Summary of environmental impacts caused by using energy resources

Stage of utilisation	Type of impacts	Examples
Exploitation (production)	Fuelwood: deforestation leading to soil erosion and desertification	Sahel region of sub-Saharan Africa
	Oil/gas well leaks from drill points: damages ecosystems, kills wildlife, leaks methane	BP's Deepwater Horizon, Gulf of Mexico oil spill
	Open cast mining for coal and oil sands scars landscape and disrupts drainage	Athabasca oil sands, Alberta, Canada
	Hydraulic fracturing (fracking) may lead to subsidence and water contamination	Pennsylvania, Ohio and Colorado in USA
Transportation/ Trade	Oil and gas leaks from ruptured pipelines	Niger Delta, Nigeria
	Oil spills from large oil tankers at sea	Exxon Valdez spill, Alaska
Consumption	Combustion of fossil fuels leading to: **Acid rain** from carbon, nitrogen and sulphur compounds causing damage to wildlife, vegetation and to buildings **Enhanced greenhouse effect** causing global climate change	 Scandinavia; Northeast USA; South-east China and Eastern Asia Global effects

Key terms

Acid rain – Rain that has been made acidic by certain pollutants in the air.

Enhanced greenhouse effect – The atmosphere absorbs long-wave radiation from the Earth and is warmed. The effect is increasing because of the release of gases such as carbon dioxide into the atmosphere by human activity.

Acid rain

During the 1980s and 1990s a major environmental crisis confronting developed regions of Western Europe and Eastern North America was acid rain. The negative impacts were notable because the areas most affected were not the source areas of the pollution. Acid rain is a 'migrating' pollution. Scandinavia suffered the effects of acid rain from polluting sources in the UK and Germany. It is less of an environmental issue now in the US and Europe because of effective government regulation and control. Acid rain is now prevalent in South and East Asia with the pollution originating from industrialisation in India and China.

- **Causes:** The main sources of acid deposition are:
  - burning of fossil fuels in power stations
  - exhaust fumes from motor vehicles.
- **Formation:** Acid rain consists of the deposition of pollutants, such as sulphur dioxide and nitrogen oxide, released by burning fossil fuels. These chemicals react with precipitation, mist and clouds in the atmosphere to produce wet deposits of sulphuric acid (H_2SO_4), nitric acid (HNO_3) and compounds of ammonia when it rains (Figure 11.54).
- **Environmental impacts:** Acid rain:
  - damages and kills trees, particularly conifers (causing yellowing of needles)
  - leaches toxic metals (aluminium) from soils which then accumulate in rivers/lakes
  - kills fish stocks and damages ecosystems
  - damages buildings and monuments by accelerating weathering, particularly of limestone
  - has harmful effects on human health causing respiratory complaints (for example, bronchitis).
- **Reducing impacts:** Various ways of reducing acid deposition exist:
  - use of catalytic converters on cars reduces NOx emissions from exhausts
  - burning fossil fuels with a lower sulphur content
  - transition from coal to cleaner sources of energy for electricity generation
  - use of 'scrubbers' and flue gas desulphurisation (FGD) equipment to remove sulphur from emissions at coal-fired power stations
  - energy conservation – reducing use of and therefore demand for electricity
  - 'liming' – adding lime to neutralise the acidity in acidic lakes – this is only a short-term measure and does not solve the problem (used in Norway and Sweden).

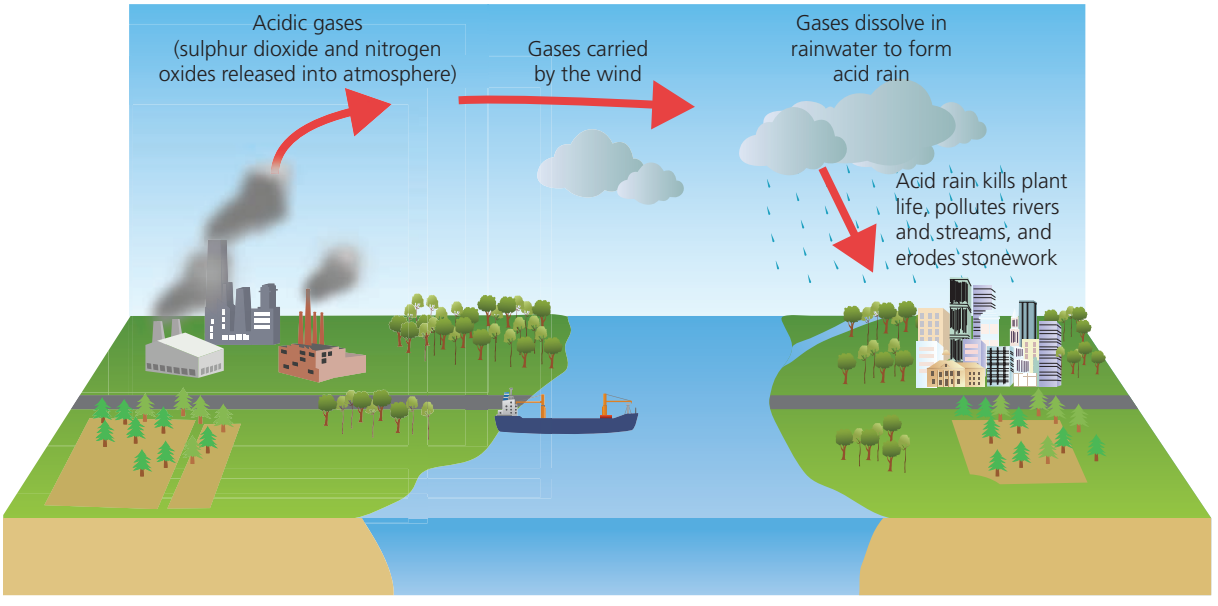


Figure 11.54 The process of acid rain formation

Enhanced greenhouse effect

The Earth's natural greenhouse effect means that solar insolation from the Sun is trapped in the lower atmosphere (troposphere) by greenhouse gases when it is re-radiated from the Earth's surface. This keeps temperatures at a level that makes the planet habitable. The rise in consumption of fossil fuels over the last two hundred years has increased the quantity of greenhouse gases in the atmosphere, for example, current CO_2 levels are 40 per cent higher than before the Industrial Revolution. As a result, the Earth's greenhouse effect is enhanced as more heat is trapped and global temperatures rise (Figure 11.55). Scientists think that global temperatures will rise on average by 1 to 2°C over the next century.

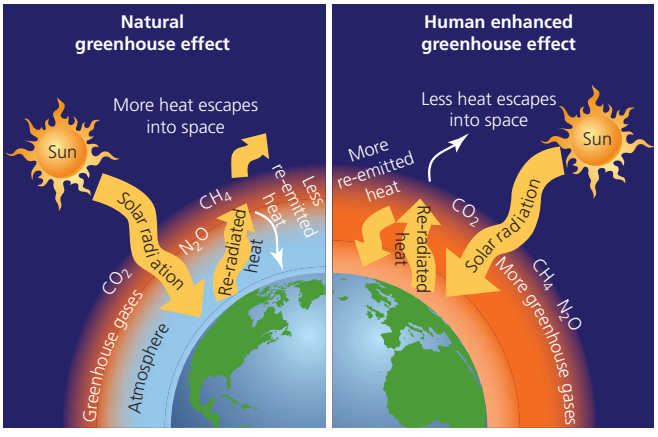


Figure 11.55 The natural and the enhanced greenhouse effect
Source: National Park Service, USA

There is a growing consensus that the burning of fossil fuels has released CO_2, methane and nitrous oxides into the atmosphere and this is a major factor causing climate change. The main environmental impacts and concerns for the future are:

- melting of polar land ice
- rising sea levels
- changing ecosystems/biomes
- extreme weather events (droughts, floods, cyclones).

Most observers agree that the less developed world is more susceptible to the effects of climate change. Poorer countries do not always have the infrastructure and resources to counter the effects. It is not uncommon for single weather events, such as tropical cyclones and floods, to kill thousands of people in regions such as South Asia, Southern China and Central America. In Africa the UN reports a 40–60 per cent decrease in total water in the large catchment basins of Niger, Lake Chad and Senegal. In a continent already struggling with poverty and famine, climate change is a matter of life and death.

CO_2 emissions continue to rise at between 2 to 3 per cent each year, though they slowed significantly to 0.4 per cent from 2013 to 2016, as a result of an economic slowdown in both China and the USA. Since 2017, emission rates started to climb again, growing by 2.6 per cent in 2018, mainly because of renewed industrial growth in China, increased emissions in India and an increase in the use of gas as a source of energy (Figure 11.56). The global COVID-19 pandemic in 2020 is predicted to cause a fall of CO_2 emissions of at least 2.5 billion tonnes, the largest fall in recorded history.

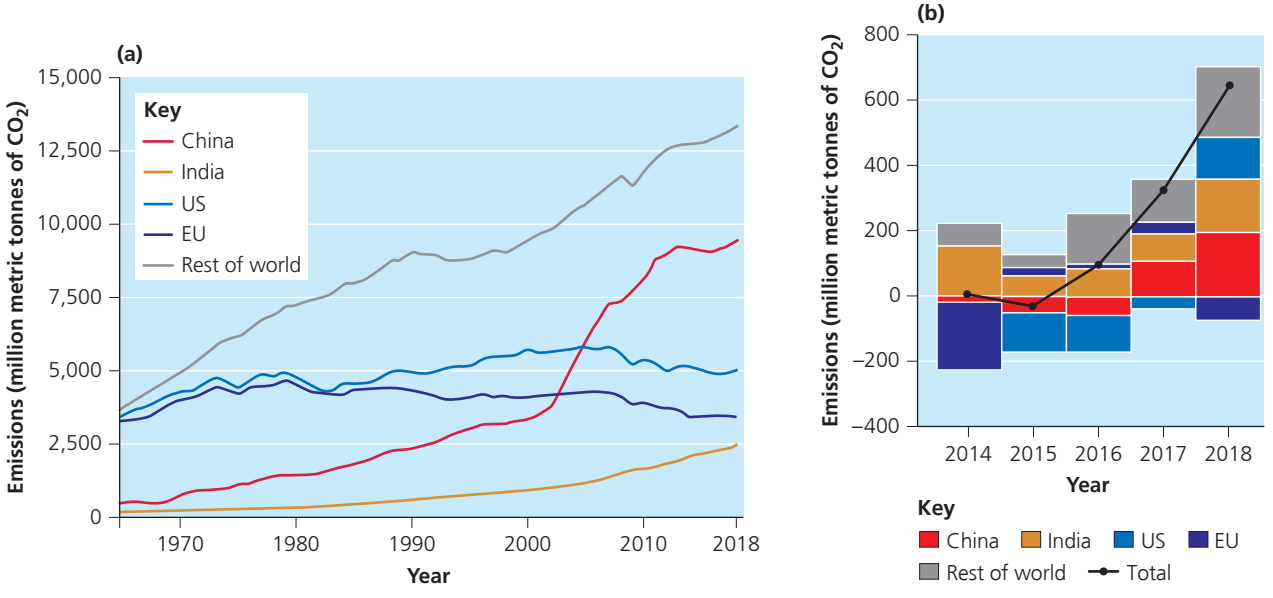


Figure 11.56 **(a)** Rate of increase of CO_2 emissions by region, 1965–2018 plus **(b)** annual changes, 2014–2018
Source: BP Statistical Review of World Energy 2019 and Carbon Brief analysis

Nuclear waste management

There are a number of concerns regarding the development of nuclear energy:

- risks associated with nuclear accidents
- disposal of radioactive waste
- health and safety concerns (particularly leukaemia and other cancers)
- potential contamination of water supplies
- concern over crops and grazing animals' safety.

Nuclear waste has a very long half-life (the measure of how long it takes to lose half its radioactivity). It remains highly radioactive for thousands of years, which must be considered when disposing of it safely. Nuclear power stations produce high-level radioactive waste in the form of spent fuel rods and fission products which have been removed from reactors. These are taken for nuclear re-processing, where reusable uranium and plutonium are separated out to leave unusable radioactive waste. This is vitrified into solid blocks of glass and stored in steel-clad or lead-lined glass containers underground.

In the UK, the Nuclear Decommissioning Authority (NDA) has responsibility for disposing of radioactive waste. The NDA gives site licences for the safe disposal and containment of low level waste (LLW) and intermediate level waste (ILW). There is a need for a geological disposal facility (GDF) site to dispose of high level waste (HLW) and they need to consider:

- **Physical conditions:** It needs to be buried 200–1,000 m deep in geologically stable rock types.
- **Economic factors:**
  - expensive to purchase land for this purpose
  - if located near to existing power stations the economy may become too dependent on the nuclear industry
  - negative impacts on tourism in the area.
- **Transport:** It requires safe transport links for the transport of waste from power stations.
- **Political and social barriers:**
  - strength of local pressure groups
  - concerns that the site may become the target of terrorism.

Review questions

15 Suggest reasons why the energy mix varies between different countries.

16 Describe the ideal physical geography conditions (climate, geology or drainage) for exploiting one primary energy source.

17 Outline the role of TNCs in energy production, processing and distribution.

18 Identify the key environmental impacts of the Athabasca tar sand development.

19 Explain how strategies used to manage energy consumption will improve energy sustainability.

11.7 Mineral security

Metals and industrial minerals represent a major category of non-renewable resources that humans extract from, and return to, the natural ecosystem. The rapid growth in metal and mineral use has serious implications for both the security of future resources and for the health of the environment.

Copper

Copper (Cu) is the third most-used metal in the world (after steel and aluminium). It is a base metal with important properties that give it value for use in a range of industries.

Sources of copper

Copper ores are found primarily in **igneous rocks** as hydrothermal deposits but also in **sedimentary rocks**. Copper mines are only developed where there is more than 5 kg of copper per tonne of rock (0.5 per cent of mass – the '**cut-off ore grade**') but ideally the content should be near to 2 per cent to establish a proven reserve.

Copper is found in many mineral ores which occur in deposits large enough to mine. The three main ores are bornite, **chalcopyrite** and malachite. Currently, the most common source of copper ore is the mineral chalcopyrite ($CuFeS_2$), which accounts for about 50 per cent of copper production.

According to the US Geological Survey (USGS) (2019), remaining discovered deposits of copper globally amount to 2.1 billion tonnes. It is estimated that there may be

a further 3.5 billion tonnes of undiscovered reserves including those available in deep sea nodules. Mineral rich nodules of copper, magnesium and other metals are formed as a product of deep-sea volcanic activity. Retrieving these nodules from the sea floor is, as yet, too expensive to allow this to be a major source of copper.

Copper ores are formed when geothermal solutions (which are hot and under pressure) bring copper, dissolved from deep underground, to cool near surface environments. When the solution reaches favourable temperature and pressure conditions, the dissolved minerals are released in veins and **disseminations** (spread deposits) within the rock. This is a process known as **mineral precipitation**. Copper is usually deposited as copper sulphide or sometimes as native copper metal (Cu).

Copper can also be deposited in sedimentary rocks such as sandstones and shales. This is the result of copper-bearing fluids moving through permeable rock strata. It is thought that the metals precipitated when the fluids reached a 'chemical trap', where the rock changed in a way that made it impossible for the metals to remain in solution. Examples of copper-bearing sandstones can be found in Zambia and the Democratic Republic of Congo.

Distribution of copper reserves and resources

Copper deposits found disseminated as crystalline intrusions in igneous rock (**porphyry** deposits) account for about 60 per cent of the world's copper resources. **Strata bound** deposits, in which copper is concentrated in layers in sedimentary rocks, account for about 20 per cent of the world's identified copper resources. Together, these deposits produced around 21 million tonnes of copper in 2019 (Figure 11.57).

South America has the largest measured and undiscovered copper resources. The world's largest igneous deposits are mined in this region. Since the 1990s, the Andean region of South America has emerged as the world's most productive copper region, producing 40 per cent of the world's copper in 2019. Chile is by far the largest producing country, although production has stabilised over the last few years, with Peru's production increasing rapidly. Six of the top ten largest mine developments are in South America (three in Chile, including the world's largest mine, Escondida, and three in Peru).

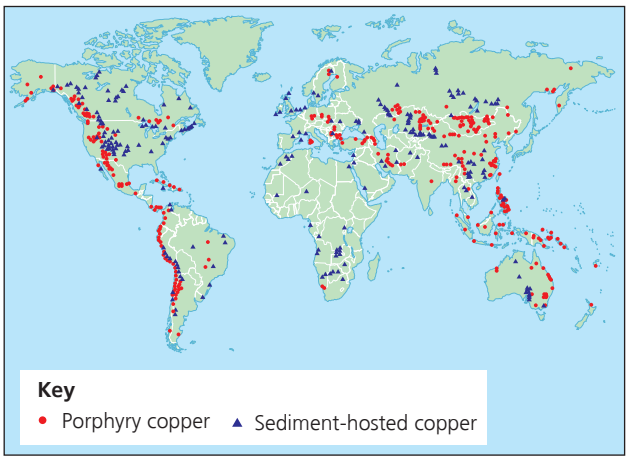


Figure 11.57 World distribution of known copper deposits. The red indicates copper associated with igneous rocks (porphyry deposits) and the blue indicates copper contained in sedimentary rocks (sediment-hosted copper).

Source: USGS

North America has reserves of porphyry copper. The Morenci mine in Arizona and the Cananea in Sonora, Mexico were respectively the sixth and eighth largest mines in the world in 2019. The majority of US copper production is from Arizona, Utah and New Mexico. In Europe, Russia and Poland are the largest producers (though much of Russia's is in Asia). The KGHM mining operations in Poland are underground rather than open-pit mines.

China, whose production has increased rapidly, is the world's largest copper consumer and third largest producer. Increasing production costs mean Chinese mining operations overseas are more economically appealing, especially in the African copper belt.

Table 11.21 World's current top ten leading copper producers

Rank	Country	2019 production (million tonnes)
1	Chile	5.60
2	Peru	2.40
3	China	1.60
4	USA	1.30
4	DR Congo	1.30
6	Australia	0.96
7	Zambia	0.79
8	Mexico	0.77
9	Russia	0.75
10	Kazakhstan	0.70

Table 11.22 Properties of copper and associated end uses

Properties	End uses
• Ductile (capable of being drawn into wire) • Strong and malleable (capable of being hammered and moulded) • A better conductor of electricity than any other metal except silver	**Electricity** • Its main uses are in the electricity supply industry and electrical and electronic products • It is used in power stations and substations and for electrical machinery such as electrical motors, electromagnets, generators and communication devices
• An excellent conductor of heat and can withstand extremes of temperature • Corrosion resistant • Easily joined by soldering or brazing	**Water, liquid vessels and pipework** • Its qualities make it extremely useful for plumbing pipework and for making sealed vessels used in water supply and in the brewing and distilling industries
• Biostatic (does not sustain bacterial growth) can be used to keep drinking water safe	**Hygiene** • Its antibacterial qualities slow down the growth of germs such as E. coli, MRSA and legionella • This is important for applications in food preparation, coinage, in hospitals and in plumbing
• Can be combined with other metals to make alloys	**Alloys** • It can be combined with zinc to make brass, with tin to make bronze and with nickel to make cupro-nickel • Alloys are harder, stronger and tougher than pure copper and are used in industrial applications
• Distinctive colour and chemical qualities	**Chemicals and agriculture** • It can be used in pigments and in agriculture, where it is used in pesticides and fungicides (although it has been largely replaced by synthetic chemicals)
• Decorative, with a sheen when polished	**Ornamental use** • It can be fashioned into ornamental objects and jewellery • It is resistant to tarnishing, making objects long-lasting
• Non-magnetic and non-sparking	• It is used in specialist machinery and tools and in military applications

Reserves in developing regions are becoming increasingly important sources of copper ores. In particular:

- **South East Asia archipelagos** have significantly large reserves with huge ore deposits in Indonesia, Papua New Guinea and the Philippines. The Grasberg mine in the Papua Highlands of Indonesia was the third largest copper producing mine in 2019.
- **Africa and the Middle East** have the world's largest accumulation of sediment-hosted copper reserves, with large deposits in the African 'copper belt' countries of the Democratic Republic of Congo and Zambia. Significant 'possible' copper resources remain to be discovered in Africa.

End uses of copper

Metals derive their usage from their physical properties and copper's most important properties make it popular for a range of end uses. These properties are listed in Table 11.22.

It is thought that the average UK home has around 180 kgs of copper for electrical wiring, plumbing and electrical appliances. Most copper used in wiring and plumbing remains in use for at least 50 years.

Components of demand for copper

The equipment and construction industries use the largest percentage of copper (Figure 11.58).

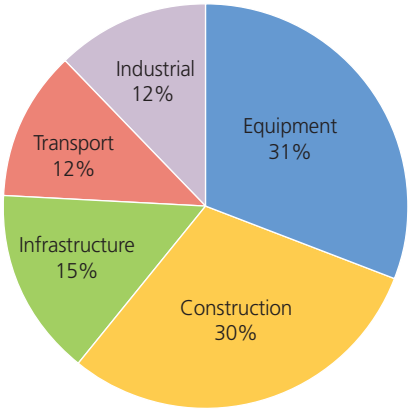


Figure 11.58 The major consumers of copper

- **Equipment**: This includes a range of capital equipment used in industry such as gears, bearings, tanks, pressure vessels and pipes. Its excellent heat transfer capabilities mean that it is widely used in producing heat exchange equipment. It is also used in power cables, transformers and other equipment for the electrical and renewable energy industries. This component includes its major use in motors to drive consumer electrical appliances.
- **Construction**: It is used in producing plumbing components such as taps, valves, pipes, fire sprinkler systems and other fittings in buildings. Brass is also used in these applications.
- **Infrastructure and communication**: Copper and copper alloys are a critical part of wires and cables used in the telecommunications industry. Fixed-line telephone cables and wires used for network connectivity contain copper although it is increasingly being substituted with fibre-optic cables. The $300 billion semi-conductor industry is also a major copper consumer.
- **Transport**: Copper contributes to several applications in the railway industry and in other forms of transportation. It is used for radiators, motors, brakes and bearings in the vehicle industry.

Role of copper in global commerce and industry

Links

There are links here with Chapter 7, Global systems and global governance – International trade and access to markets, pages 306–14, and Nature and role of TNCs, pages 314–17.

The world's production and consumption of copper have increased dramatically in the past 25 years. As large developing countries have entered the global market, demand for copper has increased. Due to its importance in construction and power transmission, the impact of any copper supply disruption would be high.

Freeport-McMoRan (US), Rio Tinto, Glencore, BHP Billiton and state-controlled Codelco (Chile) are among the top copper ore producing companies. Chinese mining companies are also major producers.

Trade:

Copper is traded internationally in various forms from raw material (ores), copper concentrates, semi-finished products (copper semis) through to refined copper.

- **Copper ores:** The main exporters are countries producing more than their domestic demand and the main importers are end users and/or those with a developed smelting industry.
- **Refined copper:** The main exporters are producers of ore with smelting operations (Chile) and the main importers are consumers using it in manufacturing industries (Table 11.23).
- Copper is virtually 100 per cent recyclable, so it is one of the most widely recycled of all metals; approximately 40 per cent of the world's demand is met by recycled copper.
- Recycled copper and its alloys can be used directly or further reprocessed to refined copper without losing any of the metal's chemical or physical properties.
- Copper is also traded as part of end-use products. This is a form of indirect copper trade, for example, importing electronic equipment also means the import of copper used in its production.

Production and consumption:

- Copper prices are the key driver for production; as they fall, producers prefer to 'leave the copper in the ground', hoping that supply shortages will raise prices.

Table 11.23 Main exporters and importers of copper ore and refined ore, 2017

Copper ore			
Main ore exporters (% of global exports)		Main ore importers (% of global imports)	
Chile	(28%)	China	(43%)
Peru	(21%)	Japan	(14%)
Australia	(7.2%)	India	(7.6%)
Indonesia	(5.4%)	South Korea	(6.8%)

Refined copper			
Main copper exporters (% of global exports)		Main copper importers (% of global imports)	
Chile	(25%)	China	(32%)
Russia	(6.1%)	USA	(8.8%)
Japan	(5.6%)	Italy	(6.8%)
Kazakhstan	(5.1%)	Germany	(5.8%)

Source: OEC

- The diversified nature of copper's end usage in electronic products, construction and vehicles makes prices a good indicator of an economy's overall health.
- Consumption across a range of markets means analysts see copper prices as a reflection of the global economy's health so it has become known as 'Dr Copper'. (Steel and aluminium have more concentrated end usage in the construction and vehicle industries, so their prices are regarded as less good indicators.)

Figure 11.59 shows the movement in copper prices from 2016–2019. Prices tend to reflect the performance of the Chinese economy. From 2015–2017, prices were generally low during China's economic slowdown. Recovery in 2018 saw prices rise again and they had been expected to reach $7,000 per metric tonne in 2020 but the COVID-19 pandemic caused a crash in the Chinese and global economies that was reflected in falling prices for all commodities.

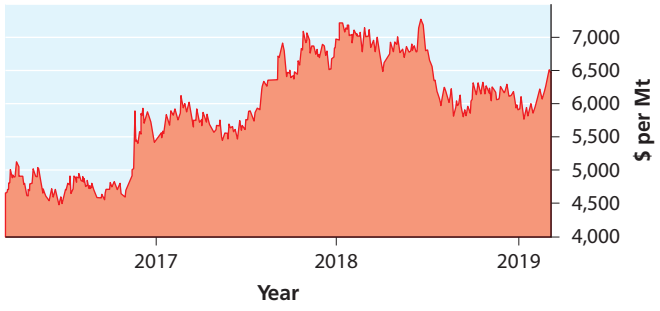


Figure 11.59 Fluctuating world copper prices reflecting the health of the global economy

Source: Al Farid Wood and Metal Trading Company

Key aspects of physical geography associated with mineral ore occurrence

If all the mineral resources in the Earth's crust were evenly distributed then none of them would be sufficiently concentrated to become viable reserves for exploitation. Geological processes have provided localised concentrations of deposits to enable exploitation to occur.

Geological conditions

Based on the process of geological formation, mineral deposits may be classified into four broad categories:

1. **Magmatic deposits**: Named because they are linked with magma which is emplaced into the crust (either continental or oceanic) and are found within rock types derived from the crystallisation of such magmas. The process of crystallisation separates ore and non-ore minerals according to their crystallisation temperature. These deposits also include ores containing nickel-copper deposits and platinum metals.
2. **Hydrothermal deposits**: These are formed by another igneous process associated with magma intrusions. The rocks around intrusions may be lifted and deformed so that cracks and fissures appear. These cracks allow hot solutions containing dissolved minerals to escape towards the surface and as these cool down, the minerals precipitate in a predictable order according to their solubility. They are found in concentrated deposits that are more viable for exploitation. Most metal ores, including iron, copper, tin and lead are found as hydrothermal deposits.
3. **Metamorphogenic deposits**: Metamorphic processes change pre-existing mineral deposits and form new ones. The changes are brought about by intense heat and pressure over a long period of time. For example, low density minerals are replaced by minerals of high volumetric mass. These are found primarily in PreCambrian formations (for example, the iron ore deposits in Ukraine; manganese deposits in Brazil and India and the gold and uranium ores in South Africa).
4. **Sedimentary deposits** (including **placer deposits**): Not all metal deposits have igneous origins – sedimentary deposits can also be a valuable source of many metals, including copper. Sedimentary deposits are generally tabular and sheet-like, horizontal in form but frequently folded and faulted. Beds containing metallic ore are generally less than 20 ft thick. Beyond copper, other metals found in sedimentary deposits are lead and zinc. **Placer** deposits are alluvial; they are carried in flowing water and deposited when the water slows down. Dense minerals are the first to settle, including tin ore and gold.

Location and working

As mineral deposits can be found in all types of geological formations, traditionally the major **physical** geographical constraints on exploration and mine development were determined by remoteness, access to processing and markets and ability to exploit the resource. Thus areas with extremes of temperature or physical environments such as deserts, densely forested areas or high mountains were largely avoided and left unexplored.

Advances in mechanical and transport technology have released these constraints and geologists can feel free to explore suitable geological formations on any

continent, with the exception of Antarctica. Trends in mine production have shifted to more remote areas such as the Atacama Desert in Chile (copper), under rainforests in Brazil (iron ore) and other locations in the continental interiors of Australia, Africa and Asia. The last remaining frontier for exploration and development is the hypothetical and speculative resources that lie in the deep seabed.

Links

There are links with Chapter 7, Global systems and governance – The 'global commons', page 333, and Threats to (pages 335–40) and Governance of (pages 340–3) Antarctica.

Physical conditions will, however, often determine the nature of the mine working itself. The shape, size, quantity and grade of the ore deposit found from exploration will define whether it is extracted through traditional deep underground mining or through open-pit (sometimes known as **open-cast**) mines.

Ore extraction: underground and open-pit mining

Mine working has undergone important changes during the twentieth century, with a major shift from underground to open-pit mining (Figure 11.60), which has become more common as mining has evolved in emerging economies. Productivity increases in the past 50 years have been achieved through the ability to access lower grade ores by more efficient processing and using large-scale capital equipment. The main cost advantage of open-pit mining is that mechanisation has allowed the rapid removal of vast amounts of material using larger digging machines and trucks. Equipment is not restricted by the size of the opening it must work in, which allows faster production and the lower cost also permits lower grades of ore to be mined.

Extraction costs will depend on the size, shape and location of the reserve. These include:

- **Overburden**: Rock and soil lying above the deposit can be expensive to remove. It is blasted away in open-pit mines. An open-pit mine is the least expensive, the first choice when a deposit is close to the surface, big enough and has little overburden.
- **Water removal**: As depth increases, more water flows into the mine; this needs to be pumped out.
- **Depth**: Costs rise as depth increases; this applies in both types of mine. Underground mining includes a wide variety of added requirements such as ventilation systems, ground support and finding a safe and cost-effective way to get the mineral ore to the surface.
- **Form of the deposit**: Costs increase if the mineral is found dispersed irregularly or found in thin layers. This is less costly and safer to access in open-pit mines.

Costs have to be balanced against what can be extracted from the mine. This will depend on:

- **The 'stripping ratio'**: This is the amount of waste rock mined relative to the amount of ore mined. A stripping ratio of 3:1 means that there will be three times as much waste rock mined as ore. To be profitable, an open-pit mine must be designed so that the cost of mining the waste rock does not exceed the value of the ore.
- **The 'cut-off ore grade'**: Processing costs depend on the purity of the ore. The **cut-off ore grade** is the lowest ore purity that can be exploited economically. This varies depending on the metal or mineral being processed.

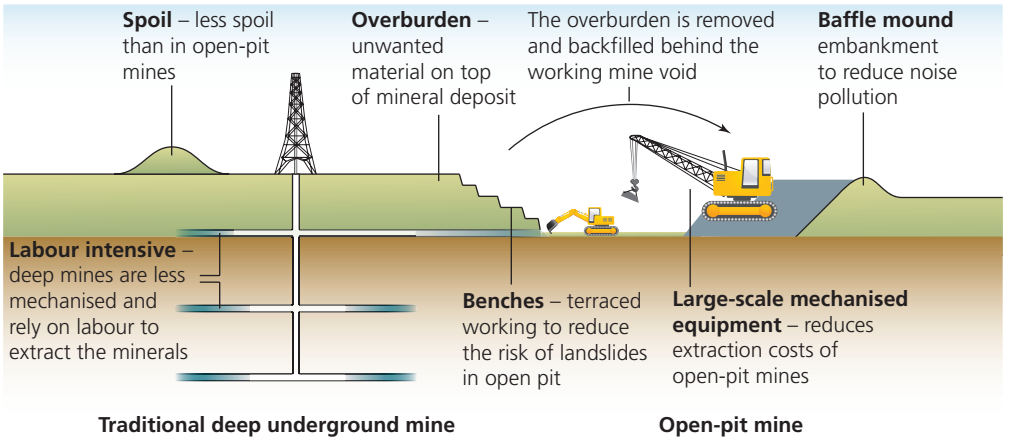


Figure 11.60 Features of traditional deep underground mine and open-pit mine

Environmental impacts of the Carajás iron ore extraction complex in Brazil

The environmental impacts of mining include:

- formation of sinkholes or surface subsidence (following the collapse of the surface layer)
- loss of biodiversity
- contamination of soils, surface water and groundwater by toxic chemicals or leachates
- significant public health effects.

Background to the Carajás scheme

The Carajás complex (Figure 11.61) is the largest single iron ore mine in the world. It is located in the state of Pará in northern Brazil, in an area naturally covered by the Amazonian rainforest.

The mine complex boasts the richest reserves and concentrations of high grade iron ore anywhere in the world, with an iron content of around 65 per cent. It is owned by Vale (CVRD), a Brazilian mining TNC, and holds 18 billion metric tonnes of iron ore in proven reserves. Carajás is not only rich in iron ore but also ores for copper, tin, aluminium, manganese and gold. The rich ores are found within Archaean iron formations to a depth of between 100 m and 150 m.

The operation uses an open-pit mining system with several giant man-made chasms. Around 3,000 workers are employed at the mine using a fleet of giant machines, such as diggers, crushers and loading trucks. The complex produces over 550,000 tonnes of iron ore per day. Around 70 per cent of this is exported as ore to China using a dedicated railway line to the Ponta da Madeira deep sea terminal near São Luís in the north-eastern state of Maranhão. Carajás is also served by its own airport. Vale was given the go-ahead to expand production at the complex by developing the giant S11D mine at Serra Sul to the south-west of the existing development. Mining commenced at S11D in 2016 and it currently contributes around 70 million tonnes of the 200 million tonnes per year mined at the Carajas complex.

Figure 11.61 The Carajás open-cast iron ore mine complex

Environmental impacts

The biggest environmental impact of the mine development in the Carajás region has been the necessity to clear large tracts of pristine rainforest to make way for the massive open-pit mines, leading to landscape and aesthetic degradation. Further clearance of rainforest has been necessary for associated infrastructure:

- the railway to the seaport which is an integrated part of the scheme
- road access for large trucks and to convey other large machinery
- the airport
- mining settlements – accommodation and services for the miners
- temporary storage of overburden (debris and soil).

Deforestation has resulted in both direct and indirect impacts on landscape, wildlife habitats, the climate, soils, relief, vegetation and drainage. Furthermore, the access into the rainforest enabled by roads has led to further forest clearance by non-related activities such as clearance for agriculture or illegal logging.

Landscape, wildlife and ecosystems

The mine is in an area of national forest but what was a landscape of dense vegetation is now a scene of bare cliffs and billowing dust. Noise pollution is another feature of the devastation, with the sounds of birds and animals in the rainforest replaced by the constant roar of massive engines and the grinding of diggers and crushing machinery.

Habitat destruction has led to endangered species being threatened with extinction. It is estimated that around 25 per cent of wildlife in the region has been lost as a direct result of the project. The development of railways and roads has fragmented habitats, reducing migration of some fauna and creating smaller gene pools, resulting in an overall loss of biodiversity.

Climate, soils and vegetation

The loss of rainforest over such a large area has had an impact on microclimate. Trees are a regulator of both temperature and humidity and their loss means that there is less rainfall, less humidity and a higher diurnal temperature range, with extremes of heat baking the red ferralitic soils. Without tree cover, the nutrient cycle is broken, the soils lose any organic input or fertility and are prone to severe erosion. It is difficult for vegetation to regenerate naturally in such harsh conditions.

Relief and drainage

The open-pit mining alters the relief and drainage of the area directly. Large masses of soil can be eroded because of storm events, which can cause landslides around the mine and the silt being washed into local rivers, which can kill some aquatic species and causes local flooding.

Water pollution

In addition to excessive siltation of rivers, leakages and leachates from tailing ponds release toxic material into surface water courses and groundwater sources.

Global environmental impacts

The Carajás project contributes to global climate change by a net addition to greenhouse gases. In addition to burning fuel to operate the trucks and large machines on site, deforestation in the area removes a significant carbon sink.

The sprays and chemicals used on the site contribute to ozone depletion and the emissions from vehicles and from waste materials will also contribute to acid rain.

Expansion and remediation efforts

Vale states that its expansion of mining at the Serra Sul (S11D) operation is good for the Brazilian economy, and that in all their operations they adopt 'green' measures in which operations are designed for minimum environmental impact. They have installed a truck-free, electrical, automated transport system to help move rock and ore around the site more efficiently and safely. Although one of the purposes is to remove fuel costs, it will also reduce emissions.

Like many transnational corporations, Vale is keen to point out that sustainability is embedded in its planning. Before it starts any new digging, it has to present a restoration plan to the state authorities and the national conservation agency. Plans aim to restore the area to its original state by using spoil to fill in the mines and reshape the topography, replanting the area with the original native species, which the company have preserved and which are growing in a separate nursery.

Vale also supports an extensive monitoring operation in the forest; they have paid for 80 additional forest rangers plus vehicles, boats and the use of a helicopter to guard against poaching of animals and illegal logging.

The company make an effort to operate the mine more sustainably, however they are a profit-making corporation and their priority is iron ore extraction; the environment may only be a secondary concern.

Sustainability issues associated with ore extraction, trade and processing

The biggest issue for mineral ores is their finite availability, which raises their value as commodities. Advances in technology can improve the supply and extend the lifespan of many metals by:

- more cost-effective exploration for new deposits
- advances in machinery and extraction methods which can convert more resources into reserves
- improving resource efficiency through, for example, better recycling methods
- substitution with the use of alternative materials can extend the lifespan of many resources.

Despite this, there is still concern about the long-term sustainability of our use of many metals. How we manage their supply and demand in the future is of critical importance.

Ore extraction

In areas of wilderness, mining may cause destruction and disturbance of ecosystems; in areas of farming, it may destroy productive arable or grazing land. In urban environments, threats to sustainability come mainly from noise and visual pollution.

Unsustainable practices of mineral ore extraction lead to a number of negative impacts. These are shown in Table 11.24 (page 624), together with measures to mitigate the impacts.

Despite mining companies being more aware of the potential for environmental impacts, there have been a number of recent environmental disasters resulting from tailing dam failures.

Tailing dam failures

Tailing dams are constructed at ore extraction sites to contain tailings in reservoirs. **Tailings** consist of a

Table 11.24 Impacts of unsustainable practices and mitigation measures

Impacts	Mitigation measures
Land take – land area needed is usually much larger than the mine itself whether underground or open-pit mining **Habitat loss** – results in loss of species	Restoration plans: using native species by capturing animals and releasing them or relocating them; habitat restoration using native plant species, (e.g. the Carajas project); creating new habitats
Pollution: **Noise pollution** – machinery disturbs wildlife **Dust pollution** – caused by blasting and vehicle movement	Baffle mounds around the periphery of open-pit mines help to absorb and deflect some noise Dust can be reduced with water sprays
Water turbidity – caused by particles and silt from mines; can block sunlight from aquatic plants and choke filter-feeding animals	Holding lagoons can help suspended solids to settle and clear water can then enter rivers
Toxic leachates – chemicals or toxic metals unearthed as part of mining can become soluble; drainage into nearby rivers kills aquatic life	Toxic metals are more soluble when acidic so passing mine drainage water through a filter of crushed limestone may immobilise some toxic metals
Spoil disposal – is often loosely compacted and unstable with potential for landslides	Drainage using pipes in the base of spoil heaps prevents waterlogging and reduces landslide risk

semi-liquid slurry mix of water, ground-up waste rock and sand, as well as toxic chemical reagents used to extract minerals from the ore. The dams themselves are constructed from rock and earth waste and are prone to failure because of overspill and collapse, which can release toxic mud into channels and water courses downhill from the site. For example:

- In August 2014, 40,000 m^3 of copper sulphate acid spilt into the Sonora river and public waterways in Sonora, north-west Mexico. Heavy rainfall caused the tailings pond to overflow at the large Buenavista copper mine operated by Grupo Mexico. The massive spill poisoned a 60-kilometre stretch of the river and caused water shortages for 20,000 in an area already experiencing water stress.
- In January 2019, a catastrophic tailings dam failure occurred at Vale's Córrego do Feijão iron ore mine near Brumadinho in Minas Gerais state in south-east Brazil. The 86-metre high dam constructed of sand and dried mud dissolved into liquid, releasing a giant mudflow that engulfed the company's offices and small settlements and farms downstream.
  - 270 people were killed as the mudflow covered vast areas of land, half of which was protected natural forest
  - 300 km of the surrounding river network was polluted, leaving thousands of dead fish in the Paraopeba river.

  It was described as the biggest environmental disaster in Brazil's history. The disaster occurred less than four years after a similar tailings dam failure at the nearby Samarco mine (operated as a joint venture by BHP Billiton and Vale), which killed 17 people (Figure 11.62). As a result of the Brumadinho disaster, Vale's chief executive resigned and a large number of employees deemed culpable were arrested.

Figure 11.62 The Samarco mine disaster in 2015. The city of Mariana in south-east Brazil, was flooded when an iron ore tailings dam failed, killing 17 people.

Trade

The comparative advantages of different countries around the world can foster the development of a mutually beneficial and more sustainable mineral resources trade. It can also mean that ore-producing nations are subject to the 'resource curse'. Encouraging these nations to establish processing industries to add value to their exports would help to prevent this.

In 2015, the World Economic Forum launched the **Mining and Metals Scenarios to 2050** project to examine the future of the minerals and metals sector.

A key element of the report is to advocate a strong transition towards circularity and recycling of metals in order to sustain trade. There are a number of reasons for this:

- There is increased demand and population growth in a world of limited resources

- Volumes of primary extracted materials are unlikely to grow in line with demand for metals.
- Metal trading companies provide the link between commodity producers and end-user industries and will become providers of recycled materials.
- More recycling also means more-efficient energy use. For example, recycling aluminium takes only five to ten per cent of the energy needed to produce primary aluminium.

The move to more recycling and circularity in the minerals and metals trade is expected but it still needs action from manufacturers, processors and governments (through regulations and incentives) to support it.

The World Economic Forum also introduced the concept of a **Green Trade Alliance (GTA)** in which environmental sustainability is fostered without compromising competitiveness. The idea would be supported by a 'Sustainable Trade Organisation' overseeing GTA agreements. Environmental standards for both recycled and primary materials would be the basis for tariffs adopted by GTA members. Trade between GTA and non-GTA members would be limited initially but it would encourage non-GTA members to adopt the same sustainability standards to enable them to join the alliance.

Processing

Metal and mineral commodities are non-renewable and their total availability is finite. However, processes and mechanisms that enable recycling or reuse of these materials increase their long-term sustainability and contribute to the production of metals. Recycling sources include old scrap materials from discarded products and waste material from industrial processes. These contribute to the 'end-of-life' (EOL) stocks of metals, most of which can be used in new materials. Recycling procedures have been established in many countries as a result of the Agenda 21 commitment. For example, the collection of lead, used predominantly in lead acid batteries, is particularly well organised.

A United Nations Environmental Programme (UNEP) status report published in 2011 estimated the end-of-life recycling rates (EOL-RR) for the following metals:

- **Iron and steel**: 70–90 per cent is recycled, which contributes to around 35 per cent of new production.
- **Aluminium**: 40–60 per cent is recycled, contributing around 24 per cent towards new production.
- **Copper**: 40–50 per cent is recycled, contributing around 18 per cent of new production.
- **Lead**: Recycling estimates vary between 60–90 per cent but this contributes to around 55 per cent of new production.
- **Tin**: Around 90 per cent is recycled.

These figures show that although secondary resources can make a significant and potentially increased contribution to production, there is still the need to extract new raw materials.

Increasing the use of secondary resources as a source of production for new materials represents a considerable challenge because:

- A large proportion of metals end up in the built infrastructure where they will remain long term.
- Some metals end up in complex products combined with other materials that inhibit recycling.

Table 11.25 Factors influencing recycling rates

Factors that make recycling more challenging	Factors that assist recycling
Mixed materials – for example, the separation of alloys	Media campaigns – to educate and raise awareness of the advantages of recycling
Dilution and dispersal – some uses are so dispersed that it is unviable to recover	Design for 'end of life' – easy dismantling to remove useful parts at point of discard
Transport – cost of transporting scrap material may make it unviable	Labelling when assembled – to state what goods are made of
Labour costs – recycling is much more labour intensive than raw material extraction	Reusing used parts – removing parts which are still of value
	Legislation – for example, EU Directives on disposal of batteries and electronic equipment
	Financial incentives – for example, landfill tax

Review questions

20 With reference to Figure 11.57 (page 617), describe the global distribution of porphyry copper reserves.

21 Explain the difference between the 'cut-off ore grade' and the 'stripping ratio' and how they affect the extraction of mineral ores.

22 Explain why the development of processing industries in ore-producing nations encourages more sustainable trade.

23 Suggest how the recycling and re-use of metals improves the sustainability of mineral ores.

11.8 Resource futures

The future of energy resources

Global energy consumption will continue to increase. Forecasting this is difficult as a range of technological, economic, environmental and political factors will affect both energy growth and mix in the future.

Different scenarios have been modelled by international energy organisations such as the World Energy Council (WEC) and the International Energy Authority (IEA). They recognise the energy transition taking place but the pace of this transition is difficult to gauge.

Whatever the future brings, a few points seem to be accepted:

- By 2050, non-carbon emitting sources will make up around 40 per cent of the global energy mix.
- Coal could remain important with increased use of clean carbon technologies such as CCS.
- Oil is likely to decline as its necessity for transport diminishes.
- Average energy use per capita will decline in developed countries as individuals adopt greater energy efficiency; this will be offset by rising per capita usage elsewhere.

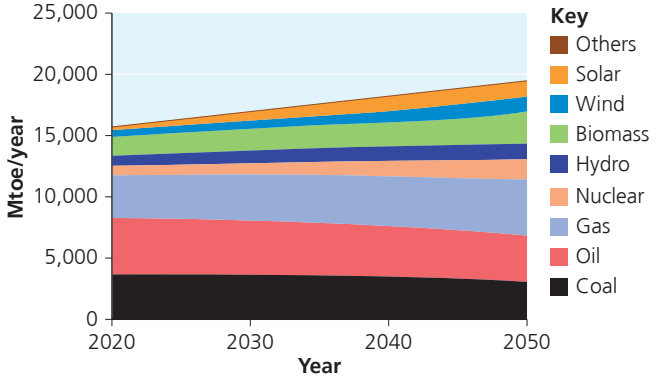


Figure 11.63 A forecast for energy consumption and energy mix from 2020 to 2050

Source: Schalk Cloete

Technology

With the problems associated with fossil fuels, new technologies are being developed to replace them:

- **Hydrogen:** Unlike many renewables, hydrogen is a high energy-density fuel and can be used to replace fossil fuels as the technology for conversion is available. The '**hydrogen economy**' is a system of delivering energy using hydrogen in different forms. The main uses being developed are:
  - to produce and store surplus energy in fuel cells by **electrolysis of water**; converting H_2O into H_2 and O releases energy
  - combustion – used as vehicle fuel, or for domestic/industrial heating, though safety measures need to be developed
  - used as chemical energy – to make hydrocarbons – which can then be used as fuel in much the same way as oil and gas.

  Hydrogen fuel cells could be the perfect alternative to fossil fuels. They generate electricity using only hydrogen and oxygen and are pollution-free. The following issues hinder further breakthroughs in this important technology:
  - Electrolysis requires a lot of other energy use and if fossil fuels are used in the process, then the environmental benefit of using hydrogen is negated. Using renewable energy for the process would overcome this issue.
  - Hydrogen cannot be compressed easily or safely, and requires large tanks for storage.

  Hydrogen may provide the key for both green (renewable) and blue (gases with CCS) energy transition, particularly if renewable energy is used in the electrolysis process, for example, geothermal energy in Iceland.

Table 11.26 Hydrogen fuels – advantages and disadvantages

Advantages	Disadvantages
• Readily available basic earth element • No emissions • Non-toxic (environmentally friendly) • Powerful and efficient energy source • Renewable	• Expensive to develop • Difficult to store • Needs a lot of energy to separate it from oxygen to create fuel cells • Need to replace existing infrastructure • Highly combustible

- **Gasification of coal:** Coal is converted into gas below ground where the deposits are deep, thin or difficult to take out as solid fuel. Bioenergy gasification can be used in conjunction with carbon capture and storage to chemically convert the coal into a synthetic natural gas (SNG).
- **Microbial enhanced oil extraction:** Secondary and tertiary extraction is already completed using steam and detergents, but using bacteria to digest heavy oil

produces lighter, easier flowing oil, enabling greater extraction rates.

- **Nuclear:** The nuclear power industry has been working on safer technology solutions since the Fukushima nuclear disaster in 2011. Ideas include:
  - smaller **modular** reactors more evenly distributed to serve smaller populations, with safety mechanisms built in to eliminate the chance of radiation release
  - **offshore nuclear reactors** – like the floating platforms the oil and gas industry already uses; constructed to withstand hurricanes and with reactor cores submerged beneath the platform, so a fresh supply of cold seawater would always be available to cool the reactor
  - **nuclear fusion** – promises limitless carbon-free power without producing dangerous nuclear waste. Trying to replicate the same atomic reaction that occurs in the Sun, in a controlled way, is still many years from maturity.
- **Offshore wind:** Promises to gain massive growth in the next 25 years as advances in turbine blade designs, developed from aeronautics technology, maximise energy from wind.
- **Geothermal energy:** Researchers in Iceland have spent several years drilling straight into volcanoes to access very hot water and magma deposits, with a view toward developing these high-temperature resources into geothermal power stations that can produce ten times more energy.
- **Solar fuels:** Converting sunlight, water and carbon dioxide into usable chemical energy that can be stored like gasoline for extended periods of time.
- **Ocean thermal energy conversion (OTEC):** Aims to generate electricity by exploiting the temperature differences between surface water, heated by the Sun, and water in the oceans' chilly depths.

Economic

- Energy demand will grow only slowly in developed OECD countries but rise rapidly in emerging economies.
- Energy trade flows towards Asian markets will continue to gather pace. Issues concerning the trade in natural gas could be eased by the increased availability of LNG.
- The pace of unconventional oil production in non-OPEC countries will start to lose momentum in the mid to late 2020s. The oil economy will again rely on supplies from Middle East OPEC nations plus Russia.
- Natural gas production will increase in every region except in Europe. Unconventional gas will account for 60 per cent of the growth in production, though shale gas outputs are likely to decline by the 2030s.
- Despite a rapidly growing economy, sub-Saharan Africa will only account for 4 per cent of the growth in global energy demand.

Environmental

- **Carbon capture and sequestration:** Although not a power generation technology, carbon capture and sequestration enables the scrubbing of carbon from power plants for storage in underground reservoirs such as exhausted oil and gas traps. This means that carbon emissions from fossil fuels and biomass can be eliminated so they can be used with less environmental impact.
- Targets for carbon emissions will become more stringent as global climate change presents more problems around the world. The Paris Climate Agreement has encouraged many governments to commit to drastically reducing emissions. For example, the UK is committed by law to net zero carbon by 2050.
- On a national scale, there is likely to be a move to more local energy generation from a combination of smaller power stations, combined heat and power plants and heat networks. There will be some energy efficiency gains and environmental benefits from this.

Political

Geopolitics will continue to be an important feature of the energy industry. A number of trends seem likely to emerge over the next thirty years until 2050.

- The influence of TNCs in resource development will continue. They will have the wealth to control the technologies of the future. However, global governance policies will attempt to dictate the pace of change on energy transition.
- Developing countries will increasingly rely on TNCs to help develop their resources. The corporations

will have a monopoly on the skills and capital needed to do this effectively.

- OPEC may re-emerge as a powerful cartel in influencing supply and pricing of both oil and gas.
- As traditional fossil fuel resources become depleted and remaining reserves more valuable, there will be a scramble to claim previously protected areas and exploit them. The Arctic is less well protected against development than Antarctica and there have been a number of territorial claims by the USA, Russia, Canada and Denmark in the race for fossil fuels in the Arctic Ocean.
- Geopolitical relationships between countries will be influenced by the rate of transition to:
  - energy equity
  - national energy security
  - environmental sustainability.

The future of water resources

The UN predicts that water shortages could affect five billion people by 2050. Water availability for the future rests upon good management of the water the world has now.

Technology

Much of the developing technology to increase the supply is focused on the largest source of water – the oceans, and the use of desalination techniques:

- **Osmotic distillation** and advanced membrane technology may, in future, require less energy intensity and thus become more sustainable.
- **Electrodialysis** is an alternative method to desalinate seawater that is being considered. This works in the opposite way to reverse osmosis by allowing sodium and chlorine ions to pass through a membrane in the presence of an electric field, leaving purified water on the other side.
- **Saltwater greenhouse technology** is well suited to arid parts of the world; it produces water for irrigation by using the cooling and humidifying power of water vapour as the seawater, which is piped into the greenhouse, evaporates over a large honeycomb surface.
- Smaller-scale **appropriate technologies** such as plastic solar stills (marketed as 'watercones') are more appropriate for distilling salt water in smaller quantities in less developed countries.

Economic

There are many countries with an abundance of water and redistribution can occur in various ways, including water shipping, water management and virtual water trade. Trading water on a larger scale will take place as water becomes more valuable than oil in some regions.

Environmental

Any action to manage water resources more sustainably will need to be co-ordinated with efforts to mitigate the effects of climate change – they cannot be considered separately. In the future, an integrated river basin management approach will be undertaken by planners, scientists and hydrological engineers.

- Impacts on water quality and drainage are key considerations when environmental impact assessments are conducted.
- Future-proofing river basin management strategies deals with these challenges.

In developed countries, the impacts of water-source development are already considered carefully before implementation goes ahead and it is thought that developing countries will follow this example.

Political

Water still presents a likely cause of international conflict in the future. The potential hotspots of conflict will need to be monitored carefully and co-operation encouraged.

The future management of global water resources features strongly in the UN's Sustainable Development Goals (SDGs).

- Proposed Goal 6 is to 'ensure availability and sustainable management of water and sanitation for all' (https://sustainabledevelopment.un.org/sdg6).
- Sustainability of water resources is also relevant to:
  - Goal 3.3 relating to diseases and epidemics (including water-borne diseases) and Goal 3.9, which relates to reducing death from water contamination
  - Goal 12, which relates to the safe management of chemicals and wastes, and reduction of their release into water
  - Goal 15, which includes targets that focus on the role of water in wildlife conservation.

The future of mineral ores

Concern surrounding the future security of mineral supplies continues. New sources of minerals are needed to meet future demand. This is particularly so for the so-called 'critical' minerals, which are those elements that are essential for new and green technologies. To ensure that future shortages do not cause serious problems, new approaches will need to be developed.

Technological future

Exploration

Exploration holds the key to the future supply of mineral ores. A range of techniques continues to be developed to **explore** new mineral deposits more cost effectively:

- **Remote sensing**: This includes techniques that collect information about mineral deposits from a distance. The most commonly used are aerial and satellite surveys, which enable large areas to be surveyed to identify topography and geological features. Infrared images may also indicate possible mineral content in surface rocks.
- **Magnetometry**: This can be used to identify magnetic rock content, such as iron ore.
- **Gravimetry**: This is used to detect the density of rocks and distinguish magmatic igneous deposits, which are usually denser than sedimentary and placer deposits.
- **Seismic surveys**: Echoes of surface vibrations are used to provide information on depth, angle and thickness of rock strata; increasingly used in underwater surveys.

Increasing deep sea exploration will be undertaken using a combination of these techniques.

Extraction and exploitation

- **Mechanisation**: Greater efficiency is found in many aspects of mining equipment. Larger excavators can extract material more rapidly and cost-effectively.
- **Electronic technologies**: The introduction of computer technology, remote-control interfaces, satellite communications and robotics has led to greater safety and productivity in mining, mineral processing, smelting and refining operations.
- **Exploiting low-grade ores**: Technological developments have made it possible to mine ores of declining grades and more complex mineralogy without increasing costs. Alternative methods have been developed to extract metals from low-grade ores:
  - Copper in leachate water from spoil heaps can be concentrated by evaporation and then separated by electrolysis.
  - Copper from low-grade copper sulphide spoil heaps can be extracted using bacteria. The sulphur is oxidised producing an acid solution that leaches out the copper, which can then be purified.

Economic future

Supply

- Mineral resources are abundant but reserves of exploitable minerals are limited. High extraction costs or a lack of suitable technology may prevent 'possible' reserves being exploited.
- Prices for the staple ores of iron, copper and nickel have been volatile but generally remain low, which will discourage extensive exploration in the near future.
- A greater proportion of recycled materials will enter the supply chain to compensate for any falls in supply of primary ores.

Demand

Future demand for mineral ores will depend on the balance between a number of factors (Table 11.27).

Table 11.27 Factors affecting the future of mineral ores

Factors causing an increase in demand	Factors causing a fall in demand
• Growing populations in developing economies • Aspirations for a better standard of living from growing affluence in emerging economies	• Increased and cheaper recycling means more reserves are left in the ground • Substitution of metals with other materials • Effects of global recession following the global COVID-19 pandemic

Environmental and political

Environmental and political developments for extracting mineral ores will be increasingly and intrinsically linked in the future.

- Exploration and development of new mines puts pressure on the natural environment.
- Protection for the environment can only come from governmental controls, at all levels, from local to global.
- EIAs may mitigate some of the worst devastation brought about by mineral exploitation; not all regions and countries will impose such restrictions on operators.
- Ore production will continue to grow in emerging economies; investments seen in Latin America, Africa and parts of Asia in recent years may escalate further over the next ten to twenty years.
- Developing nations see economic gains from foreign direct investment as an opportunity to provide wealth, jobs and improve social and physical infrastructure. In many cases, this may obscure the concerns over environmental impacts.

Globally, some sites with large mineral deposits will remain unexplored. These include legally protected areas and those that are expensive to develop because of hostile environments or infrastructure difficulties. Examples include Antarctica, the Arctic region (including Alaska, Greenland and Siberia) and the deep seas.

- Antarctica is protected from all forms of commercial mineral exploitation by the Madrid Protocol on Environmental Protection to the Antarctic Treaty.
- The right to exploit deep ocean floor deposits in international waters is regulated by the International Seabed Authority. Few companies have been willing to invest in expensive technologies to exploit them, but the first mining permits have recently been granted by Papua New Guinea for mining operations at a depth of 1,500 metres.

Links

There are links here with Chapter 7, Global systems and global governance – The 'global commons', page 333, and Governance of Antarctica, pages 340–3.

Review questions

24 Identify and investigate one technological approach to providing energy security in the future.

25 Outline the key environmental and political issues that will affect mineral ore exploration in the future.

Case study of energy resource issues: Oil and natural gas production in the Niger Delta, Nigeria

Background to Nigerian oil production

Nigeria is the largest oil producer and exporter in Africa and has been a member of OPEC since 1971. It is the most populous member country with a rapidly growing population of over 190 million. Nigeria also holds the largest natural gas reserves in Africa (though Algeria is the largest producer), but has limited infrastructure in place to develop this sector. Natural gas is associated with oil production and is mostly 'flared' in order to access the oil. Development of gas pipelines and transport as liquefied natural gas (LNG) via tanker are expected to accelerate the growth of the gas industry both for export and for use in domestic electricity production.

As one of the four MINT economies showing rapid growth in the early twenty-first century, Nigeria is heavily dependent on its hydrocarbon industries, which are the mainstay of its economy. According to the IMF, the hydrocarbon sector accounted for 93 per cent of the country's export earnings in 2018.

Resource development and production

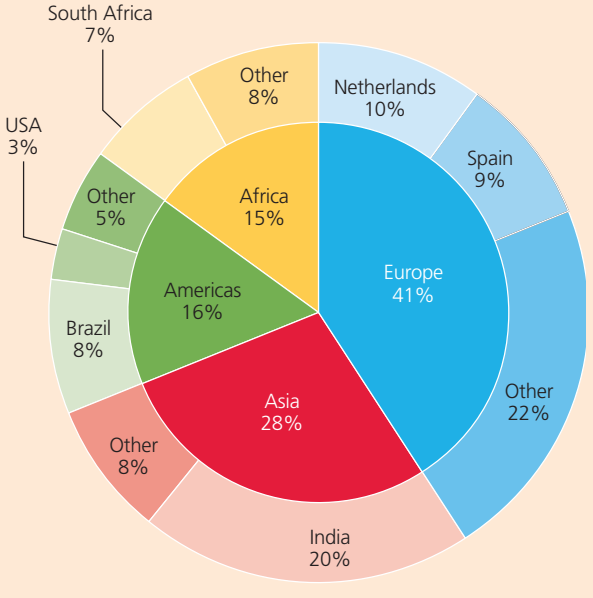
The oil industry is primarily located in the delta region of the River Niger (Figure 11.64). Oil was discovered there in 1956 by Shell-BP, who had the sole exploration rights at the time. Shell made further discoveries in shallow waters off the coast near Warri in the 1960s. Exploration rights in both onshore and offshore areas were given to other international companies such as Exxon in the 1960s. In 1977 the government established the Nigerian National Petroleum Company (NNPC), a state-owned company which both regulates and participates in the industry. This gave Nigeria more control over its oil resources but it is still reliant on the expertise and investment of TNCs such as Royal Dutch Shell, Chevron-Texaco and Exxon Mobil to explore and develop new reserves and to provide the infrastructure to distribute it for export or domestic use. These TNCs establish Nigerian-based subsidiaries (for example, the Shell Petroleum Development Company) which are 'joint ventures' with the NNPC.

The region produces around 2 million barrels per day (mbpd), though it is thought to have a production capacity of up to 3 mbpd. Production disruptions caused by local militant groups have compromised oil production in the region for many years. It is thought that light Nigerian crude oil is the best quality oil in the world and the least expensive to refine.

Trade

There are four oil refineries based in Nigeria: the two main ones are at Port Harcourt and Warri. Both only operate at 30 per cent capacity as most of the oil extracted is exported as crude by TNCs and refined overseas (Figure 11.65).

Figure 11.64 Map of Nigeria and the oilfields in the Niger Delta region



Note: Nigerian crude oil and condensate exports averaged 1.96 million barrels per day

Figure 11.65 Exports of Nigerian oil, 2015

Source: US Energy Information Administration

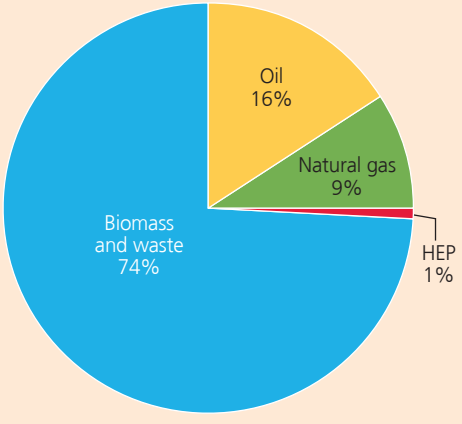
Nigeria exports its oil around the world. Until 2012, the US was the largest market for Nigerian oil exports but as America has increased its own production, it is now only tenth. Nigeria's largest regional markets are now in Asia and Europe, and India is currently the largest importer of Nigerian oil.

Links

There are links here with Chapter 7, Global systems and governance – International trade and access to markets, pages 306–14.

Nigeria's energy mix

A small proportion of the oil produced is supplied to the domestic market, although this is increasing as the country develops. Oil's contribution to the nation's energy mix increased by 23 per cent from 2012 to 2017. Sixty-four per cent of the population live in rural areas and rely on traditional biomass, such as fuelwood, as their main energy supply (Figures 11.66 and 11.67).



Note: Nigeria also consumed 29,000 short tons of coal in 2017

Figure 11.66 Nigeria's energy mix, 2017

Source: International Energy Agency

Figure 11.67 Women carrying fuelwood in southern Nigeria

Implications for resource security and human welfare in the Delta region

The region already has many ethnic and religious tensions as there are a number of local tribes living in different parts of the Delta. Around 34 million people from up to 40 different ethnic groups live in the 'South-South' region, though there are majority groups such as the Ijaw, Ogoni and Igbo tribes. Environmental devastation associated with the industry, combined with the uneven distribution of oil wealth, serve to escalate these tensions.

Conflicts

Tensions have given rise to the formation of militia groups who are largely responsible for disrupting oil production. Locals are largely excluded from the benefits of oil wealth and seek a share by attacking the oil infrastructure and staff. Oil theft (known as 'bunkering') and pipeline sabotage lead to damage, loss of production, pollution, forced shutdown of production and even piracy in the Gulf of Guinea. Government military action against these groups has been largely unsuccessful especially as the army is stretched, being engaged with Boko Haram, an Islamic extremist insurgent group in the north of Nigeria.

The Movement for the Emancipation of the Niger Delta (MEND) emerged as the most organised and visibly armed of these militia groups. However, following a government amnesty, MEND declared a ceasefire in May 2014.

In 2016, a smaller group, the Niger Delta Avengers (NDA) formed to pursue the same goals as MEND – to run the oil companies out of the region and give local people more control of oil operations. They continue to hit high value targets such as oil pipelines, causing a loss of production and revenue for the government.

The oil companies blame militant activity for the pollution but locals claim that companies are involved in illegal activities such as 'flaring' (burning off gas) and that the pipeline infrastructure is not maintained and many of the leaks are due to corrosion (Figure 11.68).

A sustainable settlement to the conflict would need to include widespread environmental clear up and compensation for damage, along with some proportion of oil earnings going to locals living in the Delta.

The implications for resource security and human welfare are summarised in Table 11.28.

Figure 11.68 Niger Delta residents pass a burning pipeline as they evacuate their homes by boat

Attempts to manage the resource

Nigeria earns some $36 billion every year from oil but Niger Delta residents remain in poverty. Less than 20 per cent of the region is accessible by good roads, even in the dry season, and hospitals and schools are seriously under-funded. The government can use oil revenue to overcome some of these issues but wealth distribution in Nigeria is very inequitable. Economic growth means more people from rural areas are moving to the cities, where there is more demand for electricity for industrial and domestic use.

- Nigeria's electricity sector is small compared to other countries with similar population size and standards of living. There is enormous potential for more electricity to be generated using natural gas and renewable sources such as HEP, solar power and biomass.

Table 11.28 Main implications for resource security and human welfare in the Niger Delta

	Positive implications	Negative implications
Resource security	Potentially high production levels Good quality oil with more efficient refining Availability of large natural gas reserves which could be used more effectively to assist development Joint venture operations between NNPC and foreign TNCs give Nigeria some control and security over an important national resource	Flaring of natural gas is illegal but satellite images suggest it has been on the increase as oil companies choose to burn the gas off rather than developing it as a resource; they argue it would be uneconomic 'Bunkering' by locals also depletes precious reserves of gas and oil Oil leaks and spills – thousands of barrels are leaked into the Delta's precious wetlands each year from oil heads and pipelines Oil wells and pipelines are open to attacks from militant groups who damage facilities and risk security Much of the wealth from the area is 'leaked' as TNC profits are repatriated or given as payments to government – an example of the **'resource curse'**
Human welfare	Many Nigerians are employed in the petrochemical industry in the area, increasing incomes, spending and skill levels Foreign investment and wealth can be used to help develop the area economically and socially in terms of infrastructure, health and education	Local tribes rarely benefit directly from the wealth, which has given rise to the militant activity Atmospheric and land pollution – families live among the oil fields, breathing in methane gas and coping with frequent oil leaks Leaked oil when heated by the Sun can become highly inflammable Flares from burning natural gas are the most visible impact of the industry – locals in the region suffer many negative social and economic consequences from gas flaring Pipelines – dense lines of unsightly pipelines cover the ground along communication routes Contaminated water – pollution means that access to safe drinking water is a major problem facing local communities Traditional agriculture and fishing – oil operations affect both farming and fishing in the area. Staple crops, such as cassavas and yams, can still be grown but the soil is losing fertility. Locals who rely on fish for protein in their diet are unable to fish because of polluted water Villages and communities – whole villages and communities in the Delta have been displaced as oil spills or forest fires mean they lose all productive land Destruction of forests – forest fires in rainforest and mangroves surrounding oil installations are fairly common and can be triggered by gas flares, explosions or by oil igniting. Fires can take weeks to extinguish and destroy much of the local farmland and natural forest

- The Nigerian government made flaring illegal in 2010, as it wastes energy and contributes to global warming. Shell committed to ending flaring in 2008, but environmental groups accuse them of reneging on this because of the expense.
- Shell has paid compensation to the Nigerian government for the damage caused by spills.
- The government is planning to transform the NNPC into a more profit-driven company that will seek more private financing and reinvest in infrastructure.
- Investment in infrastructure has established the West African Gas Pipeline (owned by a subsidiary of Chevron) which pipes natural gas to other West African states including Togo, Benin and Ghana.
- Nigeria continues to discuss a planned Trans-Saharan Gas Pipeline to carry natural gas from its oilfields to Algeria's Beni Saf gas export terminal so that they can supply gas to Europe.

Case study to illustrate how aspects of physical environment affect the availability of water: Water shortages in Shimla, northern India

Shimla is the capital city of the state of Himachal Pradesh in northern India. It is the main commercial centre and central hub for transport and trade. The city has a population of around 200,000 permanent residents but in the tourist season this can increase by a further 50 per cent. Shimla is an iconic hill city with a strong colonial legacy and distinct British Raj architecture, which contributes to its cultural tourism industry. The economy is based largely on tourism and civic administration. With two universities and many colleges, education also features strongly as an employer and there are growing IT services, technology and media publishing sectors.

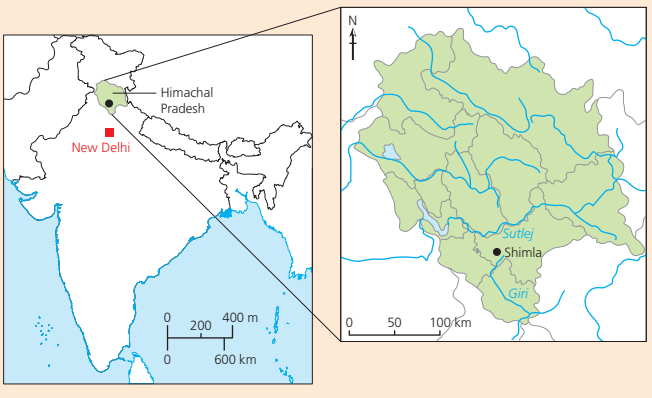


Figure 11.69 Location of Himachal Pradesh state in India, and the location of Shimla city and major rivers in the state

Physical environment of Shimla

Topography and drainage: Shimla is located in the foothills of the Himalayas. The city has an average altitude of 2,206 metres (7,238 ft) above sea level. It is built on top of a steep-sided ridge with seven elongated spurs spreading from the main ridge. There are no natural water bodies in the city and the nearest river, the Sutlej, is 20 km away. Other main rivers, including the Giri, flow through Shimla district but they are also some distance from the city. The steep slopes of the ridge cause the run-off rate to be high, so it is difficult to capture and retain any surface water on the ridge. The area around the city is surrounded by deep chasms, which contain fast-flowing streams.

Climate: Shimla has a sub-tropical highland climate with temperatures between 19°C and 29°C in the summer and -1°C to 10°C in the winter. Visitors come in summer (May–June) to escape the scorching heat of the Ganges plain to the south, where temperatures reach 45°C. Precipitation varies seasonally and is typically around 45 mm per month during winter and spring but reaches over 400 mm during the monsoon season in July and August. Altitude and cold winter temperatures mean regular annual snowfall, which used to fall in December but over the last 20 years has been falling in January and early February.

Geology and vegetation: The geology of Shimla District has been subject to intense tectonic activity. The rocks are primarily sedimentary in nature but complex faulting and folding of layers give little opportunity for groundwater storage of water.

The area around the city and on the slopes is naturally covered in forests of pine, cedar and oak. However, urban expansion has led to deforestation and a major concern is the frequent number of landslides that occur after heavy rainfall.

Figure 11.70 Shimla is located on a steep-sided ridge in the foothills of the Himalayas

Background to water supply problems in Shimla

Despite substantial monsoon rainfall, Shimla is prone to perennial water shortages because of changing climate, steep terrain, ageing infrastructure and the pressure put on the city's water supplies by visitors. In May and June 2018, Shimla experienced an acute drinking water crisis with most of the town only receiving water once in eight days. The city has been facing severe water shortages and water-borne epidemics over recent years and residents also claim that the distribution of water to people is uneven.

At the height of its tourist season in June, visitor numbers can reach up to 30,000 each day, when, according to Shimla Municipal Corporation (SMC), the city needs about 45 millon litres of water daily (MLD). The yield, even out of season, is around 32 MLD but in May 2018 it fell to less than 20 MLD, so the municipal water supplies were effectively exhausted.

Water supply: The physical environment means that the city is reliant on surface water being pumped from local streams. When the British established the settlement in the early nineteenth century, the first water came from a spring in the Dhalli catchment, which provided around 4 MLD. Dhalli is now the main water treatment plant for incoming pumped water from other sources, which were developed later.

Shimla's main water sources are from several perennial streams (known as khads or nallahs) including the Cherot Nallah and Nauti Khad. Water is transported via pumping stations (for example, the Chair pumping station). As the population grew after independence, water supply was increased with more pumped sources from the Gumma, Ashwani Khad and from the Giri River. The water is lifted from these sources and taken to the treatment plant where it is gravity fed to underground water tank reservoirs, from where the SMC supplies water to the city. (Figure 11.71)

The installed capacity of all these sources (of water) combined is 65 MLD, which should meet the demand, but the city only receives around 35 MLD on average throughout the year:

- Nauti Khad has witnessed significant reductions in water level.

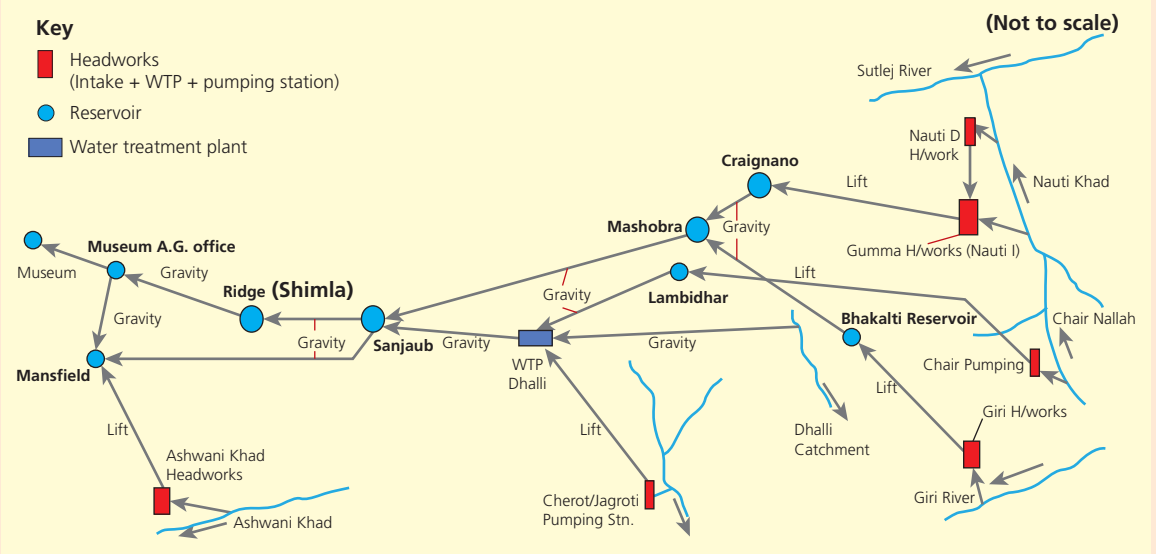


Figure 11.71 Layout of water supply system at Shimla

Source: Sharmar, Kansal, Aditya Tyagi (2015)

- The Giri and Gumma schemes have installed capacity of 20 and 24 MLD respectively but together were only pumping 20 MLD.
- Ashwani Khad, which has an installed capacity of 10 MLD, was shut in 2015 after a sewage treatment plant was found to be discharging untreated sewage into the water.

Sanitation: Around 30–40 per cent of the city's population are not covered by the sewage system and sanitation has emerged as a public health issue resulting in water-borne epidemics.

Challenges and underlying causes of water shortages in Shimla

Water has always been a scarce resource in Shimla. The topography is an underlying issue and the situation has been made worse by climate change, which has affected the physical water supply to the area. However, shortages are exacerbated by a combination of increased demand and poor water management.

Physical environment

- Shimla is located on a ridge with limited natural water supply and is reliant on artificially 'lifting' water up to considerable heights.
- Deforestation has reduced green cover in the region, which means less evapotranspiration, less cloud formation and so less rain – a meso-climatic change.
- Rapid run-off, soil erosion, flooding and landslides sometimes cause the pumping of water at source to stop due to excessive silt in the streams.
- Climate change and its impacts are profound on the city's water supply system.
  - During recent winters, precipitation has been much lower than usual so water sources have started drying up, with 50 per cent lower than average discharge.
  - In recent years, during the monsoon, rainfall has also been reduced by up to 80 per cent.
  - Rains come in shorter, heavier bursts so the water runs off quickly and is not retained below the surface, leading to flooding.
  - More frequent hailstorms have added to the problem of stream siltation.
  - A reduction in snowfall has reduced the capacity of the khads.

Human factors that have exacerbated the problem

- Failure to upgrade the city's ageing water infrastructure has resulted in leakages – around half the water is lost to leakages during pumping and distribution; some is lost to unauthorised siphoning.
- Water deficit has increased with growing population, more urbanisation and limited expansion of the water supply.
- Heavy tourist footfall places huge demands on supply.
  - 15,000–30,000 tourists visit Shimla every day during the peak tourist season.
  - Demand for water soars in the summer months when the city's resident population swells with tourist numbers needing food, drinks and water to wash.
  - Rapid and unregulated construction of hotels and other properties has occurred.
- Built infrastructure does not help with water retention.
  - More embankments are required to restrict the flow of surface water so that it percolates and recharges water bodies.
  - Catchment wells need to be constructed around water bodies to promote natural recharging.
- Farmers in the area use inefficient irrigation methods, so most water is evaporated; investments in more water-efficient technologies, such as drip irrigation, have been minimal.

Impacts on the residents of Shimla

During the 2018 crisis:

- residents were made to line up for hours each day to collect a few buckets of water from tankers supplied by the government; there were long queues at public taps
- protests broke out as large numbers of frustrated locals, politicians and activists took to the streets to voice their anger; unofficial roadblocks caused traffic jams across the city
- police officers were deployed to guard water-distribution sites and water was distributed under police protection
- people adapted their lifestyle to the shortages, for example, by switching to disposable plates and cups to save water
- demand for bottled water increased by 50 to 60 per cent due to water scarcity, the alternative was to pay steep prices to the 'tanker mafia' – private suppliers accused of siphoning water from the public supply
- water shortages forced city administrators and locals to appeal to tourists to stay away, which had a major negative impact on the local economy:
  - hotels cancelled bookings
  - the internationally famous Shimla Summer Festival was cancelled
  - local businesses and hotels suffered losses
  - local coolies and street vendors became unemployed with no income.

Figure 11.72 Shimla residents stand in queues to collect drinking water from a truck during the water shortages in May 2018

Management responses and solutions

Following the 2018 crisis, the state government decided to resolve the challenges directly by overhauling the way in which water supply and sanitation services are delivered to Shimla. They established a dedicated, autonomous utility to manage water supply and sanitation services. This formalised accountability and overcame the previously fragmented structure. The new utility, the Shimla Jal Prabandhan Nigam Ltd (SJPNL), now runs the city's water supply and sewage systems.

The World Bank supports SJPNL and the state government through development policy loans to ensure water security in the region. The state government also decided to strengthen the financial sustainability of water supply and sanitation operations as only 12 per cent of the costs needed to operate and maintain the system were actually recovered from customers.

Supply management

- SJPNL intend to restructure the Giri water scheme so that it is tapped to its full strength.
- The next phase is to bring bulk water to Shimla from a new source on the Sutlej River.
- Run-off water is to be captured in recharge drains to slow down run-off and ensure groundwater is recharged.

Demand management

- SJPNL has switched to volumetric water tariffs for the city, with subsidies targeted at low-income households.
- Around 24,000 water meters have been installed so that citizens can track the water they consume; this new approach has reduced the bills for a significant number of consumers who believe they get a fairer deal by only paying for what they use.
- City officials and volunteers are educating consumers about the changes in water delivery and helping them conserve water.
- Shimla gets abundant rainfall in summers and rainwater harvesting on a personal level will help to conserve supplies.

Other measures

SJPNL are improving sewage management for the city area to ensure that the whole population has access to good sanitation. This will be extended by expanding sewage services for the fringe areas of the growing city.

Shimla's 2018 water crisis is symptomatic of a wider crisis confronting India. With about 18 per cent of the world's population but just 4 per cent of its water resources, India is gradually running dry. On a local scale, in Shimla, there are aspects of both physical and economic water scarcity. The new approach supported by the World Bank is addressing the economic scarcity issue, but climate change may further impact the physical water scarcity.

Fieldwork opportunities

There are few specific fieldwork opportunities for this topic other than visiting energy production or water installations to gain a better appreciation of how energy is generated or how water is treated and distributed (in some cases recycled). Similarly, visits to mines or quarries will give an appreciation of mineral extraction processes but health and safety is a primary consideration on such visits.

Most power stations offer educational visits for students. The Dinorwig HEP installation inside Mynydd Gwefru (Electric Mountain) in Snowdonia is particularly interesting. For more information visit www.electricmountain.co.uk. It is also possible to visit renewable installations such as wind farms.

Primary data collection can be undertaken in river valleys containing reservoirs, both upstream and downstream of the reservoir, to establish impacts on discharge, sediment, bedload, etc. Likewise, the microclimate effects around reservoirs such as wind speed, humidity and effects on vegetation could be measured.

Further reading

General

Global Policy Forum (GPF) is an independent group that monitors work of the UN and has consultative status with the UN. Discusses many issues on its website, including resource security – www.globalpolicy.org

Energy

World Energy Outlook 2019 – an annual report published by the International Energy Agency – www.iea.org/reports/world-energy-outlook-2019

A series of factsheets based on the annual World Energy Outlook, also published by the International Energy Agency www.worldenergyoutlook.org

2020 World Energy Issues Monitor (Executive Summary) – an annual report published by the World Energy Council www.worldenergy.org

'International Energy Outlook 2019', published by the US Government Energy Information Administration – www.eia.gov/outlooks/ieo/

'Global Energy Statistical Yearbook 2019', published by Enerdata – https://yearbook.enerdata.net/

'BP Statistical Review of World Energy 2019', 68th Edition (full report) – www.bp.com/content/dam/bp/business-sites/en/global/corporate/pdfs/energy-economics/statistical-review/bp-stats-review-2019-full-report.pdf

'Renewables Now – Renewables 2019', Global Status Report – www.ren21.net/wp-content/uploads/2019/05/gsr_2019_full_report_en.pdf

Young People's Trust for the Environment, Renewable Energy factsheets available on all energy sources – http://ypte.org.uk/factsheets/renewable-energy/fossil-fuels-non-renewable

International Renewable Energy Agency, Renewable Energy Roadmap – www.irena.org/

World Nuclear Association for updates on Nuclear Power – www.world-nuclear.org

Mineral ores

'World Mineral Production 2014–2018', British Geological Survey – www.bgs.ac.uk/mineralsUK/statistics/worldStatistics.html

'Mining Technology' – periodical giving news on developments in mineral ore mining – www.mining-technology.com/

Minerals Education Coalition – informative, aimed at younger students and US based – https://mineralseducationcoalition.org/mining-minerals-information/

'World Steel in Figures 2019', published by World Steel Association – www.worldsteel.org/en/dam/jcr:96d7a585-e6b2-4d63-b943-4cd9ab621a91/World%2520Steel%2520in%2520Figures%25202019.pdf

Water

Yvonne Follows-Smith, 'Water Security in Singapore', *Geography Review*, Volume 32, Number 1, September 2018, pp 12–16

Hannah Ritchie and Max Roser, 'Water Use and Water Stress', published by 'Our World in Data' – https://ourworldindata.org/water-use-stress

Water Footprint Network – focuses on sustainable use of water – https://waterfootprint.org/en/

'World Water Development Report 2019', published yearly by UN Water (United Nations) – www.unwater.org/publication_categories/world-water-development-report/

The UN's inter-agency mechanism on all freshwater issues produces a number of topic focuses and factsheets, e.g. transboundary waters, water scarcity, water and climate change – www.unwater.org/

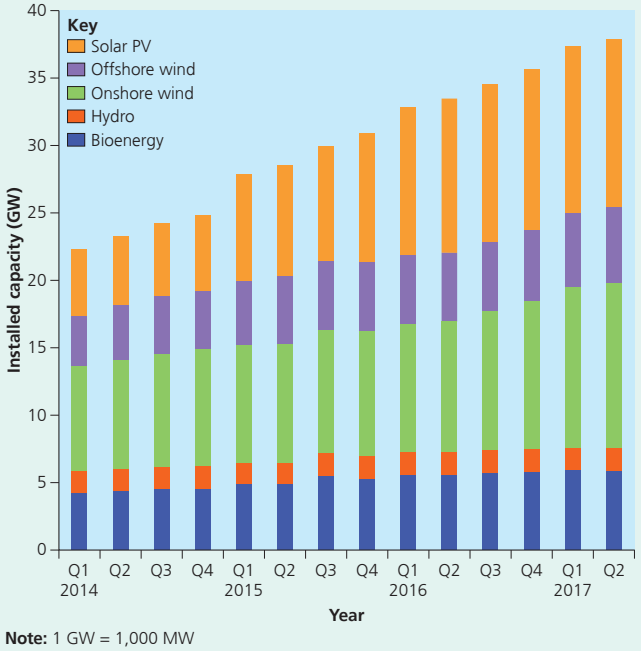
'Beyond Scarcity: Water Security in the Middle East and North Africa', published by the World Bank Group – www.worldbank.org/en/topic/water/publication/beyond-scarcity-water-security-in-the-middle-east-and-north-africa

Question practice

A-level questions

1 Explain what is meant by the virtual water trade. [4 marks]

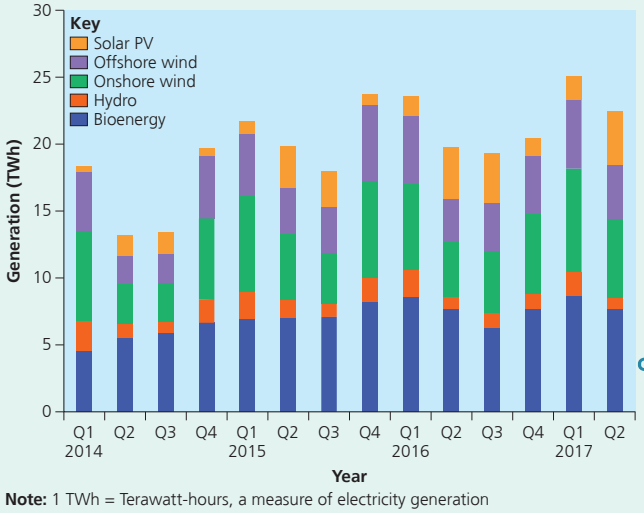
2 Analyse the data shown in Figure 11.73a and Figure 11.73b. [6 marks]



Note: 1 GW = 1,000 MW

Figure 11.73a Increase and composition of installed renewable energy capacity in the UK from 2014–2017

Source: Catch Energy



Note: 1 TWh = Terawatt-hours, a measure of electricity generation

Figure 11.73b Increase and composition of renewable energy generation in the UK from 2014–2017

Source: Biomass magazine, 2017

3 Using Figure 11.74a, Figure 11.74b and your own knowledge, assess the need to balance strategies to increase energy supply with the sustainability of energy production. [9 marks]

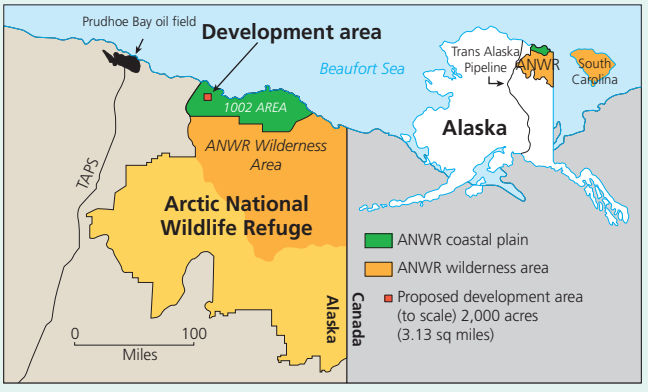


Figure 11.74a Permitted oil and gas drilling areas by the US government in the Arctic National Wildlife Refuge (ANWR) in Alaska

Source: Alaska Department of Natural Resources

Arctic National Wildlife Refuge – Drilling controversy

In Alaska, **the Arctic National Wildlife Refuge** (ANWR) has been protected from mineral exploitation for decades. It is the subject of ongoing controversy in the US, especially since the government passed legislation in 2017 that permitted oil and gas drilling in the **1002 Area** of the Refuge. Initially, the proposed development area for exploration was limited to 2,000 acres but in 2019, the Trump administration opted for the most aggressive leasing alternative by opening up the entire coastal plain for exploration.

- US Bureau of Land Management (BLM) filed a final environmental impact statement allowing development and plans to grant drilling permits on the coastal plain by 2020.
- Opponents said the BLM's final statement underestimated the impacts of the project on the climate because it viewed global warming as cyclical rather than of human origin.
- The debate rests on the amount of economically recoverable oil versus the potential harm that oil exploration would have on the natural environment.
- The area is a fragile tundra ecosystem and includes places where threatened polar bears have dens and porcupine caribou visit for calving.
- Indigenous people settled in the area rely on subsistence hunting and fishing.

Figure 11.74b Key points about the proposed development in the 1002 Area of the ANWR

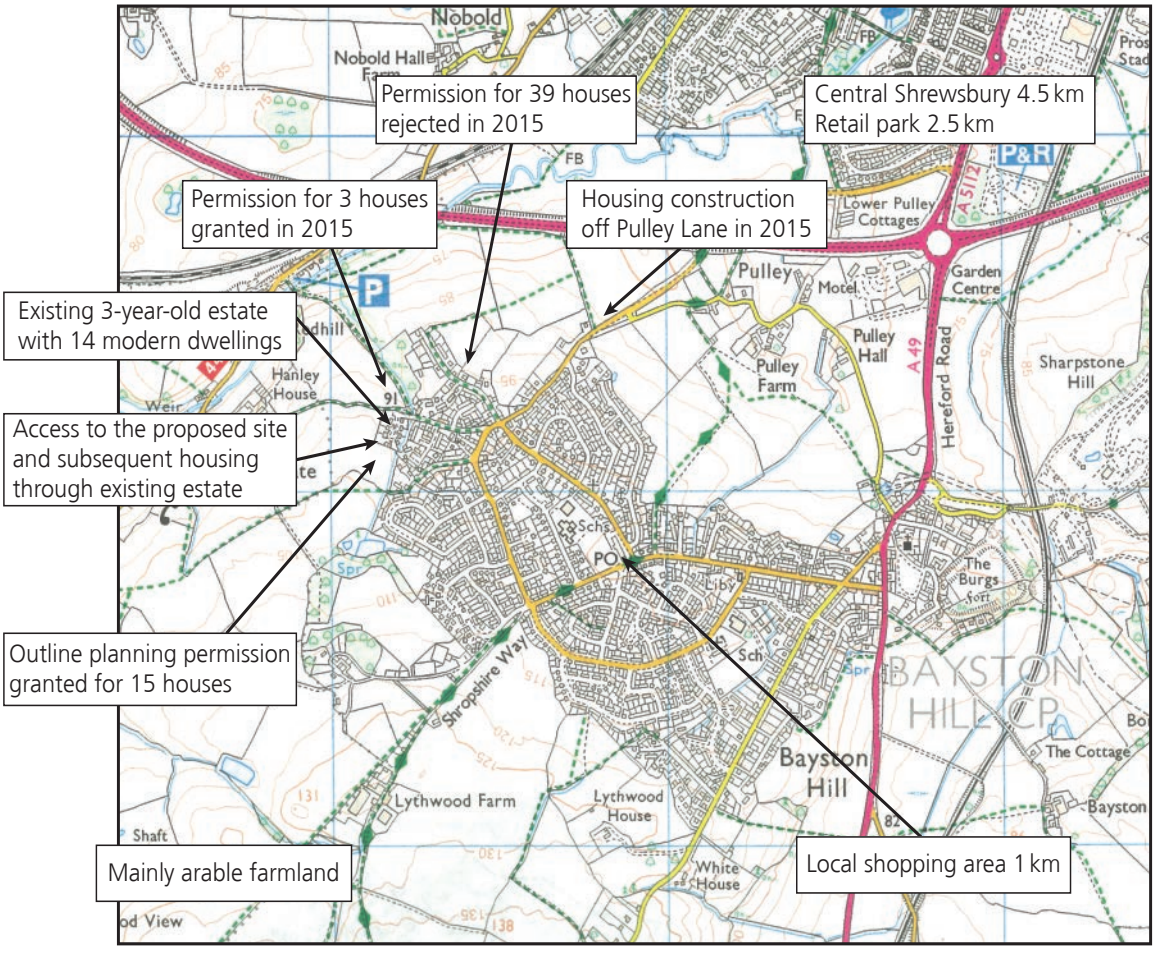
4 With reference to a named mineral ore, evaluate the extent to which global systems have influenced the role of that ore in global commerce and industry. [9 marks]

5 'Strategies used to increase water supply are a more sustainable way to achieving water security than those which focus on managing water consumption.'

To what extent do you agree with this statement? [20 marks]

Part 3

Skills and Fieldwork Investigation



© Crown copyright and database rights, 2021, Hodder Education, under licence to Ordnance Survey. Licence number 100036470

Chapter

12 Geographical skills and fieldwork

One of AQA's three assessment objectives states that you should be able to:

'Use a variety of relevant quantitative, qualitative and fieldwork skills to:

- *investigate geographical questions and issues*
- *interpret, analyse and evaluate data and evidence*
- *construct arguments and draw conclusions.'*

These skills are assessed in both the examined and non-examined (fieldwork report) components.

The AQA A-level specification states that its intention is that geographical skills should be developed in an integrated way throughout the whole course.

During the course, students should:

- understand the nature and use of different types of geographical information, including qualitative and quantitative data, primary and secondary data, images, factual text and discursive/creative material, digital data, numerical and spatial data and other forms of data, including crowd-sourced and 'big data'
- collect, analyse and interpret such information, and demonstrate the ability to understand and apply suitable analytical approaches for the different information types
- undertake informed and critical questioning of data sources, analytical methodologies, data reporting and presentation, including the ability to identify sources of error in data and to identify the misuse of data
- communicate and evaluate findings, draw well-evidenced conclusions informed by wider theory
- construct extended written argument about geographical matters.

12.1 Qualitative skills and quantitative skills

Key terms

Coding – Coding is carried out when analysing answers to interviews. The coder (person who analyses the data) looks though all the answers to a question, develops a broad classification system based on the responses, and then uses a code to categorise responses.

Qualitative data – These are non-numerical data that are used in a relatively unstructured and open-ended way. It is descriptive information, which often comes from interviews, focus groups or artistic depictions such as photographs.

Quantitative data – Quantitative data are numerical data such as metric-level measurements that are associated with the scientific and experimental approach and are criticised for not providing an in-depth description.

Sampling – To make statistically valid inferences, when it is impossible to measure the whole population, by selecting a group which will be representative of that population. Such a group is known as a sample.

Statistical population – The whole from which a sample will be chosen for a research exercise.

Data used by geographers can be divided broadly into **qualitative data** and **quantitative data**.

Qualitative methods are ways of collecting data which are concerned with describing meaning, rather than with drawing statistical inferences. What qualitative methods (for example, in-depth interviews and participant observation) lose on reliability they gain in terms of validity. They provide a more in-depth and rich description of the subject under study. It is qualitative data that is most often used in cultural geography.

Quantitative methods (or geostatistics) are those which focus on numbers and frequencies rather than on meaning and experience. Quantitative

methods (for example, till analysis, questionnaires and surveys) provide information which is easy to analyse statistically and whose reliability is known. Quantitative methods are associated with the scientific and experimental approach and are sometimes criticised for not providing an in-depth description.

Qualitative data

Interviews

One of the main ways of collecting qualitative data is through interview surveys. These can be detailed and flexible with open-ended questions and the opportunity for respondents to give their opinions without being limited to responses in option boxes. The interviewer has to be well prepared, with a specific aim for each interview, which matches the topic under consideration as well as the role of the interviewee. They can also be loosely structured so particularly interesting points can be pursued and you can adapt to follow the flow of the conversation.

Examples of where interviews would be appropriate include:

- studying the attitudes of residents to the development of a nearby housing estate
- finding out, from an entrepreneur, the reason for their choice of location of a new commercial development.

Once a series of interviews have been carried out then the technique of **coding** is used to put the qualitative data into a form in which it can be analysed using statistical methods. An example would be open-ended questionnaire data. The 'coder' (person who analyses the data) looks through all the answers to a question, develops a broad classification system based on the responses, and then uses a code to categorise responses. Often unusual or unexpected responses are left in an 'unclassified' category, and classes with very few answers may be aggregated to the nearest similar answer.

Quantitative data

One of the main skills that you are expected to use throughout fieldwork is the ability to apply geographical and the geospatial technologies (for example, GIS) that are used to collect, analyse and present geographical data. Alongside this are skills associated with descriptive statistics related to central tendency, dispersion and correlation.

Sampling

Sampling techniques are used to collect both qualitative and quantitative data.

In statistics, the entire pool of items under study is known as the **statistical population**. Sampling is a shortcut method for investigating this whole population. Data are gathered on a small part (a sample) of the whole parent population. The information obtained from the sample allows statisticians to develop hypotheses about the larger population.

Why sample?

Sampling is used when it is impossible, or simply not necessary, to collect large amounts of data. Collecting small amounts of carefully selected data will enable you to obtain a representative view of the feature as a whole. You cannot, for example, interview all the shoppers in a market town or all the inhabitants of a village, but you can look at a fraction of those populations and from that evidence indicate how the whole is likely to behave.

When you have established the need for a sample survey, you will have to decide on a method that will collect a large enough representative body of evidence. If, for example, you are interviewing the inhabitants of a village, you must ensure that your interviews cover all age ranges in the population.

Sampling considerations

Figure 12.1 summarises the things you need to take into account when deciding whether or not to use a sampling survey.

Types of sampling

Three main types of sampling strategy:

- random
- systematic
- stratified.

Within these types, you may then decide on a point, line or area method.

Random sampling

This is summarised in Figure 12.2.

The advantages of random sampling are:

- it can be used with large sample populations
- it avoids **bias**
- each member of the population has the same chance of being selected.

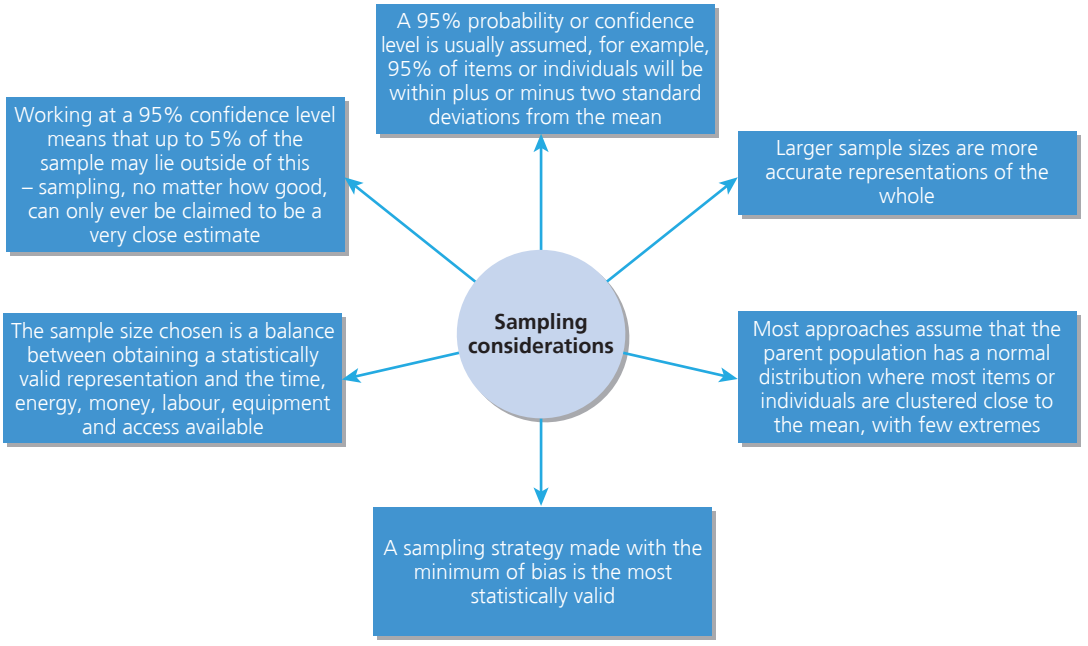


Figure 12.1 Sampling considerations

The disadvantages of random sampling are:

- it can lead to poor representation of the overall parent population or area if large areas are not hit by the random numbers generated, and is made worse if the study area is very large
- possible practical constraints in terms of time available and access to certain parts of the study area.

Systematic sampling

Samples are chosen in a systematic, or regular way. This means they can be:

- evenly/regularly distributed in a spatial context, for example, every 10 m along a transect
- at equal/regular time intervals, for example, every hour or at the same time every day
- regularly numbered, for example, every tenth person or building.

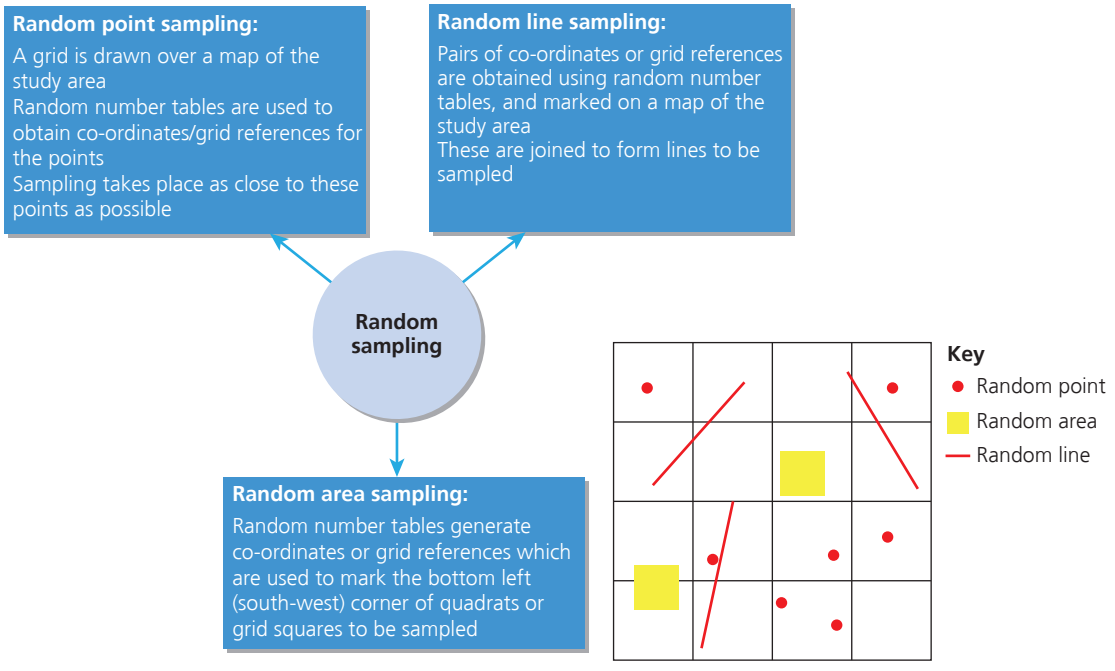


Figure 12.2 Random sampling methodology

Like random sampling, systematic sampling can be done either as a point, a line or an area. A grid can be used and the points can be at the intersections of the grid lines. If it is a transect survey then the points can be at regular intervals along that transect. Systematic line sampling could use the grid lines on a map or a series of transects (for example, a transect every 20 m along a beach). Systematic area sampling could be a series of grid squares on a map chosen to create a regular pattern.

The advantages of systematic sampling are:

- it is more straightforward than random sampling
- a grid doesn't necessarily have to be used, sampling just has to be at uniform intervals
- good coverage of the study area can be easily achieved.

Disadvantages of systematic sampling are:

- it is more biased, as not all members or points have an equal chance of being selected
- bias can lead to over or under representation of a particular pattern.

Stratified sampling

This method is based on knowing something in advance about the population or area in question. The parent population is then grouped into subsets of known size. These make up different proportions of the total population and therefore sampling should be stratified to ensure that results are proportional to the size of the subset and representative of the whole. For example, if you are surveying a population with a view to examining lived experience of a place and you know its age distribution, your sample must reflect that distribution. Choice of sample within the subsets can still be either systematic or random.

The advantages of stratified sampling are:

- it can be used with random or systematic sampling, and with point, line or area techniques
- if the proportions of the subsets are known, it can generate results which are more representative of the whole population
- it is very flexible and applicable to many geographical enquiries
- correlations and comparisons can be made between subsets.

The disadvantages of stratified sampling are:

- the proportions of the subsets must be known and accurate if it is to work properly
- it can be hard to stratify questionnaire data collection; it may be hard to identify people's age or social background effectively.

Sample size

The size of sample usually depends upon the complexity of the survey being used. When using a questionnaire it is necessary to sample sufficient people to take into account the considerable variety introduced by the range of questions. Sample size can be restricted by practical difficulties and this may affect the reliability of results. Your aim should be to keep the sampling error as small as possible. You are not a professional sampler and cannot be expected to conduct hundreds of interviews, but on the other hand, sampling only 20–30 people in an energy conservation survey is not representative of the population as a whole.

Ethical implications

You should have an awareness of ethical issues which are embedded in any study that involves the collection, analysis and representation of geographical information about human communities. The most common ethical dilemmas in human geography focus around participation, consent, safe-guarding/confidentiality of personal information as well as giving something back.

In physical geography the main ethical considerations are about consent/access to study sites and potential damage or pollution of study sites.

Review questions

1 What sort of sample method would be most appropriate for the following investigations?

- **a** A survey of differing infiltration rates within a drainage basin
- **b** A land use survey along a coast as part of an investigation into sustainable management
- **c** Pebble size and roundness on a beach
- **d** The cultural characteristics of a place
- **e** The health of the population of a place.

Justify your choices.

12.2 Specific skills

Core skills

There are certain core skills that we take for granted. These include you having sufficient literacy skills to be able to write and interpret written reports, as well as numeracy skills to measure quantitative data and make relevant calculations. Other core skills are outlined below.

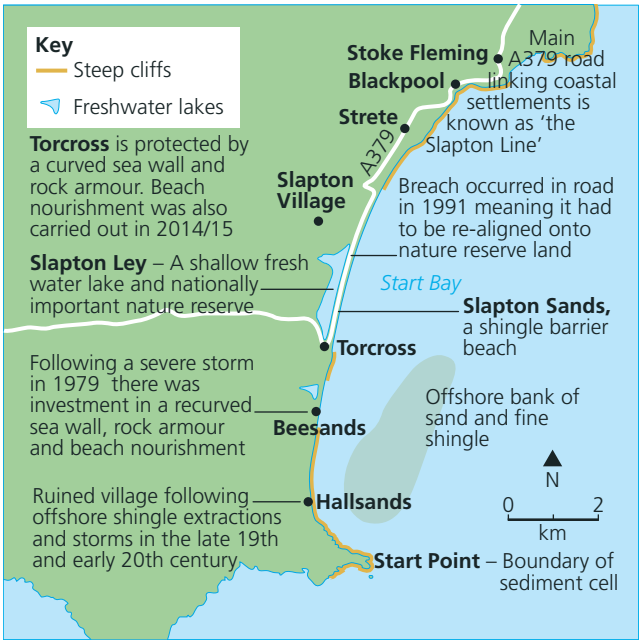


Figure 12.3 Sketch map of Start Bay

Sketch maps

Any annotated map/diagram must feature the following elements:

- a title or location of the place in the map
- a key to all symbols
- annotations or labels explaining, elaborating or emphasising particular features. Annotations are usually more than 10 words whereas labels are less.
- an indication of scale
- a north pointer.

Using field sketches and photographs

In your personal investigation, field sketches and photographs are excellent ways to record exactly what you have seen. Field sketches enable you to pick out from the landscape the features that you wish to identify and perhaps comment upon. They are particularly useful in physical geography investigations, for example, Figure 12.4 shows a field sketch of a glacial area. However, there is no reason why the technique cannot be used in an urban study. Sketching does not require artistic talent. It is far more important in geographical investigations to produce a clear drawing with good and useful labels.

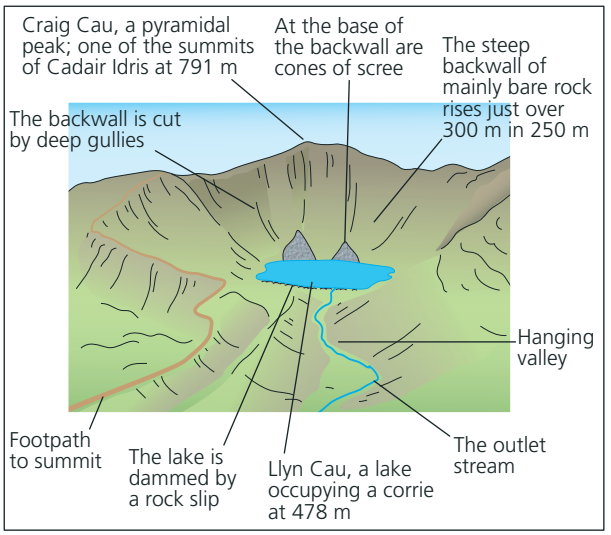


Figure 12.4 A field sketch of Llyn Cau, Cadair Idris

If you lack the confidence to produce sketches but need to show detail from the area under investigation, then taking photographs is just as valid. However, you must not use them as an attractive 'space-filler' in your report. Any photographs you include must be annotated or labelled. This can be done digitally or using an acetate overlay. A photograph is only useful when the reader is directed to the information it shows.

Ordnance Survey maps

You must be able to interpret and use Ordnance Survey (OS) maps (at a variety of scales). Each grid square of an OS map contains a huge amount of information. This can be on the physical landscape and landforms of the area (using contour lines and map symbols) and/or the human landscape. OS maps can be used as a source of information or as a base map for display of fieldwork investigation findings.

You are expected to be able to work out location (grid references), use the scale and direction indicators, interpret contour patterns, describe settlement form and function as well as inferring land use.

Cross sections of physical landscapes can be constructed from OS maps. These could be used in

fieldwork reports, though they will not be directly examined.

Geographical information systems (GIS)

Geographical information systems (GIS) are an integral part of twenty-first century geographical study. They are used extensively by environmental planners, government departments, public utility companies and commercial companies. GIS have the ability to store, retrieve, manipulate and analyse a wide range of spatially-related data. They can:

- help with questioning and understanding data
- enable multiple interrogation of complex data
- illustrate difficult abstract concepts in a dynamic visual way
- make use of three-dimensional representations
- provide opportunities for modelling and decision making.

GIS software is a set of computer tools for collecting, storing, processing and displaying sets of information linked to places on maps. Digital maps, satellite images or aerial photographs can link information. The maps are produced either as raster or vector data.

Raster maps are images made by a series of coloured dots on a screen (pixels), just like high quality digital photographs. Digital versions of OS maps are raster maps.

With vector maps, each feature (for example, roads, buildings or woods) is recorded using XY co-ordinates, enabling a GIS to link information from spreadsheets and databases to the maps. These layers can be added to a base map when and if needed. GoogleMaps is an example of vector mapping in use.

One example of GIS in action occurred in February 2020 when the Environmental Systems Research Group (Esri), a California-based supplier of geographic information system software, launched a COVID-19 GIS hub to centralise access to maps and GIS data relating to COVID-19. Its purpose was to be a central gathering point to 'aid with real-time monitoring, to catalogue GIS data that supports mapping and analysis, and to increase overall community preparedness with general public health resources and media stories related to COVID-19.' It was able to give regular updates aggregated from different health organisations, reporting confirmed cases, deaths, and recoveries from the coronavirus.

Cartographic skills

Cartographic skills fall into two main categories:

- the reading and interpretation of maps (for example, atlas and weather maps)
- the production of maps in order to present information (for example, sketch maps)

Atlas maps

An atlas is a collection of maps based on a single theme (for example, a road atlas) or a series of themes (physical, political, population, resources, etc.). Most academic atlases require you to understand and be able to use lines of latitude and longitude to identify and describe location. You need to be able to use the index and understand how colour and symbols are used to convey information.

Atlases contain huge amounts of information that can be useful when studying and comparing countries or regions. Note that it is important, in this rapidly changing world, that the atlas is up to date.

Weather maps

A **synoptic weather chart** (or map) is a summary of the current situation. In weather terms this means the pressure pattern, fronts, wind direction and speed and how they will change and evolve over the coming few days. Weather maps and synoptic charts are useful when you are planning fieldwork. They not only tell you what the weather is likely to be, but also tell you why the weather is like it is. The most common detailed type of weather chart you are likely to encounter is the Atlantic chart (Figure 12.5).

Maps with located proportional symbols

These are maps that include symbols which are proportional in area or volume to the value they represent (Figure 12.6). Symbols of representative sizes, such as squares or circles, or even small graphs, such as bar graphs or pie graphs, can be placed on a map to show spatial differences.

It is important that you take great care in placing symbols on a map. It is essential to avoid too much overlap, but it must also be clear which area or place the symbol represents.

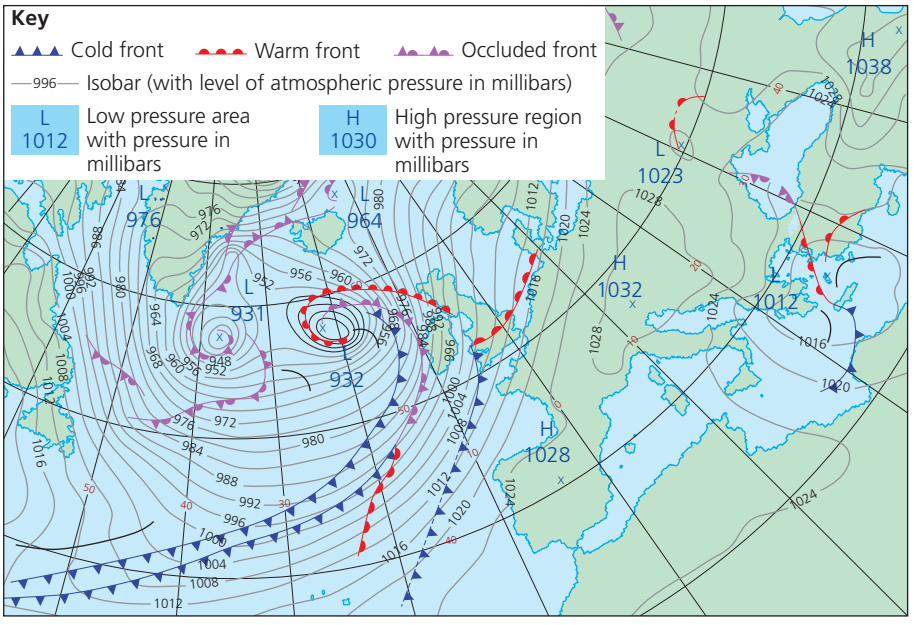


Figure 12.5 An Atlantic weather chart of the synoptic situation during named storm Dennis in February 2020

Source: Met Office

Maps showing movement – flow lines, desire lines and trip lines

Flow lines and desire lines are similar in that they both represent the volume of movement from place to place (Figure 12.7, page 648 and Figure 12.8, page 649). They are useful to show such features as:

- traffic movements along particular routes (for example, roads, railways and waterways)
- migration of populations
- movement of goods or commodities between different regions
- movements of shoppers.

In both methods the width of the line is proportional to the quantity of movement. A flow line represents the quantity of movement along an actual route, such as a train or bus route.

Figure 12.6 'Confirmed' cases of COVID-19 virus reported between 13–19 February 2020

Source: WHO

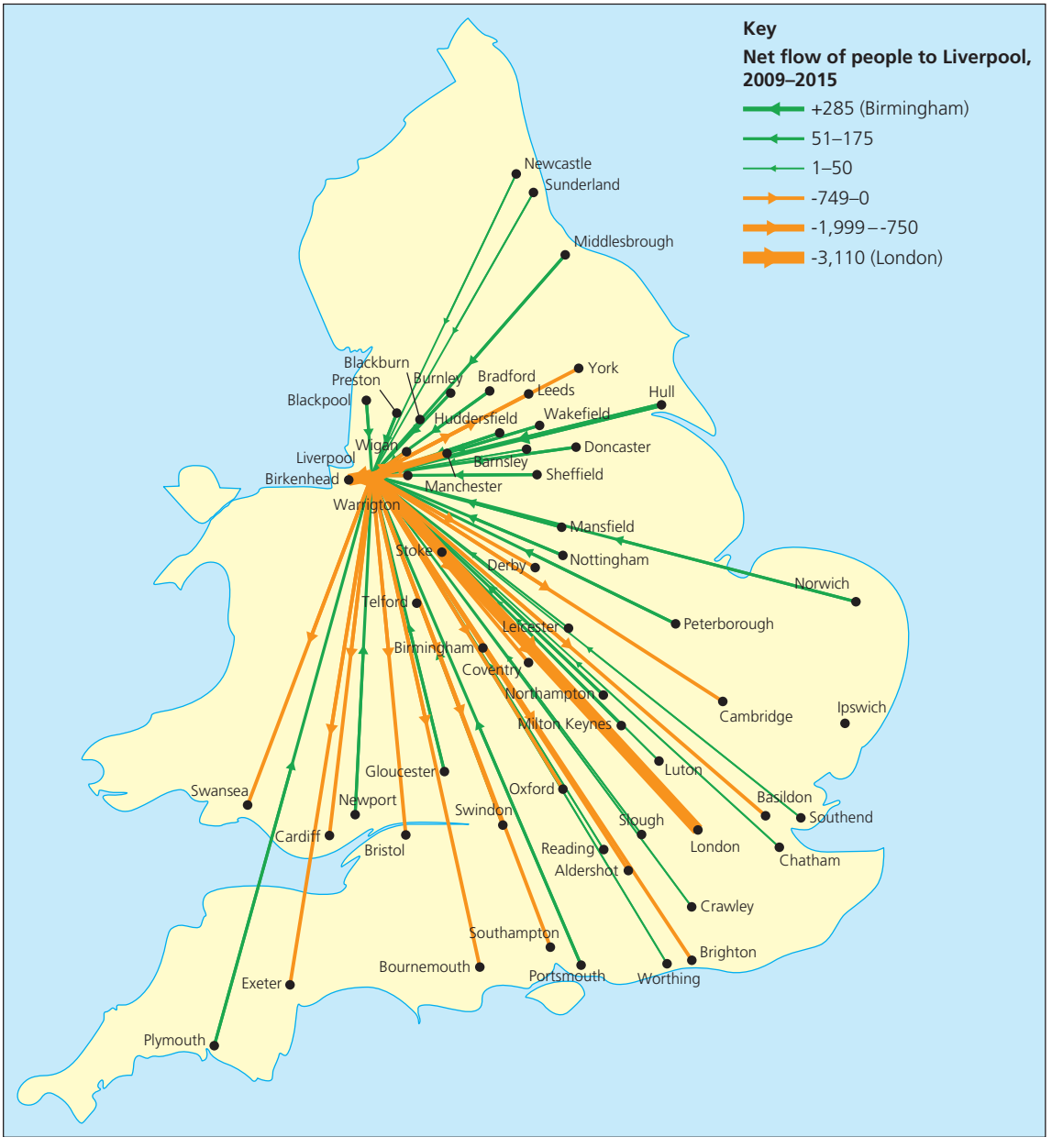


Figure 12.7 A desire line map showing net migration of people to Liverpool, 2009–2015

Source: Centre for Cities

A desire line is drawn directly from the point of origin to the destination and takes no account of a specific route.

Trip lines can be drawn to show regular trips, for example, where people shop; lines could be drawn from a town to nearby villages (Figure 12.9).

Maps showing spatial patterns – choropleth, isoline and dot maps

Choropleth maps

A choropleth is a map on which data values are represented by the density of shading within areas (Figure 12.10). The data are usually in a form that can be expressed in terms of area, such as population density per square kilometre. To produce such a map certain stages have to be followed:

- The material has to be grouped into classes. Before you can do this you have to decide on the number and range of classes required to display your data clearly.
- A range of shadings has to be devised to cover the range of the data. Darkest shades should represent the highest figures and vice versa. It is good practice not to use the two extremes of black and white because black suggests a maximum value while white implies that there is nothing in the area.

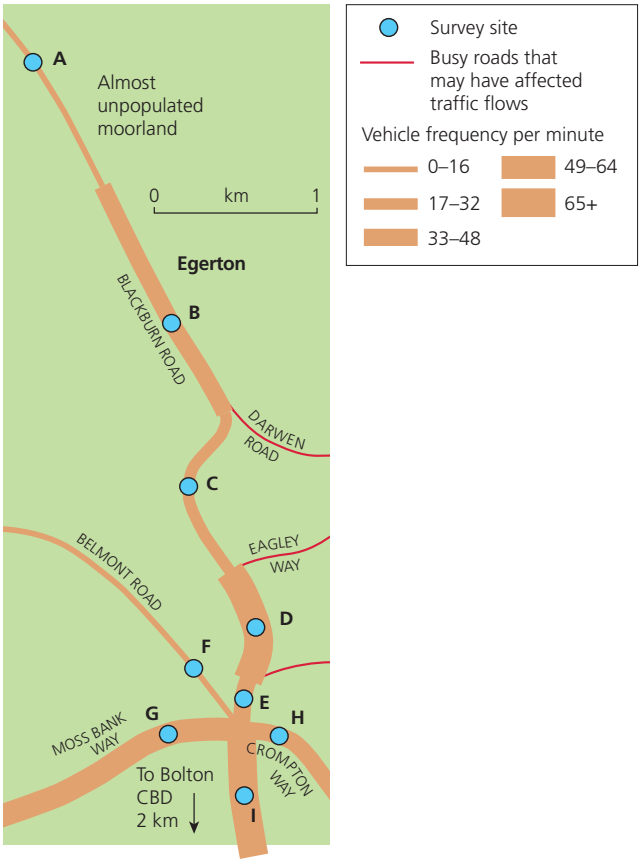


Figure 12.8 A flow line map showing traffic flows near Bolton

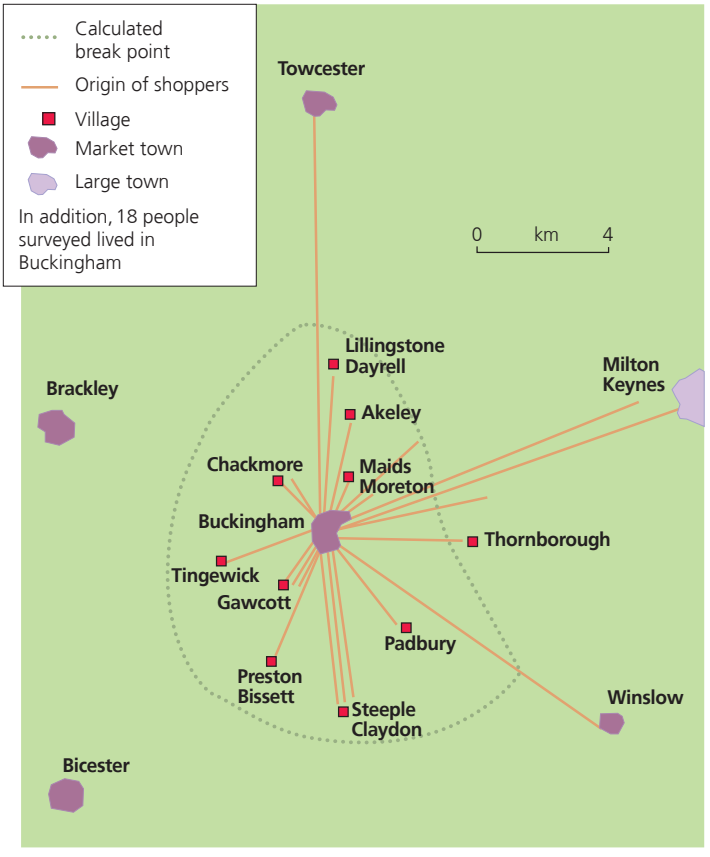


Figure 12.9 Trip line map showing the origins of shoppers on a market day in Buckingham

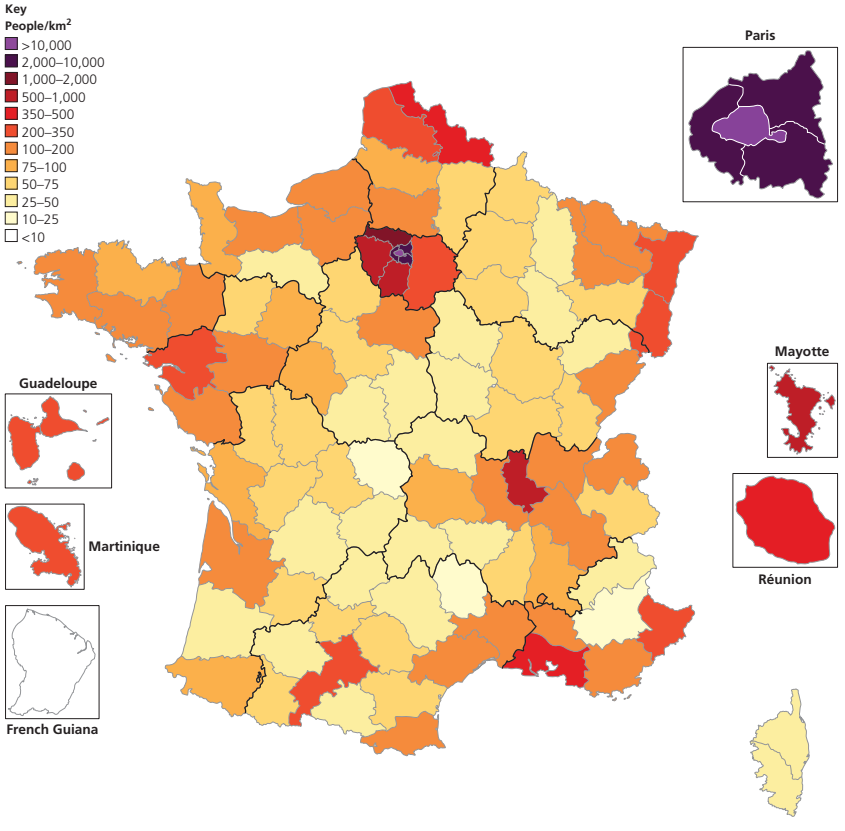


Figure 12.10 Choropleth map of population density in France and its overseas departments

Source: Thomas Brinkhoff: City Population, http://www.citypopulation.de

Choropleth maps are fairly easy to construct and are visually effective as they give the reader a chance to see general patterns in a real distribution. There are, however, a few limitations to the method:

- It assumes that the whole area under one form of shading has the same density, with no variations. For example, on maps of the UK, the whole of Scotland may be covered by one category, when it is obvious that there could be large variations between the central populated areas and the Highlands.
- The method implies abrupt changes at the drawn boundaries which will not be present in reality.

Isoline maps

It is possible to draw a map on which all points of the same value are joined by a line. This allows patterns in a distribution to be seen. The best-known example of such isolines (also called isopleths) is on Ordnance Survey maps where contour lines join places of the same height. This technique can be applied to a number of other physical factors, such as rainfall (**isohyets**), temperature (**isotherms**) and pressure (**isobars**) (see Figure 12.5, page 647), as well as human factors, such as travel times (**isochrones**) for commuters and shoppers.

Some rules regarding isoline construction:

- They connect points of equal value.
- They represent continuous surfaces (like a ground surface). It doesn't suddenly disappear.
- Isolines do not cross or touch (with the exception of vertical gradients, like cliffs.)
- Values on one side of the isoline are higher or lower than values on the other side.
- The interval between isolines is the same over the entire map (unless otherwise specified).
- Isoline maps of the environment almost always represent an overhead view, looking straight down.

Dot maps

A dot map is a map in which the spatial distribution of a geographical variable is represented by a number of dots of equal size (Figure 12.11). Each dot has the same value and is plotted on a map roughly where that variable occurs. The dot value should be high enough to avoid excessive overcrowding of dots in areas with high concentrations of the variable being mapped and low enough to prevent areas with low concentrations of the variable having no dots at all, so giving a false sense of emptiness.

Figure 12.11 World megacities in 2017

Source: The Economist

Dot maps have two limitations:

- Large numbers of dots are difficult to count. Therefore, although dot maps are good at giving an impression of distribution, they are less valuable if you need a precise idea of the values they represent.
- There must be some accompanying information about the distribution of the variable or the map could be misleading.

Review questions

2 What sort of maps would be most appropriate for displaying the following?

- **a** The population distribution within a country
- **b** The location of dairy farms in the UK
- **c** The ethnicity of different cities in the UK
- **d** The number of doctors per 1,000 of population for the UK
- **e** The main global migrations of the past 30 years.

Justify your choices.

Graphical skills

Line graphs

Simple and compound line graphs

Line graphs are appropriate when you want to show absolute changes in data. For example, they are suitable for showing changes in food production through time, population change or stream discharge. When several lines are plotted on the same graph, it is important to recognise whether it is a simple or a compound line graph:

- On a simple line graph, the line represents the actual values of whatever is being measured on the vertical axis
- On a compound line graph, the differences between the points on adjacent lines give the actual values. To show this, the areas between the lines are usually shaded or coloured and there is an accompanying key (Figure 12.12).

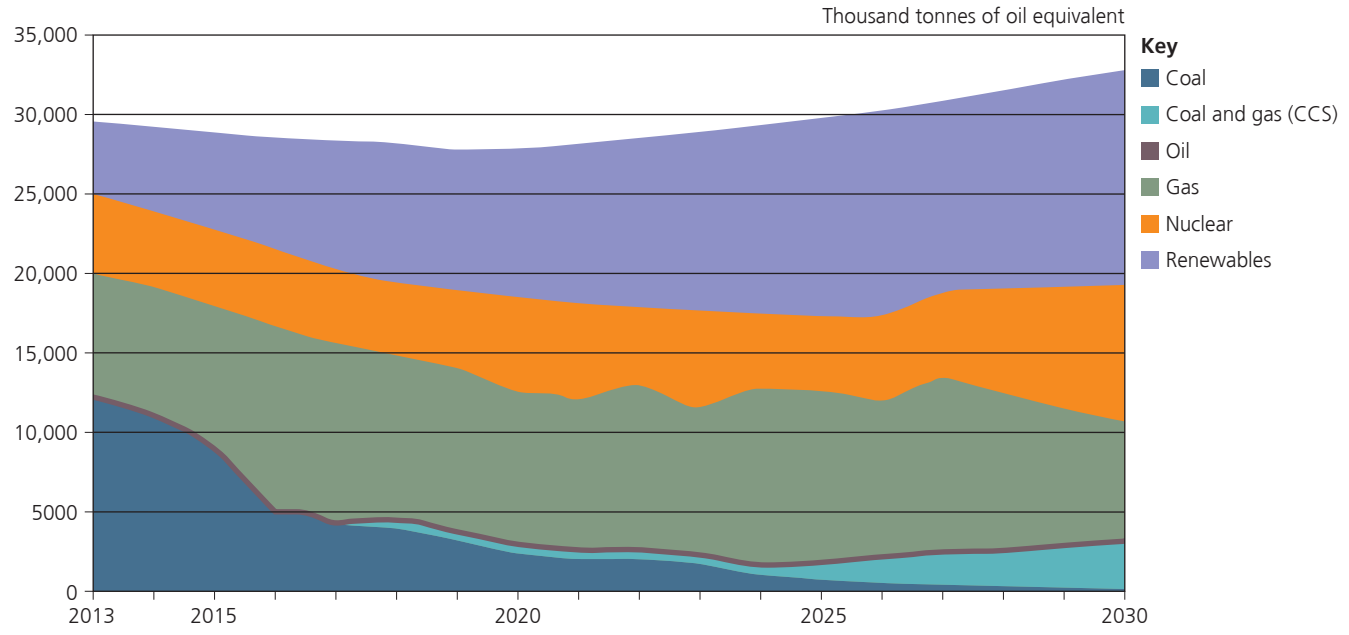


Figure 12.12 Future UK sources of energy, 2013–2030

It is possible to show two sets of data on the same graph. The left hand vertical axis can be used for one scale and the right hand vertical axis for a different scale. This can often give a useful visual impression of the connection between two sets of data.

When using line graphs you should:

- plot the independent variable on the horizontal axis and the dependent variable on the vertical axis. If you are plotting data over time, time should always be plotted on the horizontal axis
- try to avoid awkward scales and remember that the scale you choose should enable you to plot the full range of data for each variable
- clearly label the axes
- use different symbols if you are plotting more than one line.

Bar graphs – simple, comparative, compound and divergent

A bar graph (or chart) has vertical columns rising from a horizontal base (Figure 12.13). The height of each column is proportional to the value that it represents. The vertical scale can represent absolute data or figures as percentages of the whole.

Bar graphs are easy to understand. Values are obtained by reading off the height of the bar on the vertical axis. They show relative magnitudes very effectively.

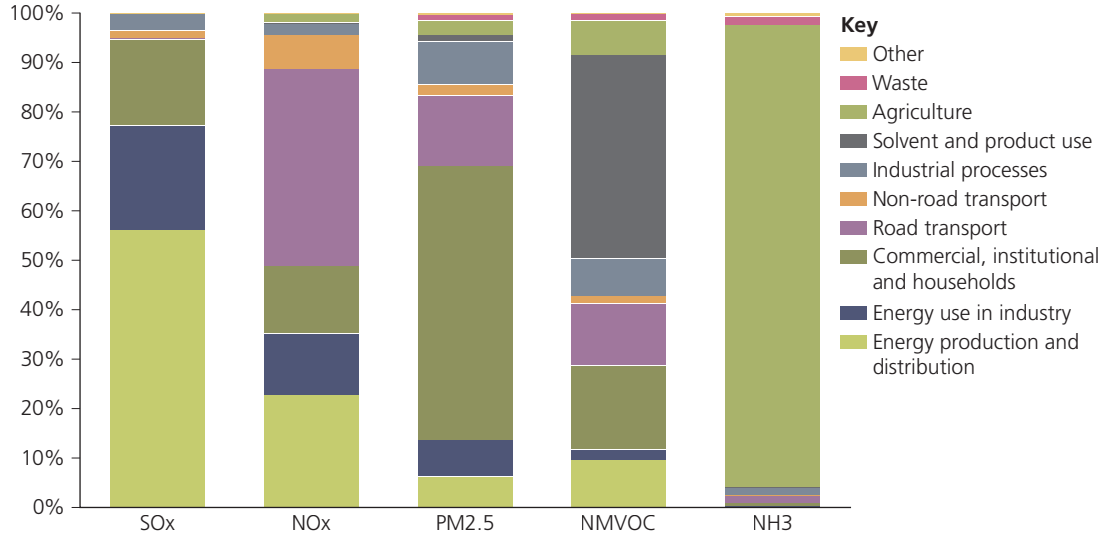


Figure 12.13 A compound **bar graph** showing share of emissions of main air pollutants by sector group for the 33 countries of the European Economic Area (EEA)

Source: European Environment Agency

Using an appropriate scale, it is also possible to show positive and negative values on the same graph – for example, profit and loss (Figure 12.14).

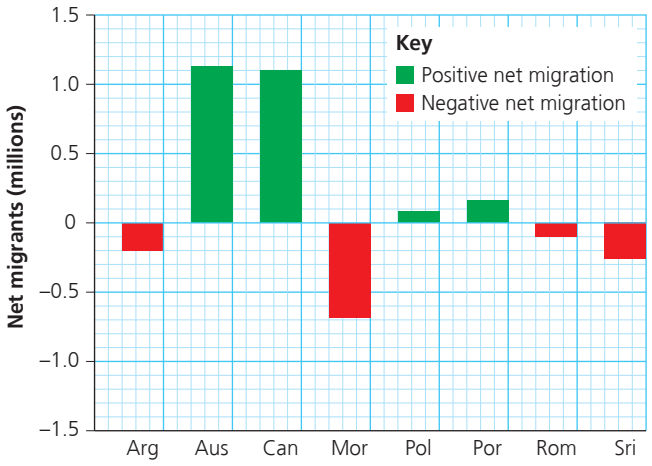


Figure 12.14 A gain–loss (divergent) bar chart showing net migration for selected countries

Scattergraphs, and the use of best-fit line

Scattergraphs are used to investigate the relationship between two sets of data (Figure 12.15). They can be used simply to present data, but they are particularly useful in identifying patterns and trends in the relationship that might lead to further inquiry.

A general trend line (**best-fit**) can be added to the graph so that the relationship can be easily observed. This is the red line in Figure 12.15. If it runs from bottom left to top right, it indicates a positive relationship; if it runs from top left to bottom right, the relationship is negative.

Other features of scattergraphs include:

- they can be plotted on arithmetic, logarithmic or semi-logarithmic graph paper
- the independent variable goes on the horizontal axis and the dependent variable on the vertical axis
- it is possible for a correlation to emerge even when a relationship is only coincidental
- points lying some distance from the best-fit line are known as **residuals** (anomalies). These can be either positive or negative. Identification of residuals may enable you to make further investigations into other factors that could have influenced the two variables.

Pie charts and proportional divided circles

The pie chart is divided into segments according to the share of the total value represented by each segment (Figure 12.16). This is visually effective – the reader is able to see the relative contribution of each segment to the whole. On the other hand, it is difficult to assess percentages or make comparisons between different pie charts if there are lots of small segments.

When a number of pie charts are drawn proportional to the value each represents in total, they are called **proportional divided circles**. The construction of a proportional circle is as follows:

- Use the formula

$$r = \sqrt{\frac{V}{\pi}}$$

where V is the value that you want the total pie chart to represent and r is the radius of the pie chart. Take the value for π as 3.142.

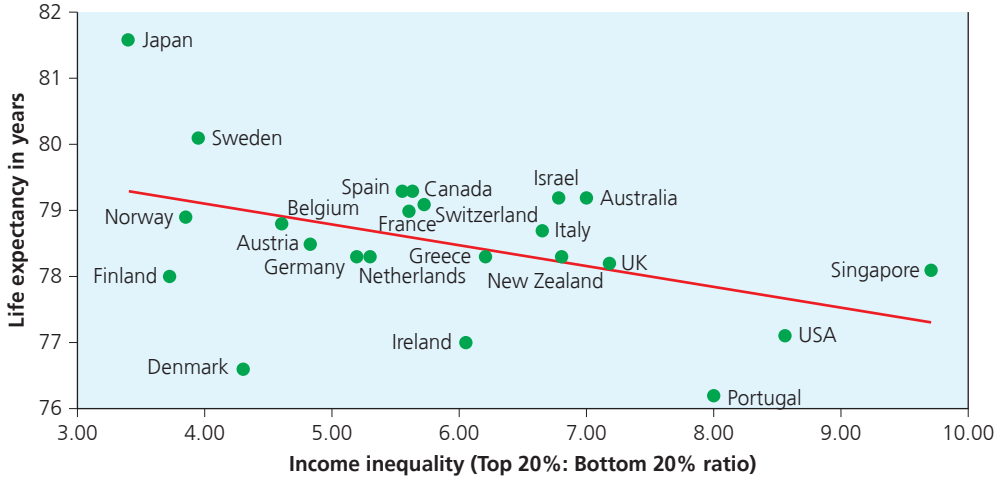


Figure 12.15 Scattergraph to show the relationship between income equality and life expectancy

Source: www.inequality.org

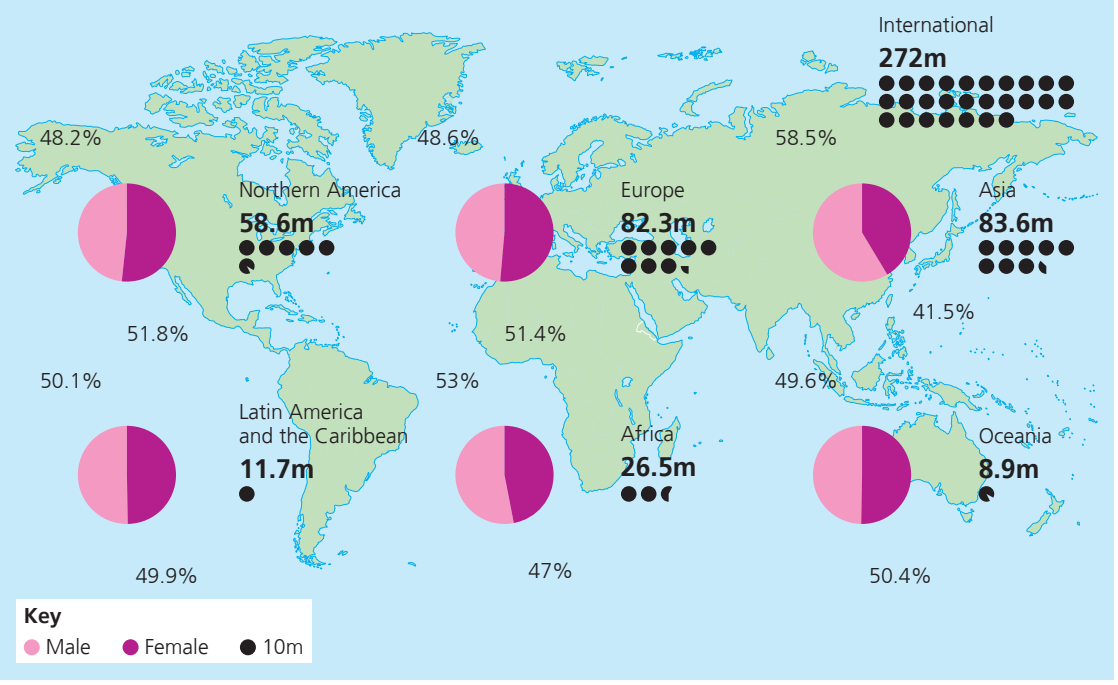


Figure 12.16 Pie charts to show total international migration by region and gender, 2019
Source: United Nations

- Draw a circle of radius *r* on graph paper

  It is important that you are able to state the scale of your proportional circle. The *area* represents the value, so the scale is in the form of:

  x units of data = 1 square unit on the graph paper.

Triangular graphs

Triangular graphs are plotted on special paper in the form of an equilateral triangle (Figure 12.17 and Figure 12.18, page 654). Although this looks on the surface to be a method that has widespread application, it is only possible to use it for a whole figure that can be broken down into three components expressed as percentages. The triangular graph cannot therefore be used for absolute data or for any figures that cannot be broken down into three components.

The advantage of using this type of graph is that the varying proportions and their relative importance can be seen. It is also possible to see the dominant variable of the three. After plotting, clusters will

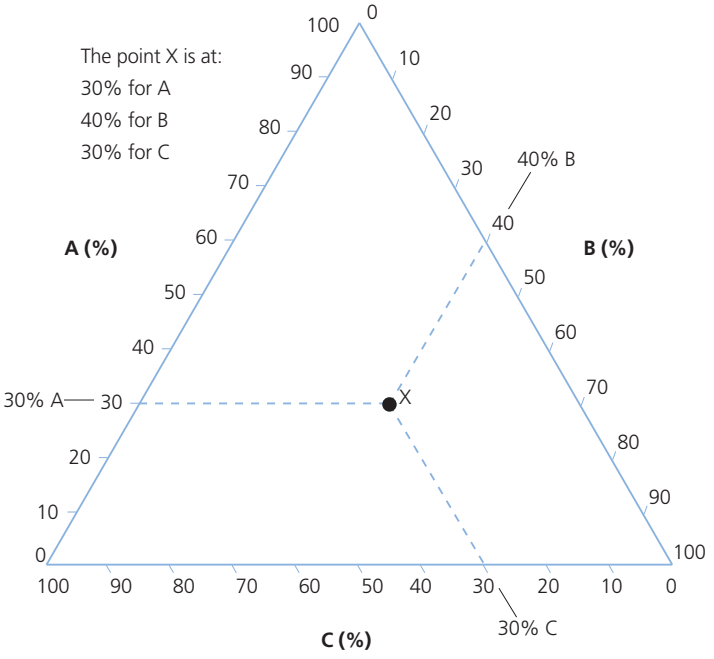


Figure 12.17 How to plot a point on a triangular graph

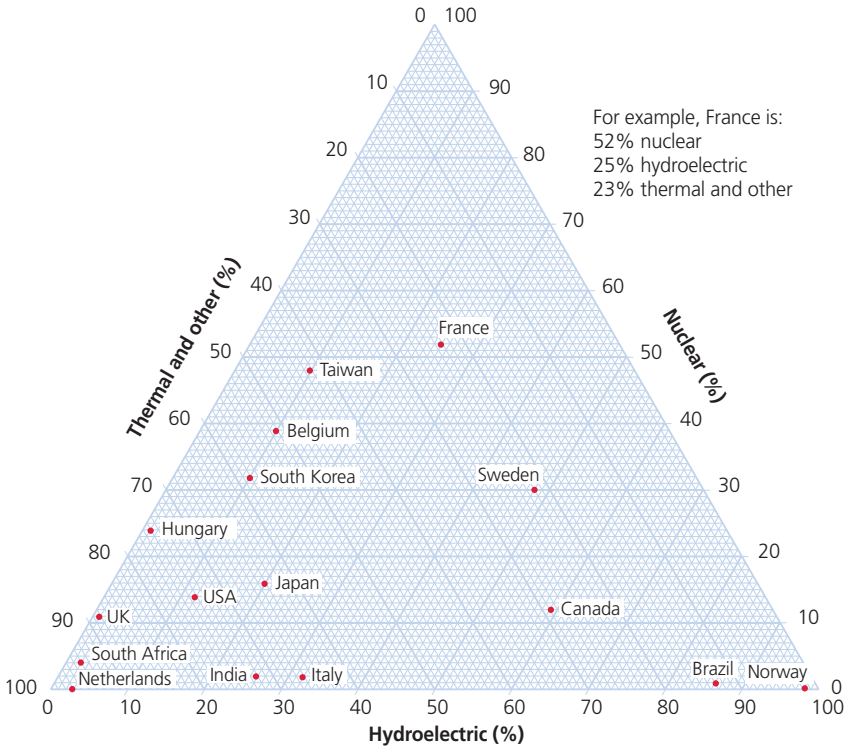


Figure 12.18 Triangular graph showing percentage of electricity produced by generating source for selected countries

sometimes emerge, enabling a classification of the items involved.

Graphs with logarithmic scales

A logarithmic graph is drawn in the same way as an arithmetic line graph except that the scales are divided into a number of cycles, each representing a tenfold increase in the range of values (Figure 12.19). If the first cycle ranges from 1 to 10, the second will extend from 10 to 100, the third from 100 to 1,000 and so on. You may start the scale at any exponent of 10, from as low as 0.0001 to as high as 1 million. The starting point depends on the range of data to be plotted.

Graph paper can be either fully logarithmic or semi-logarithmic (where one axis is on a log scale and the other is linear or arithmetic). Semi-logarithmic graphs are useful for plotting rates of change through time, where time appears on the linear axis. If the rate of change is increasing at a constant proportional rate (for example, doubling each time period) it will appear as a straight line.

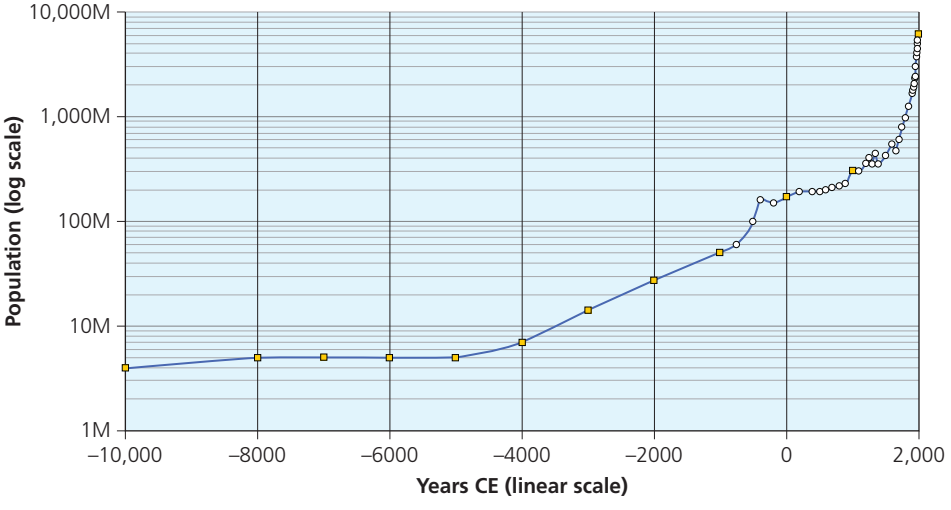


Figure 12.19 Semi-logarithmic graph showing world population growth

Table 12.1 The lip orientation in degrees of 15 corries in the Glyders of Snowdonia and 15 on the Isle of Arran

Corrie number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Glyders	30	60	45	55	75	50	80	50	10	15	10	35	45	50	85
Arran	5	5	10	55	15	30	95	5	185	70	120	40	30	115	110

Logarithmic graphs are good for showing rates of change – the steeper the line, the faster the rate. They also allow a wider range of data to be displayed.

Remember that you cannot plot positive and negative values on the same graph and that the base line of the graph is never zero, as this is impossible to plot on such a scale.

Dispersion diagrams

Dispersion graphs are used to display the main patterns in the distribution of data. The graph shows each value plotted as an individual point against a vertical scale. It shows the range of the data and the distribution of each piece of data within that range. It therefore enables comparison of the degree of bunching of two sets of data (Table 12.1). This is shown in Figure 12.20.

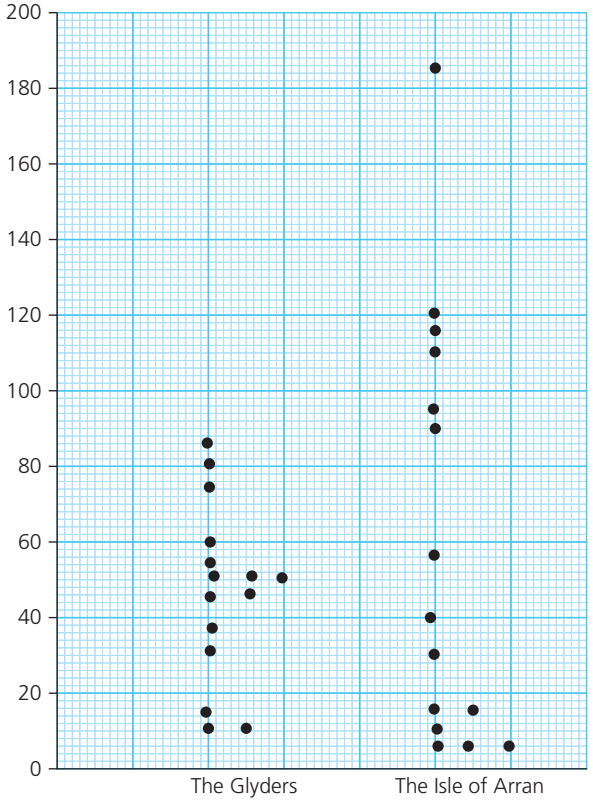


Figure 12.20 Dispersion diagram of corrie lip orientation for the Glyders (Snowdonia) and the Isle of Arran

Skills focus

Describe how the rates of human population growth change through time.

Review questions

3 What sort of graphs would you choose to display the following?

- **a** The changing length of a glacier through time
- **b** The percentage of global deaths from malaria
- **c** The energy mix for electricity production in the UK from 1990 to 2015
- **d** The top ten countries with technically recoverable shale-gas resources.

Justify your choice.

Statistical skills

Measures of central tendency

There are three such measures: arithmetic mean, mode and median.

Arithmetic mean

The mean is calculated by adding up all the values in a data set and dividing the total sum by the number of values in the data set. So,

$$\bar{X} = \frac{\sum X}{N}$$

The arithmetic mean is of little value on its own and should be supported by reference to the standard deviation of the data set.

For the data set for corrie orientation (Table 12.1) the mean orientation for:

- The Glyders corries is: 695/15 = 46.34°. This is almost exactly NE, the direction away from the Sun and warming southerly breezes.
- The Arran corries is: 890/15 = 59.34°. This shows a more easterly aspect, though still on the shady side of the mountain.

Mode

This is the value which occurs most frequently in a data set. It can only be identified if all the individual values are known.

Median

This is the middle value in a data set when the figures are arranged in rank order. There should be an equal number of values both above and below the median value. If the number of values in a data set is odd, then the median will be the

$\frac{n+1}{2}$ item in the data set.

So, for example, if the total number of items in a data set is 15, the median will be the eighth value in the rank order of the data.

If the number of values in the data set is even, the median value is the mean of the middle two values.

Distribution of the data set

It is possible that each of these measures of central tendency could give the same result, but they are more likely to give different results. For them each to give the same result the distribution of a data set would have to be perfectly 'normal', and this is extremely unlikely when using real data. It is more likely that the distribution of the data set will be skewed towards the lower end of the distribution (positive skew) or towards the upper end of the distribution (negative skew). The more it is skewed, the greater the variation in the three measures of central tendency.

None of these measures gives a reliable picture of the distribution of the data set. It is possible for two different sets of data to have the same values for mean, mode and median. Measures of the dispersion or variability of the data should therefore also be provided.

Measures of dispersion

There are three measures of dispersion or variability: range, inter-quartile range and standard deviation.

Range

This is the difference between the highest value and the lowest value in a data set. It gives a simple indication of the spread of the data.

For the data set for corrie orientation (Table 12.1, page 655) the range of orientations for:

- the Glyders corries is: 85 – 10 = 75°
- the Arran corries is: 185 – 5 = 180°

Inter-quartile range

The inter-quartile range is calculated by ranking the data in order of size and dividing them into four equal groups or quartiles. The boundary between the first and second quartiles is known as the upper quartile and the boundary between the third and fourth quartiles is the lower quartile. They can be calculated as follows:

- The upper quartile (UQ) is the value that occurs at $\frac{(n+1)}{4}$th position in the data set when arranged in rank order (from highest to lowest).
- The lower quartile (LQ) is the value that occurs at $\frac{3(n+1)}{4}$ th position in the data set.
- The difference between the upper and lower values is the inter-quartile range (IQR):
  IQR = UQ – LQ

The IQR indicates the spread of the middle 50 per cent of the data set about the *median* value, and thus gives a better indication of the degree to which the data are spread, or dispersed, on either side of the middle value.

For the data set for corrie orientation (Table 12.1, page 655) the results are shown in Table 12.2.

Table 12.2 The upper quartile, lower quartiles and interquartile range for the lip orientation of 15 corries in the Glyders of Snowdonia and 15 on the Isle of Arran

Location of corries	Upper quartile	Lower quartile	Interquartile range
The Glyders	60	30	30
Isle of Arran	110	10	100

Standard deviation

The standard deviation of a range of data is a measurement of the degree of dispersion about the *mean* value of a data set. It is calculated as follows:

- The difference between each value in the data set and the mean value is worked out.
- Each difference is squared, to eliminate negative values.
- These squared differences are totalled.
- The total is divided by the number of values in the data set, to provide the **variance** of the data.
- The square root of the variance is calculated.

$$SD = \sqrt{\frac{\sum(x-\bar{x})^2}{n}}$$

Table 12.3 Calculating standard deviation

A	B	C				
Glyder corrie orientation x	**$x - \bar{x}$**	**$(X - \bar{x})^2$**		**Arran corrie orientation Y**	**$Y - \bar{Y}$**	**$(Y - \bar{Y})^2$**
30	−16.34	267.00		5	−54.34	2,952.84
60	12.66	186.60		5	−54.34	2,952.84
45	−1.34	1.80		10	−49.34	2,434.44
55	8.66	75.00		55	−4.34	18.84
75	28.66	821.40		15	−44.34	1,966.04
50	3.66	12.40		30	−29.34	860.84
80	33.66	1,133.00		95	35.66	1,271.64
50	3.66	12.40		5	−54.34	2,952.84
10	−36.34	1,320.60		185	125.66	15,790.44
15	−31.34	982.20		70	10.66	112.64
10	−36.34	1,320.60		120	60.66	3,679.64
35	−11.4	128.6		40	19.34	374.04
45	−1.34	1.80		30	−29.34	860.84
50	3.66	12.40		115	55.66	3,098.04
85	38.66	1,494.60		110	50.66	2,566.44
$\Sigma x = 695$	$\bar{x} = 46.34$	$\Sigma(x - \bar{x})^2 = 7{,}773.4$		$\Sigma x = 890$	$\bar{x} = 59.34$	$\Sigma(x - \bar{x})^2 = 41{,}911.40$
		$\sigma = 22.76$				$\sigma = 52.86$

The standard deviation is statistically important as it links the data set to the normal distribution. Figure 12.21 shows a **normal distribution**.

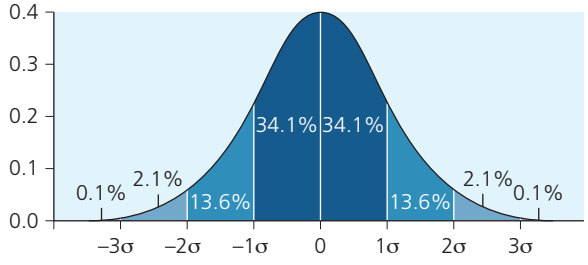


Figure 12.21 A plot of a normal distribution (or bell-shaped curve) where each band has a width of 1 standard deviation

A low standard deviation indicates that the data are clustered around the mean value and that dispersion is narrow. A high standard deviation indicates that the data are more widely spread and that dispersion is large. The standard deviation also allows comparison of the distribution of the values in a data set with a theoretical norm and is therefore of greater use than just the measures of central tendency.

Inferential and relational statistical techniques

Comparisons are made between two sets of data to see whether there is a relationship between them. Note that even if there is a relationship between two variables, this does not prove a causal link. In other words, the relationship does not prove that a change in one variable is responsible for a change in the other. For example, there may be a direct relationship between altitude and the amount of precipitation in a country such as the UK. These two variables (altitude and precipitation) are clearly linked, but a decrease in one does not automatically cause a decrease in the other – they are simply related to each other.

There are two main ways in which relationships can be shown:

- using scattergraphs
- measuring correlation using the Spearman rank correlation coefficient.

The Spearman rank correlation coefficient

This is used to measure the degree to which there is correlation between two sets of data (or variables). It provides a numerical value which summarises the degree of correlation, so it is an example of an objective indicator. Once it has been calculated, the numerical value has to be tested statistically to see how significant the result is.

The test can be used with any data set consisting of raw figures, percentages or indices which can be ranked. The formula for the calculation of the correlation coefficient is

$$R_s = 1 - \frac{6\sum d^2}{n^3 - n}$$

where d is the difference in ranking between the two sets of paired data and n is the number of sets of paired data.

The method of calculation is as follows:

- Rank one set of data from highest to lowest (highest value ranked 1, second highest 2 and so on).
- Rank the other set of data in the same way.
- Beware of tied ranks. In order to allocate a rank order for such values, calculate the 'average' rank they occupy. For example, if there are three values which should all be placed at rank 5, add together the ranks 5, 6 and 7 and divide by three, giving an 'average' rank of 6 for each one. The next value in the sequence will be allocated rank 8.
- Calculate the difference in rank (d) for each set of paired data.
- Square each difference.
- Add the squared differences together and multiply by 6 (A).
- Calculate the value of $n^3 - n$ (B).
- Divide A by B, and take the result away from 1.

The answer should be a value between +1.0 (perfect positive correlation) and –1.0 (perfect negative correlation).

$\Sigma(D - M)^2 = 562.5$

$$R_s = 1 - \frac{6 \times 562.5}{13^3 - 13}$$

$R_s = -0.55$

$\Sigma(D - H)^2 = 774$

$$R_s = 1 - \frac{6 \times 774}{15^3 - 15}$$

$R_s = -0.38$

Some words of warning

- You should have at least 10 sets of paired data, as the test is unreliable if n is less than 10.
- You should have no more than 30 sets of paired data or the calculations become complex and prone to error.

Table 12.4a Calculating Spearman's rank correlation coefficient (R_s) for doctors/1,000 of population and confirmed malaria cases

Country	Doctors/1,000 population (D)	Rank (D)	Confirmed malaria cases/1,000 population (M)	Rank (M)	(D – M)	(D – M)²
Papua New Guinea	0.6	13	42.73	1	12	144
Solomon Islands	0.22	6	42.00	2	4	16
Cambodia	0.17	10	12.77	3	7	49
Vanuatu	0.12	11	8.92	4	7	49
Myanmar	0.61	4	5.99	5	–1	1
Laos	0.18	8.5	5.60	6	2.5	6.25
Indonesia	0.2	7	1.35	7	0	0
Timor-Leste	0.07	12	0.85	8	4	16
Thailand	0.39	5	0.49	9	–4	16
Vietnam	1.19	2	0.18	10	–8	64
Malaysia	1.2	1	0.10	11	–10	100
Philippines	1.15	3	0.06	12	–9	81
Micronesia	0.18	8.5	0.00	13	–4.5	20.25

Table 12.4b Calculating Spearman's rank correlation coefficient (R_s) for COVID-19 deaths up to 12 March 2020 and percentage of GDP spent on healthcare

Country	Number of COVID-19 deaths by 12th March 2020 (D)	Rank (D)	Percentage of GDP spent on healthcare (H)	Rank (H)	(D – H)	(D – H)²
China	3,173	1	5.0	14	–13	169
Italy	827	2	8.9	10	–8	64
Iran	354	3	8.1	11	–8	64
S. Korea	66	4	7.3	13	–9	81
Spain	48	5.5	9.0	9	–3.5	12.25
France	48	5.5	11.5	3	2.5	6.25
USA	29	7	17.1	1	6	36
Japan	15	8	10.9	5	3	9
Iraq	7	9	3.3	15	–6	36
UK	6	10	9.8	8	2	4
Netherlands	5	11	10.4	6.5	4.5	20.25
Switzerland	4	12	12.2	2	10	100
Austria	3	13.5	10.4	6.5	7	49
Lebanon	3	13.5	8.0	12	1.5	2.25
Germany	2	15	11.1	4	11	121

Source: CIA World Factbook

- Too many tied ranks can interfere with the statistical validity of the exercise, although it is appreciated that there is little you can do about the 'real' data collected.
- Be careful about choosing the variables to compare – do not choose dubious or spurious sets of data.

Interpreting the results

When interpreting the results of the Spearman rank test, consider the following points:

What is the direction of the relationship? If the calculation produces a positive value, the relationship is positive, or direct. In other words, as one variable increases, so does the other. If the calculation produces a negative value, the relationship is negative, or inverse.

How statistically significant is the result? When comparing two sets of data, there is always a possibility that the relationship shown between them has occurred by chance. The figures in the data sets may just happen to have been the right ones to bring about a correlation. It is therefore necessary to assess the statistical significance of the result. In the case of the Spearman rank test, the critical values for R_s must be consulted. These can be obtained from statistical tables, though Table 12.5 also shows some examples.

According to statisticians, if there is a >5 per cent possibility of the relationship occurring by chance, the relationship is not significant. This is called the rejection level. The relationship could have occurred by chance more than 5 times in 100, and this is an unacceptable level of chance. If there is a <5 per cent possibility, the relationship is significant and therefore meaningful.

Table 12.5 Critical values for R_s

n	0.05 (5 per cent) significance level	0.01 (1 per cent) significance level
10	± 0.564	± 0.746
12	0.506	0.712
14	0.456	0.645
16	0.425	0.601
18	0.399	0.564
20	0.377	0.534
22	0.359	0.508
24	0.343	0.485
26	0.329	0.465
28	0.317	0.448
30	0.306	0.432

If there is a < 1 per cent possibility of the relationship occurring by chance, the relationship is very significant. In this case, the result could only have occurred by chance 1 in 100 times, and this is very unlikely.

How does this work? Having calculated a correlation coefficient, examine the critical values given in Table 12.5 on page 659 (ignore the positive or negative sign). If your coefficient is greater than these values, the correlation is significant at that level. If your coefficient is smaller, the relationship is not significant at that level.

In our example, suppose we have calculated an R_s value of –0.55 from 13 sets of paired data; 0.50 is greater than the critical value at the 0.05 (5 per cent) level, but not that at the 0.01 (1 per cent) level. In this case, therefore, the relationship is significant at the 0.05 (5 per cent) level, but not at the 0.01 (1 per cent) level.

Chi-square test

This technique is used to assess the degree to which there are differences between a set of collected (or observed) data and a theoretical (or expected) set of data, and the statistical significance of the differences.

The observed data (O) are those that have been collected either in the field or from secondary sources. The expected data (E) are those that would be expected according to the theoretical hypothesis being tested.

Normally, before the test is applied it is necessary to formulate a null hypothesis. In the example given here, the null hypothesis would be that there is no significant difference between the observed and expected data distribution. The alternative to this would be that there *is* a difference between the observed and expected data, and that there is some factor responsible for this. An example is given below.

A group of students investigated the orientation of corries in the mountains of Snowdonia. The students wanted to investigate whether there was a preferred orientation.

They measured the orientation of 40 corries and placed them in one of four groups. This is shown in the Observed (O) column in Table 12.6.

It looks like there is a preferred direction, but as this could be due to chance, a chi-square test was carried out.

They created a null hypothesis:

There is no significant difference between the observed orientation of the corries and the expected random orientation.

Table 12.6 was completed.

In the column headed Observed (O) are listed the numbers of corries in each of the categories of orientation. The total number of corries is 40. The column labelled Expected (E) contains the list of *expected* frequencies (E) in each of the categories, assuming that the corries are evenly orientated. In the column O – E, each of the expected frequencies is subtracted from the observed frequencies, and in the last column the result is squared. The relevant values are then inserted into the expression for chi-square,

$$\chi^2 = \sum \frac{(O-E)^2}{E}$$

and the result is 8.6.

The aim of a chi-square test, therefore, is to find out whether the observed pattern agrees with or differs from the theoretical (expected) pattern. This can be measured by comparing the calculated result of the test with its level of significance.

To do this the number of degrees of freedom must be determined using the formula $(n - 1)$, where n is the number of observations, in this case the number of cells which contain observed data (4). So, $4 - 1 = 3$. Table 12.7 gives the distribution of chi-square values for these degrees of freedom.

Table 12.6 Calculating chi-square

Orientation	Observed (O)	Expected (E)	(O – E)	$(O - E)^2$	$\frac{(O-E)^2}{E}$
316° – 045°	16	10	6	36	3.6
046° – 135°	13	10	3	9	0.9
136° – 225°	6	10	–4	16	1.6
226° – 315°	5	10	–5	25	2.5
Total	40	40			$X^2 = 8.6$

Table 12.7 Critical values of chi-square

Degrees of freedom	Significance level	
	0.05	0.01
1	3.84	6.64
2	5.99	9.21Z
3	7.82	11.34
4	9.49	12.28
5	11.07	15.09
6	12.59	16.81
7	14.07	18.48
8	15.51	20.09
9	16.92	21.67
10	18.31	23.21
11	19.68	24.72
12	21.03	26.22
13	22.36	27.69
14	23.68	29.14
15	25.00	30.58
16	26.30	32.00
17	27.59	33.41
18	28.87	34.80
19	30.14	36.19
20	31.41	37.57
21	32.67	38.93
22	33.92	40.29
23	35.17	41.64
24	36.42	42.98
25	37.65	44.31
26	38.88	45.64
27	40.11	46.96
28	41.34	48.28
29	42.56	49.59
30	43.77	50.89
40	55.76	63.69
50	67.51	76.15
60	79.08	88.38
70	90.53	100.43
80	101.88	112.33
90	112.15	124.12
100	124.34	135.81

Note: When there are A rows and B columns respectively, degrees of freedom = (A – 1) × (B – 1). If there is only one row, then there are (B – 1) degrees of freedom.

Then there are the levels of significance. There are two levels of significance: 95 per cent and 99 per cent (0.05 and 0.01 in Table 12.7). At 95 per cent there is a 1 in 20 probability that the pattern being considered occurred by chance, and at 99 per cent there is only a 1 in 100 probability that the pattern is chance. The levels of significance can be found in a book of statistical tables. They are also known as confidence levels.

If the calculated value is the same or greater than the values given in the table, then the null hypothesis can be rejected and the alternative hypothesis accepted.

In the case of our example, the value of chi-square is 8.6. This falls between the values for the 0.05 per cent and 0.01 per cent. Because the value is greater than that of the 0.05 per cent significance level then the null hypothesis can be rejected.

Some further points on this technique:

- The numbers of both observed and expected values must be large enough to ensure that the test is valid. Most experts state that there should be a minimum of five.
- The number produced by the calculation is itself meaningless. It is only of value for use in consulting statistical tables.
- As in a Spearman rank test, only significance (or confidence) levels of 95 per cent and 99 per cent should be considered when rejecting the null hypothesis. Levels of confidence greater than these simply allow the null hypothesis to be rejected with even greater confidence.
- It is strongly recommended that you do not apply the test to more than one set of observed data because the mathematics becomes too complex.

It is important to note that the chi-square test should only be used as a support for your own ideas on the geographical significance of your study. All results, and the statistical analysis of them, should be related to the original hypothesis and/or the established theory in that aspect of the subject.

Your results may support established theory or your hypothesis, or they may not. If they do not, there may be some reason or factor you can identify that is responsible. This could lead to further studies. Above all, your investigation should make geographical sense. This is far more important than demonstrating your ability to use mathematics or statistics.

Skills focus

a What conclusions can we draw from the distributions of corrie orientations shown in Table 12.1 (page 655)?

b The standard deviations for the two sets of corries are different. Using all the analysis so far, describe the two sets of data.

Review questions

4 How do the two results for R_s calculated in Tables 12.4a (page 658) and 12.4b (page 659) support the statement below?

'The Spearman's rank correlation coefficient test only tells us whether there is a correlation. It does not imply causation.'

5 Which statistical technique would be best suited to study the following?

a The frequency and latitude of local water conflicts.

b The change in air pollution away from an urban centre.

c The orientation of the long axes of drumlins in two separate areas.

d Changing pebble size from one end of a beach to another.

Justify your choice of technique.

ICT skills

Use of electronic databases

There are a variety of electronic databases that can be accessed by geographers (Figure 12.22). Schools and colleges often have access to Digimap, an interactive map database that contains Ordnance Survey maps which themselves contain vast amounts of information.

When you use the internet as a research tool you must always ask the questions:

- who published the information (organisation/agency/individual)
- who wrote the information (expert/interested individual)
- the age of the material
- why the material exists (academic research/special interest groups)
- how one-sided (biased) are the arguments in the information.

Use of innovative sources of data such as crowdsourcing and 'big data'

Crowdsourcing

It is difficult to analyse individual people's behaviour and attitudes using traditional large-scale data sources such as population censuses and social surveys, because they tend to deal with large anonymous groups rather than individuals. They focus on easily measured

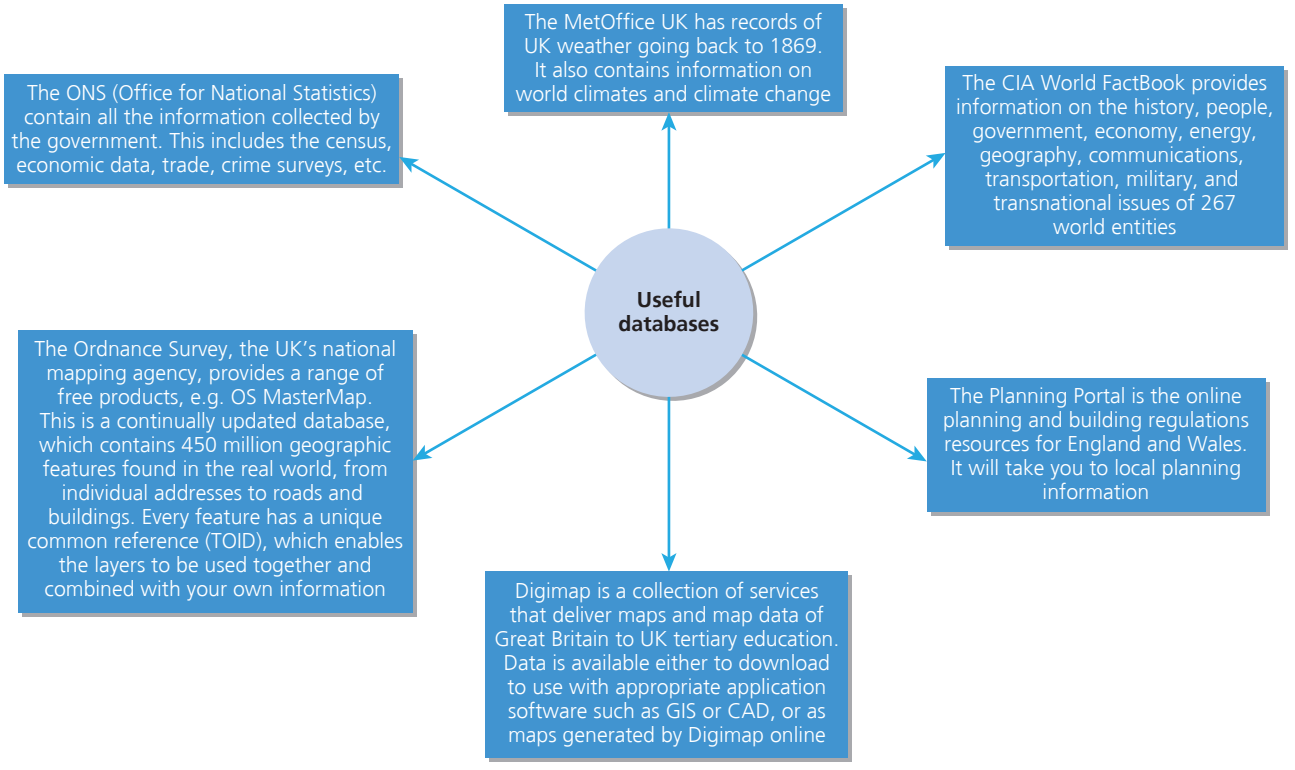


Figure 12.22 Useful databases

attributes and characteristics of a population, rather than attitudes and behaviours. They also only offer a snapshot rather than a continuously changing view. However, new data sources (such as that available through social networks) contain a wealth of information about people's geographical behaviour at an individual level; they have the potential to revolutionise our understanding of social phenomena. These sources are commonly referred to as geo-spatial crowdsourced data (GSD) or **volunteered geographical information.**

GSD is increasingly being used to study individual people's daily behaviours and explore how these dynamics can potentially lead to civil unrest. It is also being used to look at the way people respond to natural disasters.

Since the start of the twenty-first century there has been a sequence of unfortunate natural disasters, ranging from the Indian Ocean earthquake and tsunami (26 December 2004) to the wildfires in south-western Australia (2019/20). These have all focused international attention on the immediate need for maps and geospatial data on the impacted areas and the critical role of geographically-distributed information-sharing communities in providing that information.

User-generated platforms such as Twitter, Facebook, Flickr and YouTube have large volumes of information with geographic footprints or geographic associations that can be used, even though the original contributors to these platforms may not have intended the information to be geographic or to have any specific geographic purpose.

Example of crowdsourced data use – relief of Nepal earthquake (2015)

Two important questions that needed to be answered immediately after the Nepalese earthquake hit on 25 April 2015 were: Who needs help? And where? Relief efforts had to get supplies and resources to the parts of the country most desperately in need, but it was difficult to know where to deploy resources because there was no systematic plan or data in place to help make such decisions. Even some of the most basic information needs, like detailed road maps and locations of critical assets, were not available.

Ushahidi, a crowd-sourcing platform, was developed in Kenya, initially to map reports of post-election violence in 2008. Headquartered in Nairobi, it is a social enterprise that provides software and services to help improve the bottom-up flow of information.

Ushahidi helped Nepalese agencies see what was happening on the ground as the disaster unfolded. Organisations could see who needed what kind of relief, when and where. In one case, 18 trekkers from a relief centre were stranded between dangerous landslides in Dobhan. They were evacuated within 48 hours, and food supplies distributed to more than 60 others living in the area.

Elsewhere, a report called for urgent relief in Chayarsaba. Here roads were blocked or unstable and drinking water supplies were halted because the supply pipes were damaged by landslides. Using Ushahidi, the Kathmandu Living Labs team was able to verify the location of the reported needs and what exactly was needed. Roads were cleared, new water pipes were secured and installed by the community, and multiple food drops were confirmed.

Unfortunately, the sheer amount of crowdsourced information available during disasters can be difficult for aid organisations to handle; verification becomes an issue. Reports are often duplicated or inaccurate.

Big data

Geographers have always used large data sets, for example, weather data or population census data. Because of expense and logistical problems this could only be collected periodically and analysed partially. There is now a deluge of data which can be held indefinitely in cloud-based servers. This is called 'big data' (Figure 12.23, page 664).

Analysing this data has, until recently, been limited in its use. One recent development, geographical information systems (GIS), has tools that enable spatial data to be collected, stored, archived, manipulated, analysed and presented (Figure 12.24, page 664). Even this is limited in scope.

Other uses of ICT

Digital technology is becoming ever more portable and flexible in its use. A mobile device is able to help you gather field information (via GPS and geolocation information, photographs, video clips, annotated sketches, voice recorder, etc.).

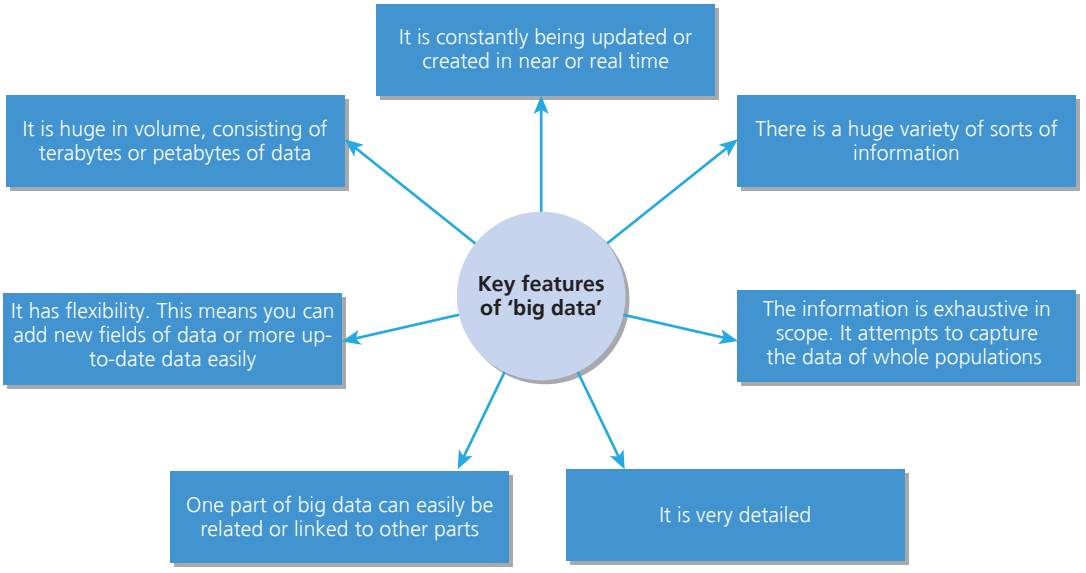


Figure 12.23 Key features of 'big data'

Some of the more 'traditional' ICT skills (for example, Excel spreadsheets) enable you to analyse data, while word processing software, along with other technology such as digital maps and remote sensing (Multimap, Memory-map, Google maps, Google Earth), allows you to present data in an interesting and meaningful way.

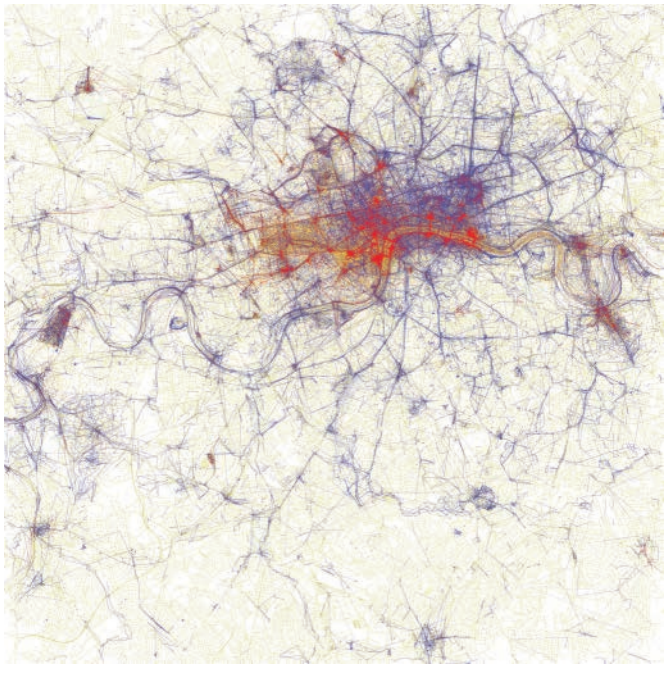


Figure 12.24 Tourist destinations made visible through mass user-generated data from the photograph network Flickr. Blue points on the map are pictures taken by locals (people who have taken pictures in this city dated over a range of a month or more). Red points are pictures taken by tourists (people who seem to be a local of a different city and who took pictures in this city for less than a month)

12.3 Fieldwork investigation

You are expected to undertake four days of fieldwork during your A-level course. Ideally this should cover topics in both human and physical geography. It will assess your ability to apply your knowledge and skills to unseen information and resources, with reference to fieldwork that you have undertaken.

You are required to undertake an independent investigation that demonstrates the required fieldwork knowledge, skills and understanding. You may incorporate field data and/or evidence from field investigations collected on your own or as part of a group.

This independent investigation must:

- be based on a question or issue that you have developed relating to any aspect of the course content
- include data and evidence collected from field investigations
- include results of your own research and relevant secondary data
- include presentation and analysis of collected data and conclusions drawn from this analysis.

You need to be able to:

- prepare for fieldwork, including background reading, drawing up aims and objectives for the enquiry, planning research in the field and from

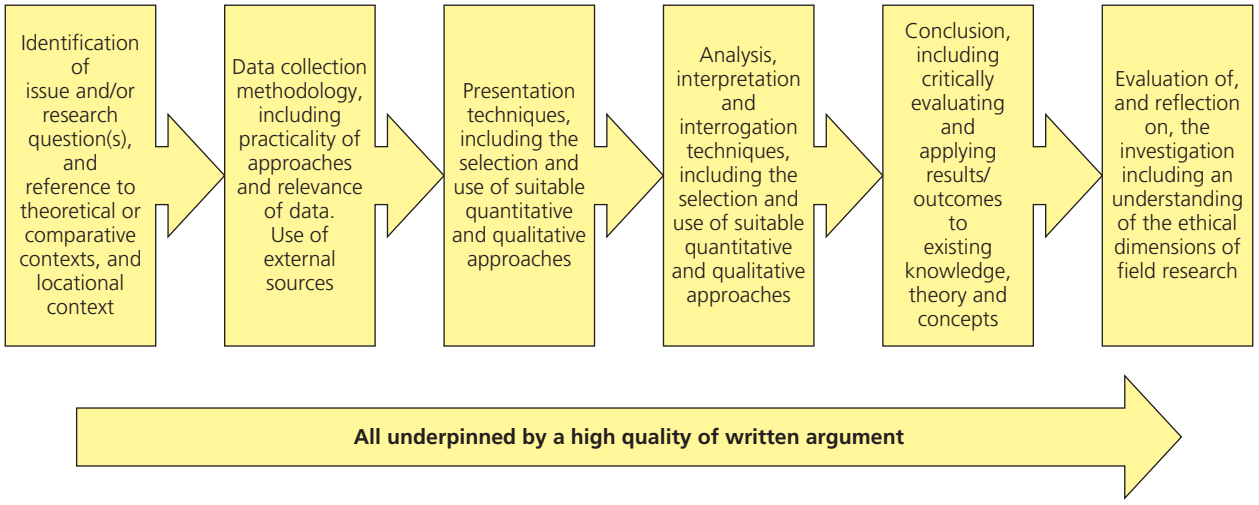


Figure 12.25 The fieldwork pathway

secondary sources, using data sampling techniques and carrying out health and safety procedures
- collect primary data in the field and use relevant secondary data sources
- process and present data using relevant graphical and cartographical techniques
- analyse data, including using statistical techniques where relevant
- draw conclusions related back to the original aims and objectives and link these conclusions to both the place studied and the general ideas forming the basis of the enquiry
- review the success, or otherwise, of all stages of the enquiry
- consider how the enquiry could be further developed.

After carrying out your fieldwork, you are required to write and submit a report of 3,000–4,000 words based on one or more aspects of your investigations. This will be marked by teachers and moderated by AQA. It is important to note that the topic chosen must be linked to an area of the AQA A-level specification.

The fieldwork pathway – Choosing your fieldwork topic

Figure 12.25 shows the fieldwork pathway. Figure 12.26 (page 666) shows some of the things you must consider when choosing your fieldwork topic. Everybody's situation is unique; you must choose a topic where the aims can be met. Clearly, advice from your teachers will be invaluable. The main consideration is that this unit must involve enquiry work outside the classroom that includes data collection in the field. It might include, for example, data collected in specialist field study centres, work experience settings, internet research and use of libraries or archives.

Geographical information systems and fieldwork

GIS are an effective mapping tool within geographical fieldwork. Figure 12.27 (page 666) summarises the key aspects of GIS and how they can be used in fieldwork.

A number of websites and associated programs have great potential for use in fieldwork. They include Multimap, Aegis, Digital Worlds, Quikmaps and Google Earth. Many allow you to use maps and photos together (Figure 12.28, page 667); others allow annotations to be added. Some types of software require a degree of training, whereas others are more straightforward to use. It all depends on your own level of ICT skills.

Global positioning systems (GPS) technology can be useful in fieldwork. Data can be recorded at points along a transect or other sampling system, and the position of each point recorded at the same time. GPS-located data can then be fed into a GIS program, bringing data recording and mapping together.

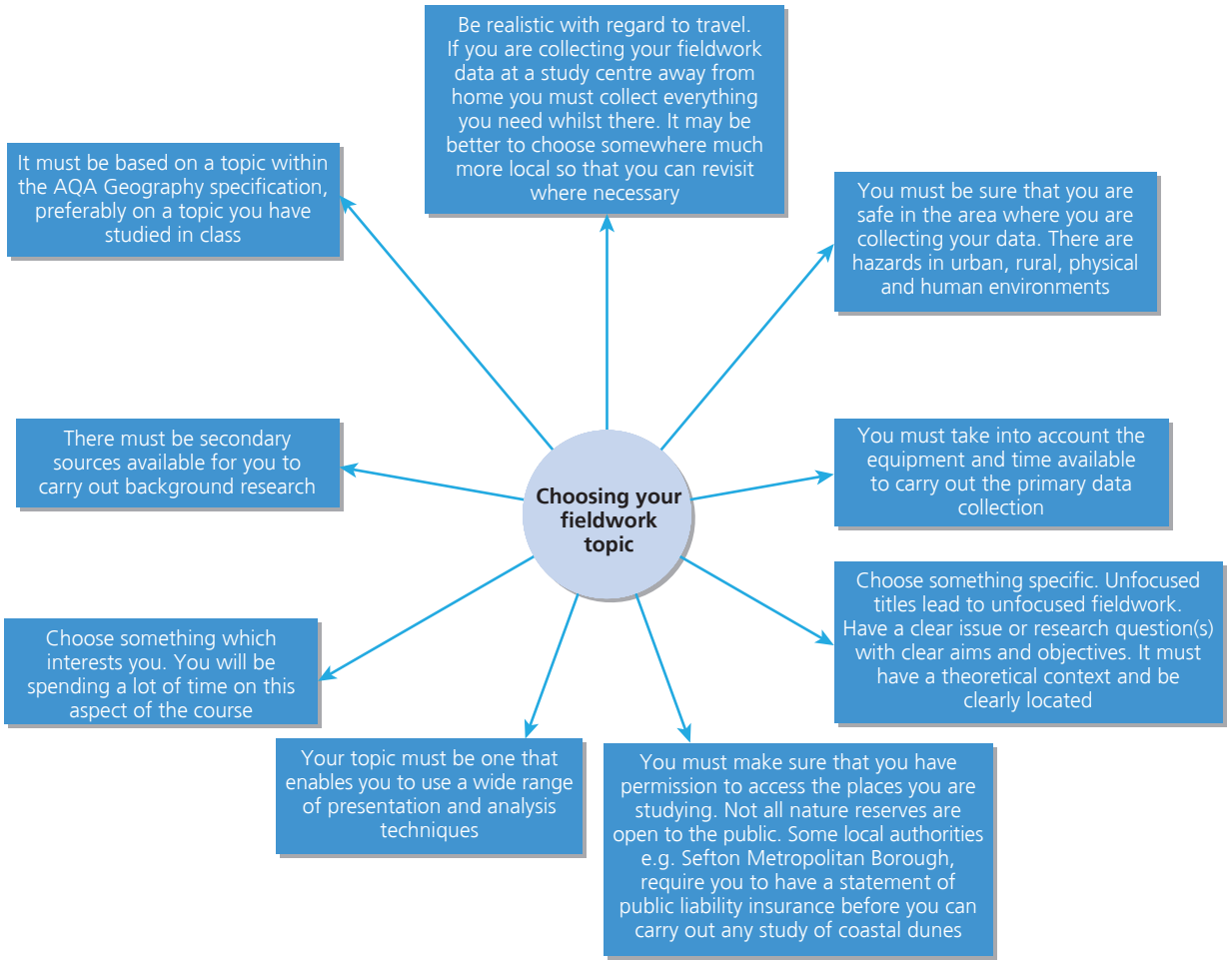


Figure 12.26 Choosing your fieldwork topic

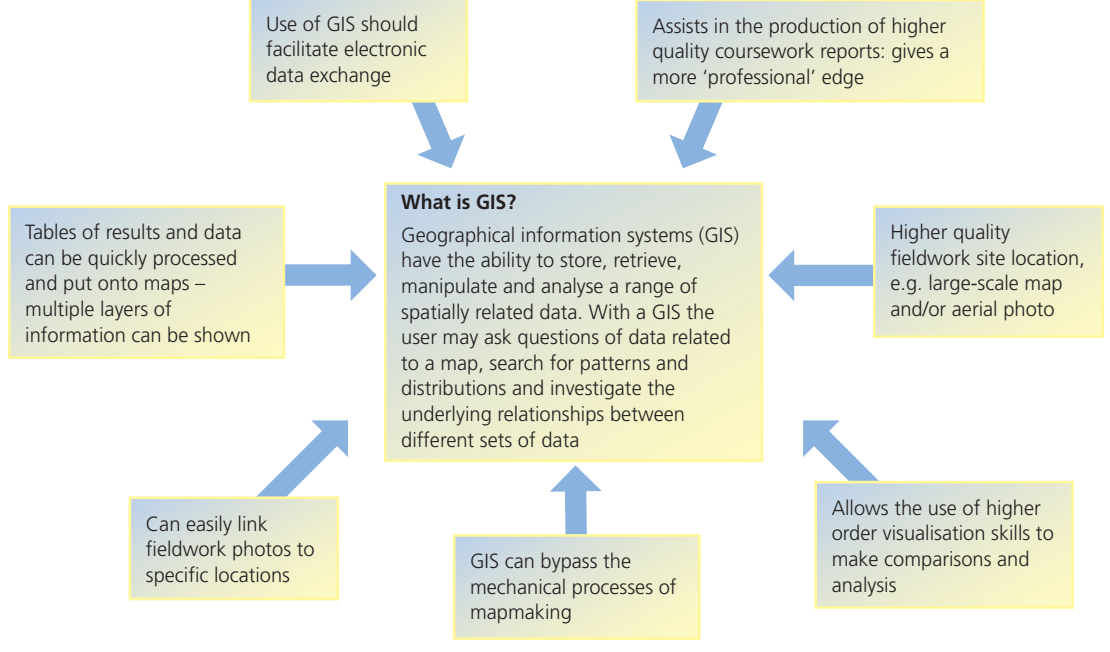


Figure 12.27 Maximising the use of GIS in fieldwork

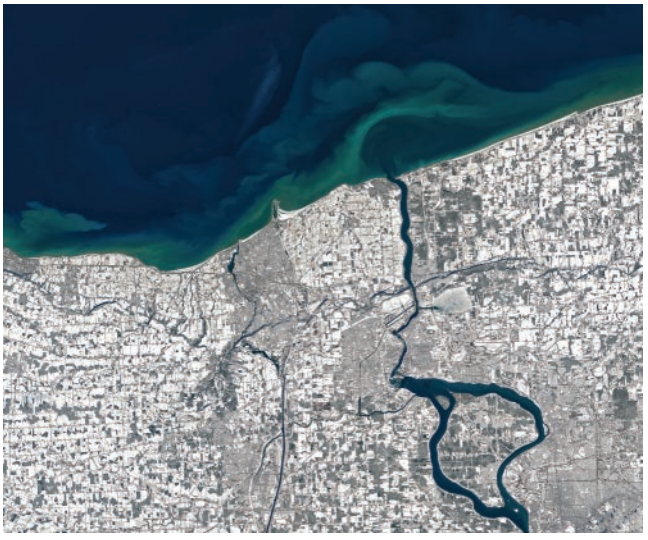


Figure 12.28 Operational Land Imager (OLI) map data superimposed on a Landsat 8 image showing the snow-covered landscape around the Niagara River, mid-winter 2021
Source: NASA

An example of fieldwork

The following suggestion for a piece of fieldwork is based on 'Changing places'.

Figure 12.29 shows the location of a planned housing development on the rural/urban fringe of Shrewsbury. Outline planning permission has been granted to build 15 houses on arable land on the western edge of the large village of Bayston Hill.

Issue/research question

In early 2012 plans to change Bayston Hill's status to a town were abandoned following concerns by residents. It is already the largest village in Shropshire, even larger than any of the county's market towns. The concerns expressed were related to the fact that many of the people in the village were concerned that the village would grow more quickly and lose its rural 'feel'. On the other hand, Shrewsbury has been described as being 'in the grip of an overwhelming housing shortage'. There is a great need for new houses.

The issue is: how can these two be reconciled?

Data collection

Secondary sources of data will include:

- the recent censuses for Bayston Hill to see how it has grown over time
- details of the housing development, which can be found through the local authority planning portal (Figure 12.30, page 668)
- Ordnance survey and other map collections including GIS.

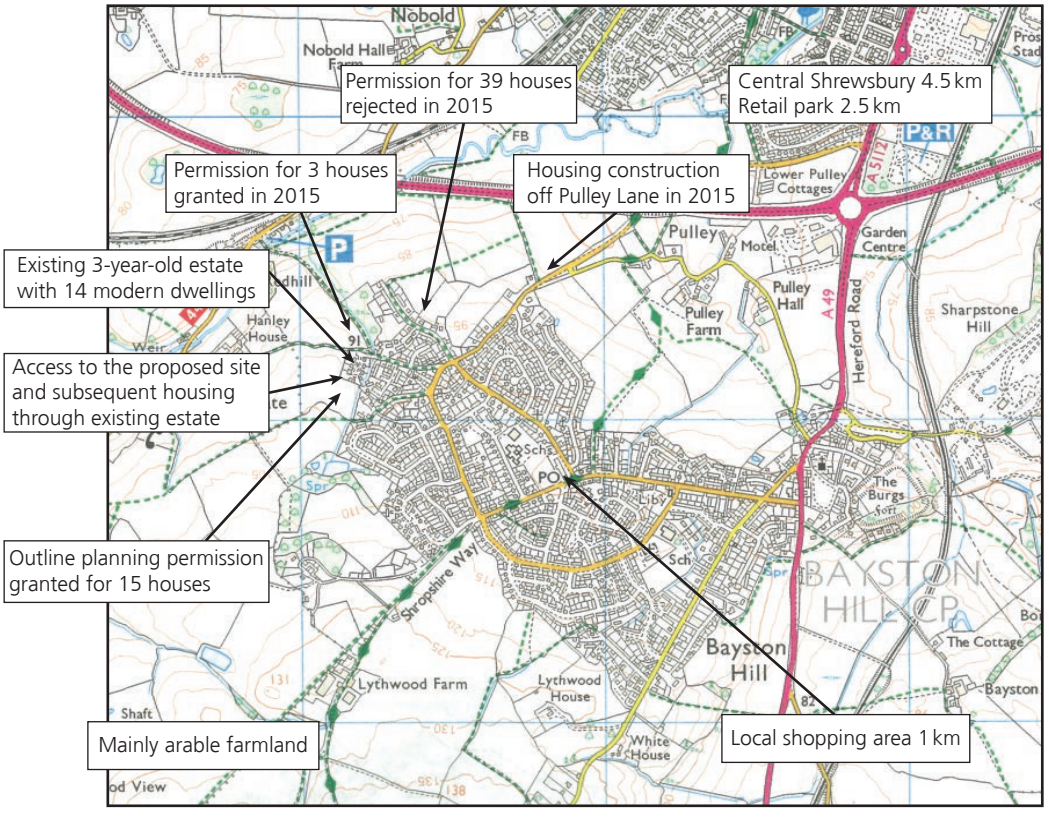


Figure 12.29 Location of the proposed housing development. © Crown copyright and database rights, 2021, Hodder Education, under licence to Ordnance Survey. Licence number 100036470

Figure 12.30 Plan of the proposed housing development

Primary data collection could include:

- carrying out surveys of the interested parties (landowner, developer, planning officer, local councillor and residents of the nearby estate)
- questionnaires about attitudes to the changes could be carried out with a larger number of the village's inhabitants. This could be in the form of an age-related stratified sample
- traffic surveys could be carried out to look at concerns over access to the site by developers and subsequent occupiers of the houses (Figure 12.31).

Presentation and analysis

The interview answers could be coded, depending on what the responses were. Questionnaire responses can be easily summarised in a table and appropriate graphs drawn. The analysis will depend on what the researcher finds.

Figure 12.31 Aerial view of new housing development after completion

Conclusions

The conclusion to this report depends on what exactly the researcher finds. Any summary should include the benefits of having more housing in this area as opposed to the perceived damage that such developments will have on this place.

Suggestions for fieldwork titles/ questions

Your choice of fieldwork topic is very important. It not only carries 20 per cent of the final course assessment, but also takes up a significant amount of time. Choose something that interests you.

Sometimes there are constraints as a result of where you live, for example, glacial work may be difficult if you live in south-east England. As well as this, some aspects of the specification lead more easily to fieldwork investigation than others.

The following list of fieldwork suggestions is not exclusive. The suggested locations are just a guide that can be changed to suit your area and circumstances.

Water and carbon cycles

- Assess the extent to which human activity has affected the runoff of the River Harbourne in south Devon.
- To what extent has the draining of blanket bogs and the burning of moorland heather above Hebden Bridge, West Yorkshire, contributed to the recent flooding of the town?

- Assess the potential for carbon sequestration and storage in the Sankey Valley area of the Mersey Forest.

Coastal systems and landscapes

- Assess the success of coastal defences at Broomhill Sands, Kent.
- To what extent has longshore drift contributed to the development of Dawlish Warren sand spit in Devon?
- Assess the extent to which geology is the main factor in determining the coastal landscape of south-west Anglesey.

Figure 12.32 The sea wall, rock revetment (rip-rap) and groynes of Broomhill Sands

Glacial systems and landscapes

- Assess the extent to which glacial erosion dominates all other processes in the shaping of the landscape of Great Langdale, Cumbria.
- To what extent has fluvioglaciation contributed to the landforms of the Fylde peninsula in Lancashire?
- Assess the evidence that suggests that Bradford Kaims in Northumberland is an esker.

Ecosystems under stress

- How do different agricultural land uses affect the biodiversity of places in north Northumberland?
- Assess the extent to which local factors have affected the ecological development of Wistman's Wood (Dartmoor, Devon) (Figure 12.33).

Figure 12.33 Wistman's Wood, Dartmoor

- Assess the extent to which biodiversity increases across coastal sand dunes from the high-water mark to the dune heath.

Global systems and global governance

- Assess the extent to which the main supermarket chains in Taunton, Somerset are reliant on overseas sources to provide year-round variety in the supply of fresh produce.
- Assess the impact of TNC retail chains on the retail environment of central Dorchester.
- To what extent are farms in the Breckland region of Norfolk reliant on migrant labour?

Changing places

- Assess the extent to which a visitor's view of Blackpool differs from that represented in the media.
- To what extent is there evidence of gentrification in Hebden Bridge, West Yorkshire?
- How far has the topography and physical geography of Totnes in Devon contributed to the character of the place?

Contemporary urban environments

- How far do urban forms affect the local climate of Croydon?
- To what extent is the SUDS development in Grovelands Park Wetlands in London successful?
- To what extent can the Jewellery Quarter be considered to have a unique urban landscape within the much larger area of central Birmingham?

Population and the environment

- Assess the extent to which there is a close relationship between soil type and land use in the Vale of Pickering, as exemplified by the parish of Snainton in North Yorkshire.
- To what extent can the village of Slapton, south Devon, be considered a declining village?

Figure 12.34 The village of Slapton, South Devon

- To what extent are the differences in the health of people living in Moss Side and Chorlton-cum-Hardy in Manchester caused by differences in the socio-economic status of the areas?

Resource security

- Assess the feasibility and impact of a proposed wind farm on Bosley Minn, Cheshire.
- Assess the impact that coal mining in the past has had on the town of Amble, Northumberland.
- Assess the extent to which the construction of Lathom Solar Farm, West Lancashire, has impacted upon both the human and physical landscapes.

Further reading

The Royal Geographical Society (RGS) produces a comprehensive guide to fieldwork, *Fieldwork techniques*, which gives information and instructions for a range of techniques that can be used to carry out geographical fieldwork investigations in different locations and settings. www.rgs.org

The Barcelona Fieldwork centre has produced a series of geography resources which are available free. http://geographyfieldwork.com

Dr Bethan Davies of AntarcticGlaciers.org has written a comprehensive guide to A-level fieldwork called 'A-Level Geography Fieldwork Investigation'. Although some of this is really aimed at university-level fieldwork, it does give sound advice on how to go about designing and carrying out fieldwork.

The Field Studies Council (FSC) has a website dedicated to A-level fieldwork. This is constantly being updated to match the changing requirements of A-level Geography. www.geography-fieldwork.org

The Geography Review published by Philip Allan for Hodder Education has a regular feature on geographical skills.

The Ordnance Survey has a large range of resources based on maps. They range from simple map skills to complex use of GIS. www.ordnancesurvey.co.uk/mapzone/gis-zone

The Office for National Statistics (ONS) produces a range of resources including the census results. This is useful background material when completing a 'Place' study. www.ons.gov.uk

AQA have produced a fieldwork toolkit comprising a package of resources that outline the tools needed to plan exciting fieldwork.

The resources are primarily designed for teachers but give students a clear understanding of the requirements of the A-level specifications, along with detailed guidance on:

- linking fieldwork to areas of the specification content
- the enquiry process
- the assessment of fieldwork.

www.aqa.org.uk/resources/geography/as-and-a-level/geography/teach/fieldwork-toolkit

Glossary

Abrasion – Where material carried by moving water or wind (or ice) hits exposed rock surfaces, thus wearing them away. Often referred to as a sandblasting or sandpapering effect.

Abstraction (water) – The removal of water available in the environment, either permanently or temporarily, from rivers, lakes, canals, reservoirs or from underground aquifers. **Over-abstraction** refers to the removal of too much water from a source, upsetting the ecological and hydrological balance of that store.

Acid rain – Rain that has been made acidic by certain pollutants in the air.

Adaptation – Any alteration or adjustment in the structure or function of an organism or system which enables it to survive better in changing environmental conditions. In the context of hazards, adaptation is the attempts by people or communities to live with hazard events. By adjusting their living conditions, people are able to reduce their levels of vulnerability. For example, they may avoid building on sites that are vulnerable to storm surges but stay within the same area.

Agents of change – These are the people who impact on a place whether through living, working or trying to improve that place. Examples would include residents, community groups, corporate bodies, multinational corporations, central and local government and the media.

Agglomeration – When companies in similar industries locate near to each other because of the benefits gained by sharing ideas and resources – called 'agglomeration economies'.

Agricultural productivity – The ratio of agricultural outputs to agricultural inputs.

Agriculture – The science or practice of farming including cultivation of the soil for the growing of crops and the rearing of animals to provide food, wool and other products.

Albedo – The reflectivity of a surface. It is the ratio between the amount of incoming insolation and the amount of energy reflected back into the atmosphere. Light surfaces reflect more than dark surfaces and so have a greater albedo.

Anaerobic (conditions) – Conditions where there is no air and little oxygen, such as waterlogging in soil.

Antarctic Convergence – A zone continuously encircling Antarctica where cold northward flowing Antarctic waters meet the relatively warmer waters of the sub-Antarctic.

Anthropogenic CO_2 – Carbon dioxide generated by human activity.

Anthropogenic factors – Those factors that are caused by human beings.

Appropriate technology – Technology that is suitable to the social and economic conditions of the geographic area in which it is to be used; it is environmentally sound and promotes self-sufficiency on the part of those using it.

Arid – The climate of an area that receives less than 250 mm of precipitation per year.

Aridity index – The ratio between precipitation (P) and potential evapotranspiration (PET).

Ash – Dust-sized particles of rock produced by the explosive eruption of some volcanoes. This material may be carried in the air for long distances from the volcano which formed it.

Asthenosphere – That part of the mantle which lies directly beneath the lithosphere. In the original plate tectonic theory, convection currents within the asthenosphere generate the movement by which crustal plates are moved.

Asylum seeker – A person who has fled their country of origin and applies for asylum under the 1951 Convention on the grounds that they cannot return to their country of origin because of a well-founded fear of death or persecution. While they wait for a decision on their application to be concluded, they are known as an asylum seeker.

Atmospheric water – Water found in the atmosphere; mainly water vapour with some liquid water (cloud and rain droplets) and ice crystals.

Autotroph – An organism capable of synthesising its own food from inorganic substances using light or chemical energy. Green plants, algae, and certain bacteria are autotrophs.

Backwash – The action of water receding back down the beach towards the sea.

Bankfull – The maximum discharge that a river channel is capable of carrying without flooding.

Base flow – This represents the normal day-to-day discharge of the river and is the consequence of slow-moving soil throughflow and groundwater seeping into the river channel.

Benthic – The ecological region at the lowest level of a body of water such as an ocean or a lake, including the sediment surface and sub-surface layers. The benthic environment comprises of a benthic community of organisms, known as **benthos**, which include worms, clams, crabs, sponges and other small organisms that live in the bottom sediments.

Bilateral (agreement) – An agreement on trade (or aid) that is negotiated between two countries or two groups of countries.

Biodiversity – The variation of life forms within a given ecosystem, biome, or for the entire Earth. Biodiversity is often used as a measure of the health of biological systems.

Biomapping – The mapping of emotions shown by people to certain places.

Biomass – The mass of living biological organisms in a given area or ecosystem at a given time. Biomass can refer to species biomass, which is the mass of one or more species, or to community biomass, which is the mass of all species in the community.

Biome – A major habitat category, based on distinct plant assemblages which depend on particular temperature and rainfall patterns, for example, tundra, temperate forest and rainforest.

Biosphere – The total sum of all living matter. The biological component of Earth systems (the others being atmosphere, hydrosphere and lithosphere).

Biotic potential – In population ecology, the natural reproductive potential of the species.

Bottom up – When local people are consulted and supported in making decisions to undertake projects or developments that meet one or more of their specific needs.

BRIC – An acronym used to describe a group of four countries – Brazil, Russia, India and China – whose economies have grown very rapidly.

Brownfield site – Brownfield is a term used in urban planning to describe land previously used for industrial purposes or some commercial uses.

Bunkering – In the Niger Delta, this is a term used for the theft of oil, mainly by local people, from pipelines or other sources.

Capital flows – The movement of money for the purpose of investment, trade or to produce goods/provide services. Usually regarded as investment into a production operation.

Carbon sequestration – The capture of carbon dioxide (CO_2) from the atmosphere or capturing anthropogenic (human) CO_2 from large-scale stationary sources like power plants before it is released into the atmosphere. Once captured, the CO_2 gas (or the carbon portion of the CO_2) is put into long-term storage.

Carbon sink – A store of carbon that absorbs more carbon than it releases.

Carrying capacity – The maximum population size that an area or environment can sustain indefinitely.

Central Business District (CBD) – The centre of an urban area containing the major shops, offices and entertainment facilities.

Chalcopyrite – The most important ore of copper.

Channelling – Wind redirected down long, straight, canyon-like streets where there is less friction. These are sometimes referred to as urban canyons.

Chemical weathering – The processes leading to the decomposition or breaking down of rocks due to chemical reactions. This most often requires the presence of water and/or exposure to the air.

Civil society – The 'third sector' of societal decision-making, distinct from government and business. It can include a wide array of non-profit organisations including NGOs, pressure groups, community groups, indigenous groups, charitable organisations, professional bodies and unions, etc.

Climate – A region's long-term weather patterns. This is measured in terms of average precipitation (that is, the amount of annual rainfall, snow, etc.), maximum and minimum temperatures throughout the seasons, sunshine hours, humidity, the frequency of extreme weather and so on.

Climatic climax – A biological community of plants and animals which, through the process of ecological succession, has reached a state of dynamic equilibrium with its climate and soils.

Clone town – A term used to describe urban retail areas dominated by national, and in some cases, international chain stores.

Coastal sediment budget – The balance between sediment being added to and removed from the coastal system; that system being defined within each individual sediment cell.

Coding – Coding is carried out when analysing answers to interviews. The coder (person who analyses the data) looks through all the answers to a question, develops a broad classification system based on the responses and then uses a code to categorise responses.

Combined cycle technology – Power plants that use both a gas and a steam turbine together to produce more electricity from the same fuel than a traditional plant. The waste heat from the gas turbine is routed to the nearby steam turbine, which generates extra power.

Common markets – A group formed by countries in geographical proximity in which trade barriers for goods and services are eliminated. (This may eventually apply to removing any labour market restrictions, as in the EU.)

Community preparedness/risk sharing – This involves prearranged measures that aim to reduce the loss of life and property damage through public education and awareness programmes, evacuation procedures, and the provision of emergency medical, food and shelter supplies.

Composite volcano – Large, steep-sided, symmetrical, cone-shaped volcano formed from alternating layers of lava flows, volcanic ash, cinders, blocks and bombs.

Condensation – The process by which water vapour changes to liquid water.

Conglomerates – A collection of different companies or organisations which may be involved in different business activities but all report to one parent company – most transnational corporations are conglomerates.

Conservative plate margin – A plate margin where two tectonic plates are moving past one another with no addition or destruction of plate material.

Constructive plate margin – A plate margin where new crust is generated as the plates pull away from each other. These are found at mid-oceanic ridges.

Constructive waves – Waves with a low wave height, but with a long wave length and low frequency of around 6–8/min. Their swash tends to be more powerful than their backwash and, as a consequence, beach material is built up.

Consumer – Organisms of an ecological food chain which receive energy by consuming other organisms.

Containerisation – A system of standardised transport that uses large standard-size steel containers to transport goods. The containers can be transferred between ships, trains and lorries, enabling cheaper, efficient transport.

Continentality – The impact of increasing distance from the coast on the climate of an area.

Conventional oil and gas – Refers to petroleum, or crude oil, and raw natural gas extracted from the ground by conventional means and methods.

Cost–benefit analysis – A systematic analysis of the advantages and disadvantages likely to result from a development project, where an objective value is allocated to all economic, social and environmental aspects affected.

Counter-mapping – This describes a bottom-up process by which people produce their own maps, informed by their own local knowledge and understanding of places.

Counter-urbanisation – The movement of people from large urban areas into smaller urban areas or into rural areas, thereby leapfrogging the rural–urban fringe. It can mean daily commuting, but could also require lifestyle changes and the increased use of ICT.

Crude birth rate – The total number of live births per 1,000 of a population per year – sometimes known as crude birth rate.

Crude death rate – The average number of deaths per 1,000 of a population per year – sometimes known as crude death rate.

Cryospheric processes – Those processes that affect the total mass of ice at any scale from local patches of frozen ground to global ice amounts. They include accumulation (the build-up of ice mass) and ablation (the loss of ice mass).

Cryospheric water – The water locked up on the Earth's surface as ice.

Cultural diversity – The existence of a variety of cultural or ethnic groups within a society.

Customs unions – A trade bloc which allows free trade with no barriers between its member states but imposes a common external tariff to trading countries outside the bloc (for example, the European Union).

Decentralisation – The movement of population and industry from the urban centre to outlying areas. The term may encompass the processes of both suburbanisation and counter-urbanisation.

Deflation – Where wind removes dry, unconsolidated (loose) sand, silt and clay particles from the surface and transports them away.

Deindustrialisation – This refers to the loss of jobs in the manufacturing sector, which occurred in the UK in the second half of the twentieth century.

Demographic dividend – The benefit a country gets when its working population outgrows its dependants such as children and the elderly. A boost in economic productivity results from growing numbers in the workforce relative to the number of dependants.

Demography (demographic) – Demography is the study of human population. A demographer is someone who studies the statistics and characteristics of populations.

De-multiplier effect – Opposite to the multiplier effect, when a withdrawal of income or investment leads to the closure of an activity, such as a factory. The loss of jobs and income for the workers means less money is spent in the economy, usually giving rise to more job and income losses in a downward spiral effect.

Density dependent/independent factors – In ecology, these are limiting factors which restrict population growth. Density dependent factors are those caused by the size of the population itself, for example, competition for limited food supplies. Independent factors are those which restrict population growth whatever the population size, for example, severe environmental conditions.

Deposition – Occurs when the velocity of the wind decreases until it can no longer transport the grains it is carrying.

Deregulation – The process of reducing or eliminating state-imposed regulations on certain activities, such as removing constraints on foreign currency transactions in financial deregulation.

Dereliction – The state of buildings having been abandoned and become dilapidated.

Desert – An arid environment receiving very low levels of rainfall.

Desertification – There have and continue to be a range of definitions of desertification, and it is generally accepted that it is not simply the 'advance of deserts', but it can include sand dunes encroaching on land. However, a UNESCO source in 2015 defines it as 'the persistent degradation of dryland ecosystems by human activities and by climate change'.

Destructive plate boundary – A plate margin where crust is destroyed as two plates converge. These are usually associated with island arcs or young fold mountains.

Destructive waves – Waves with a high wave height with a steep form and high frequency (10–14/min). Their backwash is generally stronger than their swash, so more sediment is removed than is added.

Diaspora – A group of people with a similar heritage or homeland who have settled elsewhere in the world.

Disability-adjusted life years (DALYs) – A measure of morbidity within a society. They measure the number of years of healthy life lost by being in poor health or a state of disability.

Discharge – The amount of water in a river flowing past a particular point expressed as m^3/s (cumecs).

Disseminations – In geological terms, these are ore deposits that are scattered throughout a rock.

Drainage basin – This is an area of land drained by a river and its tributaries. It includes water found on the surface, in the soil and in near-surface geology.

Dry ports – An inland terminal directly connected by road or rail to a seaport, which operates as a hub for the shipment of sea cargo for export or receives imports to inland destinations.

Earthquake – As the crust of the Earth is constantly moving, there tends to be a slow build-up of stress within the rocks. When this pressure is released, parts of the surface experience, for a short period, an intense shaking motion. This is an earthquake.

Ecological footprint – A measure of the demand placed by humans on Earth's natural ecosystems; the total area of productive land and water required to produce the resources a population consumes and absorb the waste produced.

Economic inequality – The difference between levels of living standards, income, etc. across the whole economic distribution.

Economic migrant – A person who has voluntarily left their country of origin to seek, by lawful or unlawful means, employment in another country.

Economies of scale – The cost advantages that result from the larger size, output or scale of an operation as savings are made by spreading the costs or by rationalising operations.

Ecosystem – A system in which organisms interact with each other and the environment.

Edge city – A self-contained settlement which has emerged beyond the original city boundary and developed as a city in its own right.

Electrolysis (of water) – A technique used to drive an otherwise non-spontaneous chemical reaction by passing a direct electrical current (DC) through the material involved. In the case of water, electrolysis is used to decompose water into oxygen (O_2) and hydrogen gas (H_2).

Emigrant – A person leaving their native area or country in order to settle elsewhere.

Endogenous factors/characteristics – In the context of place, this refers to the characteristics of the place itself. This would include aspects such as location, physical geography, land use and social and economic characteristics such as population size and employment rates.

Endoreic streams – Where rivers occupy drainage basins that are closed and do not flow out to the sea or other rivers, but instead end inland in lakes or swamps.

Energy – The ability to do work. In physical geography, much of this energy ultimately comes from the Sun.

Energy mix – The composition of different sources of energy used in a country or globally. 'Primary energy mix' refers to those used to produce all secondary energy. Some energy mix data refers only to sources used to produce electricity.

Energy transition – a pathway toward transformation of the global energy sector from fossil based to zero carbon by the second half of this century. At its heart is the need to reduce energy-related CO_2 emissions to limit climate change.

Enhanced greenhouse effect – The impact on the climate from the additional heat retained due to the increased amounts of carbon dioxide and other greenhouse gases that humans have released into the Earth's atmosphere since the Industrial Revolution.

Environmental resistance – A term used in population ecology to explain mortality rates controlled by environmental factors that prevent survival, for example, disease or shortage of food. Sometimes known as 'limiting factors'.

Environmental sustainability – A state in which the demands placed on the environment can be met without reducing the quality of the environment for the future.

Ephemeral streams – Streams that flow intermittently in hot desert areas following heavy thunderstorms.

Epidemiological transition – Describes changing patterns of population age distributions, mortality, fertility, life expectancy, and causes of death. It assumes that infectious diseases are replaced by chronic diseases over time due to expanded public health and sanitation.

Erosion – The wearing away of the Earth's surface by the mechanical action of processes of glaciers, wind, rivers, marine waves and wind.

Ethnography – A research method that explores what people do as well as say.

Eustatic change – A global change in sea level resulting from an actual fall or rise in the level of the sea itself.

Eutrophication – This occurs when excess fertilisers are washed off the land by rainwater into rivers and lakes. The resulting increase of nitrate or phosphate in the water encourages algal growth, which then forms a bloom over the water surface. This prevents sunlight reaching other water plants, causing them to die.

Evaporation – The process by which liquid water changes to a gas. This requires energy, which is provided by the Sun and aided by wind.

Evapotranspiration – The total output of water from the drainage basin directly back into the atmosphere.

Exfoliation – A process of mechanical weathering that results in the breaking, splitting or peeling-off of the outer rock layers. Also commonly known as 'onion skin weathering'.

Exogenous factors/characteristics – This refers to the relationship of one place with other places and the external factors which affect this. The demographic, socio-economic and cultural characteristics of a place are shaped by shifting flows of people, resources, money and investment.

Exogenous streams – Rivers that originate external to the desert in adjacent highlands and more humid environments, flow from outside of the desert and pass through it.

Exploitation – The action of using natural resources to the fullest or for the most profitable use.

Exploration – The process of searching an area with the intention of finding and mapping natural resources.

Fatalism – A view of a hazard event that suggests that people cannot influence or shape the outcome, therefore nothing can be done to mitigate against it. People with such an attitude put in place limited or no preventative measures. In some parts of the world, the outcome of a hazard event is said to be 'God's will'.

Fauna – The animals of a particular region, habitat or geological period.

Fetch – Refers to the distance of open water over which a wind blows uninterrupted by major land obstacles. The length of fetch helps to determine the magnitude (size) and energy of the waves reaching the coast.

Fjord – Former glacial valleys drowned by rising sea levels.

Flora – The plants of a particular region, habitat or geological period.

Flow/transfer – A form of linkage between one store/component and another that involves movement of energy or mass.

Flow resources – Resources that are renewable and can be replaced. Examples include fresh water and timber. They are commonly expressed in terms of the annual rates at which they are regenerated.

Food chain – An arrangement of the organisms of an ecological community according to the order in which they eat each other, with each organism using the next lower organism in the food chain as a source of energy.

Food security – Food security exists when all people at all times have access to sufficient, safe, nutritious food to maintain a healthy and active life.

Food web – A scheme of feeding relationships, resembling a web, which unites the member species of a biological community.

Footloose – Used in geography to describe an activity, such as an industry, that can be placed at any location without having to consider factors such as raw materials or transport.

Fortress landscapes – This term refers to landscapes designed around security, protection, surveillance and exclusion.

Fragile environments – Those that are easily damaged or disturbed and then difficult to restore once destroyed.

Frequency (of hazards) – The distribution of a hazard through time.

G7 – The 'Group of Seven' is an intergovernmental organisation made up of the world's seven largest advanced IMF economies: Canada, France, Germany, Italy, Japan, the UK and the US.

G20 – The G20 group is an international forum for the governments of 20 major economies. It includes the G7 countries and the EU as a single member. It was established in 1999 to give a voice to the major developing economies. This group includes China.

Gentrification – This is the buying and renovating of properties often in more run-down areas by wealthier individuals.

Geopolitics – The study of international relations, as influenced by geographical factors.

Geo-sequestration – The technology of capturing greenhouse gas emissions from power stations and pumping them into underground reservoirs.

Glacier – A large mass of ice on land, moving downhill due to the influence of gravity.

Global commons – Resource domains or areas that lie outside the political reach of any one nation state.

Global governance – A movement of political integration aimed at negotiating responses to problems that affect more than one state or region.

Globalisation – A process by which national economies, societies and cultures have become increasingly integrated through the global network of trade, communication, transportation and immigration.

Glocalisation – A term used to describe products or services that are distributed globally but which are fashioned to appeal to the consumers in a local market.

Green belt – An area of open space and/or low-density land use around a town or city where further development is strictly controlled.

Greenfield site – An area of undeveloped land.

Greenhouse gas – Any gaseous compound in the atmosphere that is capable of absorbing infrared radiation, thereby trapping and holding heat in the atmosphere.

Greenwashing – A term used to describe exaggerated environmental claims, usually made to promote the green credentials of big corporations, i.e. a form of environmental propaganda.

Groundwater – The water stored underground in the cracks and spaces in soil, sand and rock. It moves slowly through geologic formations called aquifers.

Groundwater flow – The slow movement of water through underlying rocks.

Habitat – An ecological or environmental area that is inhabited by a particular species of animal, plant, or other type of organism.

Hard engineering – Making a physical change to the coastal landscape using resistant materials, like concrete, boulders, wood and metal.

Health – Defined by the World Health Organization (WHO) as a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity.

Heliostats – Computer-controlled mirrors which keep the sun reflected on a target as the sun moves across the sky.

High-energy coast – A coastline where strong, steady prevailing winds create high-energy waves and the rate of erosion is greater than the rate of deposition.

Homogenisation – The process of making things uniform or similar, leading to places becoming indistinct from one another.

Horizons (soil) – Horizontal layers which are parallel to the soil crust. Each horizon has physical characteristics which differ from the layers above and beneath. Soil types are distinguishable by the composition of their horizon layers and each soil type usually has three or four horizons.

Hot spot – In certain places, a concentration of radioactive elements below the crust causes a hot spot to develop. From this, a plume of magma rises to eat into the plate above.

Hydrosphere – A discontinuous layer of water at or near the Earth's surface. It includes all liquid and frozen surface waters, groundwater held in soil and rock, and atmospheric water vapour.

Ice ages – A common term for periods when there were major cold phases known as glacials and ice sheets covered large areas of the world, interspersed with warmer interglacials. The last 2 million years, the Quaternary Period, contains the Pleistocene epoch lasting from about 1.8 million years to about 11,500 BP (before present). This was characterised by a series of episodes of glaciation, the last major glacial beginning around 120,000 BP reaching the Last Glacial Maximum about 18,000 BP. At that time about a third of the Earth's surface was covered by ice, compared with just a tenth today.

Ice sheet – A continental-size mass of ice, covering at least 50,000 km^2 that is dome-shaped with flows of ice outward from the centre. Ice sheets can be up to 2,000 m thick today; those in the Quaternary glaciation could have been twice that thickness.

Infrastructure – Infrastructure relates to the services considered essential to enable or enhance living conditions. These primarily consist of transport (roads, railways, canals, and/or airports), communications infrastructure (broadband and phone networks) and services such as water supply, sewers and electrical grids. They may also include infrastructure such as the local education system, healthcare provision, local government and law enforcement, as well as emergency services.

Ignition source – The means by which a wildfire can be started such as lightning strikes, falling power lines, cigarettes or agricultural burning which gets out of control.

Immigrant – A person moving into an area or country to which they are not native in order to settle there.

Improved drinking water – A source of water that by using intervening treatment is protected from outside contamination, particularly from faecal matter.

Index of Multiple Deprivation (IMD) – UK government qualitative study measuring deprivation across England, most recently carried out in 2015.

Indicator species – An organism whose presence, absence or abundance reflects a specific environmental condition. They can signal a change in the biological condition of a particular ecosystem, and thus may be used as a proxy to diagnose the health of an ecosystem.

Infiltration – The downward movement of water from the surface into soil.

Input – The addition of matter and/or energy into a system.

Insolation – The incoming solar radiation that reaches the Earth's surface.

Infant mortality rate – The number of children who die before their first birthday per 1,000 live births per year.

Integrated risk management – The process of considering the social, economic and political factors involved in risk analysis; determining the acceptability of damage/disruption and deciding on the actions to be taken to minimise damage/disruption.

Interception store – The precipitation that falls on vegetation surfaces (canopy) or human-made cover and is temporarily stored on these surfaces. Intercepted water either can be evaporated directly to the atmosphere, absorbed by the canopy surfaces, or ultimately transmitted to the ground surface.

Interculturalism – The interaction and the exchange of ideas between different cultural groups.

International trade – The exchange of capital, goods and services across international borders. Inbound trade is defined as imports and outbound trade as exports.

Island arc – A chain of volcanic islands associated with an ocean trench where subduction is taking place. During subduction, the descending plate begins to melt and the magma created rises towards the surface to form explosive volcanoes. If these eruptions take place offshore, a line of volcanic islands forms.

Isostatic change – Local changes in sea level resulting from the land rising or falling relative to the sea.

Joint venture – A commercial enterprise undertaken by two or more parties who agree to pool their resources and share ownership, risks and returns. In mining, the agreement is often between two or more large TNCs or between a large corporation and the government of the host country.

Labour – Factor of production defined as the aggregate of all human physical and mental effort used to create goods or provide services.

Lag time – The time between the peak rainfall and peak discharge.

Lahars – A flow of wet material down the side of a volcano consisting of erupted ash and water. Found when heavy rainfall occurs after a volcanic eruption. Essentially they are volcanic mudflows. In the Philippines, if a typhoon occurs after a volcanic eruption, then lahars can result.

Land remediation – The removal of pollution or contaminants from the ground which enables areas of derelict former industrial land to be brought back into commercial use.

Lava – Molten rock (magma) flowing onto the surface. Acid lava solidifies very quickly, but basic lava (basaltic) tends to flow some distance before solidifying (for example, on the Hawaiian Islands).

Leaching – The loss of soluble substances and colloids from the top layer of soil by percolating precipitation. The materials lost are carried downward (eluviated) and are generally redeposited (illuviated) in a lower layer.

Leakages (economic) – Refers to a loss of income from an economic system. It most usually refers to the profits sent back to their base country by transnational corporations – also known as profit repatriation.

Leopold matrix – An environmental impact assessment method that uses a tabulated recording sheet to assess the impact on each environmental characteristic and then the importance of that impact to the environment as a whole.

Life expectancy (at birth) – The average number of years a person born in a particular year in a location is expected to live.

Lithosphere – The crust and the uppermost mantle; this constitutes the hard and rigid outer layer of the Earth. It is this layer which is split into a number of tectonic plates.

Liveability – Aspects of urban living which make life more comfortable and endurable for city dwellers. Liveability can mean different things to different people. It may include natural amenities such as parks and green space; cultural offerings, career opportunities, economic and political stability or basic safety. It is measured by the Global Liveability Ranking.

Locale – This is the place where something happens or is set, or that has particular events associated with it.

Location – 'Where' a place is, for example, the co-ordinates on a map.
Longshore or littoral drift – Where waves approach the shore at an angle and swash and backwash then transport material along the coast in the direction of the prevailing wind and waves.
Low-energy coast – A coastline where wave energy is low and the rate of deposition often exceeds the rate of erosion of sediment.
Magma – Molten rock that is found beneath the surface of the Earth.
Magnitude – The assessment of the size of the impact of a hazard event.
Marine processes – Those that operate upon a coastline that are connected with the sea, such as waves, tides and longshore drift.
Maquiladora – A manufacturing operation (plant or factory) located in a free trade zone in Mexico. They import materials for assembly and then export the final product without any trade barriers.
Mass movement – The movement of material downhill under the influence of gravity, but may also be assisted by rainfall.
Meaning (of place) – Meaning relates to individual or collective perceptions of place.
Media – Means of communication including television, film, photography, art, newspapers, books, songs, etc. These reach or influence people widely.
Megacity – A city or urban agglomeration (urban area incorporating several large towns or cities) with a population of more than 10 million people. According to the UN, London achieved megacity status in 2013. This classification included residents in the Greater London area.
Metacity – A conurbation with more than 20 million people.
Microclimate – The small-scale variations in temperature, precipitation, humidity, wind speed and evaporation that occur in a particular environment such as an urban area.
MINT – An acronym referring to the more recently emerging economies of Mexico, Indonesia, Nigeria and Turkey.
Mitigation – Includes any actions, strategies, measures or projects undertaken (by mankind) to offset the known detrimental impacts of a process.
Monoculture – The agricultural practice of continually producing or growing a single crop, plant, or livestock species or breed in a farming system over a period of time, usually over a wide area.
Morbidity – Morbidity relates to illness and disease. It can also be used to describe the incidence of a disease within society. Some diseases are so infectious that by law they must be reported. These include malaria, rubella and tuberculosis.
Mortality – Mortality relates to death. It can be measured by death rate, infant mortality, case mortality and attack rate.
Multilateral (agreement) – An agreement negotiated between more than two countries or groups of countries at the same time.
Multilateral trade agreement – A trade agreement negotiated between more than two countries or groups of countries at the same time, usually facilitated by the WTO.
Multiplier (effect) – A situation where an initial injection of investment or capital into an economy (at any scale) in turn creates additional income by, for example, increasing employment, wages, spending, tax revenues, etc.; the overall increase in income arising from any new injection of spending or investment.
Multiscalar power – When the ability to influence (or prevent) change is vested in citizens, governments, institutions and other interested groups (players) located and acting at different geographical scales.
Mineral precipitation – The process in which dissolved minerals are released from water and form a deposit.
Naphtha – A generic term to describe a flammable liquid comprised of various hydrocarbon mixtures.
Natural change – The difference between birth rates and death rates. If birth rates are higher, there is a natural increase in population; if death rates are higher, there is a natural decrease.
Natural hazards – Events which are perceived to be a threat to people, the built environment and the natural environment. They occur in the physical environments of the atmosphere, lithosphere and the hydrosphere.
Net migration change – The difference between the total number or average rate of immigrants and emigrants in an area or country over a given period of time. More immigrants than emigrants will give a positive net migration and vice versa.
Net primary production (NPP) – The rate at which an ecosystem accumulates energy or biomass, excluding the energy it uses for the process of respiration. This typically corresponds to the rate of photosynthesis minus respiration by the photosynthesisers.
(Net) replacement rate – The number of children each woman needs to have to maintain current population levels or give zero population growth by generation. It is a measured fertility rate. In richer developed countries the replacement rate is about 2.1 but ranges from 2.5 to 3.3 in less developed countries because of higher mortality. If fertility is above the replacement rate, population will grow and if it is below, population will decline.
Non-communicable disease – A medical condition of disease that is by definition non-infectious and non-transmissible among people.
Non-government organisations (NGOs) – Any non-profit, voluntary citizens' group with a common interest, which is organised on a local, national or international level. Sometimes referred to as a 'civil society organisation' (CSO).
Norms – The shared values, traditions and customs that govern individual and group behaviour in a society. They are the perception of what is acceptable or unacceptable in societies at all scales.
Nutrient cycling – The movement of nutrients in the ecosystem between the three major stores of soil, biomass and litter.
Objective – Not influenced by personal feelings or opinions in considering and representing facts.
Oceanic water – The water contained in the Earth's oceans and seas but not including such inland seas as the Caspian Sea.
OECD – The Organisation of Economic Co-operation and Development is a group of 34 of the richest and most influential countries globally.
Outsourcing – A cost-saving strategy used by companies who arrange for goods or services to be produced or provided by other companies, usually at a location where costs are lower.
Overland flow – The tendency of water to flow horizontally across land surfaces when rainfall has exceeded the infiltration capacity of the soil and all surface stores are full to overflowing.
Overshoot – An ecological term referring to a point when the population and its associated consumption of resources exceed the long-term carrying capacity of its environment.
Palaeomagnetism – A record of the history of the Earth's magnetic field, preserved in magnetic minerals in volcanic rocks.
Particulate air pollution – A form of air pollution caused by the release of particles and noxious gases into the atmosphere. Emissions of particles can occur naturally but they are largely caused by the combustion of fossil fuels.
Peak discharge – The point on a flood hydrograph when river discharge is at its greatest.
Perception – With regard to hazards, this is the way in which an individual or a group views the threat of a hazard event. This will ultimately determine the course of action taken by individuals or the response they expect from governments and other organisations.

Perception of place – This is the way in which place is viewed or regarded by people. This can be influenced by media representation or personal experience.

Percolation – The downward movement of water within the rock under the soil surface. Rates vary depending on the nature of the rock.

Periglacial – Processes and landforms associated with the fringe of, or area near to, an ice sheet or glacier.

Photochemical pollution – A form of air pollution that occurs mainly in cities and can be dangerous to health. Exhaust fumes become trapped by temperature inversions and, in the presence of sunlight, low-level ozone forms. It is associated with high-pressure weather systems.

Place – Defined as a location with meaning. Places can be meaningful to individuals in ways that are personal or subjective. Places can also be meaningful at a social or cultural level and these meanings may be shared by different groups of people.

Placelessness – Defined by the geographer Edward Relph as the loss of uniqueness of place in the cultural landscape so that one place looks like the next.

Placemaking – The deliberate shaping of an environment to facilitate social interaction and improve a community's quality of life.

Placer deposits – A natural accumulation of minerals caused by the effects of gravity separating alluvial deposits in sedimentary processes, with the densest minerals being first to settle, including tin and gold.

Plagioclimax – The plant community that exists when human interference prevents the climatic climax vegetation being reached.

Plate tectonic theory – The theory that states that the Earth's crust is made up of several rigid plates moving relative to one another.

Plough pan – A compacted layer within the profile of a cultivated soil resulting from repeated compaction from ploughing.

Plume – A hot column of magma which rises from deep within the Earth.

Population density – The average number of people living in a specified area, usually expressed as the number of people per square kilometre.

Population distribution – The pattern of where people live. This can be considered at all scales from local to global, in an area or country.

Porphyry – A fine-grained hard igneous rock which contains large crystals and can contain hydrothermal ore deposits such as copper.

Potable (water) – Water that is suitable for drinking (it does not have to be completely pure but must not contain unacceptable levels of hazardous materials, nor look, taste or smell unpleasant).

Prediction – The ability to give warnings so that action can be taken to reduce the impact of hazard events. Improved monitoring, information and communications technology have meant that predicting hazards and issuing warnings have become more important in recent years.

Primary effects (hazard) – The effects of a hazard event that result directly from that event. For a volcanic eruption, these could include lava and pyroclastic flows. In an earthquake, ground shaking and rupturing are primary effects.

Primary energy (resources) – Energy sources obtained in their raw material or natural form, such as oil, natural gas, wind or running water; they are not usable until they are converted into heat or mechanical action to produce secondary energy.

Primary sector – The sector of the economy making direct use of natural resources, for example, agriculture, mining, forestry and fishing.

Prisere – The succession of vegetational stages that occurs in passing from bare earth or water to a climax community.

Protectionism – A deliberate policy by government to impose restrictions on trade in goods and services with other countries – usually done with the intention of protecting home-based industries from foreign competition.

Pyramid of biomass – A diagrammatic representation of the amount of organic material, measured in grams of dry mass per m^2, found in a particular habitat at ascending trophic levels of a food chain.

Pyroclastic flows – Also known as nuées ardentes, formed from a mixture of hot gas (over 800°C) and tephra. After ejection from the volcano, they can flow down the sides of a mountain at speeds of over 700 km/hr. Some vulcanologists apply the term nuées ardentes when the cloud is formed only from hot gas.

Pyrophytic vegetation – Pyrophytes are plants adapted to tolerate fire. Methods of survival include thick bark, tissue with a high moisture content and underground storage structures.

Qualitative data – These are non-numerical data that are used in a relatively unstructured and open-ended way. It is descriptive information, which often comes from interviews, focus groups or artistic depictions such as photographs.

Quantitative data – Quantitative data are numerical data such as metric-level measurements that are associated with the scientific and experimental approach and are criticised for not providing an in-depth description; data that can be quantified and verified, and are amenable to statistical manipulation.

Radiative forcing – The difference between the incoming solar energy absorbed by the Earth and energy radiated back to space.

The Ramsar Convention – An international treaty for the conservation and sustainable use of wetlands. It is also known as the Convention on Wetlands. It is named after the city of Ramsar in Iran where the Convention was signed in 1971.

Raised beach – Areas of former wave-cut platforms and their beaches which are at a level higher than the present sea level.

Recharge (groundwater) – A hydrological process involving the downward movement of water by infiltration and percolation causing the replenishment of natural groundwater. It occurs naturally but can be done artificially through human intervention and management.

Refugee – In its broader context it means a person fleeing, for example, civil war or natural disaster but not necessarily fearing persecution as defined by the 1951 Refugee Convention. Legally however, a refugee is an asylum seeker whose application claim for asylum has been successful.

Replacement fertility rate – The number of children each woman needs to have to maintain current population levels or give zero population growth by each subsequent generation. It is a measured total fertility rate. If fertility is above the replacement rate, population will grow and if it is below, population will decline.

Representation – Representation is how a place is portrayed or 'seen' in society.

Reproduction (of global systems) – The way that existing global systems remain largely unchanged and are even extended and reinforced by regulation.

Reproductive age – The age at which women can give birth. In official demographic data it is usually considered to be between 15 and 44.

Reserves – That part of a resource that is available for use under existing economic and political conditions and with available technology.

Resilience – The sustained ability of individuals or communities to be able to utilise available resources to respond to, withstand and recover from the effects of natural hazard events. Communities that are resilient are able to minimise the effects of the event, enabling them to return to normal life as soon as possible.

Resilience (ecological) – The amount of disturbance that an ecosystem can withstand without changing existing structures and processes.

Resource frontier – A newly-colonised region where resources have been discovered and are brought into production for the first time.

Resource curse – Also known as the 'paradox of plenty' or 'resource trap'; refers to the notion that (some) countries endowed with an abundance of natural resources struggle to benefit from this resource wealth and instead experience minimal economic growth or social development.

Resource peak – This marks the point in time when the maximum production rate of a resource occurs (at all scales from individual reserves to global), with production declining in subsequent years.

Retardants – Chemicals sprayed onto fires in order to slow them down. They are composed of nitrates, ammonia, phosphates and sulphates and thickening agents.

Ria – Former river valleys drowned by rising sea levels.

Rift valley – A long, deep valley found in the centre of a spreading ridge. It is formed between parallel faults where a block of the crust has sunk down.

Riparian – This describes the land adjacent to a stream or river. In water conflicts and disputes it is used as a legal term to determine the rights of those living on the banks of a river to access the water for use.

Run-off – All the water that enters a river channel and eventually flows out of the drainage basin.

Safe drinking water – Water that is safe for human consumption. The water must be free from harmful pollutants and bacteria that could make people ill.

Salinisation – The build-up of salts in soil, eventually to toxic levels for plants.

Saltation – A process where sand-sized particles are transported by bouncing and hopping along the surface.

Sampling – To make statistically valid inferences, when it is impossible to measure the whole population, by selecting a group which will be representative of that population. Such a group is known as a sample.

Saturated – This applies to any water store that has reached its maximum capacity.

Sea-floor spreading – The theory that the ocean floor is moving away from the mid-oceanic ridge and across the deep ocean basin, to disappear beneath continents and island arcs; movement of oceanic crustal plates away from divergent/constructive plate boundaries such as in the middle of the Atlantic Ocean.

Secondary effects (hazard) – These are the effects that result from the primary impact of the hazard event. In volcanic eruptions, these include flooding (from melting ice caps and glaciers) and lahars. In an earthquake, tsunamis and fires (from ruptured gas pipes) are secondary effects.

Secondary energy (resources) – Energy that is transformed or converted from primary energy sources into manufactured sources of power which are usable, such as heat, electricity or petroleum.

Secondary sector – The sector of the economy involved in processing or manufacturing to produce finished capital or consumer goods.

Sediment – Any naturally-occurring material that has been broken down by the processes of erosion and weathering and has then been transported and subsequently deposited by the action of ice, wind or water.

Sediment budget – The balance between the input and output of sediment in hot deserts.

Sediment cell – A distinct area of coastline separated from other areas by well-defined boundaries, such as headlands and stretches of deep water.

Semi-arid – The climate of an area that receives between 250 and 500 mm of precipitation per year.

Sense of place – This refers to the subjective and emotional attachment people have to a place. People develop a 'sense of place' through experience and knowledge of a particular area.

Seral stage – An individual stage within a sere, for example, colonisation or stabilisation.

Sere – The entire sequence of stages in a plant succession. Different seres are named after the starting point of the succession, for example, lithosere, hydrosere, psammosere and halosere.

Social segregation – This is when groups of people live apart from the larger population due to factors such as wealth, ethnicity, religion or age.

Soft engineering – Using natural systems for coastal defence, such as beaches, dunes and salt marshes, which can absorb and adjust to wave and tide energy.

Soil – The upper layer of earth in which plants grow, a black or dark brown material typically consisting of a mixture of organic remains, clay and rock particles.

Soil liquefaction – The process by which saturated, unconsolidated soil or sand is converted into a suspension during an earthquake. It is then able to act like a liquid and flow, particularly when under pressure.

Soil organic carbon (SOC) – The organic constituents in the soil: tissues from dead plants and animals, products produced as these decompose and the soil microbial biomass.

Statistical bias – The intentional or unintentional favouring of one group or outcome over other potential groups or outcomes in the statistical population.

Statistical population – The whole from which a sample will be chosen for a research exercise.

Stemflow – The portion of precipitation intercepted by the canopy that reaches the ground by flowing down stems, stalks or tree bole.

Stock resources – Non-renewable resources which can be permanently expended. Their quantity is expressed in absolute amounts rather than rates, for example, oil.

Store/component – A part of the system where energy/mass is stored or transformed.

Storm and rainfall event – An individual storm is defined as a rainfall period separated by dry intervals of at least 24 hours and an individual rainfall event is defined as a rainfall period separated by dry intervals of at least 4 hours (Hamilton and Rowe, 1949).

Storm flow – Discharge resulting from storm precipitation involving both overland flow, throughflow and groundwater flow.

Storm hydrograph – A graph of discharge of a river over the time period when the normal flow of the river is affected by a storm event.

Storm surge – A rapid rise in sea level in which water is piled up against a coastline to a level far in excess of the normal conditions at high tide. They are produced during the passage of a tropical storm when wind-driven waves pile up water against a coastline combined with the ocean heaving upwards as a result of much lower air pressure.

Sub-aerial processes – Includes processes that slowly (usually) break down the coastline, weaken the underlying rocks and allow sudden movements or erosion to happen more easily. Material is broken down in situ, remaining in, or near, its original

position. These may affect the shape of the coastline and include weathering, mass movement and run-off.

Sub-climax – The development of an ecological community to a stage short of the expected climax because of some factor, such as repeated fires in a forest, which arrests the normal succession.

Subduction – The process by which one plate is driven beneath another plate to form such features as ocean trenches at destructive/convergent plate boundaries.

Subjective – Based on or influenced by personal feelings, tastes or opinions.

Suburbanisation – The movement of people from living in the inner parts of a city to living on the outer edges. It has been facilitated by the development of transport networks and the increase in ownership of private cars. These have allowed people to commute to work.

Succession – The series of changes in an ecological community that occur over time.

Surface creep – Where saltating particles return to the surface and hit larger particles that are too heavy to hop; they slowly creep (slide or roll) along the surface from a combination of the push of the saltating grain and the movement of the wind.

Suspension – Transportation by wind where the smallest particles, generally less than 0.2 mm, are held in the air.

Sustainable city – A city which provides employment, a high standard of living, a clean, healthy environment and fair governance for all its residents.

Sustainable development – Development which recognises that the needs of the present have to be met but doing this without affecting the needs of future generations.

Sustainable Urban Drainage Systems (SUDS) – This is a relatively new approach to managing rainfall by using natural processes in the landscape to reduce flooding, control flooding and provide amenities for the community.

Swash – The rush of water up the beach after a wave breaks.

System – A set of interrelated components working together towards some kind of process.

Tailings – Tailings consist of the slurry mix of crushed rock and effluents generated by the mechanical and chemical processes used to extract the desired mineral from the ore. (It produces a waste stream known as mine tailings.)

Tariffs – A tax or duty placed on imported goods with the intention of making them more expensive to consumers so that they do not sell at a lower price than home-based goods – a strategy of protectionism.

Tectonic change – The rise or fall in sea level resulting from tectonic processes.

Tectonic plate – One of a series of rigid sections of the earth's crust. They float on the upper mantle and move relative to one another.

Temperature inversion – An atmospheric condition in which temperature, unusually, increases with height. As inversions are extremely stable conditions and do not allow convection, they trap pollution in the lower layer of the atmosphere.

Tephra – The solid matter ejected by a volcano into the air. It ranges from volcanic bombs (large) to ash (fine).

Terrestrial water – This consists of groundwater, soil moisture, lakes, wetlands and rivers.

Tertiary sector – The sector of the economy concerned with providing services.

Thermal fracture – The weathering of rock resulting from its rapid and repeated heating and cooling.

Throughfall – The portion of the precipitation that reaches the ground directly through gaps in the vegetation canopy and drips from leaves, twigs and stems. This occurs when the canopy-surface rainwater storage exceeds its storage capacity.

Throughflow – The movement of water downslope through the subsoil under the influence of gravity. It is particularly effective when underlying permeable rock prevents further downward movement.

Tides – The periodic rise and fall of the level of the sea in response to the gravitational pull of the sun and moon.

Top down – When the decision to undertake projects or developments is made by a central authority such as government with little or no consultation with the local people whom it will affect.

Topophilia – The affective bond between people and place or setting (Tuan, 1974).

Topography – The relief and drainage of an area.

Total fertility rate – This is the average number of children born per woman in an area or country if all women lived to the end of their child-bearing years. It is considered to be a more direct and accurate measure of fertility than birth rate as it refers to the births per woman.

Total productive bio-capacity – All the food, water and energy resources produced by the Earth's natural systems annually to sustain us.

Transpiration – The loss of water from vegetation through pores (stomata) on their surfaces.

Transportation – The processes that move material from the site where erosion took place to the site of deposition.

Trophic level – An organism's position in the food chain. Level 1 is formed of autotrophs which produce their own food. Level 2 consume level 1 and level 3 consume level 2, etc.

Tsunami – Giant sea waves generated by shallow-focus underwater earthquakes, violent volcanic eruptions, underwater debris slides and landslides into the sea.

Tundra – A Lapland term for the climate and vegetation type found across extensive areas of northern Siberia, Scandinavia, Canada and Alaska. Tundra lies between the region of continual snow and ice of the Polar regions and the northern limit of tree growth (or the taiga environments) and at high altitudes above the tree line in the Alps, Rockies, Andes and Himalayas. Similar ecosystems are found around the ice-free fringes of Antarctica. It is characterised by a stony or marshy surface with mosses, lichens and other low-lying vegetation including shrubs and dwarf trees. Tundra has short summers but mean monthly temperatures do not exceed 10°C, which is just warm enough for snow and the surface layer of the permafrost (active layer) to melt. Currently tundra covers about 8 per cent of the land area of Earth, but during the glacials of the Quaternary glaciation, tundra covered huge areas of central North America and Europe, including the British Isles.

Unconventional (oil and gas) reserves – Hydrocarbon reservoirs that have low permeability and porosity and so are difficult to exploit. They require enhanced recovery techniques such as fracture stimulation; for example, shale deposits, tar sands and heavy oils.

Under-employment – The practice in which a person is not doing work that makes full use of their skills and abilities.

United Nations – An international organisation founded in 1945 made up of 193 member states whose aim is to promote international peace and cooperation.

Urban growth – An increase in the number of urban dwellers. Classifications of urban dwellers depend on the census definitions

of urban areas, which vary from country to country. They usually include one or more of the following criteria: population size, population density, average distance between buildings within a settlement, legal and/or administrative boundaries.

Urban heat island – The zone around and above an urban area which has higher temperatures than the surrounding rural areas.

Urbanisation – An increase in the proportion of a country's population that lives in towns and cities. The two main causes of urbanisation are natural population growth and migration into urban areas from rural areas.

Urban resilience – The capacity of individuals, communities, institutions, businesses and systems within a city to survive, adapt and grow no matter what kinds of chronic stresses and acute shocks they experience.

Urban morphology – The spatial structure and organisation of an urban area.

Urban policy – Strategies chosen by local or central government to manage the development of urban areas and reduce urban problems.

Urban resurgence – Urban resurgence refers to the regeneration, both economic and structural, of an urban area which has suffered a period of decline. This is often initiated by redevelopment schemes but is also due to wider social, economic and demographic processes.

Urban social exclusion – Economic and social problems faced by residents in areas of multiple deprivation.

Urban sprawl – The spread of an urban area into the surrounding countryside.

Venturi effect – The squeezing of wind into an increasingly narrow gap, resulting in a pressure decrease and velocity increase.

Volcanic bombs – Rocks that are more than 5 mm in diameter that are thrown into the air by a volcanic eruption.

Volcanic explosive index – A scale used to measure the explosiveness of volcanoes.

Water balance – The balance between inputs (precipitation) and outputs (run-off, evapotranspiration, soil and groundwater storage) in a drainage basin.

Water stress – Water stress occurs when the demand for water exceeds the amount of water available during a certain period, or when poor quality of water restricts its availability for human use.

Water surplus (areas) – Areas where there is more than sufficient water available to meet human needs.

Wave-cut platform – A gently sloping (less than 5°), relatively smooth, marine platform caused by abrasion at the base of the cliff.

Wave refraction – When waves approach a coastline that is not a regular shape, they are refracted and become increasingly parallel to the coastline. The overall effect is that the wave energy becomes concentrated on the headland, causing greater erosion. The low-energy waves spill into the bay, resulting in beach deposition.

Weathering – The breakdown and/or decay of rock at or near the Earth's surface creating regolith that remains *in situ* until it is moved by later erosional processes. Weathering can be mechanical, biological/organic or chemical.

Well-being – The state of being comfortable, healthy, or happy.

World city – These are cities which have great influence on a global scale, because of their financial status and worldwide commercial power. Three cities which have traditionally sat at the top of the global hierarchy: New York, London and Tokyo are now being joined by the likes of Beijing, Shanghai and Mumbai. These cities house the headquarters of many transnational corporations (TNCs), are centres of world finance and provide international consumer services.

Zonal soil – A soil which has experienced the maximum effect of climate and natural vegetation upon the parent rock, assuming there are no extremes of weathering, relief or drainage.

Index

abiotic factors 231–32, 240
ablation 9, 144, 145–6
abrasion 61, 67, 93, 99, 150, 478
accumulation 9, 144–8
acid rain 102, 199, 445, 609, 613–14
active layers 139, 159
adaptation 86, 134, 171–3, 180–1, 186–7, 201, 209, 215, 222, 269, 276–7, 339–40, 475
adret slopes 144
aeolian processes 61, 63, 67–73, 75, 84–6, 101–23
afforestation 593
aflatoxins 483–4
Africa 304–5, 312, 395, 465, 542–3, 573, 618
  Sub-Saharan 311–2, 465, 473–4, 485, 543, 582
African Continent Free Trade Area (AfCFTA) 311–12
age factors 497, 508, 512–14
Agenda 21 125, 331, 625
agents of change 364, 366
agglomeration 307, 316, 318
agricultural productivity 466, 470–1, 474, 537–9
agricultural systems 31–2, 468–70, 473, 531
agriculture 19–20, 22–3, 79, 86–7, 270, 272–3, 277–9, 297–8, 397–8, 463, 466, 468, 471–81, 585, 590–2, 616, 633, 636
agro-forestry 475
aid 203, 286, 301
air density 605
air pollution 221, 271, 398–9, 422–5, 427, 429–32, 439, 441–6, 457, 462, 492–3, 500, 563, 605, 609, 633
  particulate 422, 42931, 443, 492
  photochemical 422, 431, 492, 535
albedo 138, 141–2, 422, 424, 426
algae 260, 261, 535, 592
alluvial fans 73
alpine environments 6, 139
altitude 138–9, 231–2, 534
aluminium 625
Amazon 6–7, 46–9, 266–73
Amsterdam 368, 442–3
anaerobic conditions 479
angle of incidence 58, 62, 67
animals 27, 169, 176, 180–1, 221, 223–4, 230, 234–5, 241, 243, 245–8, 250–1, 267–9, 271, 273, 277–80
  see also livestock
Antarctic Circumpolar Current 334, 337–8
Antarctic Convergence Zone 334
Antarctic Divergence 334
Antarctic Ice Sheet 5–6, 39, 139, 147, 170, 335, 337–8
Antarctic and Southern Ocean Coalition (ASOC) 344–5
Antarctic Treaty 175, 333, 334, 338, 340–2
Antarctic Treaty System (ATS) 175, 327, 333, 341–2, 344–5
Antarctica 5, 138–9, 141–2, 146–7, 173–6, 334–45, 628, 630
anthropogenic factors 257
  see also human factors
Apple Inc. 317–21
aquaculture 483
aquifers 12, 23, 562, 579–80, 582–3, 587–8, 593–95
arches 106–7, 116
architecture 209, 216, 376, 426–30
Arctic 5, 38, 138, 146–7, 169–72, 174, 472, 628
arêtes 151, 179
arid regions 54–5, 57, 58–60, 70–7, 84–6, 581–2
aridisols 56–7
aridity index 54, 55
arroyos 70
art 374–6, 385–7
artificial intelligence (AI) 547
ash 197, 199–203
assessments v–xi
Aswan High Dam 591–2
asylum seekers 517
Athabasca tar sands 607–10
athenosphere 191, 192
atlases 379, 646
atmosphere 3, 4, 8, 21, 27–8, 39–41, 58–9, 185
atolls 260
attributes 2
attrition 99, 478
Australian bushfire season 223–4
autotrophs 232, 245
avalanches 206
aviation 45–6, 219, 368–9
awareness-raising 447, 568
backshore 91
backwash 93–4, 101, 107
badlands 76–7
bajadas 73
banana industry 321–4
bankfull 15, 19
bar graphs 651–2
barrages 119
bars 109–10
base flow 15, 18
bays 104–5
Bayston Hill, Shrewsbury 667–8
beach nourishment 121, 122
beach profiling 130–1
beach surveys 129
beaches 94, 97, 99, 107–8, 110, 121, 128–9
  raised 113–14, 116–17
bed nets 498
Belfast 365
Belize Barrier Reef Reserve System 262–5
belonging, sense of 357, 358–9
Belt and Road Initiative (BRI) 305–6
berms 93, 107
bias 642
bid-rent theory 410
'big data' 663–4
bio-molecules 25
biodiversity 79–81, 176, 230, 252–7, 273, 278–9, 443
biological pumps 30
biomapping 371
biomass 30–1, 230, 233–4, 239, 262, 597–8, 601, 611, 613, 632
biomes 167, 230, 238, 265, 265–79
bioprospecting 176
biosecurity 294
biosphere 3, 26–7, 524
biotic factors 232, 240–1
biotic potential 521, 522
Birmingham 403
birth rate 506–10, 510, 515, 522
  crude 506, 508
bitumen 605, 607–9
Blakenhall, Wolverhampton 548–51
blanket bog 253–7
blockfield (Felsenmeer) 160
blogs 389
blowholes 106–7
blowouts 111
Blue Revolution 483
Booth, Charles 362, 378–9
Boserup, Ester 531–2
'bottom-up' 81, 329, 331, 370, 407
boulder clay 150
boundaries 2
Boundary Dam CCS plant 43
barchans 72
brands 290
Brazil 286, 291, 622–3, 624
breakwaters 120, 215
Brexit 292, 297, 309, 346
BRIC countries 286, 291
Brick Lane, East London 377–81, 383–5, 387–9, 391
Brock River 49–51
bromeliads 269
brownfield sites 448
business parks 411
buttes 74
bypassing 129
calcium carbonate 25, 28, 37, 260–1
calcrete 66

Canada 511, 607–10
canopy 267, 268
capacity 64, 73
capillary movement 56, 65
capital 285
capital flows 284, 285–8, 300–11
Carajás ore extraction complex 622–3
carbon budget 35–41
carbon capture 43–4, 627
carbon cycle 2, 24–41, 41–52, 668–9
carbon dioxide 25, 27–47, 52, 168–9, 271, 609, 615
  and the carbon cycle 25, 27–47, 44, 52
carbon monoxide 430
carbon pumps 29
carbon sequestration 25, 34–5, 43–4, 627
carbon sinks 25, 28, 31
carbon stores 25–8, 31–2, 36–41, 44–5, 48, 254
Carbon Trust 612
carbonation 102
carbonic acid 28, 36, 338
care services 547
carnivores 233–4, 245, 277
carrying capacity 524–7, 530, 532
cartels 571
cartographic skills 646–8
catchment management 433–6, 588
cattle grazing 270, 272–3, 277–8
caves 76, 106–7, 116
cavitation 99
cement manufacture 31
central business district (CBD) 410–12
central tendency 655–6
chalcopyrite 616
chalklands 24, 240–2
channel flows 12–13
channelling 428
Cheonggyecheon River 435–6
chernozems 476–7
chi-square test 660–1
Chile 594
China 304–6, 310, 312–14, 320–1, 395, 417–18, 438, 465, 509, 573–7, 583–4, 589, 603, 617
chlorophyll 232
cholera 492, 493
chotts 74
cities 395–8, 401, 408–9, 414, 419–20, 450–60
city centres 410–12
City Challenge 406
City Deals 407
civil society 330, 331, 333
Clean Air Acts 431, 457
cliffs 64, 98–100, 105–6, 116, 119
climate 41–52, 55–6, 138–41, 231, 240, 265–6, 274–5, 335, 422–32, 460, 463, 466, 471–5, 490–2, 500, 580–1, 587, 605–6, 622, 635–6
climate change 41–2, 44, 48, 78–82, 114–15, 117–18, 167–70, 200, 213–15, 242–3, 253, 256, 262, 300, 332–3, 337–8, 345, 453, 462–3, 473–5, 530, 534–9, 537–8, 580, 609, 615, 623, 635, 636
  see also global warming
climate emergency 332
Climate Smart Agriculture (CSA) 474–5
climatic climax 230, 237–8, 265, 278
cloud 8
coal 564–5, 568, 570, 596, 598, 603–4, 626
coastal systems/landscapes 90–136, 592, 669
coastlines 92, 94–6, 100, 104–5
coding 641, 642
cold environments 138–43, 536
colonisers 249
combined cycle technology 568
combined heat and power (CHP) 612
combustion 30–1
command words vii–viii
commerce 585, 590, 619–20
commodities 296
common market 295
communications technology 292–3, 301–2
community action 187–8, 222, 369, 449
commuting 398, 412
comparative advantage 306, 307
competence 73
competition 344
computer modelling 222
condensation 9, 10
conflict 299, 347, 361, 397, 554, 570, 583–4, 594–5, 632
congelifluction 160
congelifraction 66
conglomerates 290
connection-making viii–xi
conservation 174–5, 247–8, 251, 262–5, 278, 344–5, 593
construction industry 619
consumer awareness campaigns 485
consumers (heterotrophs) 232–4, 245
consumption 244, 292, 438, 467–8, 482–3, 524–6, 543, 564, 567–9, 572, 574, 578, 590–2, 601–3, 611–13, 619–20
containerisation 289
context 377–8
continental convergence 196
continental drift 191–2
continentality 54, 58, 59
continuity 362
contraception 485, 546
Contracts for Difference (CfD) scheme 611
Copenhagen 452
copper 616–20, 625, 629
coral reefs 28, 36–7, 259–63
core (Earth) 190
core–periphery model 285–6, 291, 559–60
Coriolis force 95, 212
cornucopian theories 531–2
coronary heart disease (CHD) 499–503
corporate bodies 368–9
corporate social responsibility 578
corrasion 99
corries (cirques) 148, 151–2, 179
cost–benefit analysis 186, 560
costs 289, 314–15, 412, 498, 546
counter-mapping 370
counter-urbanisation 401, 402–3
COVID-19 283, 298–300, 307, 333, 357, 439, 487–8, 521, 561, 571, 575, 646–7
creep 61, 68, 102, 148, 478
crime 383, 413
croplands 45
crops 220, 272–3, 277, 463, 470–1, 473, 475, 477, 483–4, 537–8, 591
crowdsourcing 367, 662–3
crust (Earth's) 191, 192–5, 198, 204, 206, 208
crusts 66
cryospheric processes 4, 5–6, 9–10, 160
crystal growth 65
cultural diversity 415, 419–22, 455
cultural homogeneity 326
cultural quarters 412
culture 299, 326, 361, 508–10
  see also multiculturalism
currencies, local 356–7, 454
currents 60, 94–5, 101, 108, 127, 130–1, 193–4
cusps 107–8
customs unions 295
cut-off ore grade 616, 621
cybercrime 294
cyberspace 333
cyclones 134, 212–19, 280–1
Dalmatian coast 116
dams 584, 586–8, 591–2, 599, 606, 623–4
databases 662
de-industrialisation 291, 401, 403–7
'dead heart syndrome' 404
death rate 506–7, 510, 515, 522
  crude 506, 508
decentralisation 398, 401, 404–7
decomposers 29, 232, 233, 246
decomposition 26, 29–31, 33, 233, 247
deep-sea trenches 195
deflation 61, 67
deforestation 19–21, 32–4, 183, 244, 269–73, 300, 477, 609, 622, 633, 636
Defra, Biodiversity Action Plan (BAP) 244, 247
deltas 130
demand 482–3, 485, 554, 562, 576, 600–3, 611–13, 618–19, 627, 629, 637
demographic dividend 514–16
demographic transition model (DTM) 510–14, 545
demographics 361, 380–1, 506, 508, 519–21
dengue fever 536–7
denudation 64
dependency 297, 512–16
dependency ratio 513, 515, 546
deposition 61, 64, 68–70, 90, 97, 101–2, 104, 107–10, 138, 147, 150, 153–5, 157–8, 165
deprivation 382–3, 415–19, 421, 500, 550
deregulation 301
dereliction 446–9, 563
desalination 590
descriptive approaches 353
desert varnish 66
desertification 53, 77–82, 86–8, 479
deserts 53–89, 75, 77, 274–5, 335
desire lines 647–8
Detroit 449
developed countries 538, 564, 577, 628
developing countries 300–3, 315, 573, 576, 577, 613, 630
development 261, 278, 464–5, 600
  low-impact (LID) 446
  see also sustainable development
Devonport, Plymouth 363
Dhaka 444
Dharavi slum, Mumbai 399, 459–60
diabetes 501
diarrhoea 493, 494
diaspora 286–7, 360–1, 419

diet 501, 502, 538, 546
digital trade 298
digital/'augmented' places 376–7
dips 100
disability-adjusted life years (DALYs) 486, 495
disaster/response curves 189–90
disaster/risk management cycle 188–9
discharge 4, 6–7, 14–21, 48, 51, 64, 586
  flashy 18–19, 433
  peak 15, 18
disease 463, 490–5, 521, 592
  agricultural 324, 538
  airborne 492
  degenerative 488–90
  infectious 488, 489–93, 521
  non-communicable 486–93, 499–503, 533–7, 546
  vector-borne 446, 491, 492, 494, 495–9, 536–8
  water-borne 492, 493
diseconomies of scale 325
disintegration, block/granular 66–7
Dismaland, Weston-Super-Mare 376
dispersion 655–7
'displaced' persons 218, 376
distance-decay theory 410
distillation 590
distributaries 69
distribution 292, 295, 656–7
diversification 316
Doha Development Round 330
domestic settings 585, 590, 603, 611–12
draa landscapes 73
drainage 240, 432–7, 478, 491, 623, 634
drainage basins 10–14, 606
drumlins 151, 153, 154
dry ports 294
Dubai 419–20, 429
duricrust 66
dynamic equilibrium 3, 53, 61, 90–2, 94, 108, 137, 147
Earth Summits 125, 331
earthquakes 188–90, 194, 196, 204–12, 224–5, 227–8, 491–2, 663
Earth's structure 190–1
ecological footprint 449, 524–7
ecological management 593
economic change 361–2, 390–1
economic development 271–2, 278, 312–13, 346–7, 437, 451–3, 458, 488–90, 495–7, 500–1, 511–12
economic factors 173–6, 382–3, 415–22, 508, 520, 546
economic growth 346, 347
economic impacts 325, 519, 591–2, 627–9
economic inequality 415–19, 455–7, 459–60
economic interdependence 298–9
economic migrants 517–19, 548
economic viability 557
economic well-being 455–7, 459–60, 497–8
economic/monetary unions 295
economically active populations 513, 521
economies
  emerging 286, 310, 316, 325
  global/world 283
  see also highly developed economies; less developed economies
economies of scale 290, 297, 314, 316, 325
ecoregions 279–80
ecosystem-based approaches 126
ecosystem services 257, 258–9, 434
ecosystems 79–80, 111–13, 132, 170, 230–79, 340, 528–9, 592, 622, 669
edge to area ratio 241
education 221–2, 299, 383, 408, 418–20, 447, 458, 497, 502, 530, 541, 546
Ehrlich, Paul 531
electricity 598, 618, 633
electrodialysis 628
electrolysis 626
electronic waste (e-waste) 441
elements 2
eluviation 268
embargoes 306
emergent features 114, 116–17
emigrants 506, 508
employment 291, 297, 299, 397, 404–5, 407–8, 418
en-glacial debris 150
endogenous factors 360–1
endoreic streams 61, 69
energy 2–3, 62, 234, 453, 544, 554–6, 564–72, 567–72, 600–13, 626–7
  coastal environments 92–6
  non-renewable 555, 613
  primary 596–7, 600, 602–3
  renewable 320, 430, 453–4, 555–6, 562, 564, 570, 595–602, 605–6, 610–11, 613, 626–7
energy conservation 568
energy efficiency 470, 568
energy flows 232–3
'energy from waste' 441, 442–3
energy gap 564
energy mix 596, 600–3, 611, 632
energy performance certificates (EPCs) 611–12
energy rich 564–5
energy saving 611–13
energy security 568, 595–616
energy transition 595–6, 602
engineering
  hard 118–21
  soft 118, 121, 122
enterprise 285
environment 320–1, 408, 457, 460, 463–65, 485–506, 520, 670
and population 78–82, 86–8, 462–553, 670
environmental audits 342
environmental capacity 467
environmental consequences/impacts 326, 347, 527, 558, 562–3, 572, 592, 607–10, 613–16, 622–4, 628–30
environmental degradation 610
environmental flow indicators (EFIs) 588
environmental groups 309, 332
environmental impact assessment (IEA) 343, 563
environmental interdependence 300
environmental management 451, 452, 453, 458
environmental resistance 522
environmental standards, minimum 419
environmental statements 563
environmental sustainability 327
ephemeral streams 61, 69
epidemiological transition 486, 488–90
epiphytes 269
equatorial regions 266–73
equilibrium line 145
erg 72
Ergodic model 139
erosion 22, 54, 61–2, 64, 67–70, 90, 92–3, 98–9, 104–7, 118–35, 138–9, 147, 149–53, 156–7, 165, 179, 271, 477–9, 592
erratics 151, 154
eskers 157–8
essay writing xii
estuarine mudflats 112
ethical implications 644
ethnicity 357, 381, 415, 419–21, 455, 501, 510–11
ethnography 371
European Union (EU) 296–7, 309, 321, 346
eutrophication 244, 246, 478
evacuation 200, 215, 221
evaporation 8–12, 19, 479
evapotranspiration 10, 12, 14, 56, 267, 271, 424–5, 427, 593
  potential (PET) 14, 24, 55
exchange rates 297, 314
exclusion measures 413–14
exclusive economic zones (EEZ) 343
Exeter 411
exfoliation 61, 65
exogenous factors x, 360, 361
exogenous streams 61, 69
exploitation 244, 253, 270, 297, 314, 333, 554–6, 559, 566, 613, 629–30
exploration 555, 558, 564–5, 578, 607, 610
exponential growth 522, 526
exports 289, 310
external forces 361–2
extinction 253, 262, 622
extraction 31, 559, 607–9, 620–4, 626–7, 629
fail-safes 434
famine 488
fatalism 186, 187
fault lines 204, 208–9
feedback 53
  negative 3, 90–1, 94, 527–30
  positive 3, 53, 137–8, 527–30
ferrallitisation 268
Ferrel cells 59
fertilisers 44, 87, 246, 272, 463–4, 591
fertility rates 463, 485, 488, 508–10, 539–42
  replacement 507–8
  total 506, 507, 508
fetch 92, 93, 99, 108
field sketches 645
fieldwork 366, 640, 664–70
Fitzroy drainage basin 264
film 374
financial systems 285, 292–4, 297, 308, 405
fire 30–3, 207, 219–25, 270–2, 277
firn 144, 145, 147
fishing 176, 262, 272, 279–80, 335–7, 483, 633
fjords 113, 114, 115–16, 153
flaring 632, 634
flocculation 112
floods 51, 70, 118–33, 135, 189, 199, 214–15, 219, 221, 224–6, 433–7, 457, 460, 490–1, 495, 592
  management 118–33, 135, 215, 433–7, 460, 587
Florianopolis Declaration 343

flow lines 647, 649
flow resources 555
flows/transfers 2, 10–14, 61, 90–2, 137, 144
fluvio-glacial processes 156–9
fog 427, 431, 605
foggaras system 87–8
food 172, 294, 309, 322–4, 354, 462–3, 466–8, 481–5, 530–1, 544
food chains 167, 230, 234–5, 245–6, 271
food security 453, 466, 473, 481–5
food stability 481
food waste, reduction 485–6
food webs 230, 234–5
footloose 290
foreshore 91
forests 30–2, 45–9, 181–3, 234, 238, 266–73, 477
  see also deforestation
fortress landscapes 408, 413–4
fossil fuels 26–8, 30–1, 44, 581–2, 564–5, 568–70, 595–8, 606, 615, 628
  see also coal; gas; oil
fossils 192
fragile environments 137, 142, 166–7, 173–4, 340
France 602
free markets 306
free trade 295, 301, 306, 312–13, 324–5, 330
free trade area (FTAs) 295
freeze–thaw cycles 148, 159–61, 165
friction 148, 149, 194, 227
frost creep 160
frost heave 160, 161
frost shattering 148, 160
fuel 220
G20 group 303
gabions 119, 121
gas 562, 564–6, 568–70, 596, 598, 604, 606–7, 610, 627, 631–4
  unconventional 560, 564, 604–5
gauging stations 16
gelifluction 160, 164
gender issues 359, 497, 509, 515, 521
gentrification 376, 390–1, 400–1, 412–13
geo-sequestration 39, 44
geo-spatial crowdsourced data (GSD) 663
geographical information systems (GIS) 377, 382, 646, 663, 665–6
geological certainty 557
geology 28, 99–100, 104–5, 240, 464, 566, 587, 503–6, 620, 634
geomorphological processes 63–7, 98–104, 147–51
geopolitics 301–3, 307, 324, 345–6, 570–2, 576–8, 583–4, 600, 627–30, 628–30
geos 106–7
geospatial data 382, 383
geothermal energy 399–400, 597, 601–2, 627
Gersmehl diagrams 239
ghettos 421
Gili Trawangan, Indonesia 227–8
Gini Index 302, 346
glacial advance 145–6
glacial budgets 144–6
glacial erosion 139
glacial periods 6, 9
glacial processes 143–51, 164–5
glacial retreat 145–6, 153, 157, 169
glacial systems/landscapes 137–82, 669
glacial troughs 151, 152, 153
glacials 77, 104, 113–14, 139, 142–3, 145, 165, 176
glaciers 5–6, 139, 142, 143–51, 156, 178–9, 338
Glastonbury 351–2
Global Agriculture and Food Security Program (GAFSP) 482–3
global commons 300, 327, 333–45
global financial system (GFS) 293–4
global governance 303, 327–33, 340–3, 345, 669
global positioning systems (GPS) 665
global shifts 291
global systems 283, 298–306, 669
global value chains 295
global village 283
global warming 8, 27, 31, 39, 40–2, 42, 114–15, 117–18, 167–70, 242–3, 337–8, 463, 473–4, 534–9, 615
  see also climate change
globalisation 81–2, 283–98, 325–6, 346–7, 355–6, 481, 572, 606–10
glocalisation 326, 356
goods, knowledge-intensive 290
governance 303, 327–33, 340–3, 345, 451, 452, 458, 669
government 270, 291, 315, 366–7, 600
grading 73
graffiti 376
graphical skills 650–5
grasslands 44–5, 220, 250, 273–7
grassroots 80
gravimetry 629
gravitational sliding 193–4
grazing 251, 270, 272–3, 277–8, 481
green areas 410, 411
green belts 401
green investment 454, 458
green revolution 467, 470, 531
green roads 426
green roofs 426
green spaces 260, 426, 454
Green Trade Alliance (GTA) 625
greenfield sites 448
greenhouse effect 40–1, 117, 168
  enhanced 39, 40, 41, 613, 615
greenhouse gases 4, 7, 40–2, 44, 168–9, 437–8, 440–1, 443, 609, 615, 623
  see also carbon dioxide; methane; water vapour
Greenland Ice Sheet 5–6, 39, 139, 170
greenwashing 466, 475
greywater use 593
ground contraction 161
ground rupture/shaking 206
groundwater 7–8, 10, 12–13, 22–4, 86–8, 432–3, 439, 446, 448, 479, 556, 585, 587–8, 592–4
Group of Seven (G7) 303, 595
groynes 101, 118–21, 128–9
Guiyu 441
gypcrete 66
habitats 240–1, 245, 273, 398, 443, 462, 587, 609, 622–6
Hadley cells 59
halophytes 113
halosere 113, 237
hanging valleys 151, 153
hardpans 66
harmattan 274
hazard management 187–90, 215, 222, 454, 460
hazard prevention 188, 189, 215, 222
hazard protection 188
hazards 185–229, 397, 491–2, 663
headlands 104–5
health 485–506, 520–1, 533–39, 548–51, 592
health-related data 383–4, 386
healthcare 299, 408, 419, 508, 521, 546–7
heat 62, 190
heat networks 612
heat recovery systems 612
heating, domestic 424–5
heatwaves 535–6
hedgerows 479
heliostats 598
Helvellyn area, Lake District 178–9
herbivores 233, 234, 245, 277
heritage quarters 412
heterotrophs 234
high pressure systems 70, 141, 274
higher-income countries 400–1, 467, 490
highly developed economies (HDEs) 285–6, 291–2, 300–3, 309, 313, 315, 330
Historic Sites and Monuments (HSMs) 175
HIV/AIDS 487, 490, 503–505
home ownership, second homes 402
homelessness 218, 413
homeostatic regulation 528–9
homogenisation 355–6, 398
hot spot theory 196–7, 198
household composition 412–13
housing 398–402, 413, 496, 501, 667–8
  affordable 399–400, 413, 415, 418, 454, 456
Hubbert curves 560–1
human capacity 467
Human Development Index (HDI) 80
human impact 19–20, 25, 78–9, 99, 118–26, 166–76, 180–3, 185–9, 198–212, 214–23, 238, 244, 250–1, 253, 257–9, 261–2, 269–73, 277–81, 338–42, 345, 472, 581–2, 636
humidity 9, 56, 60, 423
hummocky ground 162, 173
hunger 467, 468
hunter-gatherers 272
hurricanes 188, 189, 212–19
hydration 65–6, 102
hydraulic action 99
hydraulic fracturing (fracking) 205, 315, 561, 566, 569, 604–5, 610
hydrocarbons 25, 31, 40
hydroelectric power (HEP) 16, 180, 567, 591, 596–7, 599, 601–2, 606
hydrogen economy 626
hydrology, urban 432–3
hydrolysis 66, 102
hydropolitics 583
hydrosere 237, 246–7
hydrosphere 3, 4, 26
hydrothermal deposits 620
hypotheses, null 660–1
ice 8–9, 39, 144–8, 159–65, 169–70, 176, 335, 337–8
  movement 145–6, 148–50, 152–3
  see also glaciers
ice ages 139, 142, 156–7, 165

ice caps 6, 146
ice sheets 5–6, 39, 117, 138–9, 142–3, 146–7, 159
ice shelves 5, 334, 338
ice veins 161
ice wedges 161
ice-albedo feedback 138
ice-wedge casts 162–3
Iceland 602
ICT skills 662–4
ideas, unequal flow of 301
identity 354–5, 358
ignition sources 219–20
immigration/immigrants 346, 506, 508, 548
imports 289, 294, 306, 309, 310
income 287, 457, 497, 515, 568
  average income per person 244
  gross national income per capita 511
  inequality 302
  see also higher-income countries; low income countries
Index of Multiple Deprivation (IMD) 382–3, 415–16
India 310, 395, 398–400, 446, 459–60, 465, 473, 475, 509–10, 584, 603, 634–7
Indian Agricultural Research Institute (IARI) 475
indicator species 252
indigenous people 171–3, 180–3, 271–2, 472–3, 477
Indus River Commission 584
industrialisation 299, 564, 568, 574
industry 411, 432, 445, 568, 585, 590, 603, 607–10, 612, 619–20
inequality 300–5, 327, 346–7, 361–2, 83–5, 415–19, 421, 453, 455–6, 467, 526
infiltration 10, 11–12, 19, 20
influenza 492
informal sector 397
informal settlements 411
information, flows of 290, 301
infrastructure 291, 305–6, 360, 398, 400, 453–4, 559, 578, 619, 634, 636
injustice 325, 327, 347
inner cities 402–3, 405–6, 411, 413, 414–18
inputs 2–3, 10–11, 14, 53, 61–3, 67, 90–7, 104, 137, 144, 147, 232–3, 450, 468–71, 506
insecticides 498–9
inselbergs 74
inshore 91
institutions 327
integrated coastal zone management (ICZM) 125–6
integrated risk management 187–8
integration 347
  horizontal 316
  vertical 316, 607
intellectual property 298
inter-quartile range 656
inter-tropical convergence zones (ITCZs) 58, 267, 274
interception stores 10, 11
interculturalism 422
interdependence 298–302
interglacial periods 6, 10, 77, 104, 139, 142–3, 145, 165, 176
Intergovernmental Panel of Climate Change (IPCC) 25, 243, 337, 460, 535
International Bank of Reconstruction and Development (IBRD) 329
International Development Association 329
International Energy Agency (IEA) 562, 626
International Monetary Fund (IMF) 286, 294, 299, 325, 329
international organisations 181, 299, 303, 325, 327–30, 503–6, 586
international status 297
International Union for Conservation of Nature and Natural Resources (IUCN) 344
International Whaling Commission (IWC) 336, 343
internet 292, 379
interviews 127, 371, 642, 668
Inuit 171–3, 472–3
invasive species 246
investment 289, 291, 297, 306–8, 453–4, 515–16
  foreign 284, 286, 297, 301, 307, 311, 573, 577
Iran 570
Ireland 319
iron 625
irrigation 23, 45, 475, 589, 590–1
island arcs 195, 198
isobars/isochrones/isohyets/isotherms 650
Israel 594–5
Jamaica coral reef 262
Janicki omniprocessor 447
Japan 507, 514, 545–8
joint ventures 316, 572, 578, 607
just-in-time (JIT) 295
kames 157, 158
karst 164
Keeling Curve 27–8
kettle holes 158
Kinder Scout, Derbyshire 254–7
krill 173, 233, 334–8
labour 285, 286–90, 315, 320
lag times 15, 18–19, 69
lahars 197, 199–200, 224–5
lakes 7, 11, 151, 152–3, 156–8, 179, 586
land 285
land degradation 79
land remediation 448
land systems approach 165
land take 624
land use 20–4, 32–3, 44–5, 122, 215, 244, 246, 409–10, 470, 495, 524
land value 409–10
landfill 400, 437, 439–43
landforms 70–3, 75, 156–9, 161–6
Landsat 82–3
landscapes 70–7, 79–80, 90–82, 398, 622, 669
  see also geomorphological processes
landslides 103, 199, 206–7, 219, 224–6
Last Glacial Maximum 142–3, 166
laterite 275
Latin America 310–11, 322–3
latitude 138–9, 140, 265–7, 491
latosols 268, 271, 477
lava 187, 188, 196–203
leaching 57, 268, 275, 276
lead 625
leakage (economic) 286, 301
least developed countries (LDCs) 80, 313, 538, 543
Legionellosis 494
legislation 297, 327, 446–7, 502–3
leisure 242, 410
less developed economies (LDEs) 286, 292, 301–2, 310–13, 315, 330, 483, 519, 521, 534
Lever, William Hesketh 362
lichen 181–82
life on Earth 41–52
life expectancy (at birth) 386, 457, 488–9, 508, 522, 540, 542, 545
limiting factors 471
line of best-fit 652
line graphs 650–1
links, following viii–xi
lithology 67, 99
lithosere 237
lithosphere 3, 25–6, 185, 192–5
litter 26, 239
liveability 450
lived experience 354–5, 359, 370, 389
livelihood assets 133–4
livestock 20, 22, 44, 220, 270, 272–3, 277–8, 463, 470–3, 481, 483
Living Planet Index (LPI) 252–3
living standards 437–8, 546, 568
load 64
loans 300, 301, 329
local groups 369
locale 351, 352, 353, 368–70
localisation 356–8
localism 354
location 314, 351, 353, 378–80, 620–1
loess 63, 240
logarithmic scale graphs 654–5
logging 32, 183
  reduced-impact (RIL) 273
Lombok, Indonesia 210–11
Lomé Convention 323
London 448, 455–8
London Basin 23–4
low income countries 209, 303, 399, 467, 490, 493, 564
low pressure areas 212, 274, 427
Lyme disease 537
Madrid Protocol 339, 341–3
magma 192–8, 203, 620
magnetometry 629
malaria 494, 495–9, 536–7
Malthus, Thomas 530–1
managed retreat 118, 122
management and information systems 294–5
mangroves 131–4, 261, 279–81
mantle 191–4
manufacturing 291, 298, 301, 308, 319, 321, 325, 346, 404–5, 407
  see also de-industrialisation
manure 44, 45, 478
maps 371–2, 378–9, 381, 387, 425, 645–50
  Ordnance Survey (OS) 378, 645–6
maquiladora 298
marine ecosystems 252, 259–65, 592
marine platforms 117
marine processes 98–102
Marine Protection Areas (MPAs) 344, 345
maritime spatial planning (MSP) 126

marketing 290, 295, 316–17, 321
markets 306–26, 515
marriage 509, 546
marshes 109, 112–13, 247
Masdar City, Abu Dhabi 430
mass movement 61, 63–4, 76, 92, 102–3, 160
Mayon, Mount, Philippines 201–3, 225
McKelbey boxes/diagrams 556–7
mean, arithmetic 655
meaning 351, 353, 364–9
meat 485, 538
mechanisation 344
Médecins Sans Frontières (MSF) 505
Medellin, Colombia 363–4
media 351, 353, 354–60, 371–6, 388–9
median 656
megacities 396, 408–9, 450
Mekong Agreement 584
meltwater channels 156, 157–9
  braided 158, 159, 164
  overflow 156
Mercalli scale 205–6
Mercator projection 370
Mercosur 310–11
mesas 74
metacities 396, 444
metals 441, 445, 572–7, 616–21, 623–5, 629
  ferrous/ferroalloy 572, 573, 575
metamorphogenic deposits 620
methane 25, 40, 41, 168–9, 331, 437–8, 443, 615
micro-irrigation 590–1
microbial pollutants 445
microclimates 422, 587
Mid-Atlantic Ridge (MAR) 194–5
migration 286–8, 297, 299–300, 346, 359, 390–1, 395, 397, 402–4, 419, 506, 58, 516–21
Millennium Development Goals (MDGs) 257, 328, 399, 467, 494, 504, 579, 582
Millennium Ecosystem Assessment (MEA) 257
Million Ponds Project 247–8
mineral ores 572–84, 616–25, 629–30
mineral precipitation 617
mineral resources 338–9, 341–2, 464, 544, 554, 556–7, 572–84, 616–25, 629–30
mining 342, 440, 578, 620–4, 629–30
Mining and Metals Scenarios to 2050 project 624–5
MINT countries 286
mitigation 41–2, 44, 48, 86–8, 133–4, 180, 200, 209, 215, 222, 339–40, 454, 460, 498–9, 502–3
mobile phones 292
mode 656
models 2
Mohawk River 16–17
moment magnitude scale 205
monocultures 86, 272, 322
monsoons 473
moraines 150–1, 153–5
morbidity 486–8
mortality 485–90, 499–500, 535–6
  infant mortality rates 456, 463, 488, 506, 508, 515, 545–6
mountains 195–6, 232, 335
movement, maps showing 647–9
mudflats 112
mudflows 103, 197, 199–200, 206, 224–5, 624
mulching 45
multiculturalism 301, 357, 455, 549
multilateral trade agreements 306
multiplier effect 316–17, 404
Mumbai 459–60, 473
music 374
myxomatosis 241, 250
Nasser, Lake 591–2
national interest 606
National Park Cities 458
nationalism 292, 304, 346, 354
nearshore 92
neo-Malthusian perspectives 531–2
Nepalese earthquake 663
net balance 145
New York City High Line 404
newly industrialised countries (NICs) 291–2
newspapers 379
Niger 507, 511, 514, 601–2
Niger Delta, Nigeria 631–4
Nile Basin Initiative 584
nitrate loss 22
nitrogen oxides 168–9, 440–2, 443, 457, 614–15
nivation 148, 152
non-governmental organisations (NGOs) 87, 134, 330–2, 344–5, 504–6
normal distribution 657
norms 327
nuclear energy 566, 596–7, 602, 610, 616, 627
nunataks 335
nutrient cycling 29, 230, 239, 246, 268–9, 276, 475
nutritional standards 538
oases 86–8
obesity 501, 550, 551
objectivity 369
ocean floor deposits 630
ocean ridge systems 194–5, 197
ocean sediments 26, 28
ocean thermal energy conversion (OTEC) 627
ocean/ocean convergence 195
oceanic carbon pump 29–30
oceanic circulation, global 60
oceanic/continental convergence 195
oceans 4, 13, 26, 36–41, 117, 261–2, 338
Ock River 17–18
Office for National Statistics (ONS) 369, 380–1
Official Developmental Assistance (ODA) 286
offshore 91, 119
offshoring 314–15, 325
oil 554, 559–62, 564–71, 596, 598, 604–10, 626–7, 631–4
  unconventional 560, 564, 566, 604–5, 607–10, 627
oil spills 494, 569, 633, 634
oil and tar sands 605, 607–10
Omran, Abdel 488–9
opportunities 170–1, 180, 451, 453, 543–44
organic produce 324
Organisation of Economic Co-operation and Development (OECD) 303, 307
Organization of Petroleum Exporting Countries (OPEC) 296, 570, 571, 628
osmotic distillation 628
outputs 2, 3, 10–11, 15, 53, 61, 63, 90–2, 97, 104, 137, 144, 232–3, 450, 469–71, 506
outsourcing 290, 299, 314–15, 316, 318, 325
outwash plains (sandur) 157–9
overburden 621
overland flow 10, 11, 12, 19, 70
overpopulation 523
overshoot 524, 526
overweight 501
oxidation 66, 102
ozone depletion 533–4
Pacific Alliance (PA) 310–11
palaeomagnetism 191, 193
Palestine 594–5
palm oil 272–3
Pangaea 191, 192
parasites 592
Paris Agreement 332
Park's impact/response model 189–90
participant observation 127
patterned ground 161–2
peat 26–7, 253–4, 256–7
pediments 73–4
pediplanation 74
peds 480
percolation 10, 12
periglacial regions 139, 142, 159–65
periphery 285–6, 291, 559–60
permafrost 6, 10, 36, 139, 141–2, 159–60, 164, 170, 472
pests 538
Pevensey Bay, Sussex 127–31
phenomenological approaches 353
Philippines 201–3, 217–19, 224–6, 225
photography 371–2, 390, 645
photosynthesis 29, 141, 232, 234, 245
physical exercise 501, 502
phytoplankton 233, 234, 335
pie charts 652–3
pingos 163–4
'pioneer' image 412
pioneer plants 113, 141
place 351–94, 669
placelessness 356
placemaking 353, 354–5
placer deposits 620
plagioclimax 230, 237–8, 278
plate boundaries
  convergent 207
  destructive 191, 194, 195–6, 204
plate margins
  conservative (passive) 191, 194, 196
  constructive 191, 193–4, 197–8
plate tectonics 190–204
playas 74
Pleistocene 77, 137, 142–3
ploughing 19–20
plucking 150, 152
plumes 191, 196–7
plutonium 597, 616
plutons 198
pluvials 75
poetry 372–3, 388–9, 391
polar environments 138–9, 335, 472–3
  see also Antarctica; Arctic
policy 270, 406–7, 458, 502–3, 509–10, 547, 600
political factors 299, 325, 346, 508–10, 520
  see also geopolitics
pollution 441, 476, 527–30, 532, 609

dust 624
noise 443, 563, 624
see also air pollution; water pollution
pollution episodes 444
ponds 244–6
population
ageing 513, 514, 542, 545–8
and the environment 78–82, 86–8, 462–553, 670
size 462, 519, 524–44
structure 512–14, 519
population change 464–6, 504–21, 539–542
natural 506, 508, 510–14
see also population decline; population growth
population decline 395, 466, 542, 545–8
population density 398–9, 464–6, 471–5, 522, 550
population distribution 170–1, 462, 464, 542–3
population ecology 521–32
population growth 78, 86, 166, 244, 278, 395–401, 405, 437, 459, 462, 465, 474, 485, 506–7, 514, 521–32, 539–43, 554, 580
population numbers 464–6
population parameters 464
population pyramids 512–16, 519
population, resources and pollution model 527–30
populism 292, 346
porphyry deposits 617
Port Sunlight, the Wirral 362
postmodernism 414
poverty 187, 397, 415–16, 417–19, 421, 455–7, 467, 495, 497–8
poverty maps 362, 378–9
power 303–305
multiscalar 330–1
precipitation 11–23, 55–7, 59–60, 62–3, 70, 78, 138, 140–1, 144, 147, 214, 266–9, 274–5, 335, 338, 423, 427–8, 432–7, 460, 536, 634, 636
see also rainfall; snow
preparedness 200, 208–9, 214–15, 221–2, 226
pressure release 102
price volatility 572
primary data 417, 665, 666, 667–8
primary sector 315
prisere 237
privatisation 301
producers 232, 245
production 290–2, 295, 315–16, 319, 321, 564–6, 572–4, 619–20, 625, 627, 631–4
factors of 285
net primary (NPP) 230
see also exploitation
productivity 57, 233–4, 470, 470–1
products 289, 292
profits, repatriation 286
proportional divided circles 652–3
protectionism 289, 306, 309, 577
protests 355
psammoseres 111, 237, 249
public health 384–5, 498
pull factors 397, 517–18
push factors 397, 517–18
pyramidal peaks 151
pyroclastic flow 197, 199–201, 203
pyrophytes 219, 220
qualitative data 82–3, 126–7, 363, 369, 641–2
quality of life 397, 417
quantitative data 82–3, 126–7, 363, 369–71, 641–2
Quaternary glaciation 9
quaternary sector 405
questionnaires 388, 668
race 357
'race to the bottom' 324
radiative forcing 39, 40–1
rain shadows 59
rainfall 8, 10, 12, 59–60, 159, 214, 216, 218, 225, 266–9, 274–5, 279, 427–8, 433–7, 460, 463, 475, 490–1, 495, 537–8, 581–2, 586, 588, 594, 605, 634–6
convectional 274, 427
rainforests 32, 46–9, 266–73, 477
Ramsar Convention 7
range 656
rebranding 366–8
recession 307, 576
recharge 129, 556, 587, 593–4
reclamation 559
reconstruction 189
recovery (waste) 440
recycling 129, 400, 440, 454, 593, 619, 625, 629
reduction 441–3
reforestation 32, 35, 45
refugees 517–18
reg (desert pavement) 67
regeneration 366, 448
regional trade agreements (RTAs) 306
regionalisation 81–2
regionalism 354
regolith 63, 76
regulation 306, 327, 446–7
rehabilitation 189
reindeer 180–3, 472–3
rejection level 659
relationships 2
relief (hazard management) 189
relief (landscape) 409, 500, 623
religion 355, 359, 371, 509
remittances 287, 300, 519
remote sensing 629
rendzina 240
Rent Gap 412
representation 364–77, 385–9
reproduction (global systems) 327
reproductive age 506, 508
reprofiling 129
reserves 555, 556–61, 564–6
measured 557, 605
unconventional 560–1, 564, 566, 604–5
reservoirs 586–7, 588, 592, 606
residential areas 411, 412
residuals 652
resilience 86, 133–4, 180, 187–8, 339–40, 450–1, 474
resource classification 555–6
resource curse 572, 573, 576, 577, 633
resource depletion 562, 564, 628
resource development 554–63, 564, 604, 610–11
resource extraction 563
resource frontiers 555, 559–60, 605
resource management 556
resource peaks 555, 560–2
resource scarcity 554
resource security 554–639, 670
resources 133, 176, 270, 297, 308, 314, 333, 524–32, 555–8, 572, 605, 626–30
unconventional 560, 564–6, 604–5, 607–10, 627
see also energy; mineral resources; water
respiration 29, 234
retail, out-of-town 410, 411
retardants 219, 222
retirement age 547
retrofitting 209, 216
revegetation 45
reverse osmosis 590
revetments 119, 121
rias 113, 114, 115
rice 473
ridge push 193–4
ridges 107
rift valleys 191, 195, 197
riparians 584
ripples 108
risk 170–1, 185–8, 201, 206, 210–11, 218, 223, 227, 460, 558
risk assessments 200
river basins 583, 593, 595, 628
rivers 6–7, 14–20, 48, 69, 433–6, 586, 588–9, 606, 609
flow 15–18, 586, 606
regimes 14, 16–18, 49–51
roads, green 426
robots 547, 629
roches moutonnées 151, 153, 179
rock 99–100, 192–4
igneous 616, 617, 620
permeability 587, 593, 604
porosity 587
sedimentary 603, 604, 616–17, 620, 634
rock armour (rip-rap) 119–21
rock basins 116, 152, 179
rock falls 63–4, 103, 160
rock slides 63
rotational slip (slumping) 103
Royal Dutch Shell 607
run-off (overland flow) 8, 10, 12, 15–20, 62, 103, 271, 432, 445, 445–6, 446, 478, 636
runnels 107
rural areas 360, 402–3, 497, 581
rural–urban migration 397
Russia 304, 570
safe-to-fail systems 434–7
Saffir-Simpson scale 213
salinisation 466, 479–80, 538, 592
salinity 245, 261, 592
salt lakes 74
salt marshes 109, 112–13
saltation 23, 61, 64, 68, 100, 102, 118, 478
salts 65–6, 74
saltwater greenhouse technology 628
Sami 173, 180–3
Sámpi region 180–3
sample size 644
sampling 641, 642–3
sanctions 306, 313
sand 107–8
sand dunes 72–3, 84–6, 109–11, 122–3, 131–2, 248–51

sand seas 57, 72
sand-blasting 70
sanitation 399, 408, 416–17, 446, 453, 487–90, 493–4, 497, 522, 546, 579, 636–7
satellite technology 82–3
saturation 8, 10, 12, 268
savanna grasslands 220, 273–7
scattergraphs 652
schistosomiasis (bilharzia) 494, 592
schools 418, 421
science parks 411
scientific research 339–41
scree 74, 160
scrub 250, 251
sculpture 375
sea ice 5, 38, 138, 169–70, 337–8
sea level change 39, 96, 114–18, 169, 212, 279, 338, 538
sea surface temperatures (SSTs) 260–1
sea walls 119–21, 215
sea-floor spreading 191, 193
seasonal affective disorder (SAD) 490–1
seasonality 495
seawater 585
secondary data 417
secondary sector 315
  see also manufacturing
security 294, 299
sediment 61–3, 67–70, 96–8, 100–2, 107–8, 127–32, 477–8, 587, 592, 620
sediment budget 61, 63, 96–8
sediment cells 96–7, 124
seeds 475
Sefton Coast sand dunes 248–51
segregation 402
seismic gaps 209
seismic hazards 204–12
seismic surveys 629
semi-arid 54–5, 57, 76–7, 84–6, 275
seral stages 230, 237
seres 230, 237
services 289–90, 292, 316, 397, 400, 404–9, 416–17, 453, 505
Severn River 156–7
sewage 261, 400, 445–7, 491, 494, 496
sex composition 512–14
sheet flow 12
Shimla, India 634–7
shingle 107, 108, 121, 128
shocks, acute 451
Shoreline Management Plans (SMPs) 124
shrubs 268, 277
sierozems 56, 57
Simon, Julian 531, 532
Sites of Special Scientific Interest (SSSIs) 121, 241, 248
sketching skills 645
skills 640–70
sky view factor 426–7
slap pull 194
slums 398–9, 415, 418, 453, 459–60
smogs 431, 457, 492, 535, 605
smoking 500, 502
snow 138, 140–1, 144, 147–8, 170, 338, 605, 634
social constructionism 353
social development 451–3, 458, 488–90, 495–7, 500–1
social exclusion 415, 416
social factors 299, 326, 346, 383–5, 390–1, 415–22, 508, 520, 546, 591–2
social segregation 414, 415, 420–2
social welfare 506
social well-being 313, 455–7, 459–60
soil 11–15, 19–22, 26–7, 31–2, 48, 56–7, 141–2, 221, 232–3, 238–40, 265–6, 268–9, 271–2, 275–6, 463–4, 466, 471, 475–81, 538, 592, 622
soil creep 103
soil liquefaction 189, 190, 206
soil moisture 8, 14–15, 265, 268, 275, 476, 591
soil organic carbon (SOC) 39, 44–5
soil profiles 56
solar energy 8–9, 40–1, 58, 267, 430, 597–8, 601, 605, 611, 627
solar insolation 58, 61, 62, 140–2, 615
solifluction 160, 164
solution (corrosion) 64, 99–00, 102
South Downs, UK 240–2
Southern (Antarctic) Ocean 138, 173, 176, 253, 334–8, 340–1, 343–5
sovereignty 297, 346
soya production 270
spatial pattern maps 648–50
spatial structure 409–10
Spearman rank correlation coefficient 657, 658–60
Special Area of Conservation (SAC) 121
special and differential agreements (SDTs) 313–14, 323
Special Economic Zones (SEZs) 319
species diversity 237
species richness 252
Spitalfields, East London 377–81, 383–5, 387–91
spits 108–10
stability 299, 301, 347
stacks 106–7, 116
standard deviation 656–7
statistics 369–70, 641, 642, 655–62
steel 577, 625
stemflow 10, 11
stereotypes 359, 361
stock resources 555
stores/components 2, 8–11, 61, 90–2, 104, 137, 144
storm (flood) hydrographs 15, 18–20, 433
storm flow 15, 18
storm surges 212–16, 218
storms 10, 12, 212–19, 224–5, 245, 338, 428, 445–6
stress, chronic 451
stripping ratio 621
stumps 106–7
sub-aerial processes 98, 102–4, 107
sub-cells 97, 124
sub-climax 230, 237–8
subduction 193, 194, 195–6, 198, 210, 227
subjectivity 351, 353, 369
sublimation 147
submergence (waste) 441
submergent features 114–16
subsidence 117–18, 173–4
subsidies 306
subsistence cultivators 270
suburbs 398, 401–2, 414, 416
succession 230, 235–7, 246–7, 249, 265
sulphur dioxide 31, 430
Sun 232, 605
  see also solar ...
Sundarbans, The 131–5, 279–81
supply 562, 576, 610–11, 629, 637
supply chains 294, 299, 316
supra-glacial debris 150
surface creep 61, 68, 102
surface storage 12
surges (glacial) 149
surveillance 413–14
surveys 389, 668
suspension 61, 64, 68, 100, 478
sustainability 124–33, 135, 167, 171, 252–9, 273, 278, 324, 395, 482, 543, 593–4, 614, 614–16, 623–25, 628
sustainable development 86–8, 174, 176, 181, 328, 331, 343, 449–54, 554, 556, 562–3
Sustainable Development Goals (SDGs) 328, 494, 541, 579, 582, 628
sustainable urban drainage systems (SUDS) 422, 434–7
swamps 247
swarms 154
swash 92–3, 101, 107–8
systems 2–3, 53, 468–71, 506, 527–30
  cascading 3, 10
  closed 3
  global 283, 298–306, 669
  isolated 3
  open 3, 10, 53, 61, 90–104, 137, 143–51, 231–2
tailings 609, 623–4
talik 159
talus 160
tariffs 289, 291, 306, 309, 312–14, 321, 323, 325, 347, 577
tarns 151, 179
tax 270, 308, 319–20, 442, 547
technology 189, 291, 298, 301–2, 447, 467, 475, 483, 488, 530–1, 613, 620–1, 626–9
tectonics 28, 191, 192–3
telecommunications 619
television 374, 388
temperature 8–9, 14, 40–1, 55–6, 65, 75, 77–9, 138–41, 144, 147–8, 159–60, 170, 232, 260–1, 265–6, 271, 274–5, 279, 332, 335, 337–8, 422–27, 457, 463, 472, 491, 495, 500, 536–8, 615, 634
temperature inversion 422, 430
tephra 197, 198, 199
terracettes 164
territorial claims 345
terrorism 294
tertiary sector 316, 404–7
  see also services
textual sources 372, 379, 388–91
Thames basin 17
Thames Tideway Tunnel 447
thermal expansion 39, 117
thermal fracture 61, 65
thermal stress 535–6
thermokarst 164, 173–4
Thompson, Warren 510
throughfall 10, 11, 12
thunderstorms 428

Thwaites Glacier 338
tidal power 597, 599, 601
tidal range 95
tidal (storm) surges 96
tides 95–6, 107
'tiger' economies 291
till 150, 153, 154
till plains 151, 155
tillage 45
timber 220, 273
tin 625
tombolos 109
'top-down' approaches 81, 329, 331
topography 231, 240, 265, 466, 477, 491, 500, 587, 606, 634
topsoil 19, 22, 48, 271, 275, 477–9
Torquay 377, 379–89
total productive bio-capacity 524, 526
tourism 174, 176, 227–8, 261–2, 339–40, 365, 368, 371–2, 636, 664
towns 410–12
  clone 356
traces, material/non-material 353
traction 64, 100
trade 210, 286, 291, 298–9, 301, 314–25, 440–1, 569–70, 574–7, 619, 624–5, 631–2
  fair 210, 324, 325
  free 295, 301, 306, 309, 312–13, 324–5, 330
  international 283–5, 294, 295–7, 306–27, 346, 467
  intra-regional 307–9, 311
trade agreements 295–7, 309–14
  bilateral 309, 314, 330, 331
  multilateral 330, 331
trade barriers 295, 306, 309, 330
trade diversion 297
trade wars 323
trading blocs, regional 289, 295–7, 310–11, 325
Trafalgar Square, London 353–4
traffic 436, 551
'tragedy of the commons' 333
train companies 368–9
transects 130
Transition Network 358
transnational corporations (TNCs) 286, 291, 299, 301, 314–26, 342, 467, 575–8, 606–7, 623, 627–8, 631
transpiration 8–12, 19, 141
transport 172, 174, 289, 294, 398, 400–1, 458, 613, 619
  public 418–19, 431–2, 454, 458
transportation 61, 64, 67–8, 70, 84–6, 100–2, 104, 147, 150
traps, structural/stratigraphic 604
trees 268, 276–7, 431, 444, 478
triangular graphs 653–4
trip lines 647–9
trophic levels 230, 234–5, 245
trough ends 152
Trump, Donald 347
truncated spurs 151, 153
tsunamis 188, 190, 197, 199, 206–9, 225
tundra 139, 141–2, 170–4, 180–83, 472
typhoons 212–19, 224–5
ubac slopes 144
ultraviolet radiation 533–4
under employment 400
underpopulation 523
understorey 268
unemployment 297–8, 400
United Nations (UN) 32, 299, 308, 328–9, 333, 341–2, 468, 473, 477, 482–3, 504
United States 309, 321, 347, 560–1, 579–80, 582, 503, 606
unsustainable practices 300
upwelling 95
uranium 566, 616
Urban Development Corporations (UDCs) 406
urban environments 20, 360, 366–7, 395–461, 497, 581, 669
urban growth 33, 395–401
urban heat islands 422, 423–5
urban resurgence 401, 403–4
urban sprawl 398, 401–2, 414
urbanisation 395–407, 415–22, 501, 568
US-Mexico-Canada Agreement (USMCA) 296, 297–8, 314
utilidors 172
vaccines 499
variance 656
varves 156, 158
vegetation 11, 12, 19, 26, 48, 57–8, 111, 113, 122, 132, 134, 141–2, 169, 176, 219–23, 230, 234, 240–1, 244–7, 246–9, 250–2, 265–71, 275–8, 592–3, 622, 634
vehicles 427, 431–2, 432, 613
ventifacts 70–1
Venturi effect 428
vertical deep mixing 29
villages, suburbanised 403
volcanic activity 31, 196–9
volcanic bombs 197, 199
volcanic explosivity index (VEI) 197, 198
volcanic gases 199
volcanic hazards 197–204, 224–5
volcanoes 190, 195, 197–9, 201–3, 224
volunteered geographical information 663
vulnerability 186–7, 189, 201, 206, 210–11, 214, 218, 223, 227, 340, 460, 534–5
wadis 70, 73, 76
wages
  living 417, 458
  minimum 417–18
waste disposal 395, 397–400, 437–43, 453–4, 563, 6∂∂∂16
waste generation 437–9
waste hierarchy 440
water 68–70, 73–6, 78, 86–8, 398–400, 425, 438, 462, 464, 477–8, 483–4, 538, 544, 554, 556, 562, 578–84, 588–95, 628
  abstraction 20, 22–4, 556, 582, 586, 588, 609
  availability 579, 580–4, 634–7
  demand 554, 578, 580–4, 585, 588, 594, 637
  freshwater 4, 6–7, 22–3, 244, 556, 579, 581, 585–6, 593, 636
  security 585–95
  sources 585–8
  stores 4–10
  surface 6–7, 11, 585, 586, 588
  types 4, 5–6, 6–8, 8, 8–9
water balance 10, 14–15, 54, 55
water cycle 2–24, 41–52, 114, 432–3, 593, 668
water pollution 246, 399–400, 432–3, 442, 445–8, 462, 493–4, 563, 609, 623–4, 633
water stress 579, 582, 586, 589
water surplus (areas) 579, 581
water table 8, 23
water vapour 2, 8–9, 427
wave pounding/quarrying 99
wave power 399, 597, 601
wave-cut notches 105
wave-cut platforms 95–6, 105–6, 107, 116
waves 93–5, 98–101, 129–31
wealth 185, 216, 402, 508
weather 444, 534–7, 646–7
weathering 25, 28, 57, 61–7, 74–5, 92, 95, 102, 147–8, 165, 268
Wegener, Albert 191
welfare 632, 633
well-being 313, 357, 455–7, 459–60, 485–506
West Nile Virus (WNV) 537
wetlands 7
whaling 335–6, 341, 343
Whittaker, Robert 266
wilderness areas 174–6
wildfires 219–24, 458, 659
wind 59, 62, 67–8, 70–1, 75, 84–6, 92–3, 96, 101–2, 108, 140–1, 212–16, 219, 232, 334, 338, 423, 427–9, 605
  katabatic 141
  prevailing 92–3, 605
  trade 59, 267, 274
  see also aeolian processes
wind energy 597–8, 601, 605–6, 611, 627
women 509, 515, 541, 546
woodlands 247, 250, 251
working conditions 320
working population 513, 543, 547
World Bank 299, 300, 301, 325, 329, 395, 438, 446, 504
World Health Organization (WHO) 299, 328, 443, 446, 486, 503–4
world peace 296, 299
World Trade Organization (WTO) 289, 296, 299, 306, 313, 324, 325, 329–30
xerophytes 57, 276
yardangs 71
zeugens 71
zika virus 492, 537

Acknowledgements

The Publishers would like to thank the following for permission to reproduce copyright material.

Text/artwork credits

Figure 1.21 © Shoothill Ltd; **Figure 1.25** © UKCEH Based on data provided by the UK National River Flow Archive, hosted by the UK Centre for Ecology & Hydrology; **Figure 1.33** British Geological Survey © NERC 2016. All rights reserved; **Figure 1.43** © Statista 2020; **Figure 1.45** The Natural Fix? – The Role of Ecosystems in Climate Mitigation. https://www.grida.no/resources/7570; **Figure 1.46** © Copyright 2020 IOP Publishing; **Figures 1.58, 1.60, 1.61** TS.1 from IPCC, 2005: Technical Summary. In: IPCC Special Report on Carbon Dioxide Capture and Storage. Prepared by Working Group III of the Intergovernmental Panel on Climate Change [Metz, B., O. Davidson, H. C. de Coninck, M. Loos, and L. A. Meyer (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 442 pp; **Figure 1.59** SaskPower; **Figures 1.68, 1.69** © UKCEH Based on data provided by the UK National River Flow Archive, hosted by the UK Centre for Ecology & Hydrology; **Table 2.53** Chinowsky, P., Hayles, C., Schweiker, A., Strzepek, N., Strzepek, K., Schlosser, A. (2011) Climate change: comparative impact on developing and developed countries, Engineering Project Organization Journal (March 2011), 1, 67–80; **Figure 3.4** Waugh, D. (2000). Geography: An integrated approach. Oxford University Press; **Figure 3.64** © Copyright 2018 IOP Publishing; **Figure 5.27** © Earth Observatory of Singapore; **Figure 5.37** Reprinted by permission from Springer Nature, NEWS FEATURE. The race to decipher how climate change influenced Australia's record fires, Nicky Phillips & Bianca Nogrady © 2020; **Figure 7.7** Based on estimates from Abel, G. J., & Cohen, J. E. (2019). Bilateral international migration flow estimates for 200 countries. Scientific Data, 6(1), 82. https://doi.org/10.1038/s41597-019-0089-3; **Figure 7.8** From KEY STATISTICS AND TRENDS in International Trade 2018 by United Nations Copyright ©2019 United Nations. Reprinted with the permission of the United Nations; **Figure 7.11** Top 10 countries by share of global manufacturing output 2018 derived from Statista https://www.statista.com/chart/20858/top-10-countries-by-share-of-global-manufacturing-output; **Figure 7.18** © 2020 International Monetary Fund. All Rights Reserved; **Figure 7.20** © 2020 Guardian News & Media Limited. All rights reserved; **Figure 7.21** From International Trade Statistics Yearbook. © 2017 United Nations. Used with the permission of the United Nations; **Figure 7.30** © Banana Link; **Figure 7.46** © The Economist Newspaper Limited 2020. All rights reserved; **Figure 7.47** © International Whaling Commission; **Figure 7.48** © International Association of Antarctica Tour Operators; **Figures 8.18, 8.19** © 2000–2020 Barcelona Field Studies Centre S.L. All Rights Reserved; **Figure 9.4** © 2020 The New Climate Economy; **Figure 9.12** © Stephen Codrington; **Figures 9.22, 9.24** © 2020 Guardian News & Media Limited. All rights reserved; **Figure 9.59** Rogers, R. (1997). Cities for a small planet. Faber and Faber; **Figure 9.60** © 2013 United Nations. Reprinted with the permission of the United Nations; **Figure 9.65** Aldridge, H., Barry Born, T., Tinson, A. & MacInnes, T. (2015) London's Poverty Profile 2015. Trust for London & New Policy Institute; **Table 9.11** © 2013 United Nations. Reprinted with the permission of the United Nations; **Figure 10.2** Center for International Earth Science Information Network (CIESIN), Columbia University. 2018. Gridded Population of the World, Version 4 (GPWv4): Population Density, Revision 11. Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC). https://doi.org/10.7927/H49C6VHW; **Figure 10.11** Waugh, D. (2000). Geography: An integrated approach. Oxford University Press; **Figure 10.31** © Norman Rosenthal; **Figure 10.34** © MARA/ARMA (Mapping Malaria Risk in Africa/Atlas du Risque de la Malaria en Afrique); **Figure 10.68** © 2020 American Cancer Society, Inc. All rights reserved; **Figure 10.75** From Future Population Growth by Max Roser, © 2017 United Nations. Reprinted with the permission of the United Nations; **Figure 10.85a** The state of global food security based on affordability, availability, safety and quality of food, 2019 derived from Statista; **Figure 11.9** IEA (2020), Pattern of world coal production, 2017. http://energyatlas.iea.org/#!/tellmap/2020991907. All rights reserved; **Figure 11.10a** IEA (2020), Pattern of world oil production, 2017.http://energyatlas.iea.org/#!/tellmap/-1920537974/0. All rights reserved; **Figure 11.10b** IEA (2020), Pattern of natural gas production, 2017 http://energyatlas.iea.org/#!/tellmap/-1165808390/0. All rights reserved; **Figure 11.11** BP Statistical Review of World Energy; **Figure 11.13** © 2020 Organization of the Petroleum Exporting Countries; **Figure 11.15** © 2020 Worldsteel Association; **Figure 11.17** © Engineering and Mining Journal; **Figure 11.22** © International Water Management Institute; **Figure 11.28** Reprinted with permission of the National Ground Water Association. Copyright 2017; **Figure 11.30** © Daily Mail; **Figure 11.42** Compound line graph showing France's primary energy mix, 1990–2018, 2017. https://www.iea.org/countries/france. All rights reserved; **Figure 11.52** © Reddie & Grose LLP; **Figure 11.56** © Carbon Brief Ltd; **Figure 11.53** The Association for Decentralised Energy; **Figure 11.59** © 2020 tradingwm.ae; **Figure 11.63** Republished with permission from an independent (pre-Covid) global energy forecast by Schalk Cloete (https://energypost.eu/an-independent-global-energy-forecast-to-2050-to-compare-with-the-ieas-weo-2018/); **Figure 11.71** Sharma, S. K., Kansal, M. L., & Tyagi, A. (2015). Resource assessment and strategic planning for improvement of water supply to Shimla city in India using geo-spatial techniques. The Egyptian Journal of Remote Sensing and Space Science, 18(1), 85–97; **Figure 11.73a** © Catch Energy, Snapshot of installed renewable energy; **Figure 11.73b** © 2020 – BBI International. All rights reserved; **Figure 11.74a** Alaska Department of Natural Resources; **Figure 12.7** © 2020 Centre for Cities; **Figure 12.11** © The Economist Newspaper Limited 2020. All rights reserved.

Photo credits
P.5 *l* © Anton Ivanov/Alamy Stock Photo; **p.1 and 6** © David Coleman/Alamy Stock; **p.16** © Richard Croft (www.geograph.org.uk); **p.27** © Frank Fichtmüller – stock.adobe.com; **p.30** © NASA; **p.38** ©Alexander – stock.adobe.com; **p.57** *t* ©bodikp/Thinkstock/iStock/Getty Images, *b* ©Jupiterimages/Thinkstock/iStock/Getty Images; **p.63** © Goddard Space Flight Center/NASA; **p.65** © Ann and Steve Toon/Alamy Stock Photo; **p.66** *tl* © Robert Miramontes/Thinkstock/iStock/Getty Images, *tr* © Charles O. Cecil/Alamy Stock Photo, *b* © Bruce M. Herman/Science Photo Library; **p.69** © imageBROKER/Alamy Stock; **p.72** © pum_eva/Thinkstock/iStock/Getty Images; **p.74** *l* © blickwinkel/Alamy Stock Photo; *r* © lucky-photographer/Thinkstock/iStock/Getty Images; **p.76** © Patrick Burke/Thinkstock/iStock/Getty Images; **p.83** *l t&b* NASA Earth Observatory images using Landsat data and declassified military satellite Corona data provided by the U.S. Geological Survey and Astronaut photograph ISS042-E-244403 provided by the ISS Crew Earth Observations Facility and the Earth Science and Remote Sensing Unit, NASA JSC, *r* © Peter Banos/Alamy Stock Photo; **p.84** © Ozkan Bilgin/Anadolu Agency/Getty Images; **p.86** *r* © George Steinmetz/Getty Images; **p.87** *l* © Hemis/Alamy Stock Photo, *r* © George Steinmetz; **p.88** © Hemis/Alamy Stock Photo; **p.89** © TMI/Alamy Stock Photo; **p.96** © John Worrall/Alamy Stock Photo; **p.105** © B.A.E. Inc./Alamy Stock Photo; **p.106** © imageBROKER/Alamy Stock Photo; **p.109** *l* © Mr. Nut/Alamy Stock Photo, *r* © LWM/NASA/LANDSAT/Alamy Stock Photo; **p.114** © NOAA; **p.116** © Luke Farmer/Alamy Stock Photo; **p.129** *l & r* © Pevensey Bay Coastal; **p.132** © Universal Images Group North America LLC/Alamy Stock Photo; **p.133** © Images & Stories/Alamy Stock Photo; **p.134** © Neil Cooper/Alamy Stock Photo; **p.143** *t* © NASA, *b* Reto Stöckli, NASA Earth Observatory; **p.146** *t* © Pav-Pro Photography – stock.adobe.com, *b* © Galen Rowell/Corbis via Getty Images; **p.148** © Scott Rylander Travel/Alamy Stock Photo; **p.153** ©jiggotravel – stock.adobe.com; **p.154** *l* ©Julia Katherine Rich, *r* © John Morrison/Alamy Stock Photo; **p.155** © Michael Dutton/Alamy Stock Photo; **p.158** © Barry Vincent/Alamy Stock Photo; **p.159** © Cultura Creative (RF)/Alamy Stock Photo; **p.160** © Erin Paul Donovan/Alamy Stock Photo; **p.163** © Timothy Epp/Alamy Stock Photo; **p.164** © Natural History Archive/Alamy Stock Photo; **p.172** © Bryan and Cherry Alexander Photography; **p.174** ©Vladimir Romanovsky; **p.175** © imageBROKER/Alamy Stock Photo; **p.179** *t* © Jane Hallin/Alamy Stock Photo, *b* © Dave Porter/Alamy Stock Photo; **p.180** © Bryan and Cherry Alexander Photography; **p.182** © Emma Wood/Alamy Stock Photo; **p.201** NASA/Jesse Allen; **p.202** © ZUMA Press, Inc./Alamy Stock Photo; **p.208** © Jon Sheer. All Rights Reserved. Getty Images; **p.217** © MediaPunch Inc/Alamy Stock Photo; **p.219** © JAY DIRECTO/Stringer/Getty Images; **p.223** *l* Alamy Stock Photo, *r* © Fotofritz/Alamy Stock Photo; **p.225** © Buiten-Beeld/Alamy Stock Photo; **p.226** © Pool AVENTURIER/LOVINY/Gamma-Rapho via Getty Images; **p.227** © Image Professionals GmbH/Alamy Stock Photo; **p.228** Melissa Delport/Emma Buckland; **p.237** © Reimar – stock.adobe.com; **p.240** © Robin Bush/Getty Images; **p.243** © BrianLasenby/Thinkstock/iStock/Getty Images; **p.254** © John Morrison/Alamy Stock Photo; **p.262** NOAA. How does land-based pollution threaten coral reefs? National Ocean Service website https://oceanservice.noaa.gov/facts/coral-pollution.html; **p.269** NASA – MODIS Rapid Response Team at Goddard Space Flight Center; **p.272** © dpa picture alliance/Alamy Stock Photo; **p.277** © imageBROKER/Alamy Stock Photo; **p.282** © Getty Images/iStockphoto/Thinkstock; **p.285** © Alan Gignoux/Alamy Stock Photo; **p.291** © Prisma Bildagentur AG/Alamy Stock Photo; **p.294** © Jack Hinds/Alamy Stock Photo; **p.301** © National Geographic Image Collection/Alamy Stock Photo; **p.315** © ITAR-TASS Photo Agency/Alamy Stock Photo; **p.317** © P Wei/Getty Images; **p.319** © Kin Cheung/AP/Shutterstock; **p.326** © EDU Vision/Alamy Stock Photo; **p.337** *l* © wanderworldimages/Alamy Stock Photo, *r* © Kevin Schafer/Alamy Stock Photo; **p.339** © staphy/Thinkstock/iStock/Getty Images; **p.343** © kkaplin/Thinkstock/iStock/Getty Images; **p.344** © Antarctic and Southern Ocean Coalition; **p.352** *m* © Steve Taylor ARPS/Alamy, *b* © Roger Cracknell 01/classic/Alamy; **p.353** © Alistair Laming/Alamy; **p.355** KHALED DESOUKI/AFP via Getty Images; **p.356** *l* © Sven Schermer/123RF.COM, *r* Image by Wordsmith One for Transition Town Totnes' Clone-stopping Campaign; **p.357** © Mr Standfast/Alamy Stock Photo; **p.358** Used with permission from Project for Public Places. Website: http://www.pps.org/reference/grplacefeat/; **p.362** © Countrywide Images/Alamy Stock Photo; **p.364** © Panther Media GmbH/Alamy Stock Photo; **p.365** *l* © Education Images/UIG/Getty Images, *r* © jvdwolf/123RF.COM; **p.368** *l* © bloodua/123RF.COM, *r* © Pembrokeshire Coast National Park Authority; **p.369** © Science & Society Picture Library/Contributor/Getty Images; **p.370** © DEA PICTURE LIBRARY/De Agostini/Getty Images; **p.371** © Dr Jon Anderson/http://www.spatialmanifesto.com/methodology-projects/biomapping; **p.372** *l* © Sean Pavone / Alamy Stock Photo; **p.375** *tl* © classicpaintings/Alamy Stock Photo, *bl* © eye35/Alamy Stock Photo, *r* © Alan Finlayson/123RF.COM; **p.376** Courtesy of Pest Control Office, Banksy, London, 2005. Photo © Luz Martin/Alamy Stock Photo; **p.378** From LSE Library's collections, BOOTH/E/1/5/https://creativecommons.org/publicdomain/mark/1.0/; **p.380** *l* © travellinglight/Alamy Stock Photo *r* © bombaert/123RF.COM; **p.387** *tl* © Becky Bettesworth (www.beckybettesworth.co.uk), *tr* © STIK, photo © Jamie Gladden/Alamy Stock Photo, *b* © Adam Dant 2014 all rights reserved; **p.388** *t* © stevenjfrancis/Thinkstock/iStock/Getty Image; **p.390** *t* © Topical Press Agency/Hulton Archive/Getty Images, *b* © macfromlondon/123RF.COM; **p.394** © Bridgeman Images; **p.399** © Jeffrey Isaac Greenberg 16+/Alamy Stock Photo; **p.400** © Caroline Knowles (www.flipfloptrail.com); **p.401** © Paul Hayward/Alamy Stock Photo; **p.402** © Trinity Mirror/Mirrorpix/Alamy Stock Photo; **p.415** © Adeel Halim/Bloomberg/Getty Images; **p.416** https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/833959/IoD2019_Infographic.pdf. Contains public sector information licensed under the Open Government Licence v3.0; **p.419** © Stefano Paterna/Alamy Stock Photo; **p.424** *l t&b* NASA image by Marit Jentoft-Nilsen, based on Landsat-7 data, *br* © Tony McConnell/Science Photo Library; **p.426** *t* © John S Lander/LightRocket via Getty Images, *b* © tjs11/Thinkstock/iStock/Getty Images; **p.427** © Lou Linwei/Alamy Stock Photo; **p.429** © Zoonar RF/Thinkstock/Getty Images; **p.430** © Duncan Chard/Bloomberg/Getty Images; **p.431** © Cultura RM/Alamy Stock Photo; **p.434** *l* © trekandshoot/Thinkstock/iStock/Getty Images, *r* © Saxon Holt/Alamy Stock Photo; **p.435** © CN/Lee Jae-Won LJW/REUTERS; **p.436** © Wendy Connett/Alamy Stock Photo; **p.444** *t* © ZUMA Press, Inc./Alamy Stock Photo, *b* © Lourens Smak/Alamy Stock Photo; **p.446** © Bethany Clarke/Getty Images; **p.448** *both* © Paul Talling/www.derelictlondon.com; **p.449** Heidelberg University of Education, Research Group for Earth Observation (rgeo); **p.452** *l* © Joanne Underhill/Alamy Stock Photo, *r* © Cultura RM Exclusive/Atli Mar Hafsteinsson/Getty Images; **p.455** © A.P.S. (UK)/Alamy Stock Photo; **p.456** *l* © SJ Images/Alamy Stock Photo, *r* © TravelStockCollection – Homer Sykes/Alamy Stock Photo; **p.457** *l t&b* © Worldmapper/Trust for London; **p.459** © Charles Stirling (Travel)/Alamy Stock Photo; **p.460** © Dinodia Photos/Alamy Stock

Photo; **p.472** © All Rights Reserved. Dean Swope 2009; **p.473** © Tim Gainey/Alamy Stock Photo; **p.476** © Yuri Kravchenko/Alamy Stock Photo; **p.478** © Kathy Wright/Alamy Stock Photo; **p.479** © Universal Images Group via Getty Images; **p.480** © Jim Holmes; **p.493** © Sanjeev Verma/Hindustan Times/Getty Images; **p.498** *t* © Lookinglost/Alamy Stock Photo, *b* © Irene Abdou/Alamy Stock Photo; **p.502** © ZUMA Press, Inc./Alamy Stock Photo; **p.504** © ton koene/Alamy Stock Photo; **p.512** *t* © buraktumler/Shutterstock.com, *b* © Klaus Lang/Alamy Stock Photo; **p.519** © Italian Navy/dpa/Alamy Live News; **p.523** © frans lemmens/Alamy Stock Photo; **p.525** Worldmapper.org; **p.541** © Chris Fredriksson/Alamy Stock Photo; **p.546** © David Smith/Alamy Stock Photo; **p.560** © Paul Andrew Lawrence/Alamy Stock Photo; **p.589** © Cambridge Aerial Photography/Alamy Stock Photo; **p.591** *l* © Images of Africa Photobank/Alamy Stock Photo, *r* © Design Pics Inc/Alamy Stock Photo; **p.597** © David Wootton/Alamy Stock Photo; **p.598** © Kevin Foy/Alamy Stock Photo; **p.600** © Snorri Gunnarsson/Alamy Stock Photo; **p.607** © Simon Dawson/Bloomberg via Getty Images; **p.608** © Milesy/Alamy Stock Photo; **p.609** © Aurora Photos/Alamy Stock Photo; **p.622** © Tim Graham/Alamy Stock Photo; **p.624** © Agencia Brasil/Alamy Stock Photo; **p.632** © ABP/George Esiri GE/Reuters; **p.633** © George Osodi/AP/Shutterstock; **p.635** © paul kennedy/Alamy Stock Photo; **p.637** © AFP via Getty Images; **p.664** © Eric Fischer, using data Flickr and Picasa. Base map data © OpenStreetMap contributors, CC-BY-SA; **p.667** NASA Earth Observatory image by Joshua Stevens, using Landsat data from the U.S. Geological Survey; **p.669** *l* © Steve Taylor ARPS/Alamy Stock Photo.

AQA A-level Geography: Boost

AQA A-level Geography: Boost is the digital teaching and learning solution to creating outstanding lessons that ensure understanding and progression for all your students.

Using a wide range of resources and assessments, with updates to the content every year, AQA A-level Geography: Boost:

- Helps you teach the course the way you want to, using the fully editable and flexible **Course Plans**, linked to the eBook version of the textbook, that you can customise, add your own resources to, and share with colleagues
- Saves time planning with **ready-made lessons** to support the implementation of every chapter in the new edition textbook, with 50 PowerPoints and 76 worksheets, including an extension worksheet for every chapter (taking students learning beyond the textbook and ideal for homework tasks), and answers to all the Review questions in the textbook
- Provides tools to assess understanding and track student progress throughout the course with at least two **interactive knowledge tests** for every chapter to help with knowledge recall, and bonus exam-style questions for every chapter
- Extends learning with **3 bonus case studies** for the most popular units, and **key terms sheets** to help you use the textbook content flexibly

AQA A-level Geography Fifth Edition: Boost eBook

Boost eBooks are interactive, accessible and flexible. They use the latest research and technology to provide the very best experience for students and teachers.

- **Personalise**. Easily navigate the eBook with search, zoom and an image gallery. Make it your own with notes, bookmarks and highlights.
- **Revise**. Select key facts and definitions in the text and save them as flash cards for revision.
- **Listen**. Use text-to-speech to make the content more accessible to students and to improve comprehension and pronunciation.
- **Switch**. Seamlessly move between the printed view for front-of-class teaching and the interactive view for independent study.
- **Download**. Access the eBook offline on any device – in school, at home or on the move – with the Boost eBooks app.

To subscribe or register for a free trial, visit
hoddereducation.co.uk/subjects/geography/a-level-geography-2021-editions/

Boost

The Boost course and Boost eBooks have not been entered into the AQA approval process

AQA Geography For A-level

Develop in-depth knowledge and advanced geographical skills using this Student Book from the leading A-level Geography publisher*, fully updated for 2016 with detailed content coverage, fascinating case studies and assessment preparation materials.

Written by subject specialists with examining experience, this book helps you to:

- Learn, understand and revise the content you need to know using clear explanations of every topic and definitions of key terms
- Bring your study of geographical theory and concepts to life with up-to-date examples and case studies for each physical and human topic
- Build your cartographic, graphical and statistical skills through Skills Focus activities for each chapter and a separate skills section at the end of the book
- Reflect on your learning and map your progress with key questions throughout and end-of-chapter review questions
- Approach your exams with confidence by working through a variety of AS and A-level practice questions
- Be inspired by exciting fieldwork project ideas that help you plan and carry out your AS or A-level fieldwork and investigations

This title has been approved by AQA

Authors

Malcolm Skinner is a former A-level Geography Chief Examiner and Chair of Examiners. He has written and edited many successful A-level Geography textbooks.

Philip Banks is an experienced author and teacher with examining experience. He now works as an educational consultant.

Paul Abbiss, Helen Fyfe and Ian Whittaker are all experienced Heads of Geography with examining experience at AS and A-level.

Boost

This series includes an eBook and digital teaching and learning support.

Visit hoddereducation.co.uk/boost to find out more.

*Taken from Educational Publishers Council statistics

The Boost course and Boost eBooks have not been entered into the AQA approval process

HODDER EDUCATION
t: 01235 827827
e: education@hachette.co.uk
w: hoddereducation.co.uk

Schools have a *Licence to Copy* one chapter or 5% for teaching

CLA Copyright Licensing Agency

ISBN 978-1-3983-1254-8

FSC