Soil Conditions and Plant Growth
Soil Conditions and Plant Growth

Edited by

Peter J. Gregory
Centre for Food Security
School of Agriculture, Policy & Development
University of Reading
Reading, UK
and
East Malling Research
New Road
East Malling
Kent, UK

Stephen Nortcliff
Soil Research Centre
Department of Geography and Environmental Science
University of Reading
Reading, UK
## Contents

*Contributors* viii

*Preface* x

1 The historical development of studies on soil–plant interactions  1
Stephen Nortcliff and Peter J. Gregory

1.1 Introduction  1
1.2 The search for the ‘principle’ of vegetation, 1630–1750  2
1.3 The search for plant nutrients  4
1.4 The beginnings of soil microbiology  10
1.5 The development and application of modern knowledge of soils  13
   References  20

2 Plant and crop science  22
Sayed Azam-Ali

2.1 Introductory principles of resource capture by crops  22
2.2 Growth and development  23
2.3 Solar radiation  27
2.4 Water  32
2.5 Integrating principles of resource capture by crops  42
2.6 Conclusion  45
   References  46

3 Soil fertility  49
Elizabeth A. Stockdale, Keith W.T. Goulding, Timothy S. George, and Daniel V. Murphy

3.1 Introduction  49
3.2 Concepts of soil fertility  50
3.3 Plant–soil interactions determining soil fertility  57
3.4 Assessing soil fertility  62
3.5 Managing soil fertility  72
   References  79

4 Soil organic matter  86
David Powlson, Pete Smith, and Maria De Nobili

4.1 Introduction  86
4.2 Factors influencing SOM content  88
4.3 Modelling SOM dynamics  93
4.4 Chemical structure of SOM as elucidated using new forms of spectroscopy – advances and pitfalls  100
4.5 Organo-mineral interactions 108
4.6 Fractionation methods for studying SOM 111
4.7 The living part of SOM – the soil microbial biomass 115
4.8 SOM and soil physical conditions and implications for plant growth 116
4.9 Soil C and climate change 120
References 126

5 Nitrogen and phosphorus cycles and their management 132
Phil M. Haygarth, Richard D. Bardgett, and Leo M. Condron
5.1 Introduction 132
5.2 Sources of nitrogen and phosphorus to soil-plant systems 134
5.3 Chemical and biological processes of nutrient cycling, transformations and bioavailability 137
5.4 Processes of nitrogen and phosphorus losses from soils 145
5.5 Nitrogen and phosphorus use in agricultural systems 149
5.6 Future soil nutrient cycles and environmental change 153
References 155

6 Properties and management of cationic elements for crop growth 160
Philip J. White and Duncan J. Greenwood
6.1 The movement of essential cationic elements to the root surface 160
6.2 Cationic elements in plants 163
6.3 The uptake of essential cationic elements from the rhizosphere 169
6.4 Cationic elements in the soil solution 173
6.5 Management practices increasing the acquisition of essential mineral cations 180
6.6 Genetic strategies to improve the efficiency of acquisition and physiological use of essential cationic elements 183
6.7 Concluding remarks 186
References 187

7 Managing adverse soil chemical environments 195
Mark E. Hodson and Erica Donner
7.1 Potentially toxic metals (PTMs) 195
7.2 Salt-affected soils (saline, sodic and saline-sodic soils) 206
7.3 Acid and alkaline soils 218
References 228

8 Managing the soil physical environment for plants 238
Paul D. Hallett and A. Glyn Bengough
8.1 Introduction 238
8.2 Basic physical properties of soil 238
8.3 Soil structure 241
8.4 Soil tilth and seedbeds 249
8.5 Critical soil physical conditions for plant growth 257
References 261
9 Soil water and its management
Anne Verhoef and Gregorio Egea

9.1 Soil water management – a general overview 269
9.2 Key concepts related to water in soil 270
9.3 Soil water and plants 281
9.4 Determination of soil water balance and related variables 288
9.5 Water management to alleviate soil water shortage 298
9.6 Water management to alleviate excess water 308
References 312

10 Plant-induced changes in soil processes and properties
Philippe Hinsinger

10.1 Introduction 323
10.2 Roles of plants in carbon cycling: deposition, burial and emission 324
10.3 Roles of plants in soil acidification 335
10.4 Roles of plants in the weathering of soil minerals 342
10.5 Conclusion 353
References 355

11 Sustainable management of soil and plant health by optimising soil biological function
Dominic Standing and Ken Killham

11.1 Introduction 366
11.2 Managing soil and plant health 368
11.3 Conclusion 386
References 386

12 Managing the soil–plant system for the delivery of ecosystem services
David W. Hopkins and Edward G. Gregorich

12.1 Introduction 390
12.2 Roles of soils 394
12.3 Sustainable soil use and soil quality 409
12.4 Concluding remarks 413
References 414

13 The new challenge – sustainable production in a changing environment
Peter J. Gregory and Stephen Nortcliff

13.1 Introduction 417
13.2 Human-induced environmental change 418
13.3 Food security and food production 426
13.4 Routes to sustainable intensification 434
References 442

Index 449

A colour plate section falls between pages 214 and 215
Contributors

Sayed Azam-Ali
Crops for the Future Research Centre
Selangor Darul Ehsan, Malaysia

Richard D. Bardgett
Lancaster Environment Centre
Lancaster University
Lancaster, Lancashire, UK

The Late Duncan J. Greenwood FRS
University of Warwick
Wellesbourne, Warwick, UK

A. Glyn Bengough
The James Hutton Institute
Invergowrie, Dundee, UK
Division of Civil Engineering,
University of Dundee, Dundee, UK

Leo M. Condron
Agriculture and Life Sciences
Lincoln University
Canterbury, New Zealand

Erica Donner
Centre for Environmental Risk
Assessment and Remediation
University of South Australia
Mawson Lakes, South Australia, Australia

Gregorio Egea
Area of Agro-Forestry Engineering
School of Agricultural Engineering
University of Seville
Seville, Spain

Timothy S. George
The James Hutton Institute
Invergowrie, Dundee, UK

Keith W.T. Goulding
Department of Sustainable Soils and Grassland Systems
Rothamsted Research
Harpenden, Hertfordshire, UK

Edward G. Gregorich
Agriculture and Agri-Food Canada
Central Experimental Farm
Ottawa, Ontario, Canada

Peter J. Gregory
Centre for Food Security
School of Agriculture, Policy and Development
University of Reading
Reading, UK
and
East Malling Research
East Malling, Kent, UK

Paul D. Hallett
The James Hutton Institute
Invergowrie, Dundee, UK

Philippe Hinsinger
INRA
UMR Eco&Sols
Montpellier, France

Phil M. Haygarth
Lancaster Environment Centre
Lancaster University
Lancaster, Lancashire, UK
Mark E. Hodson
Environment Department
University of York
Heslington, York, UK

David W. Hopkins
School of Life Sciences
Heriot-Watt University
Edinburgh, UK

Ken Killham
Honorary Fellow
The James Hutton Institute
Invergowrie, Dundee, UK

Daniel V. Murphy
Soil Biology Group
School of Earth and Environment
UWA Institute of Agriculture
The University of Western Australia
Crawley, Western Australia, Australia

Maria de Nobili
Dipartimento di Scienze Agrarie e Ambientali
Università degli studi di Udine
Udine, Italy

Stephen Nortcliff
Department of Geography and Environmental Science
Soil Research Centre
University of Reading
Reading, UK

David Powlson
Department of Sustainable Soils and Grassland Systems
Rothamsted Research
Harpenden, Hertfordshire, UK

Pete Smith
School of Biological Sciences
University of Aberdeen
Aberdeen, UK

Dominic Standing
School of Biological Sciences
University of Aberdeen
Aberdeen, UK

Elizabeth A. Stockdale
School of Agriculture
Food and Rural Development
Newcastle University
Newcastle-upon-Tyne, UK

Anne Verhoef
Department of Geography and Environmental Science
Soil Research Centre
University of Reading
Reading, UK

Philip J. White
The James Hutton Institute
Invergowrie, Dundee, UK
Since the last edition of *Russell’s Soil Conditions and Plant Growth* in 1988, soil, plant and crop sciences have moved on considerably. For a long time during these 24 years, there was a diminishing interest in soil science as an underpinning element of crop production largely because the higher income countries of the world were food secure and, indeed, at times awash with surplus crops. Soil science became an environmental science with a much broader remit, and courses at universities changed to meet the training of this new wave of students. One consequence of these changes was that the traditional publisher of this textbook saw no demand for a new edition.

Recently, though, the mood has changed again, and both the rising awareness of global food insecurity and the need for soils and land to deliver simultaneously food and fibre and other ecosystem goods and services have focused attention on the requirement to better understand and manage the many interactions that occur between soils and plants. What better time for a new edition of *Soil Conditions and Plant Growth* to examine these interactions?

In preparing for this book, we decided at an early stage to let go of some of the content of the previous editions. There are now many introductory textbooks on soils and soil science; hence, we have quite deliberately excluded from this edition elements of pedology, chemistry, microbiology and soil survey that featured previously. Instead, we have tried to focus on what has always been the core feature of this book – the interactions between soils and plants. We have included accounts that detail how plants respond to soil properties but also how plants themselves are key agents in soil formation and modifiers of their environment. We have also chosen writers with some experience of how soils can be managed in both agricultural and ecological contexts to promote crop production but also to deliver high-quality water supplies, cope with the warming climate and all of the many other necessities of life that we have come to expect from our soils.

The contributors to this book would particularly like to thank (1) Dr Clare Benskin (Lancaster University, UK), Dr Tim George (James Hutton Institute, UK), Dr Alan Richardson (CSIRO, Australia) and Dr Ben Turner (Smithsonian Tropical Research Institute, Republic of Panama), who all contributed helpful material or input to preparing Chapter 5; (2) the financial support of The Scottish Government Rural and Environment Research and Analysis Directorate for Chapter 6, which is dedicated to the memory of Dr Duncan J. Greenwood (1932–2010); and (3) Dr Jos Raaijmakers and Dr Paolina Garbeva for providing the bacterial strain referred to in Figure 11.10.

Finally, we would like to thank acknowledge with gratitude the encouragement that we received in our careers from Professors Walter Russel, Dennis Greenland and Alan Wild and to Wiley-Blackwell for taking on the publication of this book. We also thank our many authors for so readily agreeing to write for us and for their enthusiasm in updating what, for many of us as students, was a soil science classic. This year, 2012, is the centenary of the first edition of Russell’s book, and we hope that our readers will find the same inspiration as previous generations of students and researchers.

Peter J. Gregory
Stephen Nortcliff

March 2012
1 The historical development of studies on soil–plant interactions

Stephen Nortcliff1 and Peter J. Gregory2,3

1 Department of Geography and Environmental Science, Soil Research Centre, University of Reading, Reading, UK
2 Centre for Food Security, School of Agriculture, Policy and Development, University of Reading, Reading, UK
3 East Malling Research, Kent, UK

1.1 Introduction

How plants grow and how this growth varies through time and in response to changing conditions has been an interest of people for millennia. From the early cultivators to present-day gardeners, there has been a fascination in how a flourishing plant can be derived from a dry, apparently lifeless seed. Furthermore, there has been recognition that plant growth shows different patterns in response to weather conditions and that it varies from place to place. As global population continues to increase the need to understand the growth of plants and the role of soils in crop production becomes increasingly important. The demand for both food and biomass-derived energy from plants is increasing, so we must also seek to understand how to allocate land for multiple purposes. Soils must be used for these services and to obtain other essential services such as clean water and a diverse soil community of organisms.

Many early civilisations appear to have compiled information on plant growth and crop husbandry, and there was an extensive literature on agriculture developed during Roman times, which provided important guidance on crop growth and management for many centuries after the fall of the Roman Empire. The Roman literature was collected and condensed into one volume about the year 1309 by a senator of Bologna, Petrus de Crescentius (the book was made more widely available when published in 1471), whose book was one of the most popular treatises on agriculture of any time, being frequently copied, and in the early days of printing, passing through many editions. Many other agricultural books appeared in the fifteenth and early sixteenth centuries, notably in Italy and later in France. In some of these are found certain ingenious speculations that have been justified by later work. Such, for instance, is Palissy’s remarkable statement in 1563:

You will admit that when you bring dung into the field it is to return to the soil something that has been taken away…. When a plant is burned it is reduced to a salty ash called alcały by apothecaries and philosophers…. Every sort of plant without exception contains some kind of salt. Have you not seen certain labourers when sowing a field with wheat for the second year
in succession, burn the unused wheat straw which had been taken from the field? In the ashes will be found the salt that the straw took out of the soil; if this is put back the soil is improved. Being burnt on the ground it serves as manure because it returns to the soil those substances that had been taken away.

But while some of these speculations have been confirmed, many in other sources have not, and the beginnings of agricultural chemistry was to take place later when we had learnt the necessity for investigating possible relationships and pathways using experiments.

1.2 The search for the ‘principle’ of vegetation, 1630–1750

It was probably discovered at an early stage in agricultural development that manures, composts, dead animal bodies and parts of animals, such as blood, all increased the productivity of the land; and this was the basis of the ancient saying that ‘corruption is the mother of vegetation’. Although there was empirical evidence for this linkage, the early investigators consistently ignored this ancient wisdom when they sought for the ‘principle’ of vegetation to account for the phenomena of soil fertility and plant growth. Thus, the great Francis Bacon, Lord Verulam, believed that water formed the ‘principal nourishment’ of plants, the purpose of the soil being to keep them upright and protect them from excessive cold or heat, though he also considered that each plant drew a ‘particular juyce’ from the soil for its sustenance, thereby impoverishing the soil for that particular plant and similar ones, but not necessarily for other plants. Similarly, van Helmont (1577–1644) regarded water as the sole nutrient for plants, and his interpretation of a carefully undertaken experiment in which he grew willows concluded that water was the principal requirement for plant growth (van Helmont, 1648).

Robert Boyle (1661) repeated the experiment with ‘squash, a kind of Italian pompion’ and obtained similar results. Boyle further distilled the plants and concluded, quite justifiably from his premises, that the products obtained, ‘salt, spirit, earth, and even oil, may be produced out of water’. While these experiments were laudable, they ignored the part played by air, and in the van Helmont experiment there was a small reduction in the amount of soil present, which was ignored, although we now know this to be significant. In some respects, this might be taken as a guide for many of the future experiments undertaken in agriculture; if the hypotheses are wrong and other hypotheses are ignored, conclusions which may appear to be valid will often turn out to be incorrect because the alternatives have been ignored.

The primacy of water in plant growth was questioned by an experiment published by John Woodward in a fascinating paper (1699). Based on the experiments of van Helmont and of Boyle, he grew spearmint in water obtained from various sources and noted that all of these plants were supplied with an abundance of water so that all should have made equal growth had nothing more been needed. The amount of growth, however, increased with the impurity of the water (Table 1.1). He concluded:

Vegetables are not formed of water, but of a certain peculiar terrestrial matter. It has been shown that there is a considerable quantity of this matter contained in rain, spring and river water, that the greatest part of the fluid mass that ascends up into plants does not settle there but passes through their pores and exhales up into the atmosphere: that a great part of the terrestrial matter, mixed with the water, passes up into the plant along with it, and that the plant is more or less
The historical development of studies on soil–plant interactions

augmented in proportion as the water contains a greater or less quantity of that matter; from all of which we may reasonably infer, that earth, and not water, is the matter that constitutes vegetables.

Taking account of the results in his experiment, he discussed the use of manures and the fertility of the soil from this point of view, attributing the well-known falling off in crop yield when plants are grown for successive years on unmanured land to the circumstance that:

the vegetable matter that it at first abounded in being extracted from it by those successive crops, is most of it borne off…. The land may be brought to produce another series of the same vegetables, but not until it is supplied with a new fund of matter, of like sort with that it at first contained; which supply is made several ways, either by the ground’s being fallow some time, until the rain has poured down a fresh stock upon it; or by the tiller’s care in manuring it.

The best manures, he continued, are parts either of vegetables or of animals, which ultimately are derived from vegetables.

For a time there was little progress in relation to what plants needed in addition to water and how these needs might be met. Advances were, however, being made in agricultural practice. One of the most important was the introduction of the drill and the horse-hoe by Jethro Tull, an Oxford man of a strongly practical turn of mind, who insisted on the vital importance of getting the soil into a fine, crumbly state for plant growth. Tull (1731) was more than an inventor; he discussed in most picturesque language the sources of fertility in the soil. In his view, it was not the juices of the earth but the very minute particles of soil loosened by the action of moisture that constituted the ‘proper pabulum’ of plants. The pressure caused by the swelling of the growing roots forced these particles into the ‘lacteal mouths’ of the roots, where they entered the circulatory system. All plants lived on these particles, i.e. on the same kind of food; it was incorrect to assert, as some had done, that different kinds of plants fed as differently as horses and dogs, each taking its appropriate food and no other. Plants will take in anything that comes their way, good or bad. A rotation of crops is not a necessity, but only a convenience. Conversely, any soil will nourish any plant if the temperature and water supply are properly regulated. Hoeing increased the surface of the soil or the ‘pasture of the plant’ and also enabled the soil to better absorb the nutritious vapours condensed from the air. Dung acted in the same way, but was more costly and less efficient.

Table 1.1 Growth of spearmint using water from different sources.

<table>
<thead>
<tr>
<th>Source of water</th>
<th>Mass (g) of plants when planted</th>
<th>Mass (g) of plants when harvested</th>
<th>Mass (g) gained in 7 days</th>
<th>Expense (g) of water (transpiration)</th>
<th>Ratio increase in mass:mass water used</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rain water</td>
<td>1.83</td>
<td>2.96</td>
<td>1.13</td>
<td>220.3</td>
<td>1:195</td>
</tr>
<tr>
<td>River Thames</td>
<td>1.81</td>
<td>3.50</td>
<td>1.69</td>
<td>161.5</td>
<td>1:95.6</td>
</tr>
<tr>
<td>Hyde Park Conduit</td>
<td>7.13</td>
<td>16.14</td>
<td>9.01</td>
<td>851.5</td>
<td>1:94.5</td>
</tr>
<tr>
<td>Hyde Park Conduit plus 105g garden mould</td>
<td>5.96</td>
<td>24.36</td>
<td>18.40</td>
<td>968.8</td>
<td>1:52.7</td>
</tr>
</tbody>
</table>

Source: From Woodward (1699).

Table 1.1 Growth of spearmint using water from different sources.
The position at the end of this period cannot better be summed up than in Tull’s own words: ‘It is agreed that all the following materials contribute in some manner to the increase of plants, but it is disputed which of them is that very increase or food: (1) nitre, (2) water, (3) air, (4) fire, (5) earth’.

### 1.3 The search for plant nutrients

#### 1.3.1 The phlogistic period, 1750–1800

Great interest was taken in agriculture in the UK during the latter half of the eighteenth century. Many experiments were conducted, facts were accumulated, books written and societies formed for promoting agriculture. The Edinburgh Society, established in 1755 for the improvement of arts and manufactures, induced Francis Home ‘to try how far chymistry will go in settling the principles of agriculture’ (1757). The whole art of agriculture, he says, centres in one point: the nourishing of plants. Investigation of fertile soils showed that they contain oil, which is therefore a food of plants. But when a soil has been exhausted by cropping, it recovers its fertility on exposure to air, which therefore supplies another food. Home established pot experiments to ascertain the effect of various substances on plant growth. ‘The more they [i.e. farmers] know of the effects of different bodies on plants, the greater chance they have to discover the nourishment of plants, at least this is the only road.’ Saltpetre, Epsom salts, vitriolated tartar (i.e. potassium sulphate) all lead to increased plant growth, yet they are three distinct salts. Olive oil was also useful. It is thus clear that plant food is not one thing only, but several; he enumerates six: air, water, earth, salts of different kinds, oil and fire in a fixed state. As further proof he shows that ‘all vegetables and vegetable juices afford those very principles, and no other, by all the chymical experiments which have yet been made on them with or without fire’.

Between 1770 and 1800, work was done on the effect of vegetation on air that was destined to revolutionise the ideas of the function of plants in the economy of nature, but its agricultural significance was not recognised until later. Joseph Priestley, knowing that the atmosphere becomes vitiated by animal respiration, combustion, putrefaction, etc., and realising that some natural purification must go on, or life would no longer be possible, was led to try the effect of sprigs of living mint on vitiated air (1775). He found that the mint made the air purer and concludes ‘that plants, instead of affecting the air in the same manner with animal respiration, reverse the effects of breathing, and tend to keep the atmosphere pure and wholesome, when it is become noxious in consequence of animals either living, or breathing, or dying, and putrefying in it’. But he had not yet discovered oxygen and so could not give precision to his discovery; and when, later on, he did discover oxygen and learn how to estimate it, he unfortunately failed to confirm his earlier results because he overlooked a vital factor, the necessity for light. He was therefore unable to answer Scheele, who had insisted that plants, like animals, vitiate the air. It was Jan Ingen-Housz (1779) who reconciled both views and showed that purification goes on in light only, while vitiation takes place in the darkness. Ingen-Housz’s conclusions might be summarised as follows: (1) light is necessary for this restoration (this we would now know as photosynthesis); (2) only the green parts of the plant actually perform restoration and (3) all living parts of the plant ‘damage’ the air (respire), but the extent of air restoration by a green plant far exceeds its damaging effect. Jean Senebier (1782) working in Geneva also concluded that the plant–atmosphere interactions were significant.
1.3.2 The period 1800–1860

The foundation of plant physiology

Progress to this point had been constrained by methodologies which were known and accepted. To gain new insights and further investigate some of the speculation, particularly in relation to the plant–atmosphere links, new methods were required. The work of Nicholas Theodore de Saussure (1804) in the early nineteenth century in establishing the broad principles of the quantitative experimental method, in some respects produced the paradigm shift which proved the basis for modern agricultural chemistry. The work of Boussingault, Liebig, Lawes and Gilbert drew on this new experimental method which still provides the basis of many investigations. De Saussure grew plants in air or in known mixtures of air and carbon dioxide, and measured the gas changes by eudiometric analysis and the changes in the plant by ‘carbonisation’. He was thus able to demonstrate the central fact of plant respiration – the absorption of oxygen and the evolution of carbon dioxide – and further show the decomposition of carbon dioxide and evolution of oxygen in light. Carbon dioxide in small quantities was a vital necessity for plants, and they perished if it was artificially removed from the air. It furnished them not only with carbon, but also with some oxygen. Water is also decomposed and fixed by plants. On comparing the amount of dry matter gained from these sources with the amount of material that can enter through the roots, he concluded that even under the most favourable conditions the soil furnished only a very small part of the plant food. Small as it is, however, this part is indispensable: it supplies nitrogen which he described as an essential part of vegetation and, as he had shown, was not assimilated directly from the air, and also ash constituents, which he noted contributed to the solid parts of plants, just as with animals. Further, he showed that the root is not a mere filter allowing any and every liquid to enter the plant; it has a special action and takes in water more readily than dissolved matter, thus effecting a concentration of the solution surrounding it; different salts, also, were absorbed to different extents. Passing next to the composition of the plant ash, he showed that it was not constant, but varies with the nature of the soil and the age of the plant; it consists mainly, however, of alkalis and phosphates. All the constituents of the ash occur in humus. If a plant is grown from seed in water, there is no gain in ash: the amount found at the end of the plant’s growth is the same as was present in the seed excepting for a relatively small amount falling on the plant as dust. Thus, he disposed finally of the idea that the plant generated potash.

While in retrospect we see the considerable insight into the basis of plant growth and plant nutrition presented by de Saussure, the ideas and approaches did not gain general acceptance. The two great books on agricultural chemistry then current still belonged to the old period. A. von Thaer and Humphry Davy did not realise the fundamental change in perspective introduced by de Saussure. Thaer published his Grundsätze der rationellen Landwirtschaft in 1809–1812, which was translated into English as late as 1844 by Cuthbert Johnson. In it he adopted the prevailing view that plants draw their carbon and other nutrients from the soil humus. Humphry Davy’s book (1813) grew out of the lectures on agricultural chemistry which he gave annually at the Royal Institution between 1802 and 1812; it forms the last textbook of the older period. While no great advance was made by Davy himself, he carefully sifted the facts and hypotheses of previous writers and gives an account, which, although defective in places, represents the best accepted knowledge of the time, set out in the new chemical language. His great name gave the subject an importance it probably would not otherwise have had. He did not accept de Saussure’s conclusion that plants obtain their carbon chiefly from the carbonic acid of the air: some plants, he says,
Soil Conditions and Plant Growth

appear to be supplied with carbon chiefly from this source, but in general he supposes the carbon to be taken in through the roots. Davy presented a list of sources of this carbon, but there was little supporting experimental evidence and subsequently the sources have been shown to be false, although Davy’s reputation meant they persisted. His insistence on the importance of the physical properties of soils – their relationship to heat and to water – was more accurate and marks the beginning of soil physics. In mainland Europe, to an even greater extent than in Britain, it was held that plants drew their carbon and other nutrients from the soil humus.

The foundation of agricultural science

To this point, experiments had been conducted either in the laboratory or in small pots: around 1834, however, J. B. Boussingault, who was already known as an adventurous traveler in South America, began a series of field experiments on his farm at Pechelbronn in Alsace. These were the first of their kind: to Boussingault, therefore, belongs the honour of having introduced the method by which the new agricultural science was to be developed. He reintroduced the quantitative methods of de Saussure, weighed and analysed the manures used and the crop obtained, and at the end of the rotation drew up a balance sheet, showing how far the manures had satisfied the needs of the crop and how far other sources of supply – air, rain and soil – had been drawn upon. The results of one experiment are given in Table 1.2. At the end of the period, the soil had returned to its original state of productiveness, hence the dry matter, carbon, hydrogen and oxygen not accounted for must have been supplied by the air and rain, and not by the soil. On the other hand, the manure afforded more mineral matter than the crop took off, the balance remaining in the soil. Other things being equal, he argued that the best rotation is one which yields the greatest amount of organic matter over and above what is present in the manure.

The rotation had not impoverished the soil so he concluded that nitrogen may be taken directly in to the plant if the green parts are capable of fixing it. Boussingault’s work covers the whole range of agriculture and deals with the composition of crops at different stages of their growth, with soils and with problems in animal nutrition. Unfortunately, the classic

<table>
<thead>
<tr>
<th>Crop</th>
<th>Dry matter</th>
<th>Carbon</th>
<th>Hydrogen</th>
<th>Oxygen</th>
<th>Nitrogen</th>
<th>Mineral matter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beets</td>
<td>3172</td>
<td>1357.7</td>
<td>184</td>
<td>1367.7</td>
<td>53.9</td>
<td>199.8</td>
</tr>
<tr>
<td>Wheat</td>
<td>3006</td>
<td>1431.6</td>
<td>164.4</td>
<td>1214.9</td>
<td>31.3</td>
<td>163.8</td>
</tr>
<tr>
<td>Clover hay</td>
<td>4029</td>
<td>1909.7</td>
<td>201.5</td>
<td>1523.0</td>
<td>84.6</td>
<td>310.2</td>
</tr>
<tr>
<td>Wheat/turnips (catch crop)</td>
<td>4208</td>
<td>2004.2</td>
<td>230</td>
<td>1700.7</td>
<td>43.8</td>
<td>229.3</td>
</tr>
<tr>
<td>Oats</td>
<td>2347</td>
<td>1182.3</td>
<td>137.3</td>
<td>890.9</td>
<td>28.4</td>
<td>108.0</td>
</tr>
<tr>
<td>Total during rotation</td>
<td>17478</td>
<td>8192.7</td>
<td>956.5</td>
<td>7009.0</td>
<td>254.2</td>
<td>1065.5</td>
</tr>
<tr>
<td>Added in manure</td>
<td>10161</td>
<td>3637.6</td>
<td>426.8</td>
<td>2621.5</td>
<td>203.2</td>
<td>3271.9</td>
</tr>
<tr>
<td>Difference not accounted for</td>
<td>+7317</td>
<td>+455.1</td>
<td>+529.7</td>
<td>+4387.5</td>
<td>+51</td>
<td>−2206.4</td>
</tr>
</tbody>
</table>

Source: From Boussingault (1841).
farm of Pechelbronn did not remain a centre for agricultural research, and the experiments came to an end in 1870.

During this period (1830–1840), Carl Sprengel was studying the ash constituents of plants, which he considered were probably essential to nutrition (1832). Schübler was working at soil physics, and a good deal of other work was quietly being done. No particularly important discoveries were being made, no controversies were going on, and no great amount of interest was taken in the subject.

But all this was changed in 1840 when Liebig’s famous report to the British Association upon the state of organic chemistry, published as Chemistry in Its Application to Agriculture and Physiology in 1840 (Liebig, 1840), gave rise to the need to rethink the world of science. Liebig was highly critical of the plant physiologists of his day for their continued adhesion, in spite of accumulated evidence, to the view that plants derive their carbon from the soil and not from the carbonic acid of the air. ‘All explanations of chemists must remain without fruit, and useless, because, even to the great leaders in physiology, carbonic acid, ammonia, acids and bases are sounds without meaning, words without sense, terms of an unknown language, which awake no thoughts and no associations.’ Liebig stated that the experiments quoted by the physiologists in support of their view are all ‘valueless for the decision of any question’. Liebig’s ridicule did what neither de Saussure’s nor Boussingault’s logic had done: it finally killed the humus theory. Only the boldest would have ventured after this to assert that plants derive their carbon from any source other than carbon dioxide, and for a time carbon dioxide was considered to be the sole source of the carbon of plants. Hydrogen and oxygen came from water and nitrogen from ammonia. Certain mineral substances were essential: alkalis were needed for neutralisation of the acids made by plants in the course of their vital processes, phosphates for seed formation and potassium silicates for the development of grasses and cereals. The evidence lay in the composition of the ash: plants might absorb anything soluble from the soil, but they excreted from their roots whatever was non-essential. The fact of a substance being present was therefore sufficient proof of its necessity.

Plants, Liebig argued, have an inexhaustible supply of carbonic acid in the air. But time is saved in the early stages of plant growth if carbonic acid is being generated in the soil, for it enters the plant roots and affords extra nutrient over and above what the small leaves are taking in. Hence a supply of humus, which continuously yields carbonic acid, is advantageous. Further, the carbonic acid attacks and dissolves some of the alkali compounds of the soil and thus increases the mineral food supply. The true function of humus is to evolve carbonic acid.

Liebig further argued that the alkali compounds of the soil are not all equally soluble. A weathering process has to go on, which is facilitated by liming and cultivation, whereby the comparatively insoluble compounds are broken down to a more soluble state. The final solution is effected by acetic acid excreted by the plant roots, and the dissolved material now enters the plant.

Nitrogen is taken up as ammonia, which may come from the soil, from added manure, or from the air. In order that a soil may remain fertile, it is necessary and sufficient to return in the form of manure the mineral constituents and the nitrogen that have been taken away. When sufficient crop analyses have been made, it will be possible to draw up tables showing the farmer precisely what he must add in any particular case.

An artificial manure known as Liebig’s patent manure was made up on these lines and placed on the market.

Liebig’s book (1840) was meant to attract attention to the subject, and it did; it rapidly went through several editions, and as time went on Liebig developed his thesis and gave it
Soil Conditions and Plant Growth

In a quantitative form: ‘The crops on a field diminish or increase in exact proportion to the diminution or increase of the mineral substances conveyed to it in manure.’ He further adds what afterwards became known as the Law of the Minimum, ‘by the deficiency or absence of one necessary constituent, all the others being present, the soil is rendered barren for all those crops to the life of which that one constituent is indispensable’. These and other amplifications in the third edition, 1843, gave rise to much controversy. So much did Liebig insist, and quite rightly, on the necessity for alkalis and phosphates, and so impressed was he by the gain of nitrogen in meadow land supplied with alkalis and phosphates alone, and by the continued fertility of some of the fields of Virginia and Hungary and the meadows of Holland, that he began more and more to regard the atmosphere as the source of nitrogen for plants. Some of the passages of the first and second editions urging the necessity of ammoniacal manures were deleted from the third and later editions. ‘If the soil be suitable, if it contain a sufficient quantity of alkalis, phosphates, and sulphates, nothing will be wanting. The plants will derive their ammonia from the atmosphere as they do carbonic acid’, he writes in the Farmer’s Magazine. Ash analysis led him to consider the turnip as one of the plants ‘which contain the least amount of phosphates and therefore require the smallest quantity for their development’. These and other practical deductions were seized upon and shown to be erroneous by Lawes and Gilbert, who had for some years been conducting vegetation experiments. Lawes does not discuss the theory as such, but tests the deductions Liebig himself draws and finds them wrong. Further trouble was in store for Liebig; his patent manure when tried in practice had failed. This was unfortunate, and the impression in England at any rate was, in Philip Pusey’s words: ‘The mineral theory, too hastily adopted by Liebig, namely, that crops rise and fall in direct proportion to the quantity of mineral substances present in the soil, or to the addition or abstraction of these substances which are added in the manure, has received its death-blow from the experiments of Mr Lawes.’

And yet the failure of the patent manure was not entirely the fault of the theory, but only affords further proof of the numerous pitfalls of the subject. The manure was sound in that it contained potassium compounds and phosphates (it ought, of course, to have contained nitrogen compounds), but the compounds were rendered insoluble by fusion with lime and calcium phosphate so that the manure should not too readily wash out in the drainage water. Not until Way (1850) had shown that soil precipitates soluble salts of ammonium, potassium and phosphates was the futility of the fusion process discovered, did Liebig (1851) recognise the error he had made.

1.3.3 The second half of the nineteenth century

Meanwhile the great field experiments at Rothamsted had been started by Lawes and Gilbert in 1843. These experiments were conducted on the same general lines as those begun earlier by Boussingault, but they have the advantage that they still continue, having been on the same ground without alteration, except in occasional details, since 1852. The mass of information now accumulated is considerable and has become an invaluable source of data as we seek to understand aspects of sustainability and possible responses to environmental change (see, e.g. Leigh and Johnston, 1994). The experiments rapidly provided information and, by as early as 1855, the following points were clear:

1. Crops require phosphates and salts of the alkalis, but the composition of the ash does not afford reliable information as to the amounts of each constituent needed, for
example turnips require large amounts of phosphates, although only little is present in their ash.

2. Non-leguminous crops require a supply of some nitrogenous compounds, nitrates and ammonium salts being almost equally good. Without an adequate supply, no increases of growth are obtained, even when ash constituents are added. The amount of ammonia obtainable from the atmosphere is insufficient for the needs of crops. Leguminous crops behave abnormally.

3. Soil fertility may be maintained for some years at least by means of artificial manures.

4. The beneficial effect of fallowing lies in the increase brought about in the available nitrogen compounds in the soil.

Although many of Liebig’s statements were shown to be wrong, the main outline of his theory as first enunciated stands. It is no detraction that de Saussure had earlier published a somewhat similar but less definite view of nutrition: Liebig had brought matters to a head and made people look at their cherished, but rarely examined, convictions. The effect of the stimulus he gave can hardly be over-estimated, and before he had finished, the essential facts of plant nutrition were settled and the lines were laid down along which scientific manuring was to be developed. The water cultures of Knop and other plant physiologists showed conclusively that potassium, magnesium, calcium, iron, phosphorus, along with sulphur, carbon, nitrogen, hydrogen and oxygen are all necessary for plant life. The list differs from Liebig’s only in the addition of iron and the withdrawal of silica; but even silica, although not strictly essential for all plants, is advantageous for the nutrition of many cereals.

In two respects, however, the controversies continued for many years. Farmers were slow to believe that ‘chemical manures’ could ever do more than stimulate the crop and declared they must ultimately exhaust the ground. The Rothamsted plots falsified this assertion; manured year after year with the same substances and sown always with the same crops, they even now, after more than a 150 years of chemical manuring, continue to produce good crops, although secondary effects have sometimes set in. In France, the great missionary for artificial manures was Georges Ville, whose lectures were given at the experimental farm at Vincennes during 1867 and 1874–1875. He went even further than Lawes and Gilbert and maintained that artificial manures were not only more remunerative than dung but were the only way of keeping up fertility (Ville, 1879). In recommending mixtures of salts for manure, he was not guided by ash analysis but by field trials. For each crop, one of the four constituents, nitrogen compounds, phosphates, lime and potassium compounds (he did not consider it necessary to add any others to his manures), was found by trial to be more required than the others and was therefore called the ‘dominant’ constituent. For wheat he concluded that on his soil it required a good supply of nitrogen, less phosphate and still less potassium (Table 1.3).

Other experiments of the same kind showed that nitrogen was dominant for all cereals and beetroot, potassium for potatoes and vines, and phosphate for turnips and swedes. An excess of the dominant constituent was always added to the crop manure. The composition of the soil had to be taken into account, but soil analysis was at that time not good enough for this purpose. Instead, he drew up a simple scheme of plot trials to enable farmers to determine for themselves just what nutrient was lacking in their soil. His method was thus essentially empirical.

The second controversy dealt with the source of nitrogen in plants. Priestley had stated that a plant of Epilobium hirsutum placed in a small vessel absorbed during the course of the
Soil Conditions and Plant Growth

month seven-eighths of the air present. De Saussure, however, denied that plants assimilated gaseous nitrogen. J. B. Boussingault’s pot experiments showed that peas and clover could get nitrogen from the air while wheat could not, and his rotation experiments emphasised this distinction. While he did not make as much of this discovery as he might have done, he later fully realised its importance.

Liebig, as we have seen, maintained that ammonia, but not gaseous nitrogen, was taken up by plants, a view confirmed by Lawes et al. (1861) in the most rigid demonstration that had yet been attempted. A full summary of this work is provided in Lawes and Gilbert (1889). Plants of several natural orders, including the Leguminosae, were grown in surroundings free from ammonia or any other nitrogen compound. The soil was burnt to remove all traces of nitrogen compounds, while the plants were kept throughout the experiment under glass shades, but supplied with washed and purified air and with pure water. In spite of the ample supply of mineral food, the plants languished and died: the conclusion seemed irresistible that plants could not utilise gaseous nitrogen. For all non-leguminous crops, this conclusion agreed with the results of field trials. But there remained the very troublesome fact that leguminous crops required no nitrogenous manure, and yet they contained large quantities of nitrogen and also enriched the soil considerably in this element. Where then had the nitrogen come from? The amount of combined nitrogen brought down by the rain was found to be far too small to account for the result. For years experiments were carried on, but the problem remained unsolved. Looking back over the papers, one can see how very close some of the older investigators were to resolving the mystery: in particular, Lachmann carefully examined the structure of the nodules, which he associated with the nutrition of the plant, and showed that they contained ‘vibrionenartige’ organisms. His paper, however, was published in an obscure journal and attracted little attention (Lachmann, 1891). Atwater in 1881 and 1882 showed that peas acquired large quantities of nitrogen from the air and later suggested that they might ‘favour the action of nitrogen-fixing organisms’. But he was too busily engaged to follow the matter up, and once again an investigation in agricultural chemistry had been brought to a standstill for want of new methods of attack.

1.4 The beginnings of soil microbiology

It had been a maxim with the older agricultural chemists that ‘corruption is the mother of vegetation’. Animal and vegetable matter had long been known to decompose with the formation of nitrates: indeed nitre beds made up from such decaying matter were the recognised source of nitrates for the manufacture of gunpowder during the European wars of the seventeenth and eighteenth centuries. No satisfactory explanation of the process

<table>
<thead>
<tr>
<th>Constituent added</th>
<th>Yield t ha⁻¹</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal manure</td>
<td>2.98</td>
</tr>
<tr>
<td>Manure without lime</td>
<td>2.84</td>
</tr>
<tr>
<td>Manure without potash</td>
<td>2.14</td>
</tr>
<tr>
<td>Manure without phosphate</td>
<td>1.83</td>
</tr>
<tr>
<td>Manure without nitrogen</td>
<td>0.97</td>
</tr>
<tr>
<td>Soil without manure</td>
<td>0.83</td>
</tr>
</tbody>
</table>

Source: From Ville (1879).
had been offered, although the discussion of rival hypotheses continued until 1860, but the conditions under which it worked were known and on the whole fairly accurately described.

No connection was at first observed between nitrate formation and soil productiveness. Liebig rather diverted attention from the possibility of tracing what now seems an obvious relationship by regarding ammonia as the essential nitrogenous plant nutrient, though he admitted the possible suitability of nitrates. Way (1850, 1852) came much nearer to the truth. He showed that nitrates were formed in soils to which nitrogenous fertilisers were added. Unfortunately he failed to realise the significance of this discovery. He was still obsessed with the idea that ammonia was essential to the plant, and he believed that ammonia, unlike other nitrogen compounds, could not change to nitrate in the soil but was absorbed by the soil. But he only narrowly missed making an important advance in the subject, for after pointing out that nitrates are comparable with ammonium salts as fertilisers he writes:

Indeed the French chemists are going further, several of them now advocating the view that it is in the form of nitric acid that plants make use of compounds of nitrogen. With this view I do not myself at present concur; and it is sufficient here to admit that nitric acid in the form of nitrates has at least a very high value as a manure.

It was not until 10 years later, and as a result of work by plant physiologists, that the French view prevailed over Liebig’s, and agricultural investigators recognised the importance of nitrates to the plant and of nitrification to soil fertility. It then became necessary to discover the cause of nitrification.

During the 1860s and 1870s, great advances were being made in bacteriology, and it was definitely established that bacteria bring about putrefaction, decomposition and other changes; it was therefore conceivable that they were the active agents in the soil and that the process of decomposition was not the purely chemical ‘eremacausis’ Liebig had postulated. Pasteur himself had expressed the opinion that nitrification was a bacterial process. The new knowledge was first brought to bear on agricultural problems by Schloesing and Müntz (1877, 1879, 1882) during a study of the purification of sewage water by land filters. A continuous stream of sewage was allowed to trickle down a column of sand and limestone so slowly that it took 8 days to pass. For the first 20 days, the ammonia in the sewage was not affected, then it began to be converted into nitrate; finally all the ammonia was converted during its passage through the column, and nitrates alone were found in the issuing liquid. Why, asked the authors, was there a delay of 20 days before nitrification began? If the process were simply chemical, oxidation should begin at once. They therefore examined the possibility of bacterial action and found that the process was entirely stopped by a little chloroform vapour, but could be started again after the chloroform was removed by adding a little turbid extract of dry soil. Nitrification was thus shown to be due to micro-organisms – ‘organised ferments’, to use their expression.

Warington (1878, 1879, 1884, 1891) had been investigating the nitrates in the Rothamsted soils, and at once applied the new discovery to soil processes. He showed that nitrification in the soil is stopped by chloroform and carbon disulphide; further, that solutions of ammonium salts could be nitrified by adding a trace of soil. By a careful series of experiments described in his four papers to the Chemical Society, he found that there were two stages in the process and two distinct organisms: the ammonia was first converted into nitrite and then to nitrate. But he failed altogether to obtain the organisms, in spite of some years of study,
by the gelatin methods then in vogue. However, Winogradsky (1890a–c) isolated these two groups of organisms, showing they were bacteria. He succeeded where Warington failed because he realised that carbon dioxide should be a sufficient source of carbon for them, so that they ought to grow on silica gel plates carefully freed from all organic matter; and it was on this medium that he isolated them in 1890.

Warington also established definitely the fact that nitrogen compounds rapidly change to nitrate in the soil, so that whatever compound is supplied as manure, plants get practically nothing but nitrate as food. This closed the long discussion as to the nitrogenous food of non-leguminous plants; in natural conditions, they take up nitrate only (or at any rate chiefly), because the activities of the nitrifying organisms leave them no option. The view that plants assimilate gaseous nitrogen has from time to time been revived, but it is not generally accepted.

The apparently hopeless problem of the nitrogen nutrition of leguminous plants was soon to be solved. In a striking series of experiments in sand cultures, Hellriegel and Wilfarth (1888) showed that the growth of non-leguminous plants, barley, oats, etc., was directly proportional to the amount of nitrate supplied – the duplicate pots agreeing satisfactorily – while in the case of leguminous plants no sort of relationship existed and duplicate pots failed to agree. After the seedling stage was passed, the leguminous plants grown without nitrate made no further progress for a time, then some of them started to grow and did well, while others failed. This stagnant period was not seen where nitrate was supplied. Results from two of their experiments are given in Table 1.4.

Analysis showed that the nitrogen contained in the oat crop and sand at the end of the experiment was always a little less than that originally supplied, but was distinctly greater in the case of peas; the gain in three cases amounted to 0.910, 1.242 and 0.789 g per pot, respectively. They drew two conclusions: (1) the peas took their nitrogen from the air and (2) the process of nitrogen assimilation was conditioned by some factor that did not come into their experiment except by chance. In trying to frame an explanation, they connected two facts that were already known. Berthelot had made experiments to show that certain micro-organisms in the soil can assimilate gaseous nitrogen. It was known to botanists that the nodules on the roots of Leguminosae contained bacteria. Hellriegel and Wilfarth, therefore, supposed that the bacteria in the nodules assimilated gaseous nitrogen, and then handed on some of the resulting nitrogenous compounds to the plant. This hypothesis was shown to be well founded by the following facts:

1. In the absence of nitrate, peas made only small growth and developed no nodules in sterilised sand; when calcium nitrate was added, they behaved like oats and barley, giving regular increases in crop for each increment of nitrate (the discordant results of Table 1.4 were obtained on unsterilised sand).

### Table 1.4 Relationships between nitrogen supply and plant growth.

| Nitrogen in the calcium nitrate supplied per pot, g | None | 0.056 | 0.112 | 0.168 | 0.224 | 0.336 |
| Weight of oats obtained (grain and straw, g) | 0.390 | 5.680 | 10.961 | 15.007 | 21.357 | 30.175 |
| Weight of peas obtained (grain and straw, g) | 3.093 | 2.137 | 7.725 | 5.619 | 8.186 | 11.352 |

Source: From Hellriegel and Wilfarth (1888).
2. The peas grew well and developed nodules in sterilised sand watered with an extract of arable soil.

3. The peas sometimes did well and sometimes failed when grown without soil extract and without nitrate in unsterilised sand, which might or might not contain the necessary organisms. An extract that worked well for peas might be without effect on lupins or serradella. In other words, the organism is specific.

Hellriegel and Wilfarth read their paper and exhibited some of their plants at the Naturforscher-Versammlung at Berlin in 1886. Gilbert was present at the meeting, and on returning to Rothamsted repeated and confirmed the experiments. At a later date, Schloesing and Laurent (1892) showed that the weight of nitrogen absorbed from the air was approximately equal to the gain by the plant and the soil and thus finally clinched the argument. The organism was isolated by Beijerinck (1888a–c, 1989) and called *Bacillus radicicola*, but is now known as *Bradyrhizobium*.

Thus, another great controversy came to an end, and the discrepancy between the field trials and the laboratory experiments of Lawes, Gilbert and Pugh was cleared up. The laboratory experiments gave the result that leguminous plants, like non-leguminous plants, have themselves no power of assimilating gaseous nitrogen; this power belongs to the bacteria associated with them. This result was obtained because by excluding all traces of organic matter, and thereby ammonia, from the soil, the apparatus and the air, there was no chance of infection with the necessary bacteria and no assimilation could occur. In contrast, in the field trials the bacteria were active, and there was a gain of nitrogen.

The general conclusion that bacteria are the real makers of plant food in the soil, and are, therefore, essential to the growth of all plants, was developed by Wollny and Berthelot. It was supposed to be proved by Laurent’s experiments. He grew buckwheat on humus, obtained from well-rotted dung, and found that plants grew well on the untreated humus, but only badly on the humus sterilised by heat. When, however, soil bacteria were added to the sterilised humus (by adding an aqueous extract of unsterilised soil), good growth took place. The experiment looks convincing, but is really unsound. When an organic-rich soil is heated, some substances are formed that are toxic to plants. The failure of the plants on the sterilised humus was, therefore, not due to absence of bacteria, but to the presence of a toxin.

1.5 The development and application of modern knowledge of soils

Our understanding of the physical, chemical and biological factors that control the fertility of soils has advanced greatly since the time of Gilbert and Lawes. The application of this knowledge has resulted in great increases in productivity. The most useful gains have been from new understanding of the storage and movement of water in soils, the value and valuation of reserves of plant nutrients, the physical, chemical and biological conditions in the rhizosphere, and from the role of cultivations in modern production systems. Both research and its application have been greatly aided by developments in the basic sciences, in optical and electronic instruments and in computers. Experimental work in the field with crops provides a practical assessment of advances in exploiting basic soil studies. In this context, the experimental designs initiated by R. A. Fisher proved to be invaluable by providing methods of solving problems in soil management and cropping systems which could not be attempted before. The development of statistical methods of assessing error in biological
experiments was also a major advance. Field experimentation became a major technique in research on soil fertility and, because the precision of the results could be estimated, they were readily accepted by other scientists. Designs in which several factors were tested simultaneously – and their interactions measured – were an important advance. They opened the way to multidisciplinary research; for example, the effects of biological, physical and management factors on crop nutrition were then widely investigated.

During the twentieth century, agricultural research became recognised as important to national interest and attempts were made to bring some organisation and funding to research. For example, in the UK the Agricultural Research Council was established in 1931, and while it has changed its focus and remit, parts of the original aims of the Council are still within the much broader remit of the Biotechnology and Biological Sciences Research Council. The application of the research to develop agriculture was given to the newly established National Agricultural Advisory Service (NAAS) in 1946. This government-funded body later became the Agricultural Development and Advisory Service (ADAS) in 1971 and in 1997 was privatised. It now has a remit well beyond its original focus in agriculture as a provider of environmental solutions, rural development services and policy advice. This pattern of government-funded support for guidance in land management being progressively privatised is not unique to the UK and is now a common pattern in large parts of the western world.

Internationally, agricultural research to aid food production and agricultural development in less developed parts of the world has developed under the Consultative Group for International Agricultural Research (CGIAR). CGIAR has established research centres around the world which have in some cases a specific crop focus (e.g. the International Rice Research Institute, IRRI) and in others a regional agricultural focus (e.g. the International Centre for Research in the Dry Areas, ICARDA). These centres have provided a well-resourced focus for international collaboration on agricultural research issues.

1.5.1 Advances in soil science

Developments in our knowledge of soil science will be made clearer in the succeeding chapters of this book. Important advances were made last century resulting from the application of new methodologies, such as X-ray and spectrographic analyses, which led to our understanding of the crystalline structures of layered aluminosilicate minerals and information on the distribution of micronutrients.

Work on the forms of nutrient reserves in soils, their mobilities and availabilities, and the fate of nutrients applied in fertilisers continued throughout the last century. Work on phosphorus was advanced when the radioisotope $^{32}$P became available. Other advances were made by applying thermodynamic concepts to the solubilities of nutrient ions. Understanding of cation relationships took a notable step forward when Schofield’s Ratio Law was proposed in 1947; later advances came from the use of the quantity ($Q$) and intensity ($I$) factors. Concepts of nutrient ion mobility, first developed in the USA, were used in the mathematical modelling of processes of nutrient uptake by Nye and Tinker (1977) and subsequently by Barber (1995). These models of nutrient processes in soil have become the basis for management of nutrients in farming systems.

Much research has been done on the role of nitrogen in soil/crop systems. This is justifiable as, in many conditions, the supply of nitrogen has a greater effect on crop performance than do the supplies of other nutrients, so that on the world scale this nutrient dominates fertiliser markets. In past years, there was a period in which only about 30% of the nitrogen
applied was, on average, taken up by the crops grown. The 70% that was not recovered represented a serious loss to farmers and was also the cause of environmental pollution – nitrate leached into waters used for public supply and nitrous oxide formed by denitrification contributed to greenhouse gas emissions and damage to the ozone layer of the lower part of the stratosphere. Improved technologies, including the use of the stable isotope $^{15}$N, have made it possible to show in experiments that up to 90% or more of the nitrogen applied as fertiliser can be accounted for in uptake by the crop plus that nitrogen stored in the soil which may benefit future crops. Through improved nitrogen management, we are in sight of securing much higher efficiency of fertiliser nitrogen applied to crops, ranging from wheat in Europe to rice in Southeast Asia. The investigations described previously of the roles of microorganisms in the fixation of nitrogen by leguminous crops have led to methods of preparing cultures of the organisms (Rhizobium spp.) which are specifically associated with particular legumes. These cultures have been made available to farmers for inoculating crops which are to be grown on soils where the appropriate species of Bradyrhizobium is lacking.

The management of other major nutrients, notably phosphorus and potassium but also calcium and magnesium, has been greatly aided by studies of the soils and crops in long-term experiments. The classical experiments at Rothamsted begun by Lawes in the nineteenth century have been invaluable in these studies; they have shown that reserves of phosphate and potassium accumulated in soils from fertiliser additions have considerable value in crop production. Long-term experiments also provide the best basis for relating soluble nutrients in soils to crop performance and to the need for fertilisers; they also lead to calculations of nutrient cycles which are essential for the efficient management of crop nutrition. Increasingly, the records provided by long-term experiments have provided both the underpinning knowledge to build models of how the soil system operates (for instance the Rothamsted Model of Soil Carbon, see, e.g., Jenkinson (1988) and Chapter 4 in this volume) and the data for the testing of model simulations.

The management of soil has been much improved as a result of scientific work. The cultivations necessary for good crop growth have been defined; minimum cultivation systems lessen the energy required and conserve soil structure and soil organic matter. Erosion of the cultivated layer of soil by water and by wind is a serious threat to efficient agriculture in many parts of the world. Studies of the mode and extent of losses of soil under practical conditions have led to recommendations for improved management of cropping systems which avoid these losses. Another serious hazard is the damage done to plant growth, and also to soil structure, by the salts which accumulate in saline soils. Salinity may occur naturally, or it may be the result of irrigating with unsuitable water. Investigations of this problem have led to definitions of the water quality that is required to remove soluble salts in drainage and to the use of gypsum for reclaiming saline soils. With good quality irrigation water and correctly managed drainage, salinity need not now be a problem for efficient crop production.

1.5.2 Soil surveys

The first proposals for the classification and mapping of soils were made by Russian workers in the middle of the nineteenth century. Surveys were made in Southeast England at the end of the nineteenth century, with other local surveys made early in the twentieth century (see, e.g., Kay, 1939). While the practice of soil survey has been widely undertaken in many parts of the world, Great Britain provides an example of many of the trends which
have occurred. Soil mapping was undertaken principally in the period of 1945–1990 when maps on the scale of 1: 250 000 were produced across Great Britain and maps at scales of 1:63 360 and 1:25 000 were also produced for some areas. As part of the link between soil survey and agriculture, land capability maps were also derived. These maps use soil survey data, climate data and landscape data to show the limitations exhibited by areas of land for arable agricultural production. They had their origins in the USA and where an eight-class system was first introduced by USDA in 1961 (Klingebiel and Montgomery, 1961), with subsequent modifications for use in the national contexts; for example, Bibby and Mackney (1969) produced a modification of the USDA scheme for application in Great Britain with seven classes. In all systems, Class One land is land with no limitations and high potential productivity for arable crops. Class Two land has some minor limitations for arable agricultural production which limits the choice of crops and may restrict timing of cultivations (e.g. moderate or imperfect drainage and less than ideal rooting depth and moderate slopes). Class Three land exhibits moderate limitations which restrict the choice of crops but is still considered suitable for arable production with appropriate and good management. Class Four land is considered to have moderately severe limitations (e.g. poor drainage, shallow and/or very stony soils, etc.) which restrict choice of crops and need very careful management. Land in lower classes is not considered suitable for other than low productivity agricultural uses and non-agricultural uses such as pasture, forestry and recreation. Most land capability systems identified subclasses of Classes Two to Four, where the major limitations to agricultural production were identified. These often included limitations due to drainage, erosion, wetness, climate and specific soil limitations. For specific discussion of land capability in relation to soil fertility, see Section 3.2.1.

Today in Great Britain, there is now no systematic mapping programme. The few soil surveys undertaken are for specific purposes, but databases containing soil information derived principally from earlier soil survey work and information gathered for specific tasks are available and can be interrogated.

One of the key contributions of soil surveys and soil classification was in the context of ‘technology transfer’. Technology transfer involves combining information about soil properties gained from soil survey with the results of experimental work at that site to provide guidance on land use such as fertiliser recommendations; this knowledge is transferred to a distant site where the soil classification is the same. In addition to soil maps, advisers require computer-based information services giving the capability for cropping and physical, chemical and biological properties of the soil. Such information is becoming essential to modern management systems which aim to promote the productivity of soil. On a world scale, the most widely used tool for technology transfer is probably the Soil Taxonomy (1999) system developed in the USA. More recently, under the auspices of the International Union of Soil Science, the World Reference Base for Soil Resources (IUSS Working Group WRB, 2006) has been introduced. While soil surveys are still undertaken, there have been rapid changes in the manner in which soil data are recorded and managed. A recent significant shift has been the management and presentation of data using digital soil mapping.

The IUSS Working Group on Digital Soil Mapping (WG-DSM) defines ‘Digital Soil Mapping’ as ‘the creation and the population of a geographically referenced soil databases generated at a given resolution using field and laboratory observation methods coupled with environmental data through quantitative relationships’. In addition, data may be added through inferred spatial and non-spatial relationships between soil and other environmental properties. The development of pedometrics (the application of statistical and mathematical
methods for the study and understanding of soils, their distribution and development) has greatly enhanced the development of digital soil mapping.

DSM can rely upon, but is distinct from, traditional soil mapping or soil survey which involves the manual delineation of soil boundaries by field soil scientists. This digitised and georeferenced soil survey information does not become DSM until the Geographical Information System layer is used to derive other soil-related information within a GIS or similar information software application.

Digital Soil Mapping makes extensive use of previously collected soil survey data and progress has been considerable because of rapid development in computing and the increasing ease with which data are gathered and managed electronically.

1.5.3 Precision Farming

In recent years, as the costs of fertilisers and other agrochemicals have increased and there have been increasing controls on agriculture in terms of leaching of fertilisers, nutrients and pesticides to water courses and groundwater, there has been a shift towards precision agriculture or precision farming. Precision agriculture, or information-based management of agricultural production systems, has developed over the last two decades and is based on the recognition that soils and crops will vary within a field. Initially the focus was to adapt fertiliser distribution to varying conditions across a field, but increasingly with the increased use of new technologies such as global navigation systems and geographic information systems has resulted in a much wider set of applications. Initially these uses were extended to the effective and efficient use of other agrochemicals such as herbicides and pesticides and the correct timing and placement of irrigation water, but with time other practices have evolved such as automatic guidance of agricultural vehicles and implements, autonomous machinery and processes, and product traceability and software for the overall management of agricultural production systems. One consequence is that there is much more efficiency in the use of resources for crop production such as fertiliser, irrigation water, pesticides, etc. This is in sharp contrast to earlier practices where the same fertiliser or pesticide rate was applied to the whole field, often focusing on higher levels of application to address the demands of the poorer soils within that field, with a consequent over application in some parts of the field. This resulted in poor use of valuable resources and often resulted in environmental damage. While early approaches of precision agriculture were based on traditional soil survey maps and the expertise of the farmer, the availability of new technologies – such as global positioning systems, digital soil maps, remote and proximal soil and crop sensors, satellites or other aerial images – and information management tools such as geographical information systems to assess and understand variations has radically changed the practice of precision farming greatly increasing the precision of placement of the agrochemicals and other external resources. The increasing availability of digital soil maps and soil information now provides a wealth of information which is a further contribution to the efficient and effective management of resources. These various sources of information integrated with agricultural machinery, which is able to place agrochemicals and other crop needs to a high degree of resolution, has radically changed the amount of agrochemicals used in a field, introducing economic and environmental benefits as a result of better placement. Precision agriculture enables the efficient use of resources which provides cost savings and assists in the reduction of environmental damage. Gebbers and Adamchuk (2010) have briefly reviewed the considerable impact the use of precision agriculture has had on food production and food security.
Soil Conditions and Plant Growth

1.5.4 Fertilisers

The global use of fertilisers expanded greatly during the twentieth century and the first years of the twenty-first century, resulting in large increases in crop yield. According to Cooke (1982), global use of N, P and K fertilisers by 1913 was 1.4, 0.9 and 0.7 Mt, respectively. By 1998, the consumption had increased to 82.1, 14.2 and 18.0 (FAO, 2010). By 2007, the total global consumption of fertilisers was estimated at 179 Mt (FAO, 2010). These increases have resulted from the application by farmers’ advisers of the research work which has identified the deficiencies of the main nutrients in soils and crops.

Both in the world as a whole, and in the UK, the amounts of nitrogen used dominate compared with applications of other nutrients; this nutrient is responsible for the major part of the cost of fertilisers to the world’s farmers (in the UK about three-quarters of the total spent on fertilisers is for nitrogen). The extent to which farmers change their fertiliser practice on particular crops is shown by the Surveys of Fertiliser Practice which were initiated in the 1940s and still continue. Similar surveys are made in a few other countries. While fertiliser use during the twentieth century showed a steady increase in the amount used per hectare, data on the use of fertilisers in Great Britain between 1998 and 2010 show a pattern of a ‘levelling off in use’ and, in the most recent years, an overall reduction in the amounts of N, P and K applied per hectare of land (Table 1.5). In part this can be ascribed to our increasing knowledge about crop management and increased precision in fertiliser management, but the introduction of regulations enforcing nitrate-vulnerable zones have had a major effect (see DEFRA, 2011).

1.5.5 Ecosystem services

The publication of the Millennium Ecosystem Assessment (2005) brought the concept of ecosystem services to the fore. These services were broadly defined as ‘provisioning services’ such as food and water; ‘regulating services’ such as regulation of floods, droughts, land degradation and disease; ‘supporting services’ such as soil formation and nutrient cycling and ‘cultural services’ such as recreational, spiritual, religious and other non-material benefits.

Soils are widely recognised as an essential part of natural and agricultural ecosystems. Increasingly it is recognised that the role of soils is key to the provision of ecosystem services. These ecosystem services play key roles in the functioning of the ecosystems of which the soil is a part, in both natural and managed production systems. Soils provide beneficial services through the roles played in soil formation, nutrient cycling and primary production. In some respects, the regulatory roles of soils provide key services to other components of

Table 1.5 Fertiliser use in arable and grassland in Great Britain 1998–2010 (kg/ha).

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Total N</td>
<td>126</td>
<td>125</td>
<td>123</td>
<td>116</td>
<td>117</td>
<td>113</td>
<td>110</td>
<td>109</td>
<td>107</td>
<td>105</td>
<td>95</td>
<td>97</td>
<td>102</td>
</tr>
<tr>
<td>Total K₂O</td>
<td>45</td>
<td>42</td>
<td>40</td>
<td>37</td>
<td>40</td>
<td>36</td>
<td>37</td>
<td>35</td>
<td>34</td>
<td>32</td>
<td>27</td>
<td>22</td>
<td>25</td>
</tr>
<tr>
<td>Total P₂O₅</td>
<td>35</td>
<td>32</td>
<td>32</td>
<td>29</td>
<td>31</td>
<td>28</td>
<td>28</td>
<td>27</td>
<td>25</td>
<td>24</td>
<td>20</td>
<td>15</td>
<td>19</td>
</tr>
</tbody>
</table>

the system; yet these are frequently overlooked. These include buffering water fluxes between inputs in precipitation and outputs as run-off, throughflow and groundwater recharge; playing a key role in the carbon cycle through CO₂ emissions from respiration and carbon sequestration and the filtering and detoxification of harmful substances added to the soil. Soils, of course, provide cultural services in terms of supporting recreational space, support for buildings and infrastructure and indeed many human activities.

1.5.6 The needs for future research on soil–plant growth interactions

All agricultural production has its origin in plants which grow in soils. The ultimate potential yield is set by the genetic capacity of the crop which is grown and the solar radiation which it receives at the site. These potentials have been established by models based on the physiology of plants. Average yields throughout the world, as recorded by FAO, are below the recognised potentials because of a variety of constraints. Some constraints result from attacks by pests and diseases but many are associated with soil conditions. Chemical constraints may be due to shortages of nutrients or the presence of toxins; physical constraints arise from shortages or excesses of water or from mechanical impedance to root growth. Biological constraints arise where adverse conditions inhibit the survival and beneficial operation of micro-organisms which are key to the transformation and availability of nutrients. Scientific work to improve further agricultural production must therefore concentrate on the identification of constraints and the further development of inputs and/or management practices which overcome these constraints. The multifactorial field experiments which test these means of overcoming constraints must involve all the disciplines concerned with crop growth, and the objective must be to produce reliably the recognised potential yield at all sites.

In considering the nature of future experimental work, we should note the advice given by Sir Humphrey Davy in the 1821 edition of his book, for this is still relevant. He discussed the need to substitute sound and rational principles for ‘vague, popular prejudices…. Nothing is more wanting in agriculture than experiments in which all the circumstances are minutely and scientifically detailed. The results of truly philosophical experiments in agricultural chemistry would be of more value in enlightening and benefiting the farmer than the greatest possible accumulation of imperfect trials conducted merely in the empirical spirit.’

The multidisciplinary experimentation in which soil science has a fundamental role requires adequate support from those responsible for agricultural improvement. It is therefore appropriate to quote from the writings of another giant of the nineteenth century: John Bennett Lawes concluded his paper published in 1850 to report the first work ever done on the effect of plant nutrients on water use efficiency with these words:

The interest and progress of agriculture would be more surely and permanently served if its great patron Societies were to permit to their scientific officers a wider range of discretion, and more liberal means for the selection and carrying out of definite questions of research. Results of this kind promise, it is true, but little prospect of immediate and practical application, but by their aid the uncertain dictates, whether of common experience, theory, or speculation, may, ere long, be replaced by the unerring guidance of principles; and then alone can it reasonably be anticipated that miscellaneous and departmental analyses may find their true interpretation, and acquire a due and practical value.
References


Kay, F.F. (1939) A Soil Survey of the Strawberry District of South Hampshire. Faculty of Agriculture and Horticulture, University of Reading, Reading, UK.


Van Helmont, J.B. (1648) *Ortus medicinae*, pp.84–90. Complexium atque mistonium elementalium figmentum, Amsterdam, the Netherlands.


2 Plant and crop science

Sayed Azam-Ali
Crops for the Future Research Centre
Selangor Darul Ehsan, Malaysia

2.1 Introductory principles of resource capture by crops

Communities of plants capture and convert environmental resources into carbon-based products. Their absolute capture and conversion of resources will determine their total productivity (biomass) and, where a proportion of this biomass is useful to human beings, their harvestable yield. Their relative capture and conversion of resources will determine the suitability of one crop over others at any location and in any growing season.

The principles of resource capture and conversion (hereafter, simply resource capture) can be considered as a framework for analysing the performance of crops and cropping systems. In this analysis, groups of plants (crops) act as vegetation ‘factories’ in which environmental resources are sequestered (‘captured’) and fixed (‘converted’) into chemical products. By using this physical analogy, we can apply concepts of ‘work’, ‘energy’ and ‘efficiency’ to the whole vegetation system and compare the performance of one vegetation system with another.

The resource capture framework allows quantitative decisions on how best we can manage the choice of crop (species, variety) and cropping system (architectural arrangement of one or more species) and the use of available resources naturally, from the environment, or applied, as irrigation, fertilisers, etc. For the framework to be applicable across environments, species and systems, it must be robust yet simple. The vegetation factory uses fuel (principally water, nutrients, solar radiation and carbon dioxide) from which it makes products (principally carbon-based compounds differentially composed of lipids, proteins and carbohydrates). The efficiency of the factory can be estimated from its productivity relative to the amount of ‘fuel’ it uses over the same period. When considering the efficiency of the crop factory, we should be aware that it must capture resources in two media, soil and air, and two dimensions, space and time. Hence, our resource capture framework requires an understanding of the soil and atmospheric processes that interact with a crop from the time it is sown or planted until such time that it is harvested.
2.1.1 The resource capture framework

Environmental resources required by plants must be sequestered from the atmosphere (principally carbon dioxide (CO$_2$), light and temperature) and the soil (principally water and nutrients from organic or inorganic matter and/or fertilisers). Of these resources, temperature and water (in the atmosphere and the soil) tend to limit the actual productivity of any particular crop below the potential set by its genetic characteristics and the availability of solar radiation and CO$_2$ at any geographical location. Irrespective of their composition, all green plants depend on their capture and use of these environmental resources which are finite in space and time and must be shared between individual plants for a crop community to progress through the processes of growth and development towards maturity and economic yield.

Radiant energy is only available for plant growth when it is absorbed in the leaf canopy in the wavelength range 0.4–0.7 μm (Squire, 1990; Daughtry et al., 1992; Azam-Ali and Squire, 2002). This photosynthetically active radiation (PAR) or ‘light’ provides the high energy required for the process of photosynthesis. The final products of photosynthesis are carbon-based compounds (carbohydrates, lipids and proteins) which have different energy contents.

Longer wavelengths (above 3.0 μm), which are insufficiently energy-intense to drive photosynthesis, provide the thermal energy or ‘heat’ required to drive the process of development. In this way, plants progress through their life cycle, not in response to chronological time, but through the accumulation of thermal time or the amount of heat absorbed above a base temperature at which the development process begins.

In some plants, the process of flowering (a specialised example of development) requires the accumulation of a minimum (‘short-day’) or maximum (‘long-day’) daylength (photoperiod). Such daylength-sensitive plants require the additional photoperiod signal (0.66–0.73 μm) to switch from vegetative to reproductive development.

2.2 Growth and development

Irrespective of their other characteristics, the performance of all plants can be considered in terms of the twin but distinct processes of development and growth. In most arable crops, it is a specific fraction of growth (usually seeds or ‘grain’) that constitutes their economic yield. For example, cereals are harvested when the crop is physiologically mature and the final weight of grains (the reproductive fraction of the crop) is at a maximum. In other crops, it is not the reproductive or seed fraction but the vegetative fraction of growth that determines yield. For example, sugar is extracted from the stems of sugar cane and from the roots of sugar beet but not from the seeds of either species. Thus, except in plant breeding programmes, the production of seeds in sugar crops is an inconvenience because it curtails the process of sucrose accumulation.

2.2.1 Development

Development describes the life cycle of the crop from the start to the end of its duration. We can define development as ‘progression of a plant through a series of discrete changes in the structure or number of individual organs’ (Azam-Ali and Squire, 2002). For example, the germination of seeds, the initiation of leaves, roots and branches, the onset of flowering and the number of seeds produced are all developmental events within the duration of a
crop. The duration of a specific developmental process can be measured by recording the time interval between two discrete events, e.g. the initiation of two successive leaves on a branch. The rate at which that process occurs is simply the reciprocal of its duration and can be expressed as $1/t$ where $t$ is the chronological time for that event to occur. When a crop is adequately supplied with nutrients, the rate of each developmental event is usually determined by the prevailing temperature.

**Thermal time**

Because the progress of any developmental event depends both on temperature and time, the event can be expressed in terms of ‘thermal time’, measured in units called degree days (°C d). The germination of seeds is a good example of how temperature dominates a specific developmental event. For germination to occur, a seed must be viable, non-dormant and adequately supplied with water and oxygen. In addition, seeds of any particular variety require a minimum temperature ($T_b$) below which germination cannot occur. Above $T_b$, the germination rate increases linearly with temperature to an optimum ($T_o$) above which the rate declines linearly to a maximum temperature ($T_m$) at which germination stops (Squire, 1990).

In practice, the value of $T_b$ is usually similar for different developmental processes within a crop variety, and thus the thermal time ($\Theta$; °C d) for any developmental event can be generalised as:

$$\Theta = \sum_{i=1}^{n} (T - T_b)$$

where $n$ is the number of days experienced by the plant at a mean daily temperature, $T$, above $T_b$. So, on any day, accumulated thermal time is simply the difference between the mean daily temperature and the base temperature of the crop, i.e. $T - T_b$. We can consider the values of $T_b$, $T_o$ and $T_m$ as cardinal temperatures for a particular crop variety.

Because crop species and their varieties have different cardinal temperatures, their accumulation of chronological time is a poor indicator of their rate of development. For example, on a day where $T = 20°C$, Crop A with $T_b = 0°C$ will experience a development rate ($T - T_b = 20°C$ d), i.e. twice that of Crop B with $T_b = 10°C$ ($T - T_b = 10°C$ d). These differences can be scaled up to calculate the development rates of both crops over their whole duration (see Azam-Ali and Squire, 2002).

Because the concept of thermal time remains valid over different locations and environments, it can be used to normalise chronological durations of different crop varieties between sites and seasons. The same crop grown at a number of different sites and seasons may require very different amounts of chronological time but the same amount of thermal time to reach maturity.

**Flowering**

Apart from the accumulation of thermal time, in some crop cultivars two other signals, photoperiod and vernalisation, may be required for the switch from vegetative to reproductive development through the process of flowering. In most crops, there is usually an initial vegetative phase followed by a reproductive phase until crop maturity. Flowering marks the switch between these two phases. Different crops can be classified as determinate or indeterminate, depending on their reproductive habit and the point at which they are harvested. In
some determinate crops, such as wheat, there is a sharp transition from vegetative to reproductive growth with the main axis being terminated by an inflorescence (flowering). In such crops, flowering and the subsequent development of seeds occur synchronously. Therefore, the whole crop reaches physiological maturity over a short period of time when all the seeds are ripe. In contrast, in indeterminate crops, such as groundnut, vegetative and reproductive phases occur concurrently because flowers are borne on a series of lateral shoots whilst the main axis continues to develop vegetatively. In such cases, seeds are produced throughout the reproductive phase and may be at different stages of maturity on the same plant.

In some crops, a signal is required for the switch from vegetative to reproductive activity via flowering. When this signal is provided by changes in daylength, the phenomenon is known as photoperiodism. Thus, crop varieties can be defined as photoperiod-sensitive or photoperiod-insensitive. The species that are sensitive to changes in daylength (or strictly photoperiod) can be further divided into those which respond to short days (‘short-day’ plants) and those which respond to long days (‘long-day’ plants). Species not affected by daylength are known as ‘day-neutral’ plants. Most tropical species are short-day plants, i.e. the vegetative period continues until the daylength declines below a critical minimum for a sequence of days before flowering can occur.

Some temperate species, such as winter-sown cereals, require an additional period of low temperature to induce flowering. This phenomenon is known as vernalisation. Table 2.1 summarises the photoperiod and vernalisation requirements of a range of major arable species.

The duration of a crop is the chronological time between when it is planted and harvested. Often, but not always, this will involve the sowing of seeds in soil and the subsequent harvesting of new seeds from the resultant plants. It is important to distinguish duration from the growing season of a crop which is that part of the year when atmospheric and

<table>
<thead>
<tr>
<th></th>
<th>Photoperiod-sensitive</th>
<th>Vernalisation requirement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wheat</td>
<td>Spring</td>
<td>Long</td>
</tr>
<tr>
<td></td>
<td>Winter</td>
<td>✓</td>
</tr>
<tr>
<td>Rice</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Maize</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Barley</td>
<td>Spring</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>Winter</td>
<td>✓</td>
</tr>
<tr>
<td>Sorghum</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Oat</td>
<td>Spring</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>Winter</td>
<td>✓</td>
</tr>
<tr>
<td>Millet</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Rye</td>
<td>Spring</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>Winter</td>
<td>✓</td>
</tr>
<tr>
<td>Soyabean</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Groundnut</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Potato</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Cassava</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Sweet potato</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Sugar cane</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Sugar beet</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

Soil conditions permit the growth of a particular species. In general, the maximum yield of a crop species in a given environment is most likely to be obtained from a cultivar with a duration that is close to the length of the growing season.

In the humid tropics, crops can often be grown continuously because atmospheric and soil conditions rarely limit growth. Where modern short-duration varieties are cultivated, such as rice in South-East Asia, as many as three crops can be harvested annually. In contrast, in dry tropical regions, the length of the growing season is primarily determined by the period when the supply of water from rainfall and/or irrigation exceeds the demand from evaporation at the same location. In humid temperate regions, whilst the supply of water may exceed evaporative demand, the temperature at any location may set the beginning and end of the growing season where this falls below the value of $T_b$ for any particular crop variety.

### 2.2.2 Growth

Within the duration of each developmental event, set by its thermal time (and, where appropriate, daylength and/or vernalisation requirement), the process of growth occurs. Plant growth can be defined as ‘quantitative and irreversible changes in the length, area or weight of individual organs’ (Azam-Ali and Squire, 2002). The area of individual leaves and the dry weights of leaves, roots, stems and grains are all examples of growth, whilst the number of leaves, roots, stems and grains are all examples of development.

Together, the duration and rates of different developmental and growth processes govern both the total productivity of a crop and the balance between the vegetative and reproductive fractions. Ultimately, growth depends on the division and enlargement of individual cells. However, at any particular time, active growth is restricted to certain embryonic regions of the plant called meristems. There are several different types of meristems. For example, axial organs, such as stems and roots, have apical meristems in which growth is restricted to the tip region. Apical meristems remain capable of growth over long periods (in perennial crops for many years) and are thus referred to as indeterminate meristems. In contrast, organs such as leaves, flowers and fruits grow only for a limited period before the whole organ reaches maturity. The growing regions of these organs are therefore termed determinate meristems.

**Solar radiation and water: engines for crop growth**

We have seen that between the developmental windows set by sowing and harvest, plants must accumulate biomass and partition this into different organs (roots, stems, leaves, flowers, seeds).

Development is controlled principally by the accumulation of temperature (thermal time) by plant organs. In some cases, daylength (photoperiod) and/or vernalisation may set the switch from vegetative (roots, shoots and leaves) to reproductive (seeds or grain) development through flowering.

Growth, on the other hand, is controlled principally by the accumulation of solar radiation, mainly by plant leaves, in which $\text{CO}_2$ is fixed through the process of photosynthesis into carbon-based compounds (carbohydrates, proteins and lipids).

Whilst the capture and conversion of solar radiation sets the potential productivity within the genetic limits of a species, the supply and availability of water to plant roots often determines the actual productivity of vegetation at any location. In such circumstances, we need to understand processes that govern both the above-ground capture and conversion of solar radiation and $\text{CO}_2$ by plant leaves and the below-ground capture of water and nutrients by plant roots. Together, these linked processes determine the potential and actual productivity
of a crop system for any species and location. Later, we consider how a crop can be described as part of the ‘soil-plant-atmosphere-continuum’ (SPAC) that links the below-ground and the above-ground capture and conversion of resources by plant roots and shoots.

2.3 Solar radiation

The production of dry matter by crops is the consequence of leaves chemically fixing CO$_2$ by intercepting PAR, and the efficiency with which that intercepted PAR and CO$_2$ is used for dry matter production (Black and Ong, 2000). Crop ‘canopies’ represent an investment in dry matter, principally in photosynthetically active structures (e.g. leaves), that can intercept solar energy and CO$_2$ to produce new dry matter, some of which can be allocated to non-photosynthetically active structures (e.g. roots, stems and seeds) (Campbell and van Ever, 1994).

2.3.1 Radiation capture by crops

Because it is the surface that must intercept radiation, the size and longevity of a crop leaf canopy determine its rate and duration of dry matter accumulation. The size of the intercepting surface, or leaf area index ($L$), depends on the number of plants per unit area of ground ($N_p$), the number of leaves per plant ($N_l$) and the mean area of leaves per plant ($A_s$) (Azam-Ali and Squire, 2002):

$$L = N_p \times N_l \times A_s.$$  

Irrespective of its total area, light must penetrate through the layers of a canopy for it to be effectively captured. In terms of light transmission, a crop can be considered as a number of horizontal layers (Azam-Ali and Squire, 2002). Thus, if radiation is measured at different levels in the canopy profile, each of $L=1$, then the amount of light transmitted to any level is a function of the distribution of leaves above that level.

To drive photosynthesis, individual leaves within a crop stand must absorb or ‘capture’ the PAR component of radiation as it passes down through the canopy. Whilst the individual leaves of different species have different abilities to absorb, reflect and transmit radiation, at the crop scale these differences are less significant. For practical purposes, the most important factor is the fraction ($f$) of total radiation received by a crop that is intercepted before it hits the ground.

If we assume that a crop canopy is a homogenous block of vegetation in which leaves are randomly distributed without any row structure or clumping, the pattern of light transmission through the canopy obeys Beer’s Law of exponential decay (Azam-Ali and Squire, 2002). In such circumstances, light transmission can be calculated using the Monsi–Saeki (1953) equation:

$$\frac{I}{I_o} = e^{-kl}.$$  

where

$I_o$ is the irradiance above the crop canopy
$I$ is the irradiance at a level within the canopy below a leaf area index $L$
$k$ is the typical extinction coefficient of radiation for any particular species
Strictly, for light transmission to obey Beer’s law, the leaves should be black, i.e. opaque to radiation. However, Goudriaan (1977) demonstrated Beer’s Law to be a good approximation in many real crop canopies by plotting \( \ln(II/I_o) \) against \( L \), giving a straight line with a gradient, \( k \) which depends on the architecture of the crop. The estimated \( k \) relates to the fraction of radiation intercepted per unit leaf area index and can be used to compare the pattern of light interception between cultivars, species and environments (Hay and Walker, 1989).

Essentially, \( k \) is an estimate of the erectness or flatness of leaves in a crop canopy. Monteith (1981) reported \( k \) values of about 0.3 for crops with vertical leaves, about 0.6 for an intermediate leaf arrangement and 0.9 for crops with very horizontal leaves. For example, groundnut with intermediate leaf angle has \( k \) values in the range of 0.40–0.66 (Azam-Ali and Squire, 2002), whilst cereals, such as pearl millet, with erect leaves have \( k \) values ranging between 0.3 and 0.45 (Squire, 1990).

In practice, canopies are not always made up of uniformly distributed elements. For example, in row crops leaves are clearly clumped into lines of vegetation before complete canopy closure. At any time, the value of \( k \) depends on the size of the canopy and may increase as canopies expand and become more homogeneous and their leaves are more randomly oriented. It may also decrease as leaves become more vertically oriented as a canopy becomes denser with age, an effect enhanced at higher population density (Squire, 1990). However, \( k \) remains reasonably stable for a genotype over most of its life and across a wide range of conditions and can therefore be used to characterise differences between canopies with different architecture.

There is a functional link between \( f \) and \( L \) for crops with different canopy architectures. At any time, the value of \( f \) for a given crop canopy is determined by the actual value of \( L \) and the typical value of \( k \). The relation between \( f \) and \( L \) can be expressed as:

\[
f = 1 - \exp(-kL).
\]

On any day, the radiation intercepted by a crop is the product of \( f \) and the irradiance, \( S \) (MJ m\(^{-2}\)), received that day. Over a growing season, seasonal intercepted radiation is therefore the sum of daily values of \( fS \), i.e. \( \Sigma fS \) (MJ m\(^{-2}\)). Manipulating seasonal intercepted radiation is a critical feature of agronomic practices irrespective of species, location or environment. In practice (since we cannot manipulate \( S \)), successful crop management involves agronomic options that influence \( f \) and/or the crop duration. For example, directly or indirectly, sowing date, sowing rate, planting arrangement, irrigation and fertiliser applications all influence the rate of expansion, maximum value and seasonal total of \( f \) for any crop canopy. Similarly, agronomic options that influence crop duration, by selecting appropriate crop cultivars and sowing/harvest dates, can ensure that fractional interception best matches the seasonal pattern of irradiance at any site. In other words, it makes sense for growers to produce crops that achieve their maximum value of \( f \) in the brightest days of the growing season when most total light is available.

### 2.3.2 Radiation conversion by crops

**Photosynthesis**

Photosynthesis is a photochemical process that depends on the availability and interception of PAR by green leaves to drive biochemical pathways within them. The regulatory biochemistry of photosynthesis requires \( \text{CO}_2 \), temperature and the supply of water and mineral nutrients. Photosynthesis involves light-driven photochemical processes, enzymatic
processes that need no light (so-called ‘dark’ respiration) and processes of gaseous diffusion into and out of leaves via apertures called stomata. The stomatal entry of CO\(_2\) into a leaf depends on the concentration gradient between that in the ambient atmosphere and at the internal site of CO\(_2\) fixation in the leaves. Through this common pathway, CO\(_2\) and oxygen are exchanged between chloroplasts and the atmosphere (Larcher, 2003). Plants exert a physiological influence on this concentration gradient mainly through the opening and closing of stomata which regulate both the uptake of CO\(_2\) and the loss of water from the plant (Azam-Ali and Squire, 2002; Larcher, 2003). Since leaves are the primary agents of crop photosynthesis, their efficiency of capture and use of PAR determine the overall productivity of a crop (Loomis and Connor, 1992).

The major pathway of photosynthesis is the conversion of atmospheric CO\(_2\) and water into carbohydrates and oxygen, viz.:

\[
6\text{CO}_2 + 6\text{H}_2\text{O} \xrightarrow{\text{Light}} \text{C}_6\text{H}_{12}\text{O}_6 + 6\text{O}_2.
\]

Through the energy supplied by solar radiation, two energy-poor compounds are converted into two energy-rich compounds and, by this process, photosynthesis converts light energy into chemical energy (Azam-Ali and Squire, 2002).

**Mechanisms of photosynthesis**

All plant species exhibit one of three types of photosynthesis, viz. \(C_3\), \(C_4\) and Crassulacean Acid Metabolism (CAM) (Loomis and Connor, 1992; Larcher, 2003). The only significant crop species identified as using the CAM mechanism is pineapple; other CAM species include succulents such as cacti.

Most agricultural species employ the \(C_3\) photosynthetic pathway in which CO\(_2\) is fixed and the resulting molecules are processed in a series of enzyme-driven reactions known as the ‘Calvin–Benson Cycle’ (see Raghavendra, 2000). The cycle generates additional molecules to react with CO\(_2\), the first products of which are molecules with three carbon atoms (hence ‘\(C_3\)’ photosynthesis), some of which are used as the first step in making sugars and starch.

\(C_4\) photosynthesis is an adaptation of the Calvin–Benson cycle which uses a CO\(_2\) concentrating system that, in effect, increases the availability of CO\(_2\) at the active sites of photosynthesis and results in molecules with four carbon atoms. As a result, \(C_4\) plants have faster rates of photosynthesis than \(C_3\) plants and include crops such as maize, sorghum, millet and sugar cane which are amongst the most photosynthetically efficient species in the world.

**Respiration**

Respiration is important in the carbon budget of crops because it results in the loss of already fixed CO\(_2\) from the plants through the process of growth. Respiration can be separated into that which results from growth of new tissues (photorespiration) and that which results from maintenance of existing tissues (‘dark’ respiration).

Since photorespiration acts on the CO\(_2\) initially fixed by photosynthesis, its rate is closely linked to the CO\(_2\) fixation rate. In \(C_3\) crops, photorespiration increases with temperature, resulting in a reduction in the initial efficiency of light use by individual leaves at high temperatures. Because of their ability to concentrate CO\(_2\) at the sites of photosynthesis, \(C_4\) species do not undergo photorespiration – an advantage over \(C_3\) species that is further enhanced at higher temperatures.
Irrespective of their photosynthesis system, all green plants undergo ‘dark’ respiration. Strictly, the term ‘dark respiration’ is a misnomer because the process continues in daylight at a comparable rate to that in the dark. However, because it must continue in daylight alongside photorespiration, the two components are difficult to distinguish except at night. During dark respiration, atmospheric O₂ is used by plants to convert carbohydrates into CO₂ and H₂O, with simultaneous liberation of energy (Azam-Ali and Squire, 2002). Plants use this energy to build more complex molecules from the initial products of photosynthesis which are used for the growth and maintenance of tissues. This explains why there is a higher energy cost for plants to produce unit mass of more complex lipids and proteins than simpler carbohydrates.

During the life of a crop, the relative contributions of the growth and maintenance components of respiration change with the age and weight of the crop. Maintenance respiration becomes more important later in the life of a crop because a larger amount of biomass requiring maintenance exists. Amthor (1989) reports that respiration can account for between 25% and 66% of the total assimilation of a crop.

**Photosynthesis and light interception at leaf level**

The photosynthetic response of individual leaves to their receipt of solar radiation and CO₂ is often described in terms of a ‘light response curve’ (Figure 2.1). In the light response curve, the initial photosynthesis (\(P_n\)) of an individual leaf responds to increasing irradiance (\(I\)). In very weak light, the relation for both C₃ and C₄ plant systems is linear because photosynthesis is limited almost exclusively by the absorption of light quanta (Loomis and Connor, 1992; Azam-Ali and Squire, 2002). This quantum efficiency is a measure of the amount of CO₂ absorbed per unit increase in irradiance. At \(I=0\), the leaf evolves CO₂ by dark respiration at the rate \(R_d\). At \(I=I_c\), the light compensation point, \(P_n=0\) because photosynthesis exactly balances any losses through dark respiration. Below the light compensation point, the leaf is actually losing more CO₂ through respiration than it is gaining through photosynthesis.

**Figure 2.1** The idealised light response curve for a C₃ and C₄ leaf. Photosynthetic rate \(P_n\) is plotted against irradiance \(I\), and the response curve shows the light compensation point \(I_c\), saturating irradiance \(I_s\) and dark respiration \(R_d\) for a C₃ leaf. The initial, linear ‘quantum efficiency’ is shown for the C₄ (Q₁) and C₃ (Q₂) leaf types.
As $C_4$ plants do not undergo photorespiration, they have a lower light compensation point than $C_3$ plants. In $C_3$ species, as $I$ increases above $I_c$, $P_n$ increases in a curvilinear manner until it becomes light-saturated at a maximum net photosynthesis ($P_{\text{max}}$). In most $C_3$ species, $P_n$ becomes light-saturated well below full sunlight and the maximum value of this saturating irradiance declines with leaf age and stresses, such as leaf water deficits (Azam-Ali and Squire, 2002). At saturating irradiance, photosynthesis is controlled by the rate at which CO$_2$ from the atmosphere is reduced to carbohydrates. In $C_4$ species, leaves rarely if ever reach $P_{\text{max}}$ and, as a consequence, individual leaves of $C_4$ species continue to photosynthesize in strong light at which $C_3$ leaves are already photosaturated.

**Photosynthesis and light interception at crop level**

We have seen that the photosynthetic efficiency of single plant leaves can be measured by calculating their individual light response curves (Figure 2.1). However, light response curves for individual leaves are of limited value in crop situations where we need to scale up to the photosynthetic efficiency of the whole canopy. Monteith (1994) observed that the rate at which field crops accumulate dry matter is proportional to the rate at which their canopies absorb radiant energy. Figure 2.2 shows that, for any crop, the relation between total dry weight and the seasonal interception of solar radiation is often linear with a slope, $\varepsilon_s$ (g MJ$^{-1}$), known as the dry matter/radiation ratio or often the ‘radiation use efficiency’ (RUE) (Azam-Ali and Squire, 2002).

The adjustment from the curvilinear responses at leaf level to the linear response to radiation at the canopy level can again be explained in terms of the Beer’s Law analogy previously described for radiation capture. Effectively, the ‘stacking’ of leaves within a crop stand means that most operate at the linear part of the light response curve and therefore any additional radiation intercepted by each new leaf is proportionally converted into a photosynthetic product (Biscoe et al., 1975; Asseng and Hsiao, 2000).

Typically, $C_3$ species (most agricultural crops) have values of RUE ($\varepsilon_s$) of about 1.5 g MJ$^{-1}$, whilst $C_4$ crops have values closer to about 2.4 g MJ$^{-1}$. Whilst the value of $\varepsilon_s$ for any cultivar remains fairly constant during much of the cropping season, it declines and eventu-

![Figure 2.2](image-url) The idealised relation between total dry matter (including roots, leaves, stems and seeds) of a crop and accumulated intercepted radiation. The slope of the relation $\varepsilon_s$ (g MJ$^{-1}$), known as the dry matter/radiation ratio, is shown for $C_3$ and $C_4$ crop species.
ally levels off during reproductive growth, in response to drought and/or pests and diseases. Also, as it is a crude measure of biomass produced per unit of radiation intercepted, $\varepsilon_s$ should be ‘energy adjusted’ to account for the differential energy contents of the carbohydrate, lipid and protein components of the crop (Penning de Vries et al., 1989). Biomass composition also influences both the growth and maintenance respiration requirements of crops (Penning de Vries et al., 1974; Amthor, 1989).

It is also important to note that $\varepsilon_s$ represents the efficiency with which total biological productivity (including roots) is produced in relation to intercepted radiation. In most circumstances, only a fraction of total biomass is represented as yield. Therefore, to estimate the efficiency with which crop yields are produced, we need to know the ‘harvest index’ (grain yield/total biomass) of any particular crop.

From these considerations, we can now see that the biomass of a crop ($W$) can be related to its seasonal capture of radiation ($\sum fS$) and the efficiency with which it converts that captured radiation into biomass ($\varepsilon_s$), i.e.:

$$w = \varepsilon_s \sum f S.$$  

Where crop roots are well supplied with water, the total radiation intercepted during the growing season and a knowledge of $\varepsilon_s$ for any cultivar can often be used to estimate its productivity and yield. However, where drought and/or nutrient stress restrict the uptake of CO$_2$ through stomatal regulation, seasonal radiation interception is a poor indicator of productivity because $\varepsilon_s$ does not remain constant (Figure 2.3).

### 2.4 Water

#### 2.4.1 The soil-plant-atmosphere-continuum (SPAC)

Terrestrial plants must take up water from the soil through their roots, transport it upwards through roots, stems and leaves as a liquid from where it must be transferred into the atmosphere as vapour. The movement of water from the soil to the atmosphere can be described
in terms of a SPAC. Water lost through transpiration drives the SPAC and must be replenished by further extraction from the soil by roots for the system to remain in equilibrium during the life cycle of a crop. Short-term water deficits result in symptoms of plant water stress which reduces plant growth rates by restricting CO₂ uptake with consequences for final crop yield. Long-term deficits between supply and demand may result in crop failure and plant death when water deficits prevent plants from maintaining metabolic functions and structural integrity.

To minimise water stress and avoid crop failure, growers must manage the water requirements of their crops to best match the availability of soil water to atmospheric demand and minimise water losses from the system that do not result in transpiration.

An understanding of the SPAC can help provide both a theoretical and practical basis for crop water management and a predictive basis to calculate the productivity of crops. In this section, we consider the individual components of the SPAC and how they contribute to the water use and productivity of crop plants.

### 2.4.2 Soil water

More comprehensive analyses of soil water characteristics appear in Chapter 9 and in Brady and Weil (1999). For our purposes, water in soils can be considered both in terms of its content, i.e. how much is there, and its energy status, i.e. how easy it is for plants to access.

The relation between soil water content and soil water energy

The amount of water available to plant roots is always less than either the actual soil water content (SWC) or its free energy status. Three important forces reduce the energy status of water in soil below that in its pure state. First, the particles that make up the soil exert an adhesive attraction for water within the soil matrix, i.e. a matric force that acts against its removal by plant roots. Second, the cohesive attraction of ions and other solutes in the soil water medium results in osmotic forces which reduce the water energy status below that of pure water. Third, gravitational forces impose a downward force on soil water because the energy status of water always declines with soil depth. Plant roots must remove water against the gravitational gradient which tends to move it away from the root zone. The difference between the free energy of soil water and that of pure water is known as soil water potential, \( \Psi_s \), and can be expressed as:

\[
\Psi_s = \Psi_m + \Psi_o + \Psi_g
\]

where

- \( \Psi_g \) is gravitational potential
- \( \Psi_m \) is matric potential
- \( \Psi_o \) is osmotic potential in a particular soil volume

Because actual water potential is always less than that of pure water, it is expressed as a negative value relative to zero (the energy status of pure water).
We now consider various factors that influence the volumetric SWC of soils (i.e. what percentage of the soil volume is made up of water) and its energy status, i.e. $\psi_s$. For different soil types (e.g. sandy, loamy and clay soils), there is a characteristic relationship between SWC and $\psi_s$, known as the soil moisture release curve (SMRC).

Figure 2.4 shows the key components of the SMRC for a clay soil. Also marked on the diagram are the ranges over which instruments commonly used to measure SWC or soil water potential operate.

**Saturation ratio ($\phi_s$)**

This represents the maximum amount of water that a soil can hold when water replaces all the available space in the soil volume. A soil that is saturated is at a water content equal to its saturation ratio ($\phi_s$). Because of the various forces acting on it, especially gravity, $\phi_s$ is an unstable condition as water drains through the soil (deep drainage), runs across the soil surface (run-off) or evaporates from it (surface evaporation, $E_s$). In each case, water moves from a point of high energy status to a point of lower energy status.

When soils are at or close to saturation, they are considered to be waterlogged. This condition often occurs first at low spots in the field or when the drainage of the whole field is impaired by poor soil structure. Even well-drained soils can become waterlogged during a heavy rainstorm where infiltration rate is slower than rainfall (or irrigation) intensity. Soils compacted by ploughing or machinery can also encourage waterlogging by reducing soil porosity, i.e. the space available in the soil which can absorb water before saturation.
It is important to recognise that the infiltration of water into soil results in the simultaneous removal of air from the same soil volume. Soil aeration controls the level of O₂ and CO₂ in the soil volume. In turn, both gases influence the respiration of roots and soil microorganisms. For soil respiration, O₂ must be supplied and CO₂ removed. For most crop plants, O₂ in the soil air must exceed 10% (Brady and Weil, 1999). Poor soil aeration occurs when water displaces gases and when the exchange of gases with the atmosphere is too slow, even if sufficient air space is available in the soil.

**Field capacity**

In a saturated soil, rapid drainage ceases as macroflow (saturated flow) slows because water in the macropores is increasingly replaced by air (unsaturated flow). At this stage, the soil is said to be at field capacity (FC). The matric potential of a soil at FC typically ranges between −0.01 and −0.03 MPa and is not a steady state because unsaturated flow continues, albeit more slowly than saturated flow.

**Permanent wilting point**

As its content continues to decline without being replenished by rainfall or irrigation, the water remaining in the soil volume becomes progressively more difficult for roots to extract. Plants eventually experience symptoms of water stress, most often exhibited by visible wilting of leaves, because roots are unable to replace water lost by evaporation. Plants which remain in this condition are said to be at permanent wilting point (PWP) which typically corresponds to a soil water potential of between about −1.0 and −2.0 MPa, with a mean of 0.15 MPa (Azam-Ali and Squire, 2002). The SWC at this stage within the range of plant roots is called the permanent wilting percentage. Water remaining in the soil is now too tightly held to be available to most plants as a liquid but is retained within the soil particles (matric potential), or bound to salts (osmotic potential).

**Hygroscopic coefficient**

Soil water below the PWP is so tightly held by soil colloids that it can only move as vapour. The moisture content in the soil at this stage is called the hygroscopic coefficient and is greater in colloid-rich clay soils than colloid-poor sandy soils.

**Available water content**

In terms of volumetric water content, the available water content (AWC) of a soil is often described as the difference between the FC of a soil and the PWP of the plants in it (Azam-Ali and Squire, 2002). However, we have seen that both FC and PWP are intrinsically arbitrary measures and depend on an unstable condition for an upper value (FC) and a plant-dependent lower value (PWP). In addition, AWC relates to differences in water content whilst plant roots act in response to water potential.

In terms of soil water potential, AWC can similarly be calculated as the difference between the value of \( \psi_s \) at FC and that at PWP. The presence of high levels of salts in the soil solution influences the availability of water to plants and hence the value of \( \psi_s \). In effect, these osmotic forces mean that less water can be extracted than that in the soil at PWP point because of their effect on \( \psi_s \) rather than water content *per se*. In most humid soils, osmotic
potential effects are negligible but they can be significant in soils of arid and semi-arid regions, especially those with high salt contents.

Knowledge of the SMRC for any soil provides a valuable basis to calculate how much water crops can extract at any location at any particular time. However, it ignores two important factors that influence the seasonal availability of soil water to a crop growing at that location. For this we require an understanding of the behaviour of the crop root system and the total available depth to which roots can descend. We consider the behaviour of root systems in more detail in Section 2.4.4. However, it is not simply the speed at which plant roots can descend down the soil profile but the depth of the soil profile itself. Soil depth is an important factor in many tropical soils, particularly shallow soils or those which have impenetrable layers of material within the soil profile. To compare different soil types, it is usual to present AWC in terms of millimetres of water per metre depth of soil; where the soil depth is known to be less than 1 m, the AWC should be adjusted accordingly.

2.4.3 Plant water

The relation between plant water content and plant water energy

As with soil, plant water can be considered in terms of energy and content. In metabolically active plants, cells contain 80–90% water. This water forms a liquid continuum from the surface of roots through the plant’s tissues to the leaf surfaces from where it evaporates. Between its uptake from the soil and its loss to the atmosphere, plant water must also transport nutrients and metabolites to their sites of activity as well as contribute to plant structure by maintaining the turgidity of various organs.

Unlike soil water, plant water content must remain high for metabolic activities to continue. A decline of 20–25% below that at maximum hydration arrests most growth processes, and any further depletion in tissue water content may result in their death. New plant growth must depend either on tissues that have survived dehydration or on the production of new plant material. The ability of a plant to buffer its water status between atmospheric demand and soil supply determines its metabolic success and ultimately its survival.

Plant water potential

Water movement within plants occurs through gradients in plant water potential, \( \psi_p \), from regions of high energy (abundant water) to those where the free energy is lower (less water). As with soils, the water potential within plant tissues is always less than that of pure water because of the influence of solutes on the osmotic potential (\( \psi_{o}^{'} \)) and matric forces (\( \psi_{m}^{'} \)) caused by the bonding of hydrogen ions to cell walls and proteins. Except in very tall trees, the effect of gravitational gradients is negligible. However, unlike soil water potential, plant cells must exert an additional outward hydrostatic or turgor pressure (\( \psi_{h} \)) to keep them turgid and functional both metabolically and structurally.

Plant water potential can be expressed as:

\[
\psi_p = \psi_h + \psi_o^{'} + \psi_m^{'}.
\]

For both soils and plants, water potentials are always less than those of pure water. Further, for water to move along the energy gradient between soil and atmosphere, \( \psi_p \) must always be more negative than \( \psi_s \).
In practice, crops can only extract water from the soil in relation to the distribution of their active root systems and the water available locally to metabolically active roots. This means that root systems must often operate in soils where some is close to FC (between about 0.01 and 0.03 MPa) and some may be at or below PWP (between about −1.0 and 2.0 MPa). During daylight, water movement within the plant depends on the relative values of $\psi_p$ across different parts of the plant, which in turn depends on water lost through stomata. At night, when stomata are closed, $\psi_p$ returns to equilibrium with the soil (i.e. $\psi_p = \psi_s$).

### Plant water content

Whilst water movement depends on energy (i.e. potential gradients), actual water content of tissues has a greater effect on plant metabolism and growth than any changes in $\psi_p$ or $\psi_h$. Tissue water content can be expressed in various ways as water content per unit dry weight, per unit fresh weight and per unit weight of water at full hydration (i.e. the leaf relative water content; RWC). For our purposes, RWC is the most useful means of quantifying plant water content, since it is not influenced by changes in tissue dry weight.

The value of RWC can be estimated by measuring the fresh weight ($W_2$) of leaf samples and rehydrating these on water to obtain their saturated weight ($W_3$) (Azam-Ali and Squire, 2002). The discs are then oven dried to obtain their dry weights ($W_1$). RWC can then be calculated from:

$$ \text{RWC} = \frac{W_2 - W_1}{W_3 - W_1} \times 100 $$

(2.1)

### 2.4.4 Water capture systems

#### Root system size

Plants acquire water through their roots. Although there are numerous physical and physiological factors which control and regulate this process, ultimately, the total root length density ($l_r; \text{cm cm}^{-3}$) and final depth (m) of the root system, together, play a critical role in its effectiveness in capturing water and nutrients. The final depth of the root system also allows us to calculate the potential volume of water that is available to the crop throughout its growing season. Because water moves down the soil profile, a deep rooting system is invaluable in meeting high transpiration demand when water availability is reduced. For example, Zhao et al. (2004) collated evidence that the mean rooting depth for desert vegetation may reach up to 13.4 m, with 31% of the total root biomass found below 0.3 m. They also reported that temperate forests have a maximum rooting depth of 3.7 m with 35% of the total biomass below 0.3 m and temperate grasslands have a maximum depth of 2.4 m with 17% of the total root biomass found below 0.3 m. In contrast, the effective rooting depth of annual crops rarely reaches 2 m and, in temperate crops well supplied with water, roots are largely restricted to the uppermost 30 cm or so of the soil.

The ability of a root system to take up water varies with factors such as root morphology and age as well as its total length. However, inflow rates are generally assumed to be a uniform function of $l_r$. Similarly, the distribution of a root system depends on many factors including soil type and depth as well as plant factors such as the rate of root growth and root system morphology. Gerwitz and Page (1974) concluded that a negative exponential function can usefully describe root system distribution.
So, if $W_r$ (kg dry matter of roots per m$^2$ of ground) is the root mass and $z$ (m) is the distance below the soil surface at $z=0$, then:

$$\beta W_r e^{-\beta z}$$

gives the root density per unit volume of soil at depth $z$, where $\beta$ (m$^{-1}$) is a constant.

Integrating with respect to $z$ between $z=0$ and $z$ gives:

$$W_r (1 - e^{-\beta z}).$$

Thus, $1 - e^{-\beta z}$ gives the fraction of root mass between the soil surface and depth $z$.

### Root system speed

In addition to the final size of a root system, an important consideration in its success as a seasonal capturing system is the speed at which the root system extends into soil. This can be considered in terms of a root front velocity (mm d$^{-1}$) which depends both on the species in question and the physical characteristics of the soil.

#### 2.4.5 Water movement systems

The movement of water through the plant components of the SPAC is often described as a transpirational stream. However, even when soil water supply to roots is unrestricted, diurnal changes also occur in the water potential of plant shoots because leaf surfaces must transpire at or close to the potential set by the atmosphere. The decrease in $\psi_p$ of the shoot is caused by the hydraulic resistance, $Z_p$, to water movement in the transpirational stream. We can estimate the flux, $F_p$, of water through the plant system as:

$$F = \frac{\Delta \psi_p}{Z_p}$$

where $\Delta \psi_p$ is the gradient of water potential across the plant. Often, a linear relationship (equivalent to $Z_p$) exists between $F_p$ and $\Delta \psi_p$, because osmotic forces usually play only a small role in water movement and there is a steady state whereby water absorbed by the root system is equal to that subsequently transpired by the canopy. However, such conditions may not prevail where plant material includes young, actively growing tissues that require water for cell expansion. Also, many plant organs, such as leaves, stems and fruits, may act as ‘capacitors’ which change in size as a consequence of absorbing or losing water. Soil water deficits impose a restriction on supply of water that results in dehydration and loss of cellular water and an imbalance between the loss and gain of water by the whole plant. Ultimately, water must move through a system that connects roots, stems and leaves, that themselves contain water, as well as act as a connecting pathway between the sites of water uptake and water loss.

### Root water uptake

Individual roots are made up of a cortex surrounding a central vascular cylinder (Zhao et al. 2004). Water enters the root and crosses laterally through a series of concentric layers of
cells and then axially to the rest of the plant. In living root tissue, water is transported through both the apoplastic pathway, i.e. between cell walls, and the symplastic pathway, i.e. cell to cell transport (Steudle and Peterson, 1998).

Once it is within the plant, water moves passively along a water potential gradient which is driven by transpiration from the stomatal cavities in leaves. Transpiration reduces the water potential within the xylem vessels and effectively pulls water from the roots into the aerial tissues. When transpiration is reduced or absent (e.g. during the night or under drought conditions), solutes are actively pumped into the roots, thereby increasing the osmotic potential to provide a hydrostatic root pressure. Javot and Maurel (2002) showed that this positive upward pressure forces sap along the xylem. However, because the apoplastic pathway does not distinguish between water and solutes, it is unable to maintain the osmotic gradient for extended periods. The transport of water within the rooting system is therefore mainly driven by hydrostatic forces (Steudle, 1994).

**Stem transport**

Water moves from the roots up into the xylem (rigid, dead cells) which is surrounded by flexible, living cells. The connectivity of flow paths between the living tissues and xylem is limited and creates a resistance to the flow of water. As transpiration takes place, the living tissues (cambium and phloem) shrink as water flows from them into the xylem. When transpiration slows or stops, the potential gradient is reversed and water flows back into the living cells causing them to swell. This change in size of phloem and cambium is observed as trunk diameter fluctuations. When soil water availability is less than that lost in transpiration, shrinkage of the trunk dimension occurs in response to increased water potential gradient.

### 2.4.6 Water loss and conversion by crops

**Evaporation**

All atmospheric moisture originates from the earth’s surface through the process of evaporation. This can occur from water or land surfaces as surface evaporation, or from vegetation as transpiration. In the agricultural context, surface evaporation from the soil ($E_s$) and transpiration ($E_p$) is collectively known as evapotranspiration, $E_t$, which is the total flow of water vapour from a cropped field into the atmosphere. It is important to recognise that direct evaporation and transpiration are the same physical process. In the case of $E_s$, water that evaporates must be replaced at the evaporating surface either through soil capillary action or replenishment by rain or irrigation. For $E_p$, transpired water must be replaced by upward water movement through the SPAC into the leaf surfaces from which further transpiration occurs.

Three atmospheric factors influence evaporation at any location. The first is the capacity of surrounding air to hold water vapour, i.e. the atmospheric demand for water which is a function of air temperature. The amount of water that can be absorbed by the atmosphere depends on its relative humidity at any temperature, i.e. how much water is already there relative to what it can hold at that temperature. The lower the relative humidity, the greater the potential for further evaporation at that temperature, since the atmosphere is able to absorb further water vapour until it is saturated.
The second factor is the energy or latent heat required to convert water from a liquid into vapour. This energy is mainly provided by solar radiation which, by facilitating this change in state, results in a cooling of the evaporating surface.

Thirdly, turbulence, usually measured as windspeed, is required to replace saturated air at the evaporating surface with unsaturated air from surrounding areas (Kowalik and Eckersten, 1984). For all evaporating surfaces, a boundary layer of saturated air exists between the surface and the turbulent atmosphere. Wind movements affect the stability and depth of this boundary layer and, therefore, evaporation occurs more quickly at high windspeeds that reduce the thickness of this layer than in relatively still air. For uniform surfaces, such as bare soil and water bodies, the boundary layer acts as a blanket of wet air above the location. In the case of crops, characteristics such as leaf angle and shape influence the depth and stability of the boundary layer which becomes progressively thicker down the canopy as windspeeds decrease.

**Transpiration**

Plants must absorb CO\(_2\) from the atmosphere and use radiant energy absorbed by their leaves to drive the process of photosynthesis. However, only a very small proportion of the energy absorbed by leaves is used for photosynthesis. The rest tends to heat the crop canopy above the ambient temperature. To prevent canopy temperature increasing throughout the day, plants must dissipate the excess solar energy by transpiration, convection or radiation.

Transpiration is a special form of evaporation because it occurs not across the whole canopy surface but through stomatal apertures on leaf surfaces which can regulate the rate of water vapour loss. The average size and frequency of stomata and the total area of leaf surfaces determine how much transpiration can occur. When a crop covers the ground and its roots are well supplied with water, stomatal regulation is minimal (i.e. stomata are fully open) and direct evaporation from the soil surface is negligible. In these circumstances, evapotranspiration proceeds at a rate determined by the atmospheric conditions; transpiration is effectively the same as the reference evaporation (\(E_o\)) that would have occurred from an open water surface at the same location.

However, the transpirational component of \(E_t\) is rarely equivalent to \(E_o\) because stomata exert a physiological influence on evaporation that depends on plant-related factors as well as atmospheric conditions. Physiological control of transpiration depends on the effectiveness of stomatal regulation by leaves and the total surface area (\(L\)) of the crop canopy. In practice, \(E_p\) is often limited by the supply of water to roots which in turn act on \(F_p\) and the flow of water through the SPAC. In arid climates, where air is hot and dry (large atmospheric demand) and water supply is limited (small soil water supply), stomata remain closed for part or most of the day and restrict seasonal transpiration (\(\Sigma E_p\)) to a finite proportion of \(E_o\) at the same location.

**Conversion of resource \(\varepsilon_w\)**

Plants must constantly balance the demands for CO\(_2\) absorption to support growth and the need to restrict transpiration when soil water is limited. The link between CO\(_2\) uptake and \(\text{H}_2\text{O}\) loss at the leaf level means that, at least in principle, seasonal transpiration can be used to estimate carbon assimilation at the crop scale. Figure 2.5 shows that, for any crop, the
The relation between its accumulated dry weight and its transpiration over the same period is often linear with a slope, ε\textsubscript{w} (g kg\textsuperscript{-1}), known as the ‘dry matter/transpired water’, as with radiation, the relation between productivity and water use is often referred to as water use efficiency (WUE) of a crop (Azam-Ali and Squire, 2002). Hsiao \textit{et al}. (2007) provide a comprehensive review of WUE and its improvement in agricultural systems.

Steduto (1996) and Steduto and Hsiao (2007) review the basis of the link between dry weight and transpiration. The closeness of this relation is probably because both dry matter accumulation and seasonal transpiration are closely related to the total solar energy intercepted by a crop (Monteith, 1981, 1986a). The value of ε\textsubscript{w} appears to remain independent of nutrient uptake (Tanner and Sinclair, 1983) or water stress (Azam-Ali \textit{et al}., 1994). However, when the same crop is grown at different sites or seasons, the value of ε\textsubscript{w} often varies between sites and/or seasons. There are two principal causes of this apparent variation in ε\textsubscript{w}.

The contribution of roots to total productivity

When crops are well supplied with water, the root fraction of total dry weight is small and can often be ignored in calculations of ε\textsubscript{w}. However, in arid climates and/or when soil moisture is limited, root weight may be a substantial component of total biomass. Because accurate measurements of root weight are rare, most calculations of ε\textsubscript{w} are based on above-ground biomass. However, using total (i.e. including roots) or only above-ground biomass may lead to marked variations in published values of ε\textsubscript{w}. For example, for three populations of groundnut growing on stored soil water in India, Azam-Ali \textit{et al}. (1989) showed that the calculated value of ε\textsubscript{w} increased from 1.57 to 3.00 g kg\textsuperscript{-1} when roots were included in measurements of crop dry weight.
The influence of atmospheric saturation deficit

For transpiration to occur, the sub-stomatal cavities of leaves must contain saturated air at the leaf temperature. Strictly, the driving gradient for transpiration depends on leaf temperature as this sets the saturation vapour concentration at the leaf surface, and the driving force for transpiration is therefore the difference between the saturation vapour concentration in the leaf and that in the air ($D'$). As a consequence, the value of $\varepsilon_w$ is inversely related to $D'$ during the season. When leaf and air temperatures are the same, the driving force for evaporation is the true saturation deficit of the air, i.e. $D$, and, in these circumstances, $D$ sets the potential rate for transpiration. If leaf temperature is similar to air temperature, $D$ is an acceptable approximation for the value of the driving gradient (i.e. $D'$). Similarly, although transpiration at any time is closely linked to the current saturation deficit, the seasonal mean value of $D$ is often an acceptable approximation because changes in $D$ during the season are usually small compared with differences between seasons (Monteith, 1986b).

When $D$ is included as a normalising factor (Monteith, 1990), the value $\varepsilon_w D$ (g kPa kg$^{-1}$) appears to remain conservative throughout the season (Squire, 1990) and, unlike $\varepsilon_w$, remains fairly constant from site-to-site and year-to-year. The value $\varepsilon_w D$ is known as the transpiration equivalent, $\Omega_w$, and can be treated as a quasi-constant for any particular crop (see Azam-Ali et al., 1994; Azam-Ali and Squire, 2002).

However, there are a number of drawbacks to using $D$ as a normalising factor for the relation between dry matter and transpiration. First, where canopy temperature is cooler than that of the air, the leaf vapour pressure can be substantially less than $D$. Similarly, under water stress leaf temperatures and the associated leaf vapour pressure can be substantially greater than the mean value of $D$. In each case, $D$ no longer serves as a true estimate of the driving gradient for transpiration because no account is made for the leaf vapour pressure.

Since accurate and frequent measurements of canopy temperatures are rare, the errors caused by variations in leaf to air temperature on the normalisation of $\varepsilon_w$ are difficult to quantify and are strongly dependent on the type of weather data available (Allen et al., 1998). A further complication is that normalisation by $D$ becomes very sensitive at low values of $D$ (Stockle et al., 2003).

Steduto and Hsiao (2007) present the concept of crop biomass water productivity (WP$_b$) for a number of C$_3$ and C$_4$ crops. In their analysis, they compare the normalisation of the dry matter/transpiration relation for evaporative demand through two routes: via the transpiration gradient (i.e. the atmospheric vapour pressure deficit, $D$) and via a reference transpiration flux (i.e. the reference evaporation, $E_o$). They present evidence for numerous crops to show that the normalisation of $\varepsilon_w$ by $E_o$ is more robust than normalisation by $D$ (Azam-Ali et al., 1994; Clover et al., 2001; Steduto and Albrizio, 2005) and provide theoretical and practical evidence for this improvement.

2.5 Integrating principles of resource capture by crops

2.5.1 Reassessing the solar and water engines for crop growth

The ‘Resource Capture Framework’ described in Section 2.1.1 and expounded in this chapter demonstrates that the biomass produced by a crop is a function of the amount of resources that it has captured during its life cycle and the efficiency with which it converts these
resources into dry matter (Black and Ong, 2000; Azam-Ali and Squire, 2002). Where we know the fraction of biomass allocated to the reproductive or economic fraction (i.e. the harvest index), the framework allows us to quantify crop yield.

Crop productivity and yield can be estimated through a solar-driven or a water-driven engine by relating dry matter production to either the amount of radiation captured or the amount of water transpired over the period of growth (Azam-Ali et al., 1994; Azam-Ali and Squire, 2002). Thus, the biomass of a crop \( W \) can be related to seasonal intercepted radiation \( \Sigma/S \) and dry matter/radiation coefficient \( \varepsilon_s \), i.e. \( W = \varepsilon_s \Sigma f S \), or to seasonal transpiration \( \Sigma E_p \) and the transpiration equivalent \( \Omega_w \), i.e. \( W = \Omega_w \Sigma E_p \).

When plant roots are well supplied with water, stomata remain fully open for much of the day and allow transpiration and \( \text{CO}_2 \) uptake to continue at or close to the limits set by the environment. In such circumstances, the total productivity of a crop can be estimated using the solar-driven engine since the interception of solar energy determines productivity. However, water often limits the productivity of rain-fed crops in non-temperate climates (Squire, 1990; Zhang et al., 2000; Lampurlanes et al., 2001). Even in temperate areas, water stress can still cause a significant yield loss in many crops (Jones, 1992). As a result, the water-driven engine generally predominates, because crop growth is driven by radiation only when there is no water stress. Both the capture of solar radiation (i.e. \( f \)) and its conversion efficiency (i.e. \( \varepsilon_s \)) can be significantly reduced under drought (Collinson et al., 1996, 1997; Craufurd and Wheeler, 1999).

Under drought, plants may modify their water extraction pattern from the soil, minimise water loss, reduce leaf area expansion and, in extreme cases, lose leaf area through abscission and/or senescence. All such modifications have implications on total crop productivity because each is linked with the stomatal regulation of photosynthesis.

In the semi-arid and arid tropics, rainfall is both erratic and low in some cases and most of the water received, as rainfall, is lost to run-off, drainage and evaporation – leaving only a small proportion for crop growth. Moisture stress (drought) is therefore a common phenomenon that crops in these regions have to experience during part of their growth periods, unless irrigation is applied and crops in such regions have developed mechanisms to withstand drought and ensure both survival and yield under soil moisture stress. The actual mechanisms employed vary widely both within and between species (Siddique et al., 1990; Dardanelli et al., 1997). There are three generally recognised strategies by which plants adapt themselves to drought conditions. These are classified as drought escape, drought tolerance and drought avoidance. Details of these strategies are discussed by Kramer and Bowyer (1995) and Jones (1992, 1993).

Generally, a crop with a high \( \varepsilon_w \) may give better yield than one with a low \( \varepsilon_w \) or enable it to use less water to achieve the same level of productivity. Variations in \( \varepsilon_w \) amongst species (Zhang et al., 2000; Siddique et al., 2001) and genotypes (Baigorri et al., 1999) have been reported in grain legumes, and such differences can be used to select appropriate genotypes for drought-prone environments.

### 2.5.2 Agronomic management and the resource capture framework

Ultimately, the aim of a farmer is to achieve the maximum productivity of a crop based on its genetic potential and the resources available to it at any location. In some cases, where resources are insufficient from the natural environment (e.g. through rainfall and stored soil nutrients), they may be provided as additional inputs (e.g. through irrigation and
Soil Conditions and Plant Growth

Farmers must therefore make a series of agronomic decisions on how best to match crop requirements with available and/or applied resources throughout the life cycle of a crop.

Figure 2.6 attempts to summarise the key options available to a grower at any location in the form of a ‘Crop Circle’. The factors are not definitive or necessarily chronological but provide a guide that links an inner circle of locational characteristics (site, season, system, species and supplies) in which a grower might assess available resources with a central circle of agronomic options which the grower might employ. An outer circle lists the various terms that have been presented in this chapter at the relevant position in relation to locational characteristics and agronomic options.

Site

A primary consideration in the cultivation of any crop is the existing conditions of the field site and location. The soil type and structure will influence the SMRC and moisture-related variables such as saturation ratio ($\phi$), FC and PWP. These factors linked with soil depth provide an estimate of the total AWC at the location. Available macro (NPK) and micro nutrients can similarly be calculated at this location. Factors such as aspect and slope indicate the likely variations in solar radiation and warming during the growing season and the likelihood of run-off and run-on to the site.
**Season**

The growing season at any location sets the beginning and end of potential crop duration. The weather expected during a growing season sets the likely boundaries for crop growth and development in terms of total and seasonal accumulated solar radiation ($\Sigma fS$), accumulated thermal time ($\Sigma \theta$) and associated cardinal temperatures ($T_b$, $T_m$, $T_m$). The likely amounts and pattern of rainfall, linked with the site and soil characteristics, indicate the seasonal availability of water and likely seasonal transpiration ($\Sigma E_p$).

**System**

Here, system refers to the possible choice between a perennial or annual cycle of cropping and whether crops are to be grown as monocultures of a single species (sole crops) or as combinations of species (intercrops). Further considerations include the expected sowing and harvesting dates of individual species and the architectural arrangements of plants (sowing rate and row structure) that, together, set their pattern of resource capture ($\Sigma fS$, $\Sigma E_p$), demand for nutrients and overall biomass efficiencies ($\varepsilon_s$, $\varepsilon_w$) for contrasting cropping systems.

**Species**

A distinction within perennial and annual systems is the likely duration and, therefore, the start and end of the growth cycle of particular cultivars within the overall resources determined by the growing season. The choice(s) of crops and cultivars (C$_3$, C$_4$, legumes and non-legumes) against the likely availabilities of radiation and water will also influence the respective conversion efficiencies ($\varepsilon_s$, $\varepsilon_w$) of different species individually and collectively. These various permutations will influence the individual and collective capture of radiation and transpiration ($\Sigma fS$, $\Sigma E_p$), the accumulated thermal times ($\Sigma \theta$) and minimum and maximum temperatures for development ($T_b$, $T_m$) of each species and system and the respective demands for nutrients.

An important additional consideration within the choice of species is whether these include leguminous crops which are able to fix atmospheric nitrogen ($N_{fix}$) both for their own productivity and for that of successor or companion species.

**Supplies**

In addition to the resources naturally provided by the site and its environment, there may be supplementary resources that the grower can access to support crop production. Amongst others, such resources may include irrigation, which together with available water and nutrients and applied fertilisers will influence $\Sigma fS$ and $\Sigma E_p$. Together, irrigation and fertilisers may also allow the crop to maintain photosynthesis at the genetic potential of the species (C$_3$ or C$_4$) and cultivars selected and thereby maintain $\varepsilon_s$.

### 2.6 Conclusion

The principles of resource capture provide a relatively simple yet robust framework for the analysis of crop growth, development and productivity in response to abiotic factors, principally solar radiation and water. For soil scientists and crop managers, this framework is
particularly useful where it can guide in the application of water and/or nutrients so that the total biomass of a crop can reach the potential set by its genetic composition and its seasonal capture of solar energy. It is important to recognise that this framework refers to all the products of photosynthesis including roots, stems and leaves as well as seeds or grains. The concept of a ‘soil-plant-atmosphere-continuum’ provides a link between below- and above-ground capture and conversion systems in the optimal management of a range of crops in contrasting environments.

References


3 Soil fertility

Elizabeth A. Stockdale¹, Keith W.T. Goulding², Timothy S. George³, and Daniel V. Murphy⁴

¹School of Agriculture, Food and Rural Development, Newcastle University, Newcastle-upon-Tyne, UK
²Department of Sustainable Soils and Grassland Systems, Rothamsted Research, Harpenden, UK
³The James Hutton Institute, Dundee, UK
⁴Soil Biology Group, School of Earth and Environment, UWA Institute of Agriculture, The University of Western Australia, Crawley, Western Australia, Australia

3.1 Introduction

In Soil Conditions and Plant Growth, 9th Edition, E.W. Russell (1961) stated that ‘a soil can only be fertile if it is a favourable environment for root growth’ and that ‘the root system … must also take up enough water and nutrients for the crop to give a good yield of whatever part is required’. The concept of ‘soil fertility’ is one way that soil scientists immediately engage the interest of gardeners and farmers; both hope that the scientists will answer one of their key questions: ‘How can I manage this land in order to obtain the best and most profitable yield from it year after year?’ All would agree that soil fertility has something to do with the capacity of a soil to support plant growth.

Soil fertility is a deceptively familiar term, but nonetheless an elusive one, where the disciplinary use of the term and its everyday use often do not agree (Patzel et al., 2000). To what extent is soil fertility a fixed characteristic of the site itself and to what extent is it the result of management and hence under the control of land managers? Which properties of soil (and of the plants themselves) should be measured to give a robust assessment of soil fertility? If crop yields are maintained or increasing, then can we simultaneously speak of a decline in soil fertility? As different approaches are taken in answer to these and similar questions, it has been possible to identify only an ‘observable mess of implicit, indirect and either over-stretched or over-restricted definitions of soil fertility’ (Patzel et al., 2000).

In the UK, the Agricultural Advisory Council (1970) recognised the range of definitions in use amongst agricultural scientists, advisors and farmers and hence adopted a broad definition to underpin a review of practice: ‘we have regarded the qualities which make up a soil’s fertility as those which allow crops to grow and produce optimum yields’. In this chapter, we will take this broad definition as our foundation and then briefly consider some of the frameworks within which soil fertility has been considered, highlighting the soil–plant interactions which lie at the heart of a full and integrated understanding of soil fertility. Finally, we will consider how advice on the management of soil fertility has developed and outline the challenges which remain for practical management of soil fertility in the future.
3.2 Concepts of soil fertility

3.2.1 Links between soil fertility and land capability and suitability assessment

Land use patterns and their spatial distributions at a global scale are driven largely by climate, giving broad agro-ecological zones. At regional/local scales, geology, in combination with climatic patterns, has a critical role together with associated variation in topography (gradient, aspect), hydrology, soil depth and erosion risk (FAO, 1996). To support effective and sustainable crop production, it is important to understand any permanent constraints to land use and crop yield which cannot be (economically) overcome by management and within which any farming system must be adapted. Land evaluation using a land capability classification approach was developed for farm planning in the USA by Klingebiel and Montgomery (1961) and has since been adapted and widely used across the globe. Permanent constraints (which may vary in importance in different climates/regions) are defined as limiting factors and used to assess land use capability. With this approach, it is the negative/limiting features of land which are taken into account to assign a land capability class, with the exception of location and accessibility. Few soil characteristics are routinely considered for land capability classification. The factors that are measured are those which lie beyond the control of the farmer: extremes of soil texture, soil depth and stoniness. Soil texture (the relative amounts of differently sized mineral fractions: sand silt and clay) and stoniness are inherent properties of the soil strongly related to characteristics of the parent material and cannot usually be easily or quickly changed. These inherent properties largely define the ranges within which management is able to control a number of soil properties, e.g. the formation and stability of soil structure, soil consistence/workability during tillage, infiltration capacity and depth and penetration of water and the capacity of soil to supply and retain nutrients. So, e.g., it is possible to give some general information on a range of soil characteristics based simply on the knowledge of its texture: a silty loam will tend to slake and cap at the surface; a sandy loam will hold less water and tend to become acid more quickly and its capacity to hold nutrients will be relatively low. However, even where all the permanent limiting factors are understood, land cannot be simply graded from best to worst on a single scale without taking its use into consideration (Gibbon, 1961).

The development of a Framework for Land Evaluation (FAO, 1976) was deliberately focussed on the assessment of land suitability in relation to particular defined land uses, i.e. ‘Fitness for Purpose’. In these assessments, consideration of both permanent (inherent) and manageable constraints is included. As most suitability assessments have been carried out for various types of crop production, the factors used to define subclasses are commonly those associated with definitions of soil fertility (e.g. droughtiness, soil nutrient deficiencies, waterlogging, ease of cultivation). Definitions of suitability classes are expressed in terms of expected yields and/or the level of inputs required to achieve good yields. Measuring potential productivity along a continuous scale, rather than in discrete classes, is also possible through the use of arithmetic formulae combining key limiting factors within a number of parametric approaches to land/soil appraisal, e.g. Riquier et al. (1970) and Beatty et al. (1979).

Common to all these approaches is the identification of constraints to crop production (Figure 3.1) and the development of inputs and/or management practices which overcome these constraints, so that crop yield potential defined by genetic and climatic factors can be
achieved. Russell (1937) identified the main soil constraints to crop growth that can be controlled through management as water supply, nutrient supply, soil reaction (including acidity, salinity), erosion risk and depth of soil. Factors defining many of these constraints are routinely included in suitability assessments.

### 3.2.2 Equating soil fertility with the soil nutrient reserve

As described in Chapter 1, the role of the soil in supplying plant nutrients was at the heart of early controversy in agricultural science; developing, then patenting artificial manures provided much of the income stream for early research. As a result, the availability and use of fertilisers has expanded greatly on a global scale (see Section 1.5.4). Consequently during the twentieth century, soil fertility research often focussed on ‘the study of all factors which affect the nutrition of agricultural crops’ and ‘how growth conditions for plants can be improved by manuring or fertilising’ (de V Malherbe, 1964). This is not a small area of focus (Figure 3.2). Consequently for many agricultural scientists, the term ‘soil fertility’ developed a relatively
narrow meaning restricted to consideration of a soil’s potential to supply nutrients for crop growth. This does not mean that the impact of other factors on yield, such as water availability or poor structure restricting root growth, were considered less important, but simply that these were considered separately to the management of soil fertility by fertiliser addition.

### 3.2.3 Understanding soil fertility in relation to concepts of ‘soil quality’

At the end of the twentieth century, the concept of soil quality developed alongside an increasing awareness and valuation of the goods and services provided by ecosystems globally to support human life (Costanza et al., 1997; Millenium Ecosystem Assessment, 2003) and the central role of the soil in the delivery of a range of key functions (Lal, 2001). Soil quality as defined by the Soil Science Society of America (Doran and Parkin, 1994) is ‘the capacity of the soil, within land use and ecosystem boundaries to sustain biological productivity, maintain environmental quality and promote plant and animal health’. But the standard against which soil quality is to be compared is rarely agreed or well defined and, as the definition by Doran and Parkin indicates, it must be assessed in relation to the intended use of the soil: ‘Fitness
for Purpose’ again. The utility of an overarching concept of soil quality has been debated (Karlen et al., 2001; Sojka et al., 2003) but it now stands alongside definitions of acceptable measures of air and water quality in guiding environmental public policies. These issues are considered in more depth in Chapter 12.

It is clear, however, that embedded within all definitions of soil quality there is a recognition of the role of the soil in supporting food and fibre production, i.e. soil fertility as broadly defined (Figure 3.3). The relative emphasis given to aspects of soil fertility in the assessment of soil quality at any site will depend on the intended land use and the location; e.g., if the site is within a key catchment determining water supply for a city or next to a protected wildlife habitat, the role of the soil in supporting food and fibre production would not be a key priority. Many of the early examples of soil quality assessment cited in the literature focussed on crop productivity as they were carried out in agricultural areas (e.g. Romig et al., 1995; Sinclair et al., 1996; Stenberg, 1999) and are similar to the integrating approaches used for crop suitability assessments (see Section 3.2.1). Food and fibre production is recognised as a key soil function in all current definitions of soil quality. In this context, soil fertility might be regarded as the state of soil required for the effective performance of this function. Although some authors appear to use the terms interchangeably, soil fertility is not synonymous with soil quality and it should be considered as one aspect of a more integrated definition of soil quality.

### 3.2.4 Defining soil fertility in relation to ecosystems dynamics

The large diversity of soil organisms and their interactions (predator–prey and extensive omnivory) are the driving force behind many soil functions, including nitrogen (N) and phosphorus (P) mineralisation, structure formation and stabilisation, and consequently underpin...
Soil Conditions and Plant Growth

Table 3.1  Annual average biomass (kg C ha\(^{-1}\)) and simulated contribution to overall N mineralisation (kg N ha\(^{-1}\)) for the 0–10 cm layer of a winter wheat field under integrated management for key groups of soil organisms.

<table>
<thead>
<tr>
<th>Soil organism group – by functional group if considered in the model</th>
<th>Average biomass (kg C ha(^{-1}))</th>
<th>Modelled annual N mineralisation (kg N ha(^{-1}))</th>
<th>% Reduction in amount of N mineralised if fauna group removed from the model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bacteria</td>
<td>373.1</td>
<td>52.96</td>
<td>–</td>
</tr>
<tr>
<td>Fungi</td>
<td>21.0</td>
<td>4.53</td>
<td>–</td>
</tr>
<tr>
<td>Protozoa</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amoebae</td>
<td>6.09</td>
<td>13.63</td>
<td>32.4</td>
</tr>
<tr>
<td>Flagellates</td>
<td>0.35</td>
<td>0.82</td>
<td>1.9</td>
</tr>
<tr>
<td>Nematodes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bactivorous nematodes</td>
<td>0.41</td>
<td>6.03</td>
<td>17.3</td>
</tr>
<tr>
<td>Fungivorous nematodes</td>
<td>0.02</td>
<td>0.05</td>
<td>1.2</td>
</tr>
<tr>
<td>Predatory nematodes</td>
<td>0.25</td>
<td>1.62</td>
<td>19.1</td>
</tr>
<tr>
<td>Enchytraeids</td>
<td>0.19</td>
<td>0.45</td>
<td>3.2</td>
</tr>
<tr>
<td>Earthworms</td>
<td>8.86</td>
<td>0.47</td>
<td>Not simulated</td>
</tr>
</tbody>
</table>

Data for small pools of soil microarthropods are not presented. The importance of predation in driving overall N mineralisation is shown by the relative increase in the estimated contribution made to mineralisation when it is estimated by the percentage reduction in the amount of N released through mineralisation as a result of exclusion of different groups of soil fauna.

Source: Data derived from de Ruiter et al. (1993).

...
importance of abiotic factors (climate, management, inherent site constraints) in controlling the biological interactions and driving soil processes at a range of scales is still poorly understood (Bardgett, 2002). For soil fertility, increased understanding of the interactions driving soil water balance, nutrient cycling, rooting conditions and the regulation of soil-borne
plant pathogens seem most important (Stockdale et al., 2006). A mechanistic understanding of such interactions is needed if management practices are to be robustly evaluated (Brussaard, 1998) and developed in a way that protects the ecosystem functioning at the heart of soil fertility.

Sanchez (1994) considered that the underlying rationale of soil fertility management, i.e. every constraint should be overcome through external inputs, was outdated and highlighted the need to replace this with a second paradigm to underpin soil fertility research which increased focus on the use of endogenous processes and potentials within the soil-plant system. We recommend taking an integrated ecosystem approach to the definition of soil fertility and that research into soil–plant interactions, which ultimately allow crops to grow and produce optimum yields, will be the focus for the twenty-first century rather than separate studies of soil processes in the absence of plants and/or processes in the plant root in the absence of soils. As illustrated in Figure 3.2, a wide range of plant and soil properties interact to affect nutrient supply to and uptake by plants. The study of any of these factors without an understanding of the ecosystem context in which they take place may limit the application of increased understanding to the improvement of soil fertility management in practice. Similarly for soil-borne pathogens, van der Putten et al. (2009) highlight how
modelling above- and below-ground plant–pathogen interactions could lead to new avenues for sustainable control of pests and diseases in crops, above- and below-ground pest and disease suppression, increased roles for arbuscular mycorrhizal (AM) fungi in plant nutrition and improved crop productivity.

### 3.3 Plant–soil interactions determining soil fertility

#### 3.3.1 Impacts of soil on plant roots

Soil factors control plant growth largely through their impact on plant roots; in annual cropping systems, plant roots occupy relatively little of the soil volume (5–10%), whereas in grasslands it can be up to 40% of soil volume. The supply of water and nutrients to support plant growth critically depends on uptake from the soil across root membranes, so any factor that impedes plant root growth will impact water and nutrient uptake and apparent soil fertility. The rate of supply of nutrients to any root depends on the rates of mass flow and diffusion, but of critical overall importance for water and nutrient uptake is the distribution of roots within the soil, i.e. the total surface area over which uptake can occur. Root growth shows responses and interactions with a range of chemical, physical and biological factors in the soil (Hodge, 2006). In turn, many of these interact and plants are often able to compensate, so that restricted root growth in one part of the soil may be compensated for by increased growth in other regions that are less constrained.

The ability of the plant root to find space in which to extend through the soil is often an important factor limiting plant growth. The preparation of seedbeds with a good structure (‘tilth’) and consequent network of pores, holding the right balance of air and water appropriate to the seed to be sown, is an art underpinned by the science described in Chapter 8. The effects of soil compaction on root growth are also reviewed in Chapter 8; compact zones with high bulk density may restrict root penetration and radial growth of roots and so may reduce yield. The root cap perceives touch stimuli, and roots have also been shown to modify their gravitropic growth to avoid obstacles (Massa and Gilroy, 2003). Soil pore size and organisation together with water inputs to the soil determine soil water content (Chapter 9). The range in soil water content within which plant growth is not constrained by matric pressure, aeration or mechanical resistance (the least limiting water range; Da Silva et al., 1994) is therefore one indicator of soil fertility.

The amount, form and distribution of nutrients or toxic elements in soil also influence the amount and distribution of root growth. Root branching may also be regulated by external nutrient concentrations with local stimulation of lateral root growth observed in zones of high N, P and potassium (K) availability (Drew, 1975; Forde and Lorenzo, 2001), though not all plants respond in the same way (Hodge et al., 1998). Plants will usually selectively avoid rooting into zones of high toxicity. For example, in most plants, the first visible symptom of aluminium stress (often linked with acid conditions) is the inhibition of root elongation (Matsumoto, 2002); consequently rooting depth is often restricted by acid subsoils. Similar restrictions to rooting are seen where soils are saline or have boron toxicity. Under conditions of nutrient deficiency or toxicity, a range of specific physiological changes take place caused by multigenic interactions (Figure 3.6); e.g., over 1000 genes are differentially regulated under conditions of P starvation in Arabidopsis (Hammond et al., 2003; Wu et al., 2003; Morcuende et al., 2007). Understanding the transcriptional regulation of these genes can allow the identification and selection of genotypes that are better suited to cope with conditions of low nutrient availability or elevated concentrations of potentially
Soil Conditions and Plant Growth

Toxic elements; it may also allow the use of transcriptional arrays to identify nutritional deficit or toxicity far in advance of any physical symptoms (Hammond et al., 2004).

Direct damage to plant roots by root feeding and parasitic organisms can be considerable (particularly by beetle larvae and nematodes) and can therefore limit the extent to which roots exploit the soil volume. In addition, a wide range of soil micro-organisms that can cause plant disease through root invasion, in particular soil fungi, have been identified and their ecology described (for detailed descriptions see, e.g., Bruehl, 1987). Interactions between the plant species, soil organisms and environmental conditions determine infectivity with regard to the host crop (Smith et al., 1999). Some soils have been shown to be disease suppressive (Menzies, 1959); a wide range of ecological mechanisms (including parasitism, direct and indirect antagonism) are now thought likely to contribute to this effect (Mazzola, 2002) and interactions with soil chemical and physical properties are also important (Duffy et al., 1997).

3.3.2 Association and symbiosis between plant roots and soil organisms

The majority of plants are capable of forming mycorrhizal associations with various soil fungi (see Chapter 11). The most common is the AM association, which has been shown to have
emerged very early in the evolution of land plants (Remy et al., 1994). Developments in molecular and transcriptional analysis in recent decades mean we now understand a lot about the molecular responses of non-mycorrhizal plants to nutrient deficit, primarily due to the focus of transcriptional studies on the model plant Arabidopsis. Unfortunately, very little is known about such responses in mycorrhizal plants, which make up 90% of all land plant families; only very few crop species do not form such associations, among them Brassicaceae and Chenopodiaceae. Over 60 years ago, Sir Albert Howard suggested that ‘the presence of an effective mycorrhizal symbiosis is essential to plant health’ (Howard, 1943). The wide benefits of mutualistic relationships in an AM association to crop nutrition, health, stress resistance and the development of soil structure have since been confirmed and the impacts of management on their infectivity and effectiveness studied, e.g. Bethlenfalvy and Lindermann (1992), Harrier and Watson (2003) and Gosling et al. (2006). Most soils have indigenous populations of AM fungi, but more work is needed to understand whether and how these associations can be manipulated to support integrated soil fertility management approaches.

The nitrogen-fixing symbiosis that exists between plants of the Leguminosae (legumes) and the soil heterotrophic bacteria, known generally as rhizobia, has been understood for over 100 years (Atwater, 1884; Hellriegel and Willfarth, 1888). There are a number of other nitrogen-fixing associations between soil micro-organisms and plants, those involving cyanobacteria and those in which the microbial partner is an actinomycete, which are important in natural and in agricultural systems. Rhizobia are a phylogenetically disparate group of α- and β-proteobacteria; more than 70 species have been identified (Masson-Boivin et al., 2009). Nodule formation in legumes occurs as a result of complex interactions between plant and microbial processes whose biological, biochemical and genetic aspects have been much studied (Sprent, 2001; Masson-Boivin et al., 2009). The precise details of nodule formation are still not fully understood (Martínez et al., 1990; Brewin, 1991; Povorov, 1998), although research into signalling and host plant cell responses during infection by rhizobia or arbuscular mycorrhizae (Oldroyd et al., 2009) has suggested that it may be possible to induce colonisable nodules on plants other than legumes. Rhizobia–legume interactions have been shown to be highly specific (Long, 2001) and a wide variety of morphological variations have been observed. McInnes and Haq (2003) reported the impact of a wide range of environmental, plant (host and non-host) and soil (physical, chemical, biological) factors on the persistence and effectiveness of rhizobial populations in the nodulation of legumes. Even where nodulation is effective, the achievement of high rates of nitrogen fixation depends on a very complex mixture of interactions including the plant, the bacteria species, and climatic and soil factors (Peoples et al., 2009). At the practical interface with agriculture, agronomists still struggle even to give broad ranges of N fixation associated with any particular species of crop or forage legumes, let alone give a site-specific estimate on which a rotational fertilisation plan can be based (Smil, 2001; Unkovitch and Pate, 2001; Herridge et al., 2008).

The development of molecular technologies is increasing the number of effective plant–microbial interactions that have been identified and understanding of their modes of interaction (Berg, 2009). Much of the work on these interactions to date has focussed on the understanding of the mechanism and possible isolation of the key organisms involved. Hayat et al. (2010) identify three main modes of action for plant growth promoting rhizobacteria (PGPR) species: (1) synthesising particular compounds for the plants, (2) facilitating the uptake of certain nutrients from the soil and (3) lessening the impact or protecting plants from diseases. A number of plant growth promoting fungi (PGPF) have also been isolated belonging to the genera Trichoderma, Fusarium, Penicillium, Phoma, etc. and have been
reported to be beneficial to several crop plants, not only by promoting their growth but also by protecting them from disease (e.g. Shivanna et al., 1996; Harman, 2011a). Trichoderma has been known as a fungal biocontrol agent for many plant diseases since the early twentieth century, although the mechanisms of action were not known. Much or most of the biocontrol activity of these fungi is now considered to be a result of their ability to induce systemic disease resistance, and effective strains also convey stress resistance to plants (Harman, 2011b). The basidiomycete Piriformospora indica has at least qualitatively similar effects on plants (Shoresh et al., 2010), and hence it has been suggested that mixed species approaches might have greater impact. Ardanov et al. (2011) suggest that there may be an important difference between endophytic bacteria which are able to activate basal and inducible plant defence systems and biocontrol strains which have a plant growth promotion effect which is not correlated with induced resistance.

Plant growth promotion arising from interactions between soil organisms and plant roots is complex and often cannot be attributed to a single mechanism; it may result from the synergistic interactions of a number of species, not only soil bacteria or fungi (Richardson et al., 2009). For example, many reports have documented the ability of soil micro-organisms (such as Pseudomonas, Bacillus and Rhizobium spp., actinomycetes and various fungi such as Aspergillus and Penicillium spp.) to solubilise insoluble mineral phosphate compounds (Richardson et al., 2009). The most common mechanism used by micro-organisms for solubilising mineral phosphates in neutral and alkaline soils seems to be acidification and Ca-chelation via biosynthesis and release of a wide variety of organic acids (Whitelaw, 2000 and Chapters 5 and 11). However, attempts to improve mineral phosphate solubilisation by inoculation with isolated P-solubilising bacteria and/or fungi have not been particularly successful because of limitations such as poor ecological fitness, low metabolite production, variability in inoculant-delivery systems and inconsistent performance in field applications (Shenoy and Kalgudi, 2005; Richardson et al., 2009).

### 3.3.3 Impacts of plants on soil

While soil properties have a role in controlling root distribution, plant roots also play a key role in soil structural development; roots are the most important agents creating structural pores in undisturbed soils. Roots expand by forcing a root tip into pores of 100–300 μm diameter. The young root then tends to compress soil around the pore by radial expansion, increasing the size of the pore occupied by the root but decreasing porosity and modifying pore size distribution in the rhizosphere; in the medium term, rooting tends to increase soil macroporosity through the provision of stabilised continuous channels (Angers and Caron, 1998). Differences between crops in their impact on soil structure have been noted, in part as a result of the fact that crops differ in the patterns and depth of rooting (Gregory, 2006; Chapter 10). While plant roots use pores of >100 μm diameter as points of entry, they can effectively empty pores down to 0.2 μm; this increases differences in the depth and patterns of soil wetting and drying.

The zone of soil in the immediate vicinity of the root is markedly different in its properties to the remainder of the soil as a result of the direct impacts of plants on the soil (Table 3.2). In 1904, Hiltner first used the term ‘rhizosphere’ for this interfacial zone between the plant roots and the bulk of soil, and this term is now used widely for the zone of soil surrounding the root within which the soil is directly affected by the presence of the root. The rhizosphere is a dynamic transit zone, with large fluxes of water and plant nutrients across it, and consequently small pool sizes for most nutrients, driven by patterns of plant demand and uptake (for more detail see Chapter 10).
Soil fertility

Table 3.2  The rhizosphere has markedly different physical, chemical and biological characteristics than bulk soil which is not under the direct influence of plant roots.

<table>
<thead>
<tr>
<th>Soil characteristic and observed change in the rhizosphere</th>
<th>Reference for paper where an example of these differences can be found</th>
</tr>
</thead>
<tbody>
<tr>
<td>Water retention properties</td>
<td></td>
</tr>
<tr>
<td>Smoother, less convoluted pores; larger pores on average</td>
<td>Whalley et al. (2005)</td>
</tr>
<tr>
<td>Nutrient availability</td>
<td>Marschner (1995)</td>
</tr>
<tr>
<td>Five- to tenfold accumulation or depletion depending on the balance between mass flow and uptake at the root surface</td>
<td>Neuman and Römheld (2002)</td>
</tr>
<tr>
<td>pH</td>
<td></td>
</tr>
<tr>
<td>Changes of 2–3 pH units (100–1000-fold difference in proton concentration); both increases and decreases in acidity of this magnitude are seen in the rhizosphere depending on plant species and management particularly the form of N taken up by the plant; uptake of N as NH₄ leads to proton excretion and is acidifying, NO₃ tends to increase pH</td>
<td>Neuman and Römheld (2002)</td>
</tr>
<tr>
<td>Redox conditions</td>
<td></td>
</tr>
<tr>
<td>Reductase capacity can be increased up to 100-fold in the rhizosphere</td>
<td>Römheld (1990)</td>
</tr>
<tr>
<td>Labile organic matter</td>
<td></td>
</tr>
<tr>
<td>Sugars and amino acids increased by 2–10 times in the rhizosphere as a result of root exudation</td>
<td>Uren (2000); Jones et al. (2004)</td>
</tr>
<tr>
<td>Microbial populations</td>
<td></td>
</tr>
<tr>
<td>10- to 100-fold increase in microbial activity with particular stimulation of gram -ve bacteria in the rhizosphere but reduction in microbial diversity</td>
<td>Chineau et al. (2000); Kowalchuk et al. (2002)</td>
</tr>
<tr>
<td>Soil fauna</td>
<td></td>
</tr>
<tr>
<td>Marked increase (around tenfold) in bacterial grazers (nematodes, protozoa, microarthropods)</td>
<td>Griffiths (1994); Bonkowski et al. (2000)</td>
</tr>
</tbody>
</table>

The amounts and composition of root exudates are the main driver for the large increase in size of the microbial biomass and change in the structure of communities of soil organisms found in the rhizosphere. For example, species-specific flavonoids have been identified as key signalling compounds in plant–micro-organism interactions (Shaw et al., 2006). Distinct communities of bacteria have been found within the plant root and on the root surface (rhizoplane; Lupwayi et al., 2004; Dennis et al., 2009); around 4–10% of the root surface has been shown to be covered by soil micro-organisms. Some soil organisms stimulated in the rhizosphere are plant pathogens, and there is evidence that bacterial as well as fungal pathogens have co-evolved with plants showing a high degree of host specificity (Raaijmakers et al., 2009). Plants are not the passive recipients of a pre-existing rhizosphere community that merely feed on root exudates: a series of interactions between plants and soil factors determines the structure and function of the associated rhizosphere community (Berg and Smalla, 2009).

Plants regulate and exclude elements and molecules at the root surface, manipulating the rhizosphere in a range of ways (changing pH and redox potential, hydraulic lift, exudation and secretion of organic acids and enzymes) in order to tolerate toxicities or deficiencies or manipulate microbial communities (Chapter 10). All such traits are controlled by specific genes or cascades of genes, some of which have been identified but many of which are still to be elucidated. Once understanding of the genetic control is gained, it may be possible to manipulate the rhizosphere directly by deploying crops and/or varieties in which the gene of
Soil Conditions and Plant Growth

interest is expressed, within appropriate crop rotations. Much potential exists to manipulate rhizosphere properties at a genetic scale, but such attempts need to be informed by an understanding of the controls imposed by the complex soil environment in which they need to be effective (George et al., 2004, 2005).

3.4 Assessing soil fertility

Assessment of soil fertility takes place within the context of an individual farming system, where the crop choice and crop yield and quality goals are largely defined by constraints arising from societal policies, laws and financial incentives, economic costs and the aspirations, attitudes, education and training of the land manager. In the twentieth century, the assessment of soil fertility focussed on the assessment of the need for a range of external inputs that were used to overcome the identified soil constraints. Socio-economic factors also interact with the physical constraints arising from climate and location characteristics; together these constraints also determine which, if any, management strategies are available for soil fertility management. Towards the end of the twentieth century, increased awareness of the environmental impact of agriculture led to policies and legislation in a number of countries limiting the amount and/or timing of inputs, such as fertiliser and manure applications, and cultivations, such as ploughing grassland.

During the mid-late twentieth century, the use of fertilisers to overcome nutrient limitations to crop production expanded greatly, and fertiliser purchases are now a major part of the variable costs of production in most cash-cropping systems worldwide. Alongside these changes in management, developments in soil fertility assessment were focussed on the measurement of soil nutrient status to guide the development of recommendations for lime and fertiliser application to crops (Tisdale and Nelson, 1966; Cooke, 1982; Foth, 1997). Thousands of pot and field experiments have been conducted to evaluate soil fertility (i.e. given conditions $x$, how well does the plant grow, take up nutrients and water, produce grain/fruit) and to establish management recommendations for inputs of fertiliser, animal manures or lime (i.e. given conditions $x$, with an addition of $y$, a yield increase of $z$ can be expected). However, in many developing countries, the adoption of such intensive input-focussed systems has not occurred to any great extent, in part because high yielding varieties, fertilisers and pesticides can be difficult to obtain due to poor distribution infrastructure and hence high costs. A range of renewable soil fertility replacement strategies have been proposed that are suitable for low-input systems (Ajayi et al., 2007). For instance, significant attention has been given to the development of legume agroforestry (fertiliser tree) systems, the integration of green manure and dual purpose legumes and the careful use of animal manures often integrated into conservation farming approaches; the latter integrate reduced tillage, retention of crop residues, targeted planting and farm input use and the use of legumes in the rotation (Sanchez and Salinas, 1981; Palm et al., 2001).

Increasingly in industrialised countries, farming systems are required to deliver broader societal goals, e.g. protection of wildlife habitat and/or seek alternative sustainable approaches, e.g. through certification to organic farming standards. Consequently, there has been an increased focus on practices and technologies that optimise inputs, make better use of existing soil nutrients and minimise the leakage of nutrients, especially N and P, from farms to the wider environment. This has seen farm nutrient management planning shift from a relatively small set of procedures designed to optimise fertiliser application rates for crop and animal production towards comprehensive whole-farm nutrient management
approaches that consider a range of issues to deliver both farm productivity and environmental outcomes (Goulding et al., 2008). In the twenty-first century, alongside predicted changes to climate, including the increasing frequency and intensity of extreme weather events, it seems almost certain that fossil-fuel-based inputs, such as N fertilisers and pesticides, as well as other non-renewable resources such as P fertilisers, will become increasingly expensive and scarce. These constraints are likely to necessitate the adoption of changed practices within intensive input-focused farming systems (Kirschenmann, 2007). Consequently, the uptake of practices such as carefully planned crop rotations, maintenance of soil organic matter and the recycling of local organic ‘wastes’ is likely to increase in all farming systems (Hanson et al., 2007) and be integrated with new approaches developed for lower input conditions. Within such systems, a broader integrated approach to the assessment of soil fertility that considers the integration of soil physical, biological and chemical aspects, together with inputs of water and nutrients from a range of sources, will be needed. It is therefore timely to review the assessment of soil fertility in this context.

### 3.4.1 Assessment of soil fertility and fertiliser requirements

The mere analysis of soil properties, however carefully selected, does not give a measure of soil fertility unless the selected parameters have been calibrated against plant growth. In yield response experiments, the aim is to describe the relationship between crop yield and increments of a nutrient or, in a factorial experiment, to measure the response to two or more treatments in combination. When carried out in the field, such an experiment can be used to assess the relative importance of limiting factors to crop yield at the site (Table 3.3) and may also allow the optimum soil fertility management strategy (cultivations, crop residue management strategies, inputs of organic matter, nutrients, water, lime) under the conditions of the experiment to be identified. The results for a single combination of crop, site and management factors are of very limited use alone. Consequently, farmers or their extension officers may need to carry out site-specific on-farm research to identify optimum system-adapted soil fertility management approaches. Additionally, by grouping results from similar soils and crops, it may be possible to draw out guidelines, whereby a set of measures of soil properties provide an index of soil fertility, which support the farmer to select from a limited range of soil fertility management strategies to obtain optimum economic benefit. For example, this approach has been used over many years to create and adapt the UK Fertiliser Manual (RB209) (Defra, 2010).

A number of steps are critically important in any system used to support the practical management of soil fertility:

1. Analysis of selected soil properties which have been shown to correlate with crop yield, nutrient uptake or quality.
2. Interpretation of the results to predict the degree of yield constraint as a result of expected deficiency of nutrient supply or intensity of some other constraint, e.g. compaction.
3. Development of recommendations for appropriate fertiliser, manure or other crop management practices that will enable yield optimisation.

The effective performance of any soil fertility assessment system also relies on the fact that the sample points measured in the field or presented for analysis are representative. More, deeper and perhaps more frequent, soil samples are needed for more mobile nutrients (e.g. nitrate, sulphate) or constraints (e.g. salinity). Hence, great care is taken as part of
fertiliser recommendation systems to give clear guidance on the pattern, depth, timing and frequency of sampling (Sabbe and Marx, 1987; Defra, 2010); such recommendations are fitted to different land use types and perhaps even different cultivation methods; e.g., increased use of no-till approaches requires careful attention to the depth of samples taken as increased profile stratification is seen in no-till systems (e.g. Ismail et al., 1994). Increased potential precision in fertiliser application, through the use of machinery guided by global positioning systems, has led to an increase in the use of nutrient mapping within fields on large, intensive farms. Such ‘Precision Farming’ needs to be underpinned by a good understanding of the spatial variability of the inherent soil characteristics such as texture and an appreciation of the many factors that control soil fertility and crop growth and yield. Farmers that use precision farming techniques value them for things such as savings in fertiliser costs (minimal overlapping during application). However, Anderson-Cook et al. (1999) showed no significant advantage in detailed grid sampling approaches compared to the use of stratified sampling approaches designed with an awareness of soil-type patterns to guide

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Area</th>
<th>Coefficient(^a)</th>
<th>P-value(^b)</th>
<th>% variability explained(^c)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Climate Rainfall (mm)</td>
<td>211(^a)</td>
<td>206(^a)</td>
<td>–</td>
<td>ns</td>
</tr>
<tr>
<td>Agronomy N fertiliser (kg N ha(^{-1}))</td>
<td>20(^a)</td>
<td>24(^a)</td>
<td>0.02</td>
<td>0.055*</td>
</tr>
<tr>
<td>Physical Clay content(^d) (%)</td>
<td>11.0(^a)</td>
<td>10.4(^a)</td>
<td>0.08</td>
<td>0.062*</td>
</tr>
<tr>
<td>Chemical Total carbon (T C ha(^{-1}))</td>
<td>9.0(^a)</td>
<td>10.8(^b)</td>
<td>–</td>
<td>ns</td>
</tr>
<tr>
<td>pH (CaCl(_2))</td>
<td>5.7(^a)</td>
<td>5.6(^a)</td>
<td>–</td>
<td>ns</td>
</tr>
<tr>
<td>EC(^d) (mS m(^{-1}))</td>
<td>80(^a)</td>
<td>63(^b)</td>
<td>–</td>
<td>ns</td>
</tr>
<tr>
<td>Biological Labile C (kg C ha(^{-1}))</td>
<td>83(^a)</td>
<td>118(^b)</td>
<td>0.01</td>
<td>0.041**</td>
</tr>
<tr>
<td>Microbial biomass C (kg C ha(^{-1}))</td>
<td>107(^a)</td>
<td>183(^b)</td>
<td>0.01</td>
<td>0.001***</td>
</tr>
<tr>
<td>PMN (kg N ha(^{-1}))</td>
<td>7.0(^a)</td>
<td>10.1(^b)</td>
<td>0.14</td>
<td>0.003***</td>
</tr>
</tbody>
</table>

Average grain yield was 1.76 and 3.24 t ha\(^{-1}\) in areas A and B, respectively. Mean values of attributes that contributed significantly to the explanation of wheat grain yield are presented; attributes followed by the same letter are not significantly different between the two areas. Several soil attributes were identified that did not have a significant (direct) influence on grain yield; but may have had an indirect influence through their effect on the size of the microbial biomass. In this case, 66% of the variability in microbial biomass could be explained by clay content (log transformed data), pH and labile carbon.\(^a\) The coefficient can be interpreted as t ha\(^{-1}\) grain yield change per unit change in attribute.\(^b\) The variability explained has a maximum of 100% and is not additive between individual attributes.\(^c\) Using a multiple regression model that included all parameters listed in the table, we were able to explain 42% of the yield variability.\(^d\) Clay and EC data were assessed using robust regression analysis due to unusual data points. EC, electrical conductivity.\(^*\) Significant at P < 0.10;\(^**\) Significant at P < 0.05;\(^***\) Significant at P < 0.01. ns, not significant.

Source: Unpublished data provided by D.V. Murphy.
P and K fertiliser recommendations for precision farming approaches, especially given the significantly increased costs associated with the collection and analysis of soil samples in most grid sampling approaches.

### 3.4.2 Role of plant analysis for soil fertility assessment

Where long-term assessment of soil fertility is required or where plant stress can be addressed by management within the growing season, plant analysis rather than soil testing may provide an appropriate indicator of soil fertility. Within-season testing of plant tissues at particular growth stages may permit fertiliser treatments to be used to address critical deficiencies which otherwise might limit yield (Walworth and Sumner, 1986). For example, Havlin et al. (1999) provide a clear summary of the use of plant tissue testing to guide in-season fertiliser for a wide range of crops within the USA. The availability of soil nutrients to a plant is determined both by the nutrient absorption capacity of the plant root and nutrient supply and retention characteristics of the soil. Consequently, post-harvest analysis of yield and nutrient concentrations in crops allows the evaluation of the integrated impact of both plant and soil processes and the efficacy of soil fertility management, given the actual yield, weather conditions and achieved management interventions (all of which may differ from what had been planned).

‘Smart plants’ have been developed under research conditions and it may become practical to use them in the field to diagnose nutrient deficiencies or growth limitations due to other constraints, e.g. acidity, in real time. Such plants are genetically engineered by transformation with constructs containing the promoter of a gene up-regulated specifically under the target condition (e.g. P deficiency) in an accessible tissue. For example, Arabidopsis bearing such a construct showed increased \( \beta \)-glucuronidase activity was linked to a green fluorescent protein signal molecule and was detectable, using confocal microscopy, when P was withdrawn from solution culture before growth was limited (Hammond et al., 2003). Hence, such ‘smart plants’ could have the potential to monitor plant response to previously identified limiting constraints and allow targeted precision management within the growing season before the limitation becomes a problem.

### 3.4.3 Soil analysis

While plant nutrient concentrations are useful post hoc evaluation tools, they are of little use for the advanced planning of nutrient applications to annual crops. Soil analysis can be used to provide an indication of soil fertility before any crop is planted and hence support the planning of the season’s management tactics as well as supporting long-term monitoring of soil fertility and the effectiveness of soil fertility management strategies. Which soil properties should be analysed to give a good indication of soil fertility? Put simply, those which indicate characteristics that need to be understood or modified to achieve optimum yield, namely rooting conditions, water supply, nutrient supply and risk from soil-borne pathogens. However, the identification of simple robust measures that correlate closely with nutrient uptake and crop yield for all climates, crops and soil types has proved to be problematic. In part, this is a result of the complex interaction of limiting factors and constraints to crop growth in the field (Figure 3.1).

Scientific selection of indicator measurements for soil fertility requires that they can be measured easily and economically, as well as accurately and precisely, across a wide range of soil types and soil conditions and that they are based on reliable and contemporary...
Soil Conditions and Plant Growth

The appropriateness of the indicators chosen for soil fertility assessment depends on how well they are understood with respect to the agricultural system of which they are part and, consequently, how well they can be used to guide on-farm management decisions (Lal, 2001). To date, crop responses to soil nutrient contents (and additions) and soil acidity have generally been studied more intensively than crop responses to physical condition, levels of soil organic matter and other biological indicators.

As described in Chapter 2, the uptake of nutrients by a plant is a result of both the root absorbing power and the nutrient concentration in the soil solution; both plant and soil factors are important. The key soil factors affecting nutrient cycling in soil and supply to plants are discussed in detail later (for N see Chapter 5). However, soil solution concentrations can be difficult to quantify accurately due to very small pool sizes. In addition, determination of the concentration of nutrients contained within the soil solution on one occasion has often proved to be very weakly predictive of plant growth across the season (e.g. Murphy et al., 2009). This is because soil solution concentrations are dynamic, changing from day to day, and because the buffering capacity of the soil solid phases in relation to each nutrient is critically important in determining the rate at which nutrients in the soil solution can be replenished. For example, snapshot measurements of soil mineral N (NO$_3$⁻ and NH$_4$⁺ in solution and NH$_4$⁺ held on soil surfaces) can only give a partial prediction of soil N supply for crop growth, since the bulk of soil N is found in organic forms; as part of a long-term monitoring scheme, the average amount of mineral N measured in topsoil during autumn in the UK was only 76 kg N ha$^{-1}$ compared to 7000 kg N ha$^{-1}$ in organic matter (Shepherd et al., 1996). The interlocking biological processes controlling N supply from soil organic matter (e.g. mineralisation, immobilisation, nitrification, denitrification) resulting from the interaction of soil organisms, physical, chemical and biological characteristics modified by interactions with plant roots, weather and management are still not fully understood. As such, seasonal patterns of N availability remain difficult to predict and thus include in fertiliser decision support systems.

The development of methods for assessing the size of the pool of available plant nutrients for the next season of growth has usually focussed on developing an extraction procedure which removes a ‘labile’ pool and which integrates a measure of pool size and rate of supply. However, no single method can simulate these complex dynamic processes and so measurements of ‘available soil nutrients’ are usually selected based on correlation with plant growth in otherwise non-limiting conditions. In most countries one or two tests are favoured; in many cases, the tests which are used are as much a matter of the history of their development as of rigorous comparative scientific evaluation. For example, the determination of K availability is usually directly linked to the determination of exchangeable K; a number of extractants are considered to be equally effective in displacing K from the exchange complex (Munson, 1985; Syers, 1998). In contrast, a large number of tests to assess available soil P are in use around the world. The extractants used vary in their strength and mode of action (Sibbesen, 1983). No single test is suitable for all soil types and few routinely used tests give information about the availability of organic forms of P in soil. Because of the multiplicity of methods, there is increasing interest in the development of on-the-go analyses using a range of sensing technologies. Several such approaches are now at proof-of-concept stage (as reviewed by Sinfield et al., 2010) to provide information on soil physical, chemical and biological properties rapidly in the field to support immediate, spatially variable applications of lime and/or fertiliser. Spatial patterns of some soil properties can be determined by mid-infrared sensors (Janik et al.,
Soil fertility

3.4.4 Interpretation of the results of soil analysis

The definition of justifiable target ranges or threshold values for soil nutrients and other properties relevant to soil fertility has also become a contentious issue in the development of interpretive frameworks to support soil quality monitoring (Arshad and Martin, 2002; Sparling et al., 2003). Most of the detailed work carried out on relationships between soil fertility and crop growth has focussed on the relationship between yield and nutrient supply. There are data for relationships with other chemical properties e.g., acidity, and data are increasingly available for relationships between yield and physical properties (see Chapter 8). However, relatively few data are available to define suitability limits for soil biological properties for different crops, with the exception of some soil-borne diseases. Empirical relationships between yield and constraints usually show a curve of diminishing returns, so that as the constraint reaches non-limiting status the response to change decreases. Hence, absolute targets or thresholds are difficult to determine objectively.

Figure 3.7 Spatial maps of (a) clay and silt content, (b) cation exchange capacity and (c) potentially mineralisable N (where the colours show quartiles ranges as an illustration) over a 10 ha area (200 × 500 m; 180 sampling points on a 25 × 25 m grid) where the properties were determined using mid-infrared sensors. The spatial patterns were very strongly correlated with the same properties determined by traditional laboratory-based methods ($r^2 = 0.91$, 0.91 and 0.82, respectively). Unpublished data from the central wheatbelt, Western Australia, provided by D. V. Murphy. For a colour version of this figure, please see Plate 3.3.

1998; Figure 3.7a–c). This technology has the potential to be incorporated into rapid remote sensors that could be tractor mounted.
Soil Conditions and Plant Growth

Sufficiency indices/non-limiting thresholds

Based on empirical observations of plant growth and yield responses to added nutrients, a common interpretation approach has been to categorise the range of nutrient availability into regions of low, adequate and high/excessive (Figure 3.8a). In the UK, this approach was used to develop the ‘Index System’ guidelines which provided an interpretation of soil test values with data from crop response experiments by Crowther and Yates (1941), and in the USA a similar approach was taken by Bray (1944, 1945, 1948). The use of relative yields in the analysis allows a better examination of the underlying relationship between soil supply and crop responses, even under different site and weather conditions (Evans, 1987). A range of curve-fitting approaches can then be used to support the interpretation of the data (Figure 3.8b). Because the Mitscherlich curve is asymptotic, the soil test value at a relative yield of either 0.99 or 0.95 is most often used to indicate the point at which there is no constraint to yield resulting from that nutrient under the typical management system in the field trials. The form of the response curve means that relative yields change slowly across

Figure 3.8 (a) Response of grain yield (t ha⁻¹) in spring barley to levels of available P (measured by bicarbonate extraction, Olsen) in a 5 year series of field experiments relating yields of spring barley to Olsen extractable P in the Hoo Barley Experiment at Rothamsted Research (silty clay loam soil) showing the critical ranges of soil P used to guide fertiliser recommendations, where Index 2 is considered the target range to provide sufficient soil P. (b) Response of grain yield for spring barley expressed as relative yield to normalise across years with available soil P. Response curves fitted to interpret the results and determine the optimal level of soil P using quadratic equations either with a constrained intercept = 0 (solid line) or a free intercept (dotted line) and with two straight lines with a free intercept (dotted line); these lines intersect at a relative yield of 0.898. In this case, the Olsen extract does not extract all plant available P; hence, it is not appropriate to constrain the responses to pass through the origin. (c) Soil available P values at a range of operational yield levels derived from the curve and straight line fits – operational yields of either 0.95 or 0.99 are commonly used to determine the optimum. Even small differences in the recommended level of available soil P can have large financial implications for the farmer; e.g., Defra (2010) indicates that increasing soil P by 10 mg kg⁻¹ may need additional fertiliser applications totalling 850 kg ha⁻¹ of triple superphosphate over a decade.

Sufficiency indices/non-limiting thresholds

Fitted relationship
(without constrained intercept)
used from figure 3.8b

Available soil P (mg kg⁻¹)

Operational relative yield 0.85 0.90 0.95 0.99

Curve 11.9 15.8 20.4 24.6

Two line fit (lines intersect at a relative yield = 0.898)

9.8 11.6 24.4 34.7
a wide range of soil test values (Figure 3.8c), and interpretation of the data to give sufficiency levels to inform management depends on the underlying assumptions used. When errors associated with individual field experiments are taken into account, there is often no significant difference between fitting two intersecting straight lines to the response graph rather than a curve (Boyd et al., 1976) as long as experiments are grouped by climate/soil/crop/management. The most appropriate interpretation approach also depends on the scientific understanding underpinning the development of the extraction methodology in the context of the soil-plant system; e.g., is the soil test effective in extracting all the available nutrient or does it extract only one part of plant available nutrient (hence, should the response curve be constrained to pass through the origin)? Hence decisions taken, when sufficiency ranges and thresholds are defined, are often more subjective than is realised. In the UK, expert groups (for P in 1962 and for K in 1963) were brought together to determine the most appropriate extraction methods for England and Wales and to identify soil test ranges linked to the likelihood of response to added fertilisers; the same ranges are still used in the current Fertiliser Manual (Defra, 2010; shown for P in Figure 3.8a) because no better, widely applicable method or framework for interpretation has been found. However, it is now important to consider whether, given the interactions of a number of constraints in the field and the increasing cost of inputs, it is still appropriate to set threshold values for all constraints at the highest relative yield levels with information derived from single factor response experiments. The same interpretative framework (sufficiency/non-limiting thresholds) can be used for calibration of multi-factor effects of soil constraints and management approaches – e.g., by using a response surface approach to support interpretation (Morton et al., 2001; Bossa et al., 2005).

The limiting factor/sufficiency approach can be applied to more than nutrients. For example, in Australia, recent developments in DNA-based assays, such as PredictaB® (South Australian Research and Development Institute), have enabled grain producers to determine inoculum levels of a broad range of soil-borne diseases before seeding and the data used to underpin decision support services for disease control such as the use of resistant varieties or a change in cropping patterns (Heap and McKay, 2009). Inoculum levels measured before sowing are usually correlated with root damage and crop biomass reduction and so can be interpreted as a disease risk rating. In most cases, high inoculum levels are expected to be associated with yield loss, but interactions during the growing season between weather conditions, crop nutrient status and soil type may confound this relationship (Heap and McKay, 2009).

Expert group approaches continue to be used to draw together a range of stakeholders with different focuses and expectations to define relationships for a wider range of soil properties and crop production systems and interpret them in terms of limiting or sufficiency ranges (Figure 3.9). Andrews et al. (2002) show how this was achieved for indicators of soil fertility in vegetable production systems in California and Lilburne et al. (2004) for indicators of soil quality in New Zealand. Participatory approaches with farmers are particularly important where the aim is to develop a scorecard for on-farm assessment and interpretation of soil properties affecting crop productivity (Romig et al., 1995).

The sufficiency concept is the most common framework used for interpreting soil analysis for soil fertility. Its foundation is Liebig’s law of the minimum. It is intuitively reasonable to assume that through management (irrigation, fertiliser, cultivation, etc.) the soil can be modified so that no limiting factors remain and weather conditions, together with crop genetic factors, become the primary limits to yield. Nonetheless, this interpretative approach has limitations for easily leached mobile nutrients (e.g. nitrate and sulphate) and in soils with very low cation exchange capacities that cannot retain sufficiently large
Soil Conditions and Plant Growth

pools of cations such as K or magnesium (Mg). Interpretation also needs to take account of different crops; e.g., it is well known that cereals show yield limitation at a much lower level of soil test P than potatoes. Any sufficiency levels reported are inferred and not absolute and are consequently dependent both on the amount and quality of data available for analysis within key soil/crop/climate/management combinations and the approaches taken in data analysis (Figure 3.8). Changes in crop varieties and management practices may also mean that re-interpretation of sufficiency levels is required: e.g., a decline in the residual soil test P measured in row crops in Alabama was not accompanied by a decrease in yields; instead yields were found to have increased despite low soil P levels (Evans, 1987).

Figure 3.9 Example of the approaches used to develop a soil quality indicator package, highlighting the role of expert groups. These groups drew together a wide range of stakeholders with relevant information and experience including literature review and practical experience. Through facilitated discussion, the groups derived the relationships between yield response and a range of soil properties to inform the interpretation of the indicators used within the soil quality monitoring framework (www.soilquality.org.au) which focuses on soil quality for crop production. Data presentation using the derived categories, descriptors and traffic lights shown for three soil indicators as examples. For a colour version of this figure, please see Plate 3.4.
Wortmann et al. (2009) showed that the critical level of P for maize was affected by the previous crop with a critical level of 10 mg kg⁻¹ (Bray extract) where maize followed soybean and 20 mg kg⁻¹ where maize followed maize; they recommended that soil organic matter levels and pH be used together with soil test P values to improve soil fertility management decisions.

**Base cation saturation ratios**

Interpretation of soil tests on the basis of cation saturation is a common parallel approach to the use of sufficiency thresholds – especially in the USA. The basic cation exchange phenomena in soil which control the availability of calcium, magnesium and potassium and the impacts that the degree of saturation of one cation can have on the plant availability of itself and other cations are discussed in Chapter 6. Kopittke and Menzies (2007) highlight the historical development of the approach (often known as Albrecht analysis/interpretation), which resulted in the recommendation by Bear et al. (1945) that for an ideal soil 65% of the exchange complex should be occupied by Ca, 10% by Mg, 5% by K and 20% by hydrogen; it was considered that where this cation balance was achieved, chemical, physical and biological properties would all be in an appropriate equilibrium to support optimum plant growth. Review of the literature by Eckert (1987) and Kopittke and Menzies (2007) suggests that fairly wide variations in actual ratios are of little consequence for crop yield, quality or a wide range of soil physical and biological properties, as long as gross imbalances are not created. Consequently, sole focus in the interpretation of soil fertility on a target Ca/Mg ratio often leads to recommendations for management that are extremely expensive and have little impact on crop productivity or profitability.

**Testing for mobile nutrients, e.g. nitrate**

Measurement of soil mineral N (i.e. nitrate + ammonium, but often only nitrate because amounts of ammonium present are usually small) before planting or at a specific time during the crop growing season can be interpreted under certain conditions to give an indication of likely soil N supply. In climates without excessive overwinter leaching, the use of a pre-planting test for soil nitrate is widely used; in the western USA before 1982, 16 states were using such a test (Keeney, 1982). Values obtained for nitrate before planting were compared to the estimated crop N demand to obtain the likelihood of response to additional fertiliser. A pre-sidedress nitrate test has been developed for maize (Magdoff et al., 1984), which allows a period of nitrate accumulation under field conditions before sampling takes place, and therefore also reflects the mineralisation potential under field conditions. Similar approaches are used in many countries in Europe and these were summarised by Mengel (1991). Measurement of soil mineral N can be combined directly with analysis of N already in the crop (i.e. that already supplied by the soil) to predict seasonal soil nitrogen supply and thus guide fertiliser requirements (Defra, 2010). However, research has shown that this single ‘snapshot’ in time is often very unrepresentative of soil nitrogen supply (Knight et al., 2008; Murphy et al., 2009). The UK Fertiliser Manual (Defra, 2010) consequently recommends that sampling and analysis of soil mineral N is used to give guidance ‘... where the supply of plant available nitrogen in the soil could be unusually large, and particularly where organic manures have been used regularly in recent years’, not as a precise prediction tool.
3.5 Managing soil fertility

Where yield is constrained by soil factors, perhaps diagnosed by plant or soil analysis, the next step is to determine what action might be taken by the land manager to adjust their rotation to optimise productivity/profit given the constraint and/or to implement management strategies that mitigate the constraint. Management strategies might seek to improve performance over the medium to long term (e.g. to increase effective rooting depth by improving drainage, removal of pans by cultivation and use of deep-rooted crops within the rotation) or seek an immediate impact (though to improve soil depth), ‘shattering the underlying rock or hard layers with explosives [is] not in general use’ (Russell, 1937). Local knowledge and farmer understanding of soil fertility, as indicated by crop performance, weed species and diversity and observable soil properties have been used to select sites which are most likely to respond to improved soil management practices (Suarez et al., 2001; Barrios and Trejo, 2003; Mairura et al., 2008). Farmers’ understanding of their own soils is strongly orientated towards easily observed characteristics that affect production, such as texture (heavy or light land for cultivation), waterlogging and drainage after heavy rain, nutrient retention (hungry soils) and the structural condition of the soil (tilth, heart). Such variations are often the reasons why field boundaries occur where they do, and the fields may have indicative names, e.g. Sweetlea, Hungry field, Starvall. Integration of local knowledge with technical information can provide a foundation from which to determine key local constraints or to identify which potential management practices could be used at each site (Barrios et al., 2006). Careful targeting of proposed changes in soil management approaches to key crops and the selection of rotational systems to match geographic, market and social niches is needed to maximise effectiveness of any recommended interventions and to increase uptake by farmers (Ajayi et al., 2007).

3.5.1 Development of recommendations for targeted inputs

Most of the research supporting soil fertility management has focussed on fertiliser and lime recommendations, as both of these practices seek to provide an economic return to farmers over the timescale of a crop rotation, if not more quickly. ‘If the soil test falls below the sufficiency level, how much fertiliser should be applied?’ is the primary question, but there are a wide range of further questions needed to support management decisions to inform choices about the type, timing and placement of fertiliser in practice. Fertiliser type often affects the availability of the nutrients added; so, e.g., there is a wide range in concentration and availability of P from diammonium phosphate, triple superphosphate, farmyard manure and phosphate rock (Table 3.4). Soil factors, not least pH and moisture content at application, also affect the availability of nutrients from fertiliser additions.

Often the same or a similar series of yield response experiments are used to determine optimum fertiliser application rates post hoc, as described in Section 3.2.1 for the determination of suitability ranges (Figure 3.8). However, for the development of fertiliser recommendations, economic considerations are central, with the determination of the optimum fertiliser application from a response curve determined in part by the ratio between the price of a unit of fertiliser nutrient and the saleable value of the crop; hence, actual rather than relative yields are used. Where environmental impacts are also an important consideration, the risk of nutrient loss (e.g. by leaching or denitrification) may also be taken into account alongside economic considerations. A wide range of response models can be
Soil fertility

**Table 3.4** Typical composition of some common fertilising materials; all organic fertilisers are very variable and consequently these data provide only an indication of likely composition.

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>P</th>
<th>K</th>
<th>Other notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ammonium sulphate</td>
<td>21</td>
<td></td>
<td></td>
<td>24% S</td>
</tr>
<tr>
<td>Ammonium nitrate</td>
<td>32–34.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urea</td>
<td>46</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diammonium phosphate</td>
<td>21</td>
<td>23</td>
<td></td>
<td>(soluble)</td>
</tr>
<tr>
<td>Triple superphosphate</td>
<td>20</td>
<td></td>
<td></td>
<td>(soluble)</td>
</tr>
<tr>
<td>Rock phosphate</td>
<td>12–16</td>
<td>14%</td>
<td>31–35%</td>
<td>Ca, (slowly</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>soluble)</td>
</tr>
</tbody>
</table>

|                         |       | 50    | 42    | 17% S             |
| Potassium chloride      |       |       |       |                   |
|                         |       | 0.6   | 0.08  | 0.6 (soluble)     |
|                         |       | 0.4   | 0.04  | (soluble)         |
|                         |       | 3%    | 0.65  | 0.6 (soluble)     |
|                         |       | 0.75% | 0.06  | (soluble)         |
|                         |       |       |       |                   |
|                         |       |       |       |                   |
| Cattle farmyard manure  | 0.6; 10% readily available | 0.08 (soluble) | 0.6 (soluble) | 25% DM 0.1% S |
| Pig slurry              | 0.4%, 70% readily available | 0.04 (soluble) | 0.2 (soluble) | 4% DM, 0.04% S |
| Poultry (broiler) litter| 3%, 33% readily available  | 0.65 (soluble) | 1.3 (soluble) | 60% DM, 0.3% S |
| Greenwaste compost      | 0.75%, 3% readily available | 0.06 (soluble) | 0.4 (soluble) | 60% DM, 0.05% S |

**Source:** Data adapted from Defra (2010).

used but each is likely to provide a different estimate of the optimum fertiliser rates, and it is most important that a model is chosen so that it is appropriate to the data available. A good response model should (Cooke, 1982):

1. define both the slope of the rising part of the curve and the slope beyond the peak;
2. define accurately the yield near the optimum dressing and the dressing itself.

Actual yields and profitability are likely to have been affected by a wide range of factors other than those varied in the experiment, especially if the data are collated from a series of experiments carried out at different sites and in different years. Consequently, response curves are only valid as the basis of advice to farmers under the same conditions as used in the experiments. Although it is possible to compile tables of average responses of crops to fertilisers at various soil sufficiency levels, it is not clear that these provide anything other than very general guidelines to farmers.

For less mobile nutrients (e.g. P and K), fertiliser recommendation systems now commonly use the soil sufficiency level/range as a target value and develop a fertiliser policy which seeks to achieve/maintain this value (on a rotational basis), i.e. fertilising the soil. This approach often recommends applications of small amounts (maintenance dressings) of fertiliser, even where soil test values are above the sufficiency level, to replace the amounts likely to be removed by the crop to be grown. Olson et al. (1982) reported on a long-term comparison study of fertiliser recommendations based purely on the sufficiency approach (applications only where fertiliser response was likely) and those made by a number of commercial laboratories using a build/maintenance approach. No significant difference in crop yields were recorded, but fertiliser costs where a
maintenance approach was used were on average twice those where fertiliser was applied on a sufficiency basis. Recommendations in the UK state that ‘Once deficiency has occurred, a fresh application of P and K is most unlikely to be available for uptake by roots in time to benefit the crop being grown’ (Defra, 2010). In addition, even in the low range below the soil sufficiency index, it has been observed that up to 40% of fields may not respond profitably to fertiliser application (Havlin et al., 1999), in part due to nutrient supply from pools not accounted for in the soil test, or available nutrients below the sampling depth (Davis et al., 1996), as well as the impact of other limiting factors in the year of observation. In farming systems with relatively stable crop rotations and production systems, it should be possible to couple plant and soil analysis for P and K to provide indicators of the effectiveness of management practices, rather than using such data to derive fertiliser recommendation per se. This type of farm-specific monitoring should allow modification of management underpinned by an understanding of the principles governing P and K availability (Chapters 5 and 6) rather than the application of average recommendations.

For mobile nutrients, such as N, fertiliser recommendations are usually based on yield potential; these are most often expressed as a farm/field-specific yield goal. This yield goal is used to determine the likely crop N demand in the season. A range of information can be used to guide the farmer’s yield goal, including previous yields achieved at the site. The requirement to predict the likely weather, disease risk and other factors potentially impacting on crop growth is explicitly placed with the farmer or their advisor. In dryland agriculture, it is usually considered advantageous to set the grain yield goal above that of average yields to fully take advantage of the potential for above-average growing conditions, e.g. above-average rainfall (Johnson, 1991). However, Hergert (1987) highlighted the need to improve yield goal determination for fertiliser recommendations as 50% of farmers attained yields equivalent to only 80% of their yield goal. For N fertiliser, if the yield goal is too high, fertiliser recommendations will overestimate requirements, farmers will waste money and there is an increased risk of N loss to the environment. Development of new technological approaches that can be used during the season in the field (e.g. canopy reflectance, Raun et al., 2001) may mean that it becomes increasingly possible to predict grain yield in-season and consequently adjust the fertiliser recommendations based on projected N demand.

Fertiliser N requirements are calculated from the crop requirement by adjusting for the soil N supply corrected for in-season losses:

\[(\text{Amount of N Fertiliser Needed}) = (\text{Crop N Requirement}) - (\text{Soil N Supply}) + (\text{N Losses})\]

and then adjusting for any nutrients to be supplied from manures, biosolids, etc. Improved fertiliser recommendations therefore depend in part on the ability to estimate N mineralisation and N losses (Mulvaney et al., 2001; Zebbarth et al., 2009), and also amounts and rates of nutrients supplied from manures and other organic materials, e.g. animal manures, urban green waste composts, etc. Cooke (1982) outlined how the range of factors affecting crop response to nitrogen fertiliser had been integrated into an N index system. This used information on previous cropping and manuring to predict the likely response of sites to nitrogen fertiliser, i.e. providing an adjustment for soil N supply based on data collated from many years of field experiments (e.g. MAFF, 1996).

The development of models of soil nutrient dynamics (Shaffer et al., 2001) may allow these to be integrated into decision support systems for fertiliser recommendations. This may be: (1) on a static basis, so that the model is run once to provide information on predicted
3.5.2 Developing recommendations for integrated soil fertility management

There is an increasing use of broader integrated approaches to the assessment of soil fertility which consider the impact on plant growth of the emergent properties and processes arising from ecosystem-level interactions of plants/roots with physical, chemical and biological factors in the soil. It is therefore clear that soil fertility management cannot be restricted to advice on the seasonal management of nutrient, water inputs or even the use of green manures. Soil and water conservation measures, crop breeding and rotation design, tillage, drainage and pest management strategies together with farm landscape design (presence, composition and location of shelter belts, buffer strips, hedges) may all form important components in the integrated management of the below-ground ecosystem (Figure 3.11). In the same way that soil fertility is one aspect of a more integrated definition of soil quality (Section 3.2.3), integrated soil fertility management can be considered as one aspect of broader strategies for integrated natural resource management (Vanlauwe et al., 2006).

Production practices such as the integration of reduced tillage, retention of crop residues, targeted planting and farm input use and the use of legumes in rotation (conservation farming approaches), legume agroforestry (fertiliser tree) systems, the integration of green manure and dual purpose legumes and the careful use of animal manures (Sanchez and Salinas, 1981; Palm et al., 2001) are common where more integrated approaches to soil management are used. All combine inputs of organic matter with other changes to management systems. Because of the role of decomposition in providing energy to the soil ecosystem, regular inputs of organic matter can change the biomass, activity and diversity of soil organisms driving changes to a wide range of soil physical and chemical properties. Hence, responses to additions of organic materials to cropping systems are complex and interactive and rarely can be explained by a simple causal chain. In addition, there is increasing interest in the use of soil inoculants or plant sprays developed as plant growth stimulants or biopesticides (see Section 3.3.2). Rupela et al. (2006) reported findings of an experiment (established in 1999), which investigated the role of plant biomass (rice straw or farm waste combined with hedge clippings) as an input, together with applications of rock phosphate, biopesticides and reduced tillage in a rainfed cropping system on a Vertisol in the semi-arid topics (India). After a first year in which low-cost input/biologically based systems had lower yields (attributed to the lag phase in release of N and P from complex organic inputs), crop yields were close to those in optimally managed conventional plots, and net income to farmers was significantly higher due to the lower cost of inputs. N and P balances were positive and consequently there was no nutrient depletion; soil organic matter, microbial biomass C and
Soil Conditions and Plant Growth

N levels were increased in the treatments receiving additional C inputs. Such integrated system experiments are of considerable value in demonstrating how the principles underpinning integrated soil fertility management can be implemented in practice within inherent site constraints. Best practice may well be region and even site specific, e.g. the local availability of the biomass and biopesticides in the earlier example. However, such system management experiments are often the despair of scientists seeking to understand the

<table>
<thead>
<tr>
<th>Method</th>
<th>Optimum*</th>
<th>RB209 Default weather and yield</th>
<th>SUNDIAL Actual weather and yield</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recommendation</td>
<td>139</td>
<td>182</td>
<td>162</td>
</tr>
<tr>
<td>Calculated yield*</td>
<td>9.1</td>
<td>9.5</td>
<td>9.3</td>
</tr>
<tr>
<td>%difference from optimum yield</td>
<td>4</td>
<td>2</td>
<td>0</td>
</tr>
</tbody>
</table>

*Calculated from linear plus exponential curve fitted to trial data.

Figure 3.10 Evaluation of a dynamic model (SUNDIAL) for fertiliser recommendation (Smith et al. 1996). In 1997, a project was initiated to establish a series of N response trials on working farms across the UK and to evaluate the performance of model recommendations across a range of typical crops and soils. Data are presented from one of these sites as an example. The site was a sandy clay loam soil in Suffolk, Eastern England. Winter oilseed rape (cv. Navajo) was harvested on 25 July 1998 and yielded 4.3 t ha⁻¹ (＠90% DM). Crop residues were chopped and incorporated and 40 t ha⁻¹ of farmyard manure taken from the piggery was applied just before ploughing. Winter wheat (cv. Equinox) was drilled on 3 October 1998. The farmer was expecting a yield of 11–12 t ha⁻¹. The N response curve obtained in the field (three replicate plots of 6 N rates) was fairly flat and the replicate plots quite variable due to the N supplied by the pig manure. The response curve was fitted by a linear plus exponential relationship (Genstat) with an optimum estimated at 139 with a standard error of 17 kg N ha⁻¹. The actual yield obtained in the field was much lower than the farmer’s target only reaching about 9.5 t ha⁻¹. The standard UK reference book Fertiliser Recommendations for Agricultural and Horticultural Crops (RB209; MAFF, 1996) recommended an application of 182 kg N ha⁻¹, while predictive use of the SUNDIAL model recommended 162 kg N ha⁻¹. The SUNDIAL recommendation was not significantly different from the measured optimum, while the RB209 recommendation was a significant overestimate. However, this over-application of N had little effect on the yield achieved, but may have increased the residual mineral N available for leaching after harvest. In plots that had received 250 kg N ha⁻¹, mineral N was increased very significantly above those plots which had received N applications close to the optimum. When SUNDIAL was used retrospectively, i.e. with the actual weather and the target yield set at the actual optimum achieved, the SUNDIAL recommendation almost exactly predicted the measured optimum. It is important that SUNDIAL achieves good retrospective predictions as this demonstrates that the model can be used reliably to carry over information from one year to the next without introducing significant errors.
ecological interactions underpinning soil fertility and their links to individual component management practices. In general, in such systems, scientists have been very effective in cataloguing the impact of management practice changes on the root/soil system (chemical, physical and biological properties, e.g. Mäder et al. (2002) and Cookson et al. (2008)) but much less effective in developing management systems that achieve targeted changes.
3.5.3 Challenges for soil fertility management in the twenty-first century

As already noted, climate change coupled with increased costs and declines in the availability of many of the common inputs will necessitate changes to soil fertility management in the twenty-first century. Reduced tillage, legumes and other N-fixing plants as intercrops, green manures or as hedgerow trees and the selection of crops and varieties to promote rhizosphere-mediated nutrient availability have already been integrated into so-called ‘sustainable’ agricultural systems. It is likely that these and similar approaches will become a more common part of conventional agricultural systems as the twenty-first century progresses. These changes will require soil scientists and agronomists in the developed world to learn from their counterparts working in developing countries where farmer participatory research has been successful in promoting improved soil fertility, crop yield and farmer livelihoods. The challenge for all agricultural systems is to achieve sustainable intensification; this will require increased focus on the use of biological processes and potentials within the soil-plant system.

Simple single-factor response experiments (or even multifactorials), as commonly used for fertiliser trials, have little relevance in the evaluation of integrated soil fertility management approaches. Each input (e.g. compost, inoculant or new crop variety) or management practice (e.g. change in depth of tillage, incorporation of crop residues or green manures) has interacting impacts on multiple organisms and on the soil physical and chemical properties defining their habitats. Consequently, though there is a myriad of data cataloguing the impact of many of these practices on soil physical, chemical and biological practices, there is little systematic evaluation of these integrated management systems on the maintenance of soil fertility, as broadly defined, except on a case-by-case basis. Understanding how soil-plant systems respond to management and climate will require a more ecologically based approach to experimental design and analysis. Furthermore, experiments will need to be maintained for sufficient years to enable the resistance and resilience of systems to be determined and the underpinning soil–plant interactions to establish.

Integrated soil fertility management approaches often have at their heart the incorporation of organic inputs to the soil-plant system. However, if these systems are to become more widely adopted, consideration is needed of how the organic matter inputs required can be sustained over decades. The plant is the conduit bringing energy into the soil ecosystem; hence in natural ecosystems, net primary production is often strongly linked to the biomass and activity of the soil organisms. In many environments, effective capture of solar radiation is limited by water availability; in ‘conventional’ agricultural systems, plant breeding has focussed on increasing the biomass of the marketable component (e.g. grain, fruit) at the expense of crop residues, hence limiting or reducing inputs. A combination of rotational planning (exploiting the benefits of plant diversity in space and time) and plant breeding approaches (both to manipulate rhizosphere inputs, to increase amounts or manipulate quality of residues) will be needed to overcome this potential constraint for integrated fertility management approaches in the future.

The specialisation of soil science researchers has often led to their dissociation with the practical implementation of soil fertility; one might see the application of molecular methods to the analysis of soil microbial function as a particular example of this; function must be linked to process. The soil-plant ecosystem functions as a result of the complex interactions between soil organisms and their habitats, so the effective development of soil fertility
management in the twenty-first century should be underpinned by collaborative interdisciplinary investigations emerging from dialogue amongst researchers, extension specialists and farmers/growers. Integrated soil fertility management should be at the heart of a global approach to improving the sustainability of agriculture, its impact on food security and the environmental impact of the industry. This will require the development of new plant genotypes, management systems, indicators and recommendations.

References


4 Soil organic matter

David Powlson¹, Pete Smith², and Maria De Nobili³

¹Department of Sustainable Soils and Grassland Systems, Rothamsted Research, Harpenden, Hertfordshire, UK
²School of Biological Sciences, University of Aberdeen, Aberdeen, UK
³Dipartimento di Scienze Agrarie e Ambientali, Università degli studi di Udine, Udine, Italy

4.1 Introduction

It is the presence of organic matter that makes the difference between a collection of mineral particles and a functioning soil with a recognisable physical structure that provides a suitable medium for plant growth. Soils often contain only a small quantity of organic matter (typically a few per cent by mass in agricultural soils), yet this has a major influence on soil properties. For example, organic matter is a key factor determining the interactions between mineral particles and thus the formation of aggregates of varying size and stability and the distribution of spaces (pores) between the solid particles which may be filled with either air or water. This physical structure, relying in large part on the quantity and type of organic matter, determines whether or not a soil retains water, whether it becomes flooded or is well drained, whether it becomes easily compacted or retains a structure conducive to root growth. Organic matter also provides substrate for soil organisms and, often, anchorage points for bacteria or fungi. Interactions between organic matter, organisms and mineral particles form the basis for nutrient transformations on which all plant growth depends. Organic matter is a reservoir of many plant nutrients (such as N, P and S) and also provides a substantial part of the soil’s cation exchange capacity (CEC), which is important for the retention of many nutrients. It is also the repository of organic carbon (C) in soil which, on a global scale, is the largest terrestrial stock of carbon after carbonate rocks. Thus, the influence of management practices on the quantity of organic C in soil is vitally important both for the maintenance of soil functions, and its influence on plant growth, and for its impact on global carbon cycling and climate change.

Even a small quantity of organic matter can have a profound influence on soil properties. Soil in arable cropping in a temperate climatic region may typically contain 1–5% organic carbon, yet small variations within this range have a large effect on soil properties.
4.1.1 Terminology – organic matter or organic carbon?

It is customary to refer to the total amount of organic matter (OM) in soil, in all forms, as soil organic matter (SOM). But chemical analysis normally gives a value for C content. It is common to use a conversion factor of 1.724 to convert a measured organic C value to organic matter, implying that organic matter contains 58% C, though this is based on work conducted in the nineteenth century. Values up to 2.0 for surface soils and 2.5 for subsoils have been published (Nelson and Sommers, 1982). A more recent review of data (Pribyl, 2010) shows that the value is variable but an average conversion factor of 2 (i.e. 50% C in organic matter) is more accurate in almost all cases. However, it is now strongly recommended not to quote OM values but use organic C, usually termed soil organic carbon (SOC).

When soil samples are analysed for organic C content, the result is conventionally expressed in units such as %, mg C kg\(^{-1}\) soil or μg C g\(^{-1}\) soil. These are units of concentration, though they are often incorrectly referred to as C content. To express C as a quantity or content rather than concentration, the concentration is multiplied by the mass of soil to a given depth, often determined from measurement of bulk density. SOC content can then be expressed in units such as Mg C ha\(^{-1}\) or g C m\(^{-2}\) to the defined depth. When expressed in such units, the quantity of C in 1 ha of soil can correctly be termed a C content or stock. Alternatively, the value may be multiplied by an area to give a C stock in an area of interest such as Europe, often in units of Tg or Pg. A term sometimes used in such studies is ‘C density’, which is synonymous with ‘C stock per unit area’. Figure 4.1 shows an example of a C density map for the Amazonian region of Brazil (Cerri et al., 2007). In this region there is at least a fivefold range between areas with the lowest and highest C density. From the viewpoint of maintaining SOC stocks (as discussed later in relation to climate change mitigation), such maps can be useful as a tool for policy makers in determining appropriate land management practices for each area. For areas with a large C density, there is a priority to maintain them in their current state in order to avoid C loss and thus prevent the large C stock becoming a source of CO\(_2\). For areas with a small C density, there may be potential to accumulate SOC through altered management or land use, thus creating a sink and helping to mitigate climate change. However, in many cases, a small C density reflects a small potential for SOC accumulation, either because of soil type (e.g. sandy soils which have less capacity to stabilise C than soils with a greater clay content) or through limited plant growth resulting from climatic factors such as low rainfall.

4.1.2 Analysis of SOC

It is now most common to analyse soil for C by high temperature combustion of a sample in the presence of oxygen; various commercial instruments are available for this, using a small finely ground soil sample. All C in the sample is converted to CO\(_2\) and its quantity measured by a gas chromatographic detector, infrared analysis or (less commonly) gravimetrically. If the soil contains carbonate, the C contained in this will also be converted to CO\(_2\), so the quantity derived is total soil C. To determine organic C, a separate analysis for carbonate C is made and this quantity subtracted from the total C value from combustion. Carbonate C is usually determined by treatment of a soil sample with dilute acid and measurement of the volume or mass of CO\(_2\) evolved.

Before the advent of combustion instruments, it was usual to determine organic C by chemical oxidation by heating a soil sample with a mixture of potassium dichromate, sulphuric acid and, in some methods, phosphoric acid. Either the evolved CO\(_2\) was determined gravimetrically or the amount of unused dichromate oxidising agent determined,
Soil Conditions and Plant Growth

often by titration against ferrous ammonium sulphate which was oxidised. Many variations on this principle have been published and used; see Nelson and Sommers (1982) and Kalembasa and Jenkinson (1973) for details of methods and discussion of their merits and limitations. One variation, known as the Walkley–Black method, relies on the heat naturally generated from the reaction between organic matter and the oxidising agent. In well-equipped laboratories, this method would rarely be used now, but in less well-endowed laboratories it is still in use and one sees reference to it in the literature.

An approximate method of analysis, suitable for soils with a high organic C content and where very large numbers of samples require analysis, is loss on ignition. In this method a sample is burned and the loss of mass determined. The sample is either dried first (so that water does not contribute to the mass loss) or a correction is made for the small water content of the soil which is normally air-dried prior to analysis. The assumption is that the loss of weight is due entirely to the destruction of organic matter, with the mineral components remaining.

4.2 Factors influencing SOM content

4.2.1 Inputs and outputs of organic carbon

The quantity of organic matter (or organic C) in soil is a balance between inputs and outputs (Figure 4.2). In agricultural soils, the main organic C inputs are plant roots and associated material (sloughed off roots cells, root exudates) and above ground plant parts (stems, fallen leaves, any other non-harvested plant parts) plus any manures or similar materials applied. Minor inputs might include organic matter in dust deposited at a site. The main output of C

Figure 4.1  Map of soil organic carbon stocks [carbon density] in the Brazilian Amazon. Reprinted from Cerri et al. (2007). With kind permission from Elsevier. For a colour version of this figure, please see Plate 4.1.
is decomposition by the soil biological population: under aerobic conditions, C from decomposition is evolved to the atmosphere as CO₂ but under anaerobic conditions, such as in flooded soils growing paddy rice or natural or man-made wetlands, it is methane, CH₄.

The rate and extent of decomposition is determined by environmental factors, especially temperature and soil moisture content. Decomposition will almost always be faster at higher temperature than lower assuming sufficient water is available for microbial activity. Very dry or very wet conditions will normally slow decomposition. Because of the generally increased rate of decomposition under warmer conditions, it is usual for soils in tropical regions to have a lower SOC content than those in cooler climates but under comparable land use.

### 4.2.2 Impact of management practices

Management practices, through their impacts on C inputs and decomposition rates, have a major impact on SOC content but changes often occur slowly – though faster under warmer climates. In this section several examples are presented to illustrate some of the main impacts and trends observed.

Figure 4.3 illustrates several aspects of management impacts on the SOC content of an agricultural soil in a temperate region. The data show SOC content in several treatments of the long-term Broadbalk Experiment at Rothamsted Research, UK, and also a simulation of the data using the RothC model (see Section 4.3). By the year 2005, the SOC content in the treatment given a high rate of farmyard manure (FYM) every year since the start of the experiment in 1843 was almost three times that in treatments given neither manure nor inorganic fertilisers. But the increase occurred slowly and the tendency of SOC to move asymptotically towards a new equilibrium value is clear. The annual rate of SOC increase was much greater in the early years: output from the RothC simulation shows that the average annual increase during the first 20 years was 1.0 t C ha⁻¹ but during the last 20 years was only 0.2 t C ha⁻¹. Consequently about half of the final observed increase occurred within the first 30–40 years of manure application. During the period 1926–1966, all plots on Broadbalk were bare-fallowed one year in five, to control weeds, and no manure was applied in the fallow years. Thus, organic C inputs were reduced in these years due to absence of manure and absence of inputs from crop roots and stubble. The small decrease in SOC during the fallow years gives rise to the ‘saw-tooth’ shape in the simulations. The overall effect during this 40 year period was a small decline in SOC content in all treatments but is most clearly seen in the FYM treatment. After fallowing ceased (because weeds were controlled by herbicides), the increasing trend in the FYM treatment resumed but at a slower rate as the soil tended towards a new equilibrium value. If the soil had not had fallow years during this 40 year period, the new equilibrium value would presumably have been attained earlier and the
annual rate of SOC increase in the later years would have been even less. The same trend can be seen in the later FYM treatment that started in 1885.

A surprising result in Figure 4.3 is the fact that SOC in the control treatment with no manure or fertiliser inputs did not decline but remained stable for the duration of the experiment. This is thought to be because the site had been in arable cropping for many centuries; the field is documented as arable at least 200 years before the start of the experiment and is thought to have been initially cleared of natural vegetation at least 1500 years earlier. So the high SOC content under native forest presumably declined during this long period and reached a new low equilibrium value that is maintained by the small crop inputs in the unfertilised and unmanured treatment. In the treatment receiving inorganic fertilisers, crop yields were much higher than in the control (similar to yields in the FYM treatment), so the organic inputs to soil from crop roots and stubble were also greater. This small increase in the annual input of organic C led to a small increase in SOC compared to the control treatment. This is a general trend seen in many long-term studies. Ladha et al. (2011) reviewed data on this point from 135 studies at 114 long-term sites worldwide. On average, application of inorganic N (often with P and K) led to the SOC and soil organic nitrogen (SON) contents of soils being 8% and 12% greater than in the corresponding controls. However, in many cases there was a general decline of SOC and SON over time in all treatments at the sites, and inorganic fertiliser applications did not reverse this; rather the decrease was slightly less where inorganic fertilisers were applied. These observations at so many sites contradict the claim that is sometimes made that inorganic fertilisers, especially N, cause a loss of SOM.

Figure 4.4 (redrawn from Gollany et al., 2011) shows data from a site where natural vegetation was cleared more recently than at Broadbalk, namely, the Morrow Plots in Illinois, USA. This is the oldest agricultural experiment in North America. The native prairie grassland was cleared in 1876 and a range of arable crop rotations and manure or fertiliser treatments established, though the data shown only starts in about 1905. It can be seen that
Soil organic matter

SOC declined until about 1960 in all treatments and then approximately stabilised. This decline is attributed to (a) smaller inputs of organic C as arable crops replaced perennial grasses and (b) soil disturbance due to ploughing.

SOC in the continuous maize treatment (with no manure input) declined the most, from about 22 t C ha$^{-1}$ in 1905 to 15 t C ha$^{-1}$ after 70 years. Where manure was applied to continuous maize, SOC remained higher than in the control, but still declined throughout the period shown. In the maize–oats–hay rotation (with no manure), the rate of decline in the first half of the twentieth century was slightly lower and SOC content stabilised at about 5 t C ha$^{-1}$ higher than under continuous maize. This is presumed to be due to the larger inputs of organic C in the year when hay is grown; inputs from oats are probably also greater than from maize due to the narrower row spacing. In all cases, maize stover (stubble) was removed until about 1960 but returned thereafter: this increase in C input will also have contributed to the slowing of the SOC decline that started about this time. If stover had continued to be removed, it is likely that SOC in both rotations would have stabilised at a slightly lower value. Where manure was applied to the continuous maize rotation, SOC continued to decline compared to the initial value but remained higher than in the treatment with no manure input. In this respect, the results are analogous to the effect of manure in the Broadbalk Experiments (Figure 4.3), though in the Morrow Plots the increase from manure was relative not absolute.

Figure 4.5 shows changes in SOC when tropical forest was cleared and converted to pasture (Cerri et al., 2003). The data were obtained through the use of a chronosequence rather than at specific experimental site. In this case, through historical records and interviews with land owners, it was possible to identify areas of pasture of different age. To make valid comparisons it was also necessary to establish that the soil type and management practices were very similar at all sites. Such an approach is sometimes termed a ‘space for time’ comparison. Figure 4.5
Soil Conditions and Plant Growth

shows an initial rapid decline in SOC content after clearing forest but, somewhat surprisingly, this is followed by an increase and after 90 years SOC is at a higher value than in the original forest soil. Two factors are relevant here. Under well-managed grazed pasture in a tropical climate, with perennial grasses and some legumes, inputs of organic C to the soil can be higher than under forest. It is also observed that, under tropical forests, a larger proportion of the total C stock is above ground, in the trees, than is the case for temperate forests. Thus SOC content under tropical forests is often surprisingly small. Even though the long-term result of clearing forest and establishing pasture was a small increase in SOC stock (by about 20 t C ha$^{-1}$ in this case), there was still a major loss of C from the system. It was estimated that 100–200 t C ha$^{-1}$ were lost from the trees at the time of clearing.

Figure 4.5 shows an example of two contrasting land management changes in a temperate climate at two experimental sites at Rothamsted Research, UK (Johnston et al., 2009). In one case (the Highfield Experiment), a section of an area that had been in grass for over 100 years, and contained >60 t SOC ha$^{-1}$ to a depth of 23 cm, was ploughed and converted to arable cropping. Over the next 20–30 years, its SOC content declined rapidly. Within 50 years it had lost one third of its original SOC and had reached the same content as that of a soil that had been under arable cropping for a period of well over 100 years. Figure 4.6 also shows that the section of grassland that continued under grass increased in SOC. This is because the intensity of management increased compared to that prior to the start of the experiment, including increased fertiliser N applications. At the arable site (the Fosters Experiment), the area that continued under arable cropping remained at an almost constant SOC content through the 50 years shown in Figure 4.6. The section of the arable field sown to grass and managed in the same way as the grass
area of Highfield increased in SOC from 40 to 65 t ha\(^{-1}\) within 50 years but was still below the value in the continuing grass section of Highfield.

Figure 4.7 illustrates the importance of soil texture in determining SOC content. It shows the changes in SOC content over a period of almost 100 years at a site with a sandy soil containing only 7% clay (Johnston et al., 2009). The topsoil contained about 40 t C ha\(^{-1}\) at the start of the experiment, the relatively high value probably attributable to previous periods under grass which was ploughed up a few years before the start of the experiment. But SOC decreased to less than half the initial value during 100 years of arable cropping; even the management involving a four-course rotation receiving animal manure failed to significantly slow this decline. This is because clay and silt size particles, having a very large surface area, are so central to the stabilisation of organic matter in soil. Soils dominated by sand have much less potential to retain organic matter.

### 4.3 Modelling SOM dynamics

Models of SOM dynamics are simply mathematical mechanisms for expressing our conceptual understanding of how SOM decomposes and how carbon inputs and environmental factors affect SOM transformations. As with all models, SOM models are a simplification of reality, and there are a number of ways in which decomposition can be conceptualised. The vast majority of SOM models (and there are well over 30 globally) are process-based, describing SOM dynamics through first-order rate kinetics in one or more conceptual pools of carbon. A few models have been developed, which treat organic matter decomposition, and the subsequent change in the quality of the organic matter, as a continuum, with each fresh cohort of litter being tracked separately. Another small subset of models follows the
flows of carbon and other nutrients and energy through the trophic web. These model types are described in more detail in Section 4.3.1.

4.3.1 Types of SOM model

Process-based models

Most models are process-based, i.e. they focus on the processes mediating the movement and transformations of matter or energy and usually assume first-order rate kinetics (Paustian, 1994). Early models simulated SOC as one homogeneous compartment (Jenny, 1941). A few years later, two-compartment models were proposed to reflect increased understanding (Beek and Frissel, 1973), and as computers became more accessible, multi-compartment models were developed (McGill, 1996; Molina and Smith, 1998). Of the 33 SOC models currently represented within the Global Change and Terrestrial Ecosystems (GCTE) Soil Organic Matter Network (SOMNET) database (Smith et al., 2001), 30 are multi-compartment, process-based models. Each compartment or SOC pool within a model is characterised by its position in the model’s structure and its decay rate. Decay rates are usually expressed by first-order kinetics with respect to the concentration \( C \) of the pool: \( \frac{dC}{dt} = -kC \), where \( t \) is the time. The rate constant \( k \) of first-order kinetics is related to the time required to reduce by half the concentration of the pool when there is no input. The pool’s half-life \( h = (\ln 2)/k \), or its turnover time \( \tau = 1/k \) are sometimes used instead of \( k \) to characterise a pool’s dynamics: the lower the decay rate constant, the higher the half-life, the turnover time and the stability of the organic pool (Molina and Smith, 1998).
The flows of C within most models represent a sequence of carbon going from plant and animal debris to the microbial biomass, then to soil organic pools of increasing stability. Some models also use feedback loops to account for catabolic and anabolic processes and microbial successions. The output flow from an organic pool is usually split. It is directed to a microbial biomass pool, another organic pool, and, under aerobic conditions, to CO$_2$. This split simulates the simultaneous anabolic and catabolic activities and the growth of a microbial population feeding on any given substrate. Two parameters are required to quantify the split flow. They are often defined by a microbial (utilisation) efficiency and stabilisation (humification) factor which control the flow of decayed C to the biomass and humus pools, respectively. The sum of the efficiency and humification factors must be inferior to unity to account for the release of CO$_2$. Examples of the structure of two commonly used SOM models are shown in Figure 4.8. A recent advance for allowing multi-pool models to be tested rigorously against data comes from fractionation techniques that yield fractions that are equivalent to measurable pools, so model simulations of the size of individual pools can be tested experimentally in addition to simulation of total SOC (Smith et al., 2002). Zimmermann et al. (2007) proposed an experimental fractionation scheme that yields measurable fractions that approximate to the conceptual pools in the RothC model, providing new opportunities for model initialisation and for testing the performance of models, component by component.

**Cohort models describing decomposition as a continuum**

Another approach to modelling SOC turnover is to treat each fresh addition of plant debris into the soil as a cohort (McGill, 1996). Such models consider one SOM pool that decays with a feedback loop into itself. Q-SOIL (Bosatta and Ågren, 1995), for example, is represented by a single rate equation. The SOC pool is divided into an infinite number of components, each characterised by its ‘quality’ with respect to degradability as well as impact on the physiology of the decomposer organisms. The rate equation for the model Q-SOIL represents the dynamics of each SOC component of quality $q$ and is quality-dependent. Exact solutions to the rate equations are obtained analytically (Bosatta and Ågren, 1994).

**Food-web models**

Another type of model simulates C and N transfers through a food web of soil organisms, and such models explicitly account for different trophic levels or functional groups of biota in the soil. Figure 4.9 shows the structure of two food-web models (Hunt et al., 1987; de Ruiter et al., 1993). Some models have been developed, which combine an explicit description of the soil biota with a process-based approach (McGill et al., 1981). Food-web models require a detailed knowledge of the biology of the system to be simulated and are usually parameterised for application at specific sites.

### 4.3.2 Factors affecting decomposition in SOC models

Rate ‘constants’ ($k$) are constant for a given set of biotic and abiotic conditions. For non-optimum environmental circumstances, the simplest way to modify the maximum value of $k$ is by multiplication by a reduction factor $\mu$ – ranging from 0 to 1. Environmental factors
Figure 4.8 The structure of two commonly used SOM models. (a) RothC, (b) CENTURY.
Figure 4.9  The structure of two commonly used food-web models. (a) Hunt et al. (1987), (b) de Ruiter et al. (1993). Reproduced with kind permission from Springer Science and Business Media.
considered by SOC models include temperature, soil water content, pH, nitrogen availability, oxygen concentration in soil air space, clay content, CEC, type of crop/plant cover and type or intensity of tillage (Molina and Smith, 1998).

Many studies show the effect of temperature on microbially mediated transformations in soil, either expressed as a reduction factor at lower temperatures or by the Arrhenius equation. However, the assumption that SOC decomposition was temperature-dependent was challenged by a study suggesting that old SOM in forest soils does not decompose more rapidly in soils from warmer climates than in soils from colder regions (Giardina and Ryan, 2000). This conclusion has been strongly challenged by other studies (e.g. Fang et al., 2005), with Knorr et al. (2005) suggesting that old SOC is more temperature-sensitive than young SOC.

Models represent temperature sensitivity in different ways. Burke et al. (2003) reviewed how temperature sensitivity in different biogeochemical models was represented. They showed great variation among models in the representation of temperature sensitivity, represented by the \( Q_{10} \) response. \( Q_{10} \) is defined as the change in decomposition rate for a change in temperature of 10°C. Some models such as DAYCENT show rapidly increasing \( Q_{10} \) at lower temperatures, while others such as CENTURY and RothC have a more gradual increase in \( Q_{10} \) at lower temperatures. Others, such as the 2003 versions of pNET, TEM and Biome-BGC, have a fixed \( Q_{10} \) of 2 at all temperatures. Davidson and Janssens (2006) reviewed the subject of temperature sensitivity of SOC decomposition, and it is considered further in Section 4.9.3.

### 4.3.3 Coupling of SOM dynamics to other ecosystem processes

Some models such as RothC (Coleman and Jenkinson, 1996) and ICBM (Andrén and Kätterer, 1997) only simulate SOC transformation, with no simulation of carbon inputs to the soil. Instead, values for C inputs have to be supplied by the user, based either on estimation or derived from measurements, or from fitting to initial SOC levels (Coleman and Jenkinson, 1996). Other SOC models (e.g. CENTURY, EPIC and DNDC; McGill, 1996) simulate the whole ecosystem: these have plant growth components which simulate the growth of plants and direct carbon (through roots, exudates, debris, etc.) to the soil components of the model. Models containing the greatest feedbacks between climate and soil C turnover are the coupled climate carbon cycle models (termed C4 models; Friedlingstein et al., 2006). These models not only have coupling between soil C and plant C inputs, but plants and soils also respond to, and feedback to, the climate via the atmospheric CO$_2$ concentration. C4 models tend to have a simpler representation of soils than dedicated soil or ecosystem models (Jones et al., 2005). Though C4 models can be used to examine climate impacts on soil C in a fully coupled way (Jones et al., 2005), ecosystem models can also simulate climate impacts (without climate feedback) through use of climate scenario driving data (e.g. Muller et al., 2007). Dedicated soil models can also be used, by deriving C inputs from other models or data sources (e.g. Smith et al., 2006).

### 4.3.4 Application of SOM models

SOM models are used for a variety of purposes, over a variety of timescales and spatial scales and have become an integral part of SOM research. SOM models are commonly used
to test our understanding of SOM dynamics in given environments or at given sites. Perhaps the most common type of study involves applying models to explain changes in measured SOM dynamics at a given site. Such studies can also be used to test and compare the performance of the models (e.g. Smith et al., 1997).

Numerous model applications are now using eddy covariance data to simulate diurnal, seasonal and inter-annual carbon dynamics, particularly from the soil’s perspective for examining factors such as the contribution of soil heterotrophic respiration to total net ecosystem exchange (e.g. Sus et al., 2010). At the other extreme, SOM modules in larger global models have been used to examine long-term temperature sensitivity using data over millennia. For example, Lenton and Huntingford (2003) constructed a simple model of vegetation and soil and tested various temperature sensitivities of, among other things, plant and soil respiration. They compared the terrestrial carbon sink observed since the Last Glacial Maximum to test which figures in the range of temperature sensitivities of soil respiration found in the literature could explain the observations and found that the upper limit of reported sensitivities ($Q_{10} = 3.63$) under-predicted observed carbon storage. Such studies, using earth system observations/reconstructions over many thousands of years, allow long-term soil carbon responses to be tested, facilitating greater confidence in predictions of future responses to climate change.

SOM models remain valuable at the scale of a single field or land use unit in order to simulate and explain SOM dynamics in response to management practices at a specific site. But they are increasingly applied at a global scale (e.g. Jenkinson et al., 1991), often using detailed spatial data now available from remote sensing and other data collation exercises. Global Climate Models have moved from simple single-pool representations of SOM to having fully coupled SOM models integrated into them. RothC, for example, has been integrated into the Hadley Centre Climate Model and, when run globally, changed the behaviour of the model by dampening the response of the soil to climate fluctuations (Jones et al., 2005).

### 4.3.5 Confidence in SOM models: uncertainties and areas for further development

To test the validity of any model, it is necessary to compare model simulations with measured data. This was done in a comprehensive comparison of 9 models of SOC turnover, tested against 12 measured data sets from long-term experiments representing different land use, soil type, climate zone and land management (Smith et al., 1997). Through such evaluation, the applicability and limitations of models can be assessed. Where models can simulate SOC trends across a range of environmental conditions, they can be applied to future scenarios to project likely changes in SOC (Smith et al., 2008).

Areas of uncertainty requiring further research include the inability of current SOM models to represent potentially important ecological phenomena, including ‘priming’ of soil C decomposition in response to fresh organic C additions, particularly in subsoil (Fontaine et al., 2007; but see Jenkinson and Coleman, 2008; Jenkinson et al., 2008), the ‘gadkil’ effect and non-equilibrium dynamics (Schimel and Weintraub, 2003; Fontaine and Barot, 2005; Neill and Gignoux, 2006; Wutzler and Reichstein, 2007) and assumptions about the temperature sensitivity of different soil carbon pools (Davidson and Janssens, 2006). Other challenges include understanding the nature of impacts of land management change on the stability of different pools of soil carbon (e.g. the physical protection of organic matter – Denef et al., 2007) and the inclusion of microbial feedbacks of decomposition that are not captured by first-order assumptions (e.g. soil methanogenesis and methanotrophy).
4.4 Chemical structure of SOM as elucidated using new forms of spectroscopy – advances and pitfalls

What exactly do we mean by soil organic matter?

The term soil organic matter (SOM) is generally used to indicate the total of all organic components present in soil, including living organisms. Other definitions recall that suggested earlier by Waksman (1936) for ‘humus’, which excludes undecomposed or partially decomposed plant material and living organisms (Stevenson, 1994). At first sight this latter definition may seem more rational but the practical problem is that SOM defined in this manner would not coincide with any analysable fraction. This is because soil organisms (the soil microbial biomass) and microscopic plant and root debris cannot be reliably separated from the soil and are currently analysed as part of total organic C. The old word ‘humus’ itself is poorly defined and its meaning has not been completely agreed. It is generally agreed by soil chemists that, from a chemical point of view, SOM can be divided into non-humic and humic substances (Stevenson, 1994). The non-humic materials comprise organic substances that have defined chemical structures, such as carbohydrates, proteins, hydrocarbons, alcohols, aldehydes, resins and amino acids as well as aliphatic and aromatic acids. Humic substances are a largely heterogeneous mixture comprised of yellow to black-coloured polyphenolic polycarboxylic acids exhibiting a wide range of molecular weights. Although their functional groups and reactivity may be described in great detail, their actual chemical structure is not yet clearly defined. In fact, it is likely that there is no universal structure that can be defined in a rigorous chemical sense.

4.4.1 Structural information from nuclear magnetic resonance (NMR) spectroscopy

The development of modern solid state $^{13}$C NMR techniques and instrumentation has led to substantial progress in SOM studies. The availability of high field magnets, from 600 to 1000 MHz, now permits the study of whole soil samples and allows the characterisation of SOM in situ, avoiding the need to extract organic materials and separate them from inorganic components. The extraction step, to provide a liquid sample, was unavoidable before the development of solid state $^{13}$C NMR. But it has always been the most troublesome, for two reasons. With mild extractants, the recovery of SOC is low, so it is uncertain whether or not the material obtained is representative of the total. To obtain higher recoveries, solutions such as strong alkalis have been used but it is likely that these cause chemical alteration through hydrolysis reactions.

The main limitation of solid state NMR still remains the fact that solid state spectra cannot be recorded directly on soils rich in iron oxides because of interference from paramagnetic materials, notably Fe. Pre-treatment of soil with hydrofluoric acid is sometimes used to remove mineral components, including those containing Fe. However, removal of Fe is not always complete and the treatment may alter SOM chemical composition, thus nullifying the advantage of applying NMR spectroscopy to solid soil rather than a liquid extract. Even so, solid state NMR studies have substantially changed our perception of SOM composition: the most important change being the decreased importance of aromatic structure compared to aliphatic and the greater significance given to hydrophobic interactions (Mahieu
Soil organic matter

et al., 1999; Ahmad et al., 2006). As a consequence new concepts have been developed to explain the complex chemical behaviour and intrinsic recalcitrance of SOM. One such concept is the hypothesis that a substantial fraction of SOM can be regarded as supramolecular aggregations based on hydrophobic interactions (Simpson et al., 2001) – this is discussed later.

Another limitation of $^{13}$C NMR spectroscopy is poor sensitivity, in part because NMR only records C atoms having the relatively rare $^{13}$C nucleus, comprising only 1.1% of the C in natural materials. Sensitivity is increased by a technique termed ‘cross-polarisation’ (CP) in which energy from adjacent $^1$H nuclei is used to boost the signal from $^{13}$C nuclei, but this can lead to overestimation of the quantity of C in alkyl groups (that are generally abundant in H atoms) and underestimation of C in aromatic groups that generally have fewer H atoms in comparison to C.

Solid state $^{13}$C spectra are normally obtained using a combination of cross-polarisation and ‘magic angle spinning’ (MAS), the latter to overcome complications that otherwise make it impossible to obtain spectra from solid samples. This approach is termed $^{13}$C CPMAS NMR. A comprehensive analysis of $^{13}$C CPMAS NMR data from more than 300 soils, representing a wide range of SOM contents, climates and management practices (Mahieu et al., 1999), showed the remarkable similarities of functional group abundance. The largest fraction of soil organic C was within O-alkyl groups (on average 41–46% of total visible C), followed by alkyls (25%), aromatics (20%) and finally carbonyls. Grassland soils were the poorest in alkyls and richest in carbonyls, with the opposite for forest soils. The proportion of O-alkyls increased with C content, reaching up to 60% in high C soils. The general similarity of SOM structure between soils is not surprising. In spite of the great

**Solid state nuclear magnetic resonance spectroscopy of SOM: limits and advancements**

Solid state NMR spectroscopy, although still far from being the ultimate analytical tool, has proved to be very powerful. Yet only 30% of the soil organic C may be visible by NMR, due to various interferences and overall lack of sensitivity. Comparison of spectra recorded on soil samples before and after treatment with hydrofluoric acid (used to destroy clay minerals and remove Fe which interferes with NMR spectra) suggests that the distribution of organic C in solid state spectra is reasonably representative of the total. However, results obtained using more sophisticated techniques lead to the opposite conclusion. For example, aromatic and carbonyl C signals appear more intense in so-called ‘Bloch decay’ spectra, suggesting that they are underestimated in the more usual forms of $^{13}$C NMR spectroscopy. It is possible that much of the C that is invisible in solid state CPMAS spectra could correspond to charcoal or black C originated from biomass burning or wild fires, though this remains to be seen.

Until the introduction of cross polarisation magic angle spinning (CPMAS), information gained by NMR was purely qualitative – thus these techniques represented a major move towards quantitation. However, as mentioned in Section 4.4.1, CP leads to disproportionate enhancement of C atoms located close to H atoms within molecules, thus exaggerating the abundance of alkyl chains at the expense of aromatic rings. A further inaccuracy is the formation of artefacts termed ‘spinning side bands’ that overlap with genuine signals, so appropriate measures must be taken in order to achieve the ‘total suppression of side bands’ ($^{13}$CPMAS TOSS NMR).
chemical complexity and range of input sources, above- and below-ground plant residues which constitute the main C inputs to soils comprise for the most part (>90%) a relatively narrow mixture of organic compounds: 50–60% polysaccharides (cellulose, hemicelluloses, starch, pectin), 15–20% lignin and 5% proteins. The abundance of the other compounds (tannins, waxes, lipids) is much more variable, but always limited to a few per cent.

In addition to organic entities synthesised within the soil as a result of biological processing of inputs, chemical species directly derived from plants can also be detected in soil. This is because some processes such as the action of tillage or fauna foster contact of plant residues with clay minerals and oxides (Wolters, 2000), resulting in physical stabilisation making these potential organic substrates unavailable to decomposers (Sohi et al., 2001; Six et al., 2002).

In natural ecosystems, above ground plant residues are deposited at the soil surface and enter the soil by the mixing action of invertebrates such as earthworms, insects and arthropods at various stages of their life cycle. Under temperate climatic conditions, the activity of earthworms is so crucial for the formation of stable soil aggregates that it is included in diagnostic criteria for humus forms and soil classifications (Zanella et al., 2011). Soil ingested by detritivores undergoes many alterations, including physical realignment of clay particles, which incorporate undigested fragments and favour the formation of organo-mineral complexes (Bossuyt et al., 2005). Aggregates formed by soil mixing and tillage can contain discrete particles of freshly incorporated plant residue (termed particulate organic matter, POM) and protect them from decomposition (Gregorich and Ellert, 1993). POM and the so-called light fraction (LF) are operationally defined fractions of SOM that are frequently used in soil C turnover studies. POM is isolated by sieving (size separates between 53 and 2000 μm) and makes up about 6% of soil C in cultivated soils and about 18% C in grassland and forest soils. Smaller fragments of POM, which make up the LF (isolated by density separation, suspending soil in a heavy liquid, density >1.9 g cm\(^{-3}\)), often comprise about 28% of soil C in forest soils and sometimes up to 20% in arable and grassland soils (Gregorich et al., 2006). So, in total, physically protected and non-protected chemically unaltered plant materials may sometimes comprise as much as 30–40% of SOM. On this basis, we can estimate that at least 5–8% of soil C is made up of polysaccharides of direct plant origin.

Cellulose, the most abundant natural polymer, is composed of approximately 8 × 10\(^3\) D-glucopyranose residues linked by β 1 → 4 glycosidic bonds. About 36 of these glycan chains are held together by hydrogen bonds to form a cellulose microfibril, which binds to hemicellulosic polysaccharides, such as xylans, mannans and galactans, which have a lower degree of polymerisation and more branching than cellulose (Kogel-Knaber, 2002). Besides hemicelluloses, cellulose fibrils are closely associated with lignin in secondary cell walls of plants. Other common polysaccharides in plants are starch and fructans, but being much more easily degradable their presence in soil is transient.

Solid state \(^{13}\)C CPMAS NMR allows monitoring of changes which occur during decomposition of plant materials in soil. It is observed that the proportion of alkyl C increases compared to O-alkyl C as decomposition proceeds as a result of polysaccharides being decomposed (Baldock et al., 1997). Figure 4.10 (from Webster et al., 2001) shows the ratio (alkyl C)/(O-alkyl C) in relation to the proportion of soil C in the microbial biomass in a set of soils from diverse environments and land uses. This shows a significant negative correlation between the (alkyl C)/(O-alkyl C) ratio and soil microbial biomass C content (expressed as a fraction of total organic C), consistent with the hypothesis that an increased proportion of alkyl C represents a greater degree of decomposition. These authors proposed
the (alkyl C)/(O-alkyl C) ratio as an index of resource quality of soil C for the soil microbial biomass. This ratio is also negatively correlated with the C/N ratio of soil fractions, consistent with fractions having a larger proportion of polysaccharides being low in N.

Polysaccharides in SOM are not derived exclusively from plants but are also synthesised during decomposition of plant residues by soil microorganisms. These polysaccharides provide structural components for microbial cell walls, energy storage and extracellular mucigels which help bacteria to adhere to particle surfaces and provide protection from environmental stress. All these forms of polysaccharides accumulate during decomposition, partially masking the disappearance of carbohydrates from decomposing plant residues. Bacterial cell walls, particularly those of Gram-negative bacteria, contain murein, a carbohydrate backbone of alternating units of N-acetyl muramic acid and N-acetyl glucosamine cross-linked with oligopeptides. Fungal cell walls are mostly made of a different polysaccharide: chitin (N-acetyl-\(\beta\)-D-glucosamine in \(\beta-(1-4)\)-glycosidic bonds), which is also a structural component of the exoskeleton of insects. Microbially derived polysaccharides in soil cannot be separated from those of plant origin by physical isolation or selective hydrolysis. Their accumulation or disappearance can only be detected by analysing the composition of the products of hydrolysis with respect to the relative ratios of arabinose (A) and xylose (X), which are considered to be mostly originated from plants, whereas mannose (M) and galactose (G) in soil are mainly of microbial origin (Chantigny et al., 2000). When the ratio of (M + G) to (A + X) is lower than 0.5, the contribution is considered mainly from plant materials, whereas ratios higher than 2 are an indication of a microbial origin (Oades and Waters, 1991).

The most abundant polymers in plant residues, after cellulose, are lignins. Lignins are polydisperse polymers which are synthesised by random radical coupling reactions of phenylpropane units (Figure 4.11), namely, \(p\)-coumaryl, coniferyl and sinapyl alcohols (monolignols). This mode of synthesis produces irregular three-dimensional macromolecules with no extended sequences of regularly repeating units. Lignins are generally characterised by a high abundance of \(p\)-hydroxyphenyl propane (coumaryl alcohol), guaiacyl

---

**Figure 4.10** Relationship between the biomass C to total C \(C_{\text{mic}}/C_{\text{org}}\) ratio and alkyl-C to O-alkyl-C ratio derived from \(^{13}\)C CP MAS NMR spectroscopy for 47 soils. The set includes mineral, organic and urban soils. Reproduced with permission from Webster et al. (2001).
Figure 4.11  The formation of lignin from monolignol (phenylpropane) units dimerising by oxidative radical-mediated coupling leading to several possible types of linkage. The dimers randomly react together to produce a three-dimensional network of aromatic rings linked by an array of inter-unit linkages.
Soil organic matter

Soil organic matter includes a variety of complex organic molecules, with a significant portion being lignin. Lignin is composed of coniferyl alcohol and sinapyl alcohol units, derived from each of the three primary monolignols and by the distribution of inter-unit linkages in the polymer (e.g. β-aryl ether or β-O-4, phenylcoumaran or −  β5, resinol or β-β, biphenyl or 5-5, diphenyl ether or 4-O-5). The few initial dimerisation reactions also produce hydroxycinnamyl alcohol end groups from structures such as non-cyclic α-aryl ethers and various other structures. Lignins of gymnosperms, angiosperms and grasses differ in their monolignol composition. Grass lignin is composed of equal proportions of the three primary monolignols, whereas that of gymnosperms (softwood lignin) derives in large part only from guayacyl propane monomers and that of angiosperms (hardwood lignin) from equal proportions of guayacyl and syringyl propane.

The irregularity and the aromaticity of its structure, together with the relatively large variety of linkages, makes lignin a recalcitrant and therefore a poorly degradable substrate for microorganisms. Only white rot fungi are able to completely mineralise lignin, and its decomposition in soil is probably carried out by a complex microbial community. The first step of lignin degradation, which can only occur under aerobic conditions, is demethylation of methoxyl groups, followed by oxidation of side chains. This leads to enrichment of the product in acidic functional groups (COOH and phenolic OH). Oxidation to quinones of orthohydroxy-benzene moieties resulting from demethylation favours condensation with NH₃ and amino-compounds. Lignin, however, is also depolymerised by fungi, which release dilignol components that are further transformed into coniferyl alcohol, coniferaldehyde, ferulic acid, syringaldehyde, syringic acid, vanillin, vanillic acid, etc. Lignin is therefore a source of polyphenols that in the soil environment are readily oxidised to quinones, which can eventually condense with amino-compounds.

Secondary components of plant and animal residues include a huge variety of complex organic molecules as well as relatively simple compounds, ranging from chlorophylls, sphingolipids, tannins, sterols, vitamins and nucleic acids to fatty acids and simple alkanes. Among these, a group of structurally different molecules, sharing a sufficient structural degree of hydrophobicity which makes them insoluble in water but extractable by non-polar organic solvents, are generally referred to as lipids (Bull et al., 2000). This empirically defined class of compounds is not unequivocally defined, as obviously much of its composition depends on the type of solvent used (e.g. hexane, acetone). It constitutes only from 4% to 8% of soil organic C but strongly influences the physical properties of soil aggregates (e.g. wettability, stability) and participates in many soil processes such as humus formation and clay dispersion. From a qualitative point of view, soil lipid extracts reflect the overlying vegetation and soil use. For example, the lipid extracts of the grazed grassland area of the Broadbalk Wilderness at Rothamsted are dominated by hexacosanol and other alkanol analogues (C₁₈–C₃₄); an adjacent area with herb vegetation is similar (Bull et al., 2000). Wax esters are extracted in much smaller quantities from these soils, indicating possible loss through hydrolysis or oxidation, and show an increased proportion of the lower simpler homologues. Root material, by contrast, is a predominant source of aliphatic dicarboxylic acids and of long-chain (>C₂₀) n-alkylcarboxylic acids, which derive from suberin.

4.4.2 Organic forms of nitrogen

¹⁵N NMR and other spectroscopic methods of analysis show amide N to make up some 85% of the organic N in soil with a further 10% in amino acids, amino sugars, chitin, nucleic acids and heterocyclic N-containing aromatic rings. This dominance of amide N could be the result of an oversimplification in the interpretation of spectra, different spectral
sensitivities to N in different chemical forms or by the overall poor sensitivity of $^{15}$N NMR, which depends upon the low $^{15}$N natural abundance (0.366%). However, the dominance of N in amide groups seems clear and is consistent with N contained in proteins. Proteins enter soil as a component of plant material and are also synthesised within the cells of soil microorganisms. They are located at different sites within cells, for example in cytoplasm or at different locations in cell walls, but are also released into soil solution. In fact, in order to break down plant residues and avail themselves of the energy and essential elements contained in natural macromolecules, microorganisms produce a range of extracellular enzymes which are released outside the membrane and then diffuse through soil solution to their specific substrates and hydrolyse them into small simpler units that can be utilised by the decomposing organisms (Lynd et al., 2002). However, proteins are among the most labile among natural macromolecules and, when added to soil under experimental conditions, normally persist for no more than 2 or 3 days, so it is surprising to find evidence of their dominant position within SOM. It is thought that this unexpected persistence is a result of stabilisation processes that include chemical entrapment by association with humic substances and physical adsorption on clay surfaces.

Total extractable protein concentration in soil is often around 1–2 mg protein C g$^{-1}$ C (McClougherty and Linkins, 1988). Different proteins can fulfil different functions in soil. For example, glomalin is a glycoprotein that is a major component of the cell walls of the hyphae of mycorrhizal fungi. It remains in soil after hyphal death and is thought to have an important role in stabilising aggregate structure. It can be extracted from both native and agricultural soils with citrate solutions and is relatively resistant to decomposition because of its strong tendency to bind with multi-valent cations, particularly iron and aluminium. Glomalin is also partially hydrophobic; this may increase its stability and also confer stability on other molecules with which it interacts. Glomalin-related soil protein is operationally defined by its extraction method and cannot be regarded as a discrete chemical substance. Despite much discussion in recent years of its possible significance, its concentration in soil has probably been considerably overestimated (Gillespie et al., 2011).

Proteins that are exocellular enzymes, whether excreted from living soil organisms or remnants after cell death, can be stabilised in soil through association with other soil components such as sorption on clay surfaces or chemical or physical interactions with other organic materials. This gives rise to significant abiotic enzyme activity in soil that has been extensively measured and plays an important role in soil biological activities (Nannipieri, 2006; Kemmitt et al., 2008).

### 4.4.3 Black carbon

Black carbon or biochar is defined as residues derived from the combustion of vegetation, either through natural processes such as wildfires or human burning of vegetation. The term pyrogenic carbon is also used. If combustion is complete, the product comprises elemental C, or charcoal, but more commonly it is the result of incomplete combustion such that some chemical structures from the original plant material remains though the O/C ratio is low (pure charcoal would contain no oxygen). Black carbon can also include combustion condensates such as soot, derived from the burning of fossil fuel and reaching the soil through atmospheric deposition. It is thought that black carbon may comprise a significant proportion of the highly resistant fraction of soil C; some estimates are that it is up to 30% of SOM in agricultural soils in the USA and >45% of some German soils (Knicker, 2009), though whether these large values are generally true is not known. The mean residence time
of black C in the environment has been assumed to exceed 1000 years. However, it is not inert and slow oxidation in aerobic soils occurs, either biotically or abiotically, leading to surface formation of carboxylic functional groups that can contribute to soil CEC (Knicker, 2009). Smaller fragments are preferentially found inside micro-aggregates, so that physical stabilisation probably concurs to increase persistence and retard oxidation processes (Brodowski et al., 2006).

There has been much recent discussion regarding the occurrence of black carbon in soils (Sohi et al., 2010) and its potential value as an additive to soils. Most evidence for beneficial effects comes from sites in the Amazon region with highly weathered soils where it appears that biochar was applied to soil at least 500 years ago and the impacts on soil properties persist. For example, there is circumstantial evidence from archaeological sites that the presence of biochar confers stability on other forms or organic C in soil (Liang et al., 2010) and that its modern application leads to the retention of nutrients in soil in high rainfall tropical environments (Major et al., 2010). However, there is as yet no unequivocal evidence that its addition to temperate soils leads to major improvements in soil functioning or plant growth (Powlson et al., 2011c) – research on this topic continues.

4.4.4 Modern concepts

The new insights provided by modern spectroscopic techniques have led scientists to rethink the nature of SOM and the process of humification, drawing attention towards classes of organic chemicals such as lipids and, for different reasons, charcoal that were previously considered minor constituents. In particular, the so-called supramolecular theory (see Schaumann, 2006) describes humic substances as associations of relatively small molecules arranged in micellar aggregations held together by hydrophobic interactions. The structures are hypothesised to contain some regions derived from relatively unaltered remnants of plant or microbial tissue interconnected with more degraded or newly synthesised small molecules. This theory contrasts with the classical view of SOM as comprising extremely large polymeric molecules. It is still a matter of debate as to which view of SOM is a more accurate representation – there is experimental evidence in support of both but neither completely accounts for the range of complex properties exhibited by SOM. See Swift (1999) for a critical review of evidence for both views, with a conclusion that the supramolecular theory, postulating an aggregation of relatively small molecules, is unlikely to explain the majority of SOM structure.

Another new theory, based on spectral evidence and improved knowledge of biotic processes, has recently highlighted the possibility that a class of recalcitrant secondary metabolites known as polyketides may be a significant component of the stabilised C pool that makes up the great majority of SOM residing in the smallest size fractions within soil (clay size particles and smaller) where the age of C is >1000 years (Schnitzer and Monreal, 2011). The polyketides comprise a very diverse family of natural products characterised by more than two carbonyl groups connected by single intervening carbon atoms, with diverse biological activities and pharmacological properties including antibiotics and natural insecticides. They are also very stable against biological decomposition. In addition to the stability of this class of compounds, it is hypothesised that the skeleton formed by polyketides can trap and bind other organics such as carbohydrates, peptides and lipids as well as inorganics such as clay minerals and hydrous oxides. This mechanism would confer additional stability on the trapped organic molecules, at least in part by holding them within a hydrophobic structure.
4.5 Organo-mineral interactions

Organic components in soil contain both hydrophobic and hydrophilic groups. Alkyl chains and aromatic rings are hydrophobic and thus have a strong tendency to avoid interaction with water, while hydroxyl, phenoxy and carboxyl groups are hydrophilic. Both classes of group can lead to an association between organic matter with minerals or amorphous oxides in soil that increase the stability of organic matter and also have profound impacts on soil physical properties. The main mechanism by which organic matter binds to minerals is through the formation of so-called cationic bridges (see Figure 4.12a). Clay particles have negative charge at their surface, which is balanced by association with cations such as the alkaline earth elements (most commonly Ca\(^{2+}\) or Mg\(^{2+}\)) or polycations such as Fe\(^{3+}\) and Al\(^{3+}\) derived from soil solution. These cations in turn are surrounded by water molecules which can be displaced by anionic groups (such as dissociated carboxyl or phenolic groups) on organic molecules. When organic entities bind to mineral surfaces, their stability against decomposition is increased because access to them by microorganisms or exocellular enzymes is limited. This stabilisation has been demonstrated by measuring the rate of decomposition of added organic materials to clays that are saturated with either Ca\(^{2+}\) or Na\(^{+}\) ions (Baldock and Oades, 1989). Lower evolution of CO\(_2\) is observed with Ca\(^{2+}\) because this divalent ion favours such stabilisation through formation of bridges between organic matter and clays, whereas monovalent Na\(^{+}\) does not. Stabilisation through polyvalent cations such as Al\(^{3+}\) and Fe\(^{3+}\), which can form strong coordination complexes with SOM, is even greater (Blaser et al., 1997). The decomposition of dissolved organic matter (DOM) extracted from a soil, then adsorbed on clays, was only 15–30% of the rate measured when the material was in solution (Schwezig et al., 2003). Sorption increased the mean residence time of the most stable DOM sample from 28 years to 91 years. The increase in stability due to sorption is greatest for labile DOM high in carbohydrates, and relatively small for stable DOM high in aromatic and complex molecules.

Figure 4.12b and c shows variations of organo-mineral interactions that are thought to occur. If the quantity of organic matter present is larger than in Figure 4.12a, crowding at the mineral surface can cause a different orientation of organic molecules (Figure 4.12b) such that the surface exposed to soil solution becomes more hydrophilic than in Figure 4.12a. If the quantity of organic matter becomes still larger, layers of organic molecules can build up via cationic bridges. This causes the surface to become increasingly hydrophobic (Figure 4.12c) and more like the situation in Figure 4.12a. These differences can have profound impacts on the surface properties of mineral particles influencing aggregate formation and wettability, both of which influence organic matter stabilisation (Bachmann et al., 2008).

A variation of the adsorption processes described earlier in this section is a process termed ligand exchange, which is particularly prevalent with iron and aluminium oxides, especially in tropical soils. In this process, a hydroxyl group on the mineral surface is displaced by an OH or oxygen atom belonging to an organic molecule. This group binds to an Al or Fe atom and becomes part of the mineral’s structure, while the original OH group exiting, becoming part of a water molecule or a solvated OH\(^{-}\) group. The organic molecule connected to the attached oxygen-containing group is now firmly linked to the mineral structure. The process is highly irreversible and results in permanent sequestration, so that even organic molecules that would otherwise be labile become unavailable for decomposition by microorganisms.

Another mechanism leading to binding between organic and mineral particles is through the hydrophobic sections within organic matter. These groups, and organic xenobiotics such as organic pollutants entering soil, can escape interactions with water either by binding with
Soil organic matter

hydrophobic sites on other organic molecules or by binding to solid surfaces such as clay particles and aluminium or iron oxides. The bonding between hydrophobic groups and a mineral surface is through van der Waals forces which are weak. However, the process is favoured on the basis of thermodynamics. When hydrophobic groups are present in the aqueous phase, water molecules cannot form a solvation layer around the hydrophobic structure but instead, in an effort to avoid contact, tend to form an ordered cage-like structure in the vicinity and consequently lose freedom of movement. This represents a substantial decrease in entropy. When the organic molecule is bound to a mineral surface, the energy involved in forming the cage structure of water molecules is released and entropy increased. It is these thermodynamic factors that tend to drive the binding of hydrophobic regions of organic matter to mineral surfaces, even if the actual binding forces are weak.

The coating of clays and iron or aluminium oxides by organic molecules modifies their surface properties and eventually reduces the surface area accessible for other reactions: Figure 4.13 shows some examples of this (Kaiser and Guggenberger, 2003). In the experiments illustrated, organic matter bound to the different mineral fractions was destroyed (by oxidation with sodium hypochlorite) and then the specific surface area (SSA) of the particles determined by measuring adsorption of a monolayer of N$_2$ gas using standard methodology. The particles were then treated with a solution containing organic matter that had been extracted from an organic soil horizon using water and the SSA measured again. Figure 4.13 shows that the initial quantity of sorbed organic matter caused the largest decrease in SSA, and this was interpreted as organic matter being preferentially sorbed on surfaces at the mouths of micropores and thus blocking further potential sorption sites within these pores. This provides a mechanism whereby a small increase in the organic matter content in a soil may have a disproportionately large effect on the soil’s physical properties. Preferential binding at reactive sites also favours surface accumulation of organic matter in patches rather than in a continuous coating, even on an initially uncovered mineral surface.

Increasing soil C concentration leading to additional organic C in soil solution available for interaction with mineral components.

**Figure 4.12** Different types of associations between organic molecules and mineral surfaces at (a) low, (b) medium and (c) high organic carbon: mineral ratios. Redrawn from Bachmann et al. (2008). With kind permission from John Wiley & Sons.

(a) Hydrophobic carbon backbone
(b) Hydrophilic functional groups
(c) OM molecules

Organic C to mineral ratio:
low: (<10 g kg$^{-1}$)
medium: (10 g kg$^{-1}$)
high: (>10 g kg$^{-1}$)
Figure 4.13  Effect of increasing amounts of sorbed organic C on the specific surface area (SSA) of seven mineral phases. Reprinted from Kaiser and Geggenberger (2003). With kind permission from John Wiley & Sons.
4.6 Fractionation methods for studying SOM

4.6.1 Fractionation based on solubility

An approach that has been used for many years is based on extraction of soil with alkali (e.g. 0.1 M NaOH) followed by acidification of the extract with H₂SO₄, though many variations exist. Conventionally organic matter that is precipitated after acidification is termed ‘humic acid’ while that which remains in solution is termed ‘fulvic acid’. The material that is not extracted by alkali is termed ‘humin’. These fractions have no specific significance in terms of chemical structure. This point was stated succinctly by Jenkinson (1973) in an earlier edition of this book who stated that they ‘… cannot be defined in precise chemical terms, despite the efforts of generations of chemists. They are complex mixtures, no two humic molecules being exactly alike, and cannot be crystallized or otherwise separated into classes of homogeneous molecules’. In other words, these fractions are ‘operationally defined’ and based on solubility at different pH values. There is also considerable likelihood that the extractant solutions will cause chemical changes and it was shown over 60 years ago that humic colloids in alkaline solution take up atmospheric oxygen, causing chemical change and an increase in their pH. Despite the clear limitations of this approach, one still sees reference to it in the soil science literature. In some cases, alkali extraction is simply used as a means of obtaining a liquid that can be subjected to various analyses that cannot be applied to solids; this was discussed in Section 4.4.1 in relation to obtaining material for liquid state NMR spectroscopy.

4.6.2 Physical fractionation approaches

Measuring the distribution of particles of different sizes (soil texture) is achieved by the total dispersion of soils into primary particles (usually using ultrasonics) followed by sedimentation. This is normally done after destroying organic matter using H₂O₂. But if dispersion is done without organic matter destruction, it is possible to determine the location of organic matter in relation to the different sizes. A useful way of expressing results is to use the concept of an enrichment factor, \( E_C \). \( E_C \) is defined as the ratio of organic C concentration (e.g. expressed as mg g⁻¹ or %) in a given fraction to that in whole soil. A value >1 indicated that the fraction is enriched in C compared to whole soil while a value <1 indicated depletion compared to whole soil. Table 4.1 (taken from Christensen, 1996) shows some typical results from soils from four continents. Particles in the clay and fine silt size categories are always enriched in organic C compared to whole soil: \( E_C \) values of about 2.0 or higher are common. Coarser particles, such as coarse silt (50–2 μm), are depleted in C, typically with \( E_C \) values of about 0.1. Sand size fractions are also typically low in C with \( E_C \) values <0.1, except in soils containing much macro-organic matter (such as fresh plant fragments) as these can appear in the sand-size fraction.

Figure 4.14 (also from Christensen, 1996) demonstrates the importance of clay and the finer silt particles for the retention of organic C in soil. For example, in a soil containing only 3% clay this fraction contains almost 50% of the total organic C in the soil. In the examples shown, with clay content ranging from 3% to 47%, the sand fraction contains a maximum 10% of total soil organic C, and usually much less. In soils with very low clay content, the silt fraction becomes highly significant for C retention: it contains about 40% of total soil C in the 3% clay soil. But as clay content increases the role of silt particles diminishes.
Table 4.1: Organic carbon enrichment factors, $E_c$, in different silt and clay size classes isolated by ultrasonic dispersion from predominantly cultivated soils.

<table>
<thead>
<tr>
<th>Location (sampling depth)</th>
<th>USDA soil taxonomy</th>
<th>%C</th>
<th>Carbon enrichment, $E_c$</th>
<th>Silt (μm)</th>
<th>Clay (μm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>USA (0–25 cm) Typic Argiudoll</td>
<td>1.71</td>
<td>0.2</td>
<td>2.18</td>
<td>0.2</td>
<td>3.8</td>
</tr>
<tr>
<td>Canada (Ap horizon) Typic Argiboroll</td>
<td>3.32</td>
<td>0.7</td>
<td>2.41</td>
<td>0.1</td>
<td>3.0</td>
</tr>
<tr>
<td>Canada (Ap horizon) Typic Cryoboroll</td>
<td>2.50</td>
<td>0.1</td>
<td>1.1</td>
<td>1.0</td>
<td>1.4</td>
</tr>
<tr>
<td>Canada (A horizon) Typic Haploboroll</td>
<td>4.50</td>
<td>0.2</td>
<td>0.2</td>
<td>1.2</td>
<td>2.9</td>
</tr>
<tr>
<td>Canada (A horizon) Baralic Cryoboroll</td>
<td>2.10</td>
<td>0.1</td>
<td>0.1</td>
<td>1.0</td>
<td>1.4</td>
</tr>
<tr>
<td>Canada (A horizon) Aridic Nativaroll</td>
<td>6.29</td>
<td>0.3</td>
<td>0.8</td>
<td>0.6</td>
<td>2.0</td>
</tr>
<tr>
<td>Australia (0–10 cm) Vertic Rhodoxeralf</td>
<td>1.04</td>
<td>0.2</td>
<td>2.7</td>
<td>2.6</td>
<td>11.1</td>
</tr>
<tr>
<td>Denmark (0–20 cm) Orthic Haplohumod</td>
<td>1.73</td>
<td>0.3</td>
<td>2.3</td>
<td>2.1</td>
<td>7.0</td>
</tr>
</tbody>
</table>

Source: From Christensen (1996).
In addition to association of organic C with primary mineral particles, in recent years increased emphasis has been placed on the role of secondary structures within soil in stabilising organic C. For example, Killham et al. (1993) concluded that organic matter located within small pores was partially protected from decomposition by microorganisms. Many studies reviewed by Six et al. (2002) have emphasised the role of aggregates. It was concluded that macro-aggregates (defined as those >250 μm) provide little protection but micro-aggregates (<250 μm) offer a considerable degree of protection. Evidence comes, in part, from experiments in which aggregates of different size are crushed or disrupted in other ways and the effect on C mineralisation is measured. Disrupting macro-aggregates causes only a small increase in mineralisation rate, but with micro-aggregates it was three to four times greater (Bossuyt et al., 2002). Other evidence comes from experiments in which fresh substrates are added to soil; for example, Angers et al. (1997) incubated wheat straw labelled with $^{13}$C and $^{15}$N with soil under field conditions and found that the labelled C retained in soil was predominantly held within micro-aggregates. Micro-aggregates conferring stability on organic C can either be free in soil or contained within macro-aggregates. Although it is micro-aggregates that have the greatest influence on C stabilisation, it is thought that their formation is favoured if they are held within stable macro-aggregates – so soil-management practices that favour the stability of macro-aggregates indirectly favour C

**Figure 4.14** Distribution of whole soil organic C between clay (<2 μm), silt (2–20 μm) and sand (20–2000 μm) from Ap horizons of Danish arable soils. From Christensen (1996). With permission from CRC Press.
Soil Conditions and Plant Growth

Stabilisation within the micro-aggregates. This may be a significant factor in the accumulation of organic C near the soil surface where minimum tillage is practised as the formation and stabilisation of macro-aggregates is favoured under this practice compared to conventional tillage that causes greater disruption (Six et al., 2002; Denef et al., 2007). There is also evidence that the formation of micro-aggregates occluded within macro-aggregates, and the resulting protection of organic C against decomposition, is favoured by the activity of earthworms. Some of the organic matter stabilised through the influence of micro-aggregates is relatively undecomposed material derived from plants termed particulate organic matter (POM); indeed it seems likely that discrete POM particles form nuclei around which mineral particles accumulate and confer stability through the range of mechanisms discussed in Section 4.5.

Many physical fractionation schemes include a step in which soil is dispersed in a liquid of high density and a SOM fraction termed ‘light fraction’ floats to the surface. This fraction is predominantly composed of relatively fresh plant debris that has undergone only a small degree of decomposition, although it can also contain some charcoal or other forms of black carbon or biochar. The fractionation scheme published by Sohi et al. (2001) uses a very concentrated solution of sodium iodide (NaI) with a density of 1.8 g cm$^{-3}$. Soil is first dispersed with minimal disruption such that only macro-aggregates are broken and the floating material is termed ‘free OM (FR-SOM)’ or free light fraction (Figure 4.15). After its removal, the soil is again dispersed in NaI, but this time ultrasonic dispersion is
used to break all aggregates including the more stable micro-aggregates, and it is found that additional organic matter is released and floats – this is termed ‘intra-aggregate OM (IA-SOM)’ and comprises discrete organic matter particles (sometimes termed ‘particulate organic matter’ or POM) that were previously entrapped within stable aggregates. A final fraction is the material that sinks in the high-density liquid and is mainly inorganic minerals with its associated organic matter, termed the organo-mineral fraction; this may be separated into size classes to give clay, silt and sand size fractions. Spectroscopic analyses ($^{13}$C NMR and a version of infrared spectroscopy) show that the intra-aggregate fraction contains a smaller proportion of C in O-alkyl groups and a larger proportion in alkyl groups compared to free OM fraction; i.e. IA-SOM had a smaller O-alkyl/alkyl ratio. This indicates that it has undergone greater alteration in structure due to decomposition. Other spectroscopic studies confirmed this conclusion and were consistent with the IA-SOM fraction containing a larger amount of microbial metabolites (Poirier et al., 2005). This difference in chemical properties of the two LFs (FR and IA) was also seen when the fractionation scheme was applied to soils from eight long-term experiments in widely differing environments including temperate regions in Europe and North America, an arid tropical site in Syria and two sites growing flooded rice (Sohi et al., 2005). It was therefore concluded that this difference was a general property of soils and gives further strength to the idea that physical location within soil structure is a significant factor in determining the stabilisation of organic matter in soil, in addition to chemical structure conferring an inherent resistance to decomposition.

4.7 The living part of SOM – the soil microbial biomass

Microorganisms and soil animals form a part of SOM and can be regarded as one specific fraction with extremely important properties. A common approach is to attempt the isolation of a single species or group from soil and then study its behaviour under laboratory conditions. An alternative is to regard the entire population as an entity: an analogy is the study of a whole forest rather than an individual tree. The following quotation (from Stockdale and Brookes, 2006) eloquently puts the case for the ‘whole population’ approach:

Studies of single soil organisms, while useful in specialized cases, e.g. Rhizobia and mycorrhizae, do not yield information on the functioning of the soil ecosystem. This is because most important soil processes, e.g. carbon and nitrogen mineralization, depend upon interactions between entire suites of organisms, many of which still await identification and most of which remain unculturable. For many purposes, treating the soil microbial community as a single, undifferentiated unit, the soil microbial biomass (defined as all soil organisms <5000 mm$^3$ volume), has much to commend it.

In addition to the activities of Rhizobia and mycorrhizae, there are additional soil processes that are of vital importance to soil functioning and plant growth and that are mediated by a single microbial species or small group. These include nitrification and denitrification, and in these cases it is logical to study the specific groups involved. The application of molecular methods in soil microbial ecology is a powerful approach that is bringing new insights, either on the activity of a defined group or species or on overall biodiversity in soil. But the case for the ‘whole population’ approach is extremely strong for many investigations.
During the last 30–40 years, a suite of techniques have been developed that permit the measurement and study of soil microorganisms as a whole unit, termed the soil microbial biomass. These methods are based on the finding that chloroform kills soil organisms, making their cell contents readily decomposable, and soluble in aqueous extractants, but does not alter the decomposability or extractability of non-living organic matter. In the original method soil was fumigated with chloroform and, after chloroform removal, the soil inoculated with a little fresh soil (to reintroduce living organisms) and CO$_2$ evolution measured for 10 days under standard conditions. Evidence was accumulated to demonstrate that the additional CO$_2$ evolved, compared to that from unfumigated soil, is derived from the decomposition of killed organisms by the reintroduced population (Jenkinson and Powlson, 1976a). By using a suitable conversion factor, the quantity of C held in the cells of the original population can be calculated (Jenkinson and Powlson, 1976b). This method, now termed fumigation-incubation (FI) has been largely superseded by the fumigation-extraction (FE) method in which the additional organic C extractable in 0.5 M K$_2$SO$_4$ is used as the measure of microbial biomass C (Vance et al., 1987; Wu et al., 1990).

Variations on the FE method now exist for measuring the quantities of N, P and S held in the microbial biomass – see Stockdale and Brookes (2006). Powlson (1994) also reviews the range of methods, their benefits and limitations, and the many pitfalls in applying them inappropriately and various ways in which they have been misused.

It is commonly found that the quantity of C in the microbial biomass is in the range of 2–4% of total organic C in soil. This translates to a surprisingly large quantity when considered at the field scale. In an old grassland soil, it has been estimated that the quantity of microbial biomass in topsoil is approximately equal to the mass of 100 sheep per ha! And the quantities of crop nutrients held in the microbial biomass are of significance to crop nutrition, especially because the rate of turnover of this fraction is faster than that of SOM as a whole (Kouno et al., 2002). For example, the quantity of P estimated to pass through the microbial biomass annually in a set of grassland soils was similar to or greater than annual P uptake by grass (Brookes et al., 1984).

A valuable application of biomass C measurements is to detect trends in total SOM content before they become measurable directly. For example, at two experimental sites in Denmark where the straw of spring barley had been either burned or incorporated for 18 years, no significant difference in soil total SOC or organic N content could be measured. But measurements of biomass C and N showed increases of 40% (Powlson et al., 1987). This is because the microbial biomass has a much higher turnover rate than total organic C or N and so, after a change in management practice, the biomass moves towards its new equilibrium value far quicker than total C or N. Powlson et al. (2011a) show other examples of biomass measurements acting as an ‘early warning’ of changes in total SOC before they become measureable by classical methods.

### 4.8 SOM and soil physical conditions and implications for plant growth

Soil physical properties are determined to a very great extent by soil texture but it has been well known for centuries that they are strongly modified by organic matter. For example, SOM increases water holding capacity, facilitates the formation of stable aggregate, increases formation of stable pore spaces that improve aeration and make root penetration
Soil organic matter

easier. Johnston and Dawson (2010) give numerous examples. A more recent realisation is that small changes in SOM content can have a disproportionately large impact on certain soil physical properties, almost certainly through the role of specific fractions within the totality of SOM (Powlson et al, 2011b). Some examples are given here to illustrate this.

Watts et al. (2006) used a specially designed ploughing apparatus to measure the energy required to pull the plough through soil on all the plots of the Broadbalk experiment at Rothamsted, UK. The measured factor was specific plough draught ($S$) defined as the force per cross-sectional area of worked soil in units of kPa. A larger value of $S$ indicates that greater force is required to pull the implement through the soil; for ease of working and decreased fuel requirement, a low value of $S$ is favoured. Because of spatial variability in the soil, clay content in topsoil varied between 19% and 39% between the different plots and this factor had the greatest impact on $S$. But organic matter content had a strong modifying effect.

Table 4.2 shows data for three treatments after correcting for variation in clay content. The treatment-receiving FYM contained over three times the amount of SOC than the Nil plot receiving no manure or fertiliser. This very considerable increase in SOC led to a 15% decrease in specific draught, $S$, compared to the Nil treatment, reflecting the improved soil structure. The treatment-receiving inorganic fertilisers contained only 29% more SOC than the Nil treatment, yet the value of $S$ was decreased almost as much as in the FYM treatment. Thus, a relatively small change in total SOC concentration had a disproportionately large impact on the energy required to pull an implement through the soil. In practical terms, this is significant because it is very difficult to increase SOC by a large amount in normal agricultural practice (see Section 4.2). This finding shows that even small changes can be beneficial and therefore worthwhile.

A similar finding was shown for the effect of SOC on aggregate stability and water infiltration. Blair et al. (1995) developed a technique to determine the so-called labile fraction of SOC (termed $C_L$) and soil organic N based on oxidation with dilute potassium permanganate. This reagent is relatively weak and only oxidises about 10% of total SOC, sometimes termed the readily oxidisable fraction. It is likely that this fraction comprises the soil microbial biomass (typically about 2% of total SOC in agricultural soils) plus microbial metabolites. It has been shown that the size of this labile fraction is correlated with certain soil physical characteristics, in particular aggregate stability (measured as mean weight diameter, MWD, after wet sieving) and water infiltration rate (indicated by the value of unsaturated hydraulic conductivity, $K_{unsat}$). The method was applied to plots in long-term experiments in Australia, Germany and the UK that covered a range of SOC contents due to past management practices including manure or fertiliser applications. As expected, soils at each site with higher SOC content showed greater aggregate stability and infiltration rate with a rough correlation between total SOC and MWD or $K_{unsat}$. However, the correlations between these measurements and labile C (as opposed to total SOC) were much closer. For

### Table 4.2 Effect of soil organic C concentration on specific draught in the Broadbalk Experiment at Rothamsted, UK.

<table>
<thead>
<tr>
<th>Long-term treatment</th>
<th>SOC %</th>
<th>Specific plough draught, $S$ kPa</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nil</td>
<td>0.84</td>
<td>88</td>
</tr>
<tr>
<td>Farmyard manure</td>
<td>2.80</td>
<td>75</td>
</tr>
<tr>
<td>Inorganic fertilisers</td>
<td>1.08</td>
<td>77</td>
</tr>
</tbody>
</table>

Source: From Watts et al. (2006). Copyright Blackwell Publishing Ltd.
example, at the UK site (Blair et al. 2006) the regression equations relating MWD to total SOC ($C_T$) or labile C ($C_L$), respectively, were:

$$MWD = 0.12C_T - 0.41 \quad r^2 = 0.43 \text{ (significant at P < 0.05)}$$

$$MWD = 0.63C_L - 0.12 \quad r^2 = 0.70 \text{ (significant at P < 0.01)}$$

As with the data on plough draught, this suggests that specific fractions with total SOC are particularly significant in influencing soil physical properties and that relatively small changes in total SOC content can have a disproportionately large impact. Some practical conclusions from this finding are:

1. Management practices that lead to only small increases in SOC can be worthwhile from the viewpoint of soil physical properties that influence management.
2. Even small decreases in total SOC resulting from soil management may have a greater negative impact on soil properties than might be deduced from the small magnitude of the change.

Whether or not improvements in soil physical conditions translate into improved crop growth depends on the soil type (whether or not it is inherently well structured) and the crop and environmental conditions. Clearly identifying such effects can also be difficult because organic matter in soil provides additional nutrients to crops, whether directly (e.g. from nutrients in added manure) or indirectly via mineralisation of the additional organic matter accumulated as a result of an ongoing management practice.

Figure 4.16 shows an example of additional organic matter in soil contributing to increased yields of arable crops. The data is from an experiment on sandy soil in the UK (Johnston and Poulton, 2005) where winter wheat or spring barley was grown following either (a) a long run of arable crops, (b) a 3-year grass ley or (c) a 3-year ley comprising a grass/clover mixture. Where the arable crops followed the ploughing up of the grass/clover mixture, yields with no added N fertiliser were almost 2 t ha$^{-1}$ greater than following arable cropping. SOC would have been slightly greater following the grass/clover ley compared to long-term

![Figure 4.16](image-url) Yields of (a) winter wheat and (b) spring barley in the Woburn Ley-Arable Experiment on a sandy soil in southeast UK. Graphs show yields of arable test crops (t ha$^{-1}$) following three different previous cropping sequences: 3 years arable followed by 2 years arable test crops, •; 3 years grass ley + N followed by 2 years arable test crops, □; 3 years grass/clover ley followed by 2 years test crops, ▲. From Johnston and Poulton (2005). Reproduced with kind permission of the International Fertilizer Society.
arable cropping, but, more importantly, the soil would have contained additional nitrate and readily mineralisable N as a result of the legume. Yields of wheat or barley following grass/clover were increased further if additional N fertiliser was added. When additional N fertiliser was applied to the wheat or barley grown following long-term arable cropping, yields increased to levels comparable to those following grass/clover. Thus, in this example, all yield effects can be explained on the basis of N supply.

In the Broadbalk Wheat Experiment at Rothamsted, UK, yields of winter wheat achieved with FYM can be matched by applications of inorganic fertilisers, provided sufficient N is applied. Thus, there is little evidence of the additional organic matter in the FYM treatment having an impact on crop growth other than through nutrient supply, despite the improved soil physical structure. By contrast, in the neighbouring long-term experiment with spring barley grown each year (the Hoosfield Experiment), there is evidence of a crop benefit from increased SOM content (Johnston et al., 2009) – though the actual mechanism is still an open question. Figure 4.17 shows grain yields of spring barley in two main treatments of the experiment in four periods during which different crop varieties have been grown. With both main treatments (FYM or inorganic fertilisers, termed the PK treatment) four N treatments (including zero N) are superimposed. In the earliest period shown (1976–1979) when an old variety of spring barley was grown, grain yield in the FYM treatment could be matched in the PK treatment provided sufficient N fertiliser was applied (panel A). In later periods,

![Figure 4.17](image_url)

when more modern and N-responsive varieties were grown, this was no longer the case. For example, in the two latest periods shown (panels C and D) yields with FYM were about 2 t ha\(^{-1}\) greater than in 1976–1979 and could not be matched in the PK treatment even with the highest rate of N. The likely explanation is that the improved soil physical structure in the topsoil in the treatment receiving FYM permits more rapid and extensive growth of roots from the emerging seedlings. This leads to greater exploration of the soil, greater uptake of water and nutrients, and greater resilience to any stresses such as water shortage or disease occurring later in the growing season compared to plants in the more poorly structured soil of the PK treatment. Other factors may also contribute, including the larger and different microbial community in the FYM treatment compared to PK; for example, the composition of the rhizosphere population may give roots greater protection from soil-borne pathogens.

Of course, the same arguments can be applied to the situation of winter wheat plants in the Broadbalk Experiment – but here there is no unique benefit to crop yield from the higher SOM content in the FYM treatment. The most likely explanation of the difference between spring barley and winter wheat is that spring barley, having a growing season of only 5–6 months, is more sensitive to the impacts of soil structure than winter wheat, which grows for about 10 months.

4.9 Soil C and climate change

Soils contain a very large stock of carbon in organic forms, estimated by Batjes (1996) as 700 Pg to a depth of 30 cm and 1500 Pg to a depth of 100 cm. Thus SOC to a depth of 100 cm contains twice about the amount of C currently in atmospheric CO\(_2\) (730 Pg) and nearly three times that in vegetation (500 Pg). Small changes in this large stock of SOC could have significant impacts upon future atmospheric CO\(_2\) concentrations. If the SOC stock declines due to management practices or the impact of climate change, this means that additional CO\(_2\) will be released to the atmosphere, adding to future climate change. Conversely, if management practices can be devised that cause SOC to increase over large areas, this could be a means of slowing the current rise in atmospheric CO\(_2\) concentration and contribute to efforts to mitigate climate change.

Key factors to be considered in this context are:

1. What management practices or changes in land use tend to release additional C from SOC and should thus be avoided?
2. What management practices or changes in land use tend to increase SOC stock, either by slowing SOC decomposition or through increased atmospheric C input to soils via plants, and thus should be encouraged as climate change mitigation measures?
3. How is the current SOC stock likely to be influenced by climate change? This includes consideration of the impact of increased temperature, increased atmospheric CO\(_2\) concentration and changes in soil moisture resulting from climate change.

4.9.1 Carbon losses from soil

Converting land of high SOC content, such as forest or grassland, to arable cropping will almost always lead to a loss of SOC and emission of additional CO\(_2\) to the atmosphere. This is illustrated for grassland to arable conversions in Figures 4.4 and 4.6. Table 4.3 shows estimated changes in SOC stocks in four regions of the world over a 30 year period. Data are from a
Another land use change with serious consequences for SOC stocks, and hence for increased emission of CO$_2$ to the atmosphere and the exacerbation of climate change, is the drainage of peat soils. These soils hold vast stores of C, estimated at around 500 Pg C globally (Houghton et al., 1985; Gorham, 1991). Drainage of these soils leads to very large annual SOC losses that continue for hundreds of years; this has already occurred in the fens of eastern England. A current example is the drainage of tropical peatlands to grow additional oil palm for production of biofuels. After drainage and conversion to more aerobic conditions, peat oxidises at an increased rate leading to greatly increased emission of CO$_2$ for decades or longer. Fargione et al. (2008) estimated that the CO$_2$ emitted after clearance of peatland rainforest in southeast Asia to grow oil palm for biodiesel production would outweigh the greenhouse gas benefit from biodiesel replacing fossil fuel for over 400 years and termed this a ‘carbon debt’.

### 4.9.2 Sequestration of carbon in soil to mitigate climate change

The converse of releasing C from soil to the atmosphere is the retention and stabilisation of additional C into SOC. The additional C retained is normally derived from CO$_2$ in the

---

**Table 4.3  Projected changes in SOC stocks in four regions of the world due to expected changes in land use and management**

<table>
<thead>
<tr>
<th>Region</th>
<th>Current SOC stock (Tg)</th>
<th>SOC change to 2030 (Tg)</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amazonia (Brazil)</td>
<td>30,000</td>
<td>-4200</td>
<td>Continued deforestation</td>
</tr>
<tr>
<td>Indo-Gangetic Plain (India)</td>
<td>1,300</td>
<td>Small loss</td>
<td>Some pasture to cropland conversion</td>
</tr>
<tr>
<td>Jordan</td>
<td>85</td>
<td>Small loss</td>
<td>Overgrazing of rangeland</td>
</tr>
<tr>
<td>Kenya</td>
<td>1,500</td>
<td>-114</td>
<td>Savannah to subsistence agriculture conversion</td>
</tr>
</tbody>
</table>

*a From expected changes in land use or management based on local expert knowledge. Source: From Milne et al. (2007). Reproduced with kind permission from Elsevier.*
Soil Conditions and Plant Growth

atmosphere through photosynthesis followed by transfer from plants to soil via roots or above-ground plant residues. Increasing SOC stock may be achieved by growing larger plants that deposit additional C into soil or changing land use from annual arable crops to perennials, grassland or forest. This process is commonly termed carbon sequestration and is promoted as a means of slowing climate change by decreasing atmospheric CO₂ concentration – or at least slowing the rate of increase. In some cases this is genuine, but there are nuances to the approach that are often overlooked. If an area of unproductive or degraded land can be revegetated with grass or trees, this undoubtedly leads to an increase in the stock of organic C, in both vegetation and soil, where this C would otherwise have been in the atmosphere as CO₂. However, if the same strategy were applied on land used for food production, wider implications have to be considered. It is true that SOC stock is increased at the site of land use change (arable to forest or grass), but, at least in some cases, some land is likely to be cleared elsewhere in the world to replace the lost production of food. The land use change elsewhere (forest or grassland to arable) will not be a direct response to a specific removal of agricultural land from production, but a general global trend. If the newly cleared land had native or semi-natural vegetation, and the soil a high SOC content, then the loss of C at such sites will tend to counteract the sequestration from afforestation at the initial site. This effect is termed indirect land use change (Searchinger et al., 2008); it has been discussed mainly in the context converting agricultural land to biofuel production but is equally valid when considering the global impacts of afforestation or other approaches involving removal of land from agriculture.

Soils used for agriculture, especially arable cropping, almost always have a lower SOC content than the same soil type under semi-natural vegetation. They are also subject to human management, so the potential for increasing SOC stock and mitigating climate change is large. Smith et al. (2008) summarise a range of management practices for increasing SOC stocks in arable and pasture systems. These authors point out that, in many cases, the biophysical potential for increasing SOC stock cannot be realised because the required practices are uneconomic, impractical or conflict with the demands of food security. But within arable agriculture, there are some management practices that have received attention as potentially practical measures to achieve C sequestration, one being reduced tillage. In both no-till (NT) and reduced till (non-inversion tillage) systems, organic C accumulates near the soil surface because it is not mixed into the plough layer. This effect can be seen in the near surface soil (0–10 cm) in Figure 4.18 showing data taken from a meta-analysis of 43 sites in North and South America comparing SOC in NT and ploughed soil (termed full inversion tillage, FIT, in this study). The meta-analysis (Angers and Eriksen-Hamel, 2008) showed a wide range of results. In many cases, the increased SOC content in surface soil under NT was offset by a slightly decreased amount compared to FIT in deeper layers. Many earlier studies have overlooked this offset, concentrating only on the additional SOC in surface soil, and this has led to an overemphasis on the potential of reduced tillage systems to sequester soil C and mitigate climate change. Nevertheless, it was concluded by Angers and Eriksen-Hamel (2008) that there was a small net accumulation of SOC under NT: on average, soil under NT contained 4.9t C ha⁻¹ more than that under FIT with the accumulation under NT tending to be more marked where the NT treatment had continued for at least 15 years. This is probably because continued disturbance tillage limits C stabilisation through the formation of macro-aggregates composed of micro-aggregates (Six et al., 2004 and discussed in Section 4.6.2). Thus, moving to reduced tillage in arable systems appears to have some role to play in sequestering C in soil, but less than was claimed in the past.
Reduced tillage also reduces the energy (and hence CO$_2$ emissions from fuel) used for crop establishment and in many soils, though not all, delivers a range of other benefits for soil quality and functioning including decreased risk of erosion. However, in environments with relatively high rainfall and/or poorly drained soils moving to NT can increase nitrous oxide (N$_2$O) emissions compared to conventionally tilled soil (Rochette, 2008). Because of the high greenhouse warming potential of N$_2$O (almost 300 times that of CO$_2$ on a 100 year basis), a small increase in N$_2$O emission can easily outweigh the modest climate change benefit of NT through soil C sequestration. Wetter conditions (in terms of the categories defined by Rochette (2008)) are more prevalent in northwest Europe than in the central regions of North America; so whether or not NT or reduced till has an overall beneficial impact for climate change, mitigation will be region-specific. The discussion around this issue emphasises the need to take account of management impacts on the fluxes of all greenhouse gases (CO$_2$, N$_2$O and CH$_4$) when considering climate change mitigation through soil management – not to concentrate only on SOC stocks, as pointed out by Smith et al. (2008).

As described in Section 4.2, adding organic materials such as manures and crop residues to soil increases SOC content. But this does not necessarily imply sequestration of C and climate change mitigation – whether or not this is the case depends on the alternative fate of the organic materials. For example, manure addition increases SOC as shown in Figure 4.3. But in practice virtually all animal manure is used in some way in agriculture; so if the manure applied to the field in Figure 4.3 had not been applied at that location, it would have been applied elsewhere and given an equivalent increase in SOC. So the

![Figure 4.18](image_url)
increase shown for the field in Figure 4.3 is simply a movement of C from one soil location to another, not a net transfer of additional C from atmosphere to soil. The same is generally true of crop residue additions such as straw, unless the alternative practice is to burn the material. However, for organic wastes that would otherwise be disposed of by burning or placing in landfill sites, addition to soil where a fraction of the contained organic C is retained does represent a genuine increase in the stock of C in the soil, and this C would otherwise have been emitted to the atmosphere as CO₂ or CH₄. One example of a material that was often not previously used to increase SOC stock is green waste compost derived from municipal wastes such as plant material from tree pruning, grass cutting or household waste. Better utilisation of such materials through addition to soil can contribute to C sequestration as well as to improving soil properties (Powlson et al., 2012). The concept of soil C sequestration, of where it provides genuine climate change mitigation and where it does not, is discussed by Powlson et al. (2011c).

4.9.3 Climate change impacts on soil C stocks

Because of the large stock of organic C held in the world’s soils, small losses of C from this large pool could have significant impacts upon future atmospheric carbon dioxide concentrations. Models that have coupled climate and carbon cycles show a large divergence in the size of the predicted biospheric feedback to the atmosphere. Central questions which still remain when attempting to reduce this uncertainty in the response of soils to climate change are (a) the temperature sensitivity of SOM, especially the more recalcitrant pools, (b) the balance between increased carbon inputs to the soil from increased plant production and increased losses due to increased rates of decomposition and (c) interactions between climate change and other aspects of global change including other climatic effects (e.g. changes in water balance), changes in atmospheric composition (e.g. increasing atmospheric carbon dioxide concentration) and land use change.

Climate change is likely to influence both the quantity of C entering soil from plant inputs and its rate of decomposition. An increased concentration of CO₂ in the atmosphere is likely to increase plant productivity and hence inputs of C from plant litter and roots, though plant growth will also be influenced by temperature and water availability. These influences could be either positive or negative in different regions, depending on whether climate change relieves a current constraint to growth or imposes a new constraint. Effects may well vary between seasons. It is generally thought that increased temperature will increase the rate of decomposition of SOC but there are numerous caveats to this. First, decomposition rate will be affected by other climate-related factors such as soil moisture which in turn is a result of rainfall amounts and patterns. Second, there is some debate about the actual impact of temperature on SOC decomposition, and it is possible that different SOC fractions (the more and less recalcitrant) are influenced differently. Given these uncertainties and the numerous possibilities for different factors to interact and cause opposing effects, it is not surprising that it is difficult to make unequivocal statements about the impact of climate change on SOC and the role of soil C in climate feedbacks.

In one study (Jones et al., 2005), the Hadley Centre General Circulation Model was linked with the RothC SOC dynamics model and run for a period of 240 years assuming certain climate changes but taking no account of land use change. The results showed large losses of soil C in some regions, especially the Amazon region of South America and the northwest of North America, and gains in some other regions. Overall it showed a loss of SOC of 54 Gt averaging 0.23 Gt year⁻¹. This is equivalent to an increase of about 3% on current
Soil organic matter

Soil organic matter (IPCC, 2007) and is far from negligible. For Europe, one modelling study indicates that increased plant productivity would likely counterbalance increased decomposition due to climate change up to 2080 (Smith et al., 2005).

**Especially vulnerable soils**

High-latitude regions are thought to be particularly vulnerable to SOC loss under climate change for a number of reasons. Firstly, high latitudes are projected to experience some of the greatest warming (see Mitchell et al., 2004). Secondly, at cooler temperatures, especially for those at around freezing point, a change in temperature has a much greater impact than an increase in temperate regions. This is particularly true of permafrost soils in the taiga and tundra, which hold around 500 Pg C and could lose this carbon rapidly under warming. The release of this huge stock of carbon to the atmosphere could significantly enhance further climate warming (Zimov et al., 2006).

Peatlands and wetlands are also expected to be particularly sensitive to climate warming. Peatlands also hold vast stores of carbon: estimates range from ~200 Pg C to 860 Pg C with most estimates in the region of around 500 Pg (e.g. Houghton et al., 1985; Gorham, 1991). Plant productivity in cool northern peatlands is relatively low compared to the large carbon losses that can occur when these soils are drained, with yearly losses of C of 0.8–8.3 t C ha⁻¹ y⁻¹ (Nykänen et al., 1995; Lohila et al., 2004; Maljanen et al., 2004) and losses occurring for over a hundred years (Lohila et al., 2004). If increasing temperature increases evapotranspiration (either coupled with or independently from changes in precipitation), peatlands would be expected to dry causing decomposition rate to increase further. Given the large stocks of carbon in peatland soils, losses of carbon will likely far exceed any increase in plant productivity arising from increases in temperature.

Other sensitive regions include currently hot arid areas. Though the soils in these areas tend to be low in carbon (Batjes, 1996), complete failure of plant growth in a warmer world would switch off all inputs of carbon to the soil, enhancing desertification and degradation. Even partial loss of vegetation integrity could make soils more vulnerable to degradation through other agents such as grazing and cultivation. Since soil C stocks are not as high in these regions, total potential C losses are lower, but these regions cover large areas of the arid tropics/subtropics.

**Temperature impacts on SOC decomposition**

The temperature sensitivity of litter and SOC decomposition is important, not only because of the rate of decomposition itself but also because it determines the magnitude of feedback from atmospheric CO₂. Additional CO₂ in the atmosphere would lead to increased global warming and still further additional SOC decomposition. In general, experimental results show that any soil moved to a higher temperature evolves more CO₂. However there is debate on this topic, largely due to the existence of confounding and feedback effects which may or may not be taken into account in different approaches (Davidson and Janssens, 2006; Kirschbaum, 2006; Pare et al., 2006; Briones et al., 2007; Czimczik and Trumbore, 2007). A specific issue is the sensitivities of the more recalcitrant fractions of SOC: in a warmer world, it is assumed that the more readily decomposable parts of total SOC will be selectively decomposed, leaving the more recalcitrant fractions as a greater proportion of the remaining C and thus having a larger impact on the overall sensitivity. Some studies suggest that recalcitrant C is not sensitive to temperature variation (Giardina and Ryan, 2000), others
suggest that it is more sensitive to temperature than labile pools (Fierer et al., 2005, 2006; Knorr et al., 2005) or that recalcitrant and labile pools have a similar temperature sensitivity (Fang et al., 2005; Conen et al., 2006). This issue is a matter of ongoing research and debate.

References


Smith, P., Martino, D., Cai, Z., Gwary, D., Janzen, H.H., Kumar, P., McCarl, B., Ogle, S., O’Mara, F., Rice, C., Scholes, R.J., Sirotenko, O., Howden, M., McAllister, T., Pan, G., Romanenkov, V,


5 Nitrogen and phosphorus cycles and their management

Phil M. Haygarth¹, Richard D. Bardgett¹, and Leo M. Condron²

¹Lancaster Environment Centre, Lancaster University, Lancaster, UK
²Agriculture and Life Sciences, Lincoln University, Canterbury, New Zealand

5.1 Introduction

Nitrogen (N) and phosphorus (P) are the most important macronutrients that sustain soil, ecological and human life on Earth. All plants require adequate amounts of water, light, carbon dioxide and, critically, nutrients in order to allow them to grow to their maximum potential, and a shortage of nutrients can cause serious restrictions to crop growth (Brennan and Bolland, 2007; Cassman et al., 2002; Haygarth et al., 1998; Jarvis, 1993; Jarvis, 1998; Ladha et al., 2005). They give the basic nutrition for the production of crops, animals and fibre, providing the foundations upon which human existence is built. Soil provides the critical veneer through which the N and P cycling occurs, and plant growth is highly dependent on the availability of these nutrients (Haygarth and Ritz, 2009). Humans manage the N and P cycles in agricultural systems to enhance food production and to minimise outward consequences on the environment. The focus of this chapter is primarily, but not exclusively, on N and P cycles in managed agricultural soils.

Although the fundamental global need for N and P and their functions are similar, their properties are quite different. Nitrogen is group 15 in the periodic table. Critically from a soils perspective, N is a vital component of fertilisers and foods. Nitrogen gas and liquid is colourless and odourless, with the gas being generally inert and making up about 78% of the atmosphere by volume. When mixed with oxygen and subjected to electric sparks (hence its use in explosives), N forms nitric oxide (NO) and then nitric dioxide (NO₂). When heated under pressure with hydrogen in the presence of a suitable catalyst, ammonia forms (this is called the Haber process). Nitrogen can be ‘fixed’ from the atmosphere by free-living and symbiotic bacteria that associate with the roots of certain legume plants, such as clover, and non-legumes, such as alder, that have the actinomycete Frankia as their endophyte. Phosphorus is a colourless/red/silvery white element and has the atomic number 15 and an atomic weight of 30.97, with a standard state that is solid at 298 K (Emsley, 2000). It is group 15 in the periodic table (Parkes and Mellor, 1939) and is an essential component of living systems, playing a key role in the functioning of nervous tissue, bones and cell protoplasm. Nitrogen on the other hand has the atomic number 7 and an atomic weight of 14.01,
with a standard state that is gas at 298 K (Parkes and Mellor, 1939). As a consequence of these differing properties, the behaviours and cycling of N and P through soils and plants are quite different, with the N cycle having a larger and more significant atmospheric component than the P cycle, and N being biologically renewable via N fixation.

Although plants are known to access organic N forms, most N in agricultural systems is taken up by plants as ammonium (\(\text{NH}_4^+\)) or nitrate (\(\text{NO}_3^-\)) ions, but is then reduced in the plant and, when synthesised with carbohydrates, converted to amino acids mainly in the green leaf itself. Hence, as the extra level of N supply increases compared to other nutrients, the extra protein produced allows the plant leaves to grow larger and to have a larger surface area for photosynthesis. Thus, for many crops, the amount of leaf area available for photosynthesis is proportional to the amount of N supplied. In agricultural systems, N usually has a larger effect on crop growth, yield and crop quality than any other nutrient. Nitrogen is cycled from large atmospheric pools as dinitrogen gas into the soil ecosystem, either through biological fixation by N-fixing plants and free-living bacteria or by chemical fixation to produce N fertiliser (De Clercq et al., 2001). In the soil, the largest pool of N is in the organic matter as nitrogenous compounds arising from the decomposition of plant material and microorganisms and added organic matter (e.g. manure in agricultural systems). The soil food web, including microbes and fauna, plays an important role in making organic N available to plants through mineralisation, resulting in the release of \(\text{NH}_4^+\), and via microbial immobilisation that reduces the availability of N to plants. Thus, the rates of mineralisation/immobilisation play an important role in controlling the soil N cycle and thus governing the transfer of N to the wider environment. The \(\text{NH}_4^+\) produced during the mineralisation process is converted to \(\text{NO}_3^-\) by nitrification and involves oxidation of \(\text{NH}_4^+\) to nitrite (\(\text{NO}_2^-\)) and then to \(\text{NO}_3^-\), which, as a highly mobile ion, forms the basis upon which N transfer to the atmosphere or water can potentially occur. Figure 5.1 shows a simplified schematic of the N cycle through plants and soils.

![Figure 5.1](image-url)  
Phosphorus plays an essential role in agriculture and in all forms of life: respiration, photosynthesis in green leaves, microbial turnover and decomposing litter all require adequate levels of P in specialised forms (Cole et al., 1977). In agricultural cropping systems, adequate supplies of P are essential for seed and root formation, crop quality and strength of straw in cereals and for the accumulation and release of energy during cellular metabolism (Finkl jr and Simonson, 1979). Phosphorus in natural ecosystems is, like N, usually a scarce resource and is efficiently recycled: in contrast, in agricultural systems, P is removed in the crop or animal products. This means that P in fertilisers and animal fodders/concentrates is imported to the agricultural system in order to sustain productivity, but only 5–10% of fertiliser P that is added to soil is taken up by crops (Loehr 1974). The P cycle through managed soils and plants is principally controlled by inputs from mineral fertilisers and from animal dung and manures (which will also act as a conduit for ‘new’ farm P that is brought into the soil via P in the food concentrate supplement (common in dairy cows) – animal–soil pathway). In soil P processes are generally controlled by inorganic adsorption–desorption reactions, biologically controlled mobilisation and immobilisation by the microbial biomass, and the forms of P, principally, whether it is in an organic or inorganic form. Figure 5.2 shows a simplified schematic of the P cycle through plants and soils.

### 5.2 Sources of nitrogen and phosphorus to soil-plant systems

In natural ecosystems, N inputs are dominated by biological fixation. This process involves the conversion of atmospheric N₂ gas to ammonia by bacteria (predominantly *rhizobia, cyanobacter* spp.) that can freely exist in soil or are more commonly found in symbiotic association with plants such as legumes (*Fabaceae*) (Canfield et al., 2010). Biological N fixation inputs are not well quantified in these systems and are believed to range from 10 to 150 kg N ha⁻¹ an⁻¹. In addition, smaller quantities of N (1–20 kg N ha⁻¹ an⁻¹) are added in the form of direct atmospheric inputs of ammonium, ammonia, nitrate and particulate and organic N. These are derived from a variety of sources, including lightning, dust and anthropogenic pollution. Historically, N inputs to managed agroecosystems were dominated by biological N fixation associated with crop and pasture legumes (McNeill and Unkovich, 2007). These inputs still play a significant role in meeting the N requirements of food and forage crops, although the rate of N fixation varies widely depending on a combination of factors including plant species and cultivar, soil fertility (e.g. pH and P status) and environmental conditions. For example, pea crops such as chickpea (*Cicer arietinum*) fix 40–70 kg N ha⁻¹ an⁻¹ compared with 200–250 kg N ha⁻¹ an⁻¹ fixed by lucerne (*Medicago sativa*) and soya beans (*Glycine max*), while up to 600 kg N ha⁻¹ an⁻¹ can be fixed by some clovers under favourable conditions. The nature and quantities of N inputs to agroecosystems (and by association natural ecosystems) were changed dramatically in the early twentieth century by the development of the Haber process of producing ammonia from atmospheric dinitrogen gas (Canfield et al., 2010; Vitousek et al., 1997). The subsequent large-scale manufacture and use of synthetic N fertilisers such as urea, calcium ammonium nitrate and anhydrous ammonia contributed significantly to increased agricultural production across the globe. Inputs of
fertiliser N increased steadily during the latter half of the twentieth century, in particular, to the current annual global total utilisation of 90 million tonnes, which will increase further in the future with expansion in demand linked to continued population growth. China currently consumes 25% of total N fertiliser production, followed by North America (16%), Europe (15%) and India (12%). To put these various N inputs into perspective, total annual global N inputs are estimated to be 170 million tonnes, which is sourced from a combination of industrial fixation (98), biological fixation (40), combustion (22) and lightning (10).

In contrast to N, atmospheric P inputs to natural and managed ecosystems in the form of dust are very limited at less than 1 kg P ha\(^{-1}\) an\(^{-1}\), and P in natural ecosystems is almost exclusively derived from sparingly soluble primary minerals (apatite) present in the soil parent material (Vitousek et al., 2010). The manufacture of P fertilisers from phosphate rock (deposits of apatite P) was developed in the mid-nineteenth century and resulted in dramatic increases in the productivity of agroecosystems (Smit et al., 2009). Currently,
Table 5.1  Characteristics of some continents (and N. America) with respect to area of arable land (including permanent crops), yield and area of total cereals, P fertiliser consumption and rate (on arable land). P off-take with cereals is calculated assuming a content of 3.5 kg P/t (FAO statistics 2005/6).

<table>
<thead>
<tr>
<th>Continent</th>
<th>Area arable + perm. crops (Mha)</th>
<th>Total cereals Area (Mha)</th>
<th>Area (%)</th>
<th>Production (Mt)</th>
<th>P fertiliser applied (Mt P)</th>
<th>Off-take cereals (Mt P)</th>
<th>Cereal yield (tha⁻¹)</th>
<th>Average P-rate (kg ha⁻¹)</th>
<th>Cereal P off-take (kg ha⁻¹)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Africa</td>
<td>239</td>
<td>105</td>
<td>44</td>
<td>146</td>
<td>0.4</td>
<td>0.5</td>
<td>1.4</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>Asia</td>
<td>577</td>
<td>327</td>
<td>57</td>
<td>1141</td>
<td>10.0</td>
<td>4.0</td>
<td>3.5</td>
<td>17</td>
<td>12</td>
</tr>
<tr>
<td>Europe</td>
<td>296</td>
<td>121</td>
<td>41</td>
<td>395</td>
<td>1.8</td>
<td>1.4</td>
<td>3.3</td>
<td>6</td>
<td>11</td>
</tr>
<tr>
<td>N. America</td>
<td>229</td>
<td>78</td>
<td>34</td>
<td>463</td>
<td>2.5</td>
<td>1.6</td>
<td>5.9</td>
<td>11</td>
<td>21</td>
</tr>
<tr>
<td>Oceania</td>
<td>55</td>
<td>49</td>
<td>35</td>
<td>23</td>
<td>0.4</td>
<td>0.1</td>
<td>1.2</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td>World</td>
<td>1562</td>
<td>700</td>
<td>45</td>
<td>2342</td>
<td>17.2</td>
<td>8.2</td>
<td>3.3</td>
<td>11</td>
<td>12</td>
</tr>
</tbody>
</table>

The P-rate is calculated for total arable area including permanent crops (not just cereals).

Source: Redrawn from Smit et al. (2009).
the majority of the 20 million tonnes of phosphate rock produced annually around the world are used in the manufacture of fertilisers. Most fertiliser P applied to soil is in the form of water-soluble P (e.g. monocalcium phosphate and diammonium phosphate) made by acidulation of phosphate rock. Some forms of phosphate rock can be used directly as fertiliser, although their effectiveness is limited in many situations compared with water-soluble fertiliser P. Table 5.1 shows the usage of P fertilisers around the continents of the world.

In grazed natural and managed ecosystems, a significant proportion of the N and P taken up by plants from soil are returned in the form of dung and urine. For example, in grazed pastoral systems, 70–85% of the ingested N and P are returned to the soil in excreta (Haynes and Williams, 1993). Moreover, given that the nature of grazing means that N and P in excreta are seldom if ever returned to the area of uptake, this leads to significant N and P inputs localised in small areas, which can have profound effects on nutrient dynamics and bioavailability. Dung deposits of grazing sheep represent N and P inputs of 130 and 35 kg ha⁻¹, respectively, while the corresponding N and P inputs for cattle dung deposits are 1000 and 280 kg ha⁻¹. Similarly, animal manure and slurry are important N and P inputs in intensive management systems where animals are housed indoors for part or all of the year. In some of these systems, most of the forage consumed by the housed animals is grown locally, and the N and P can therefore be effectively recycled by appropriate management and distribution of the manure and slurry resources. However, in high-intensity farming systems such as pig and poultry production in Europe and North America, almost all of the animal feed is imported from other parts of the respective continents or other countries. This large-scale dislocation of nutrients leads to significant net inputs of N and P when large quantities of manure are effectively disposed of by application to adjacent land. This results in the accumulation of large amounts of N and P in soil in excess of agronomic requirements, which in turn increases diffuse N and P transfer from the land to water bodies with consequent adverse environmental consequences linked to eutrophication.

5.3 Chemical and biological processes of nutrient cycling, transformations and bioavailability

5.3.1 Nitrogen mineralisation

Mineralisation is the process by which soil microbes break down soluble and insoluble organic matter and convert it into inorganic forms that are available for plant uptake (Figure 5.1). The majority of soil N (some 96–98%) is found in dead organic matter in the form of complex insoluble polymers such as proteins, nucleic acids and chitin. These polymers are too large to pass through microbial membranes, so microbes produce extracellular enzymes (e.g. proteinases, ribonucleases and chitinases) that break them down into smaller, water-soluble subunits that can be absorbed by microbial cells (e.g. amino acids). This material is called dissolved organic N (DON) and it makes up a large proportion of the total soluble N pool in natural ecosystems, especially those with highly organic soils (Schimel and Bennett, 2004). Even in fertilised agricultural soils, DON concentrations can be equal to, or even more than, inorganic N (Bardgett et al., 2003).

Most DON is absorbed by free-living soil microbes, which use C and N within it to support their growth. When microbes are C limited, they use the C from DON to support
Soil Conditions and Plant Growth

their energy needs and excrete plant-available ammonium (NH$_4^+$) as a waste product, i.e. N is mineralised. When DON is insufficient to meet microbial N demand, microbes absorb additional inorganic N (NH$_4^+$ and NO$_3^-$) from soil solution, i.e. N is immobilised by the microbial biomass which reduces inorganic N availability for plants. Both N immobilisation and mineralisation occur simultaneously, and a key factor regulating this balance is the relative demand by microbes for C and N, which is determined by the C:N ratio of the organic material used. Heterotrophic bacteria and fungi have C:N ratios ranging from 4:1 to 12:1, but as they break down organic matter they respire much of the C contained within it and use the remainder for biomass production (Kaye and Hart, 1997). There is a critical substrate C:N ratio of about 30:1, which is needed to meet microbial needs for N (Kaye and Hart, 1997); if substrate C:N ratios are greater, which is often the case for plant residues, microbes immobilise exogenous inorganic N and thereby reduce the availability of N for plants (Kaye and Hart, 1997).

Not all DON is utilised by microbes, and significant quantities may be leached from soil in drainage waters or taken up directly by plants as amino acids or peptides, thereby bypassing microbial mineralisation. This has been shown to be the case in many natural ecosystems, including arctic tundra (Chapin et al., 1993; Kielland, 1994) and boreal forest (Näsholm et al., 1998; Nordin et al., 2001). Agricultural plants also have the capacity to uptake amino acids (Näsholm et al., 1998; Weigelt et al., 2005) and peptides (Hill et al., 2011) directly, but this is thought to be of limited importance for plant nutrition due to rapid microbial turnover of DON in fertile agricultural soils (Bardgett et al., 2003; Owen and Jones, 2001).

5.3.2 Nitrification and denitrification

Nitrification is the process by which ammonium (NH$_4^+$) is oxidised to nitrite (NO$_2^-$) and then to nitrate (NO$_3^-$) (Figure 5.1). This process is carried out by autotrophic bacteria that are classified into two groups on the basis of whether they oxidise NH$_4^+$ to NO$_2^-$ (Nitrosomonas and other Nitros-genera) or NO$_2^-$ to NO$_3^-$ (Nitrobacter and other Nitro-genera). Heterotrophic nitrification can also occur, especially by fungi in forest soils. The process of nitrification is influenced by several soil factors, but the availability of NH$_4^+$, which is the sole energy source for autotrophic nitrifiers, is of primary importance; as a result, rates of nitrification tend to be highest in heavily fertilised agricultural soils. Oxygen is also required, so factors that influence the diffusion of O$_2$ through soil, such as soil moisture status and structure, will also affect rates of nitrification. Nitrification is also affected by soil temperature, being optimal at 30–35°C, but it can also occur in cold, but unfrozen, soils (Brooks et al., 1996). Rates of nitrification are also affected by soil moisture, being very low in dry soils where thin water films restrict the diffusion of NH$_4^+$ to nitrifiers. Nitrification is an acidifying process because the oxidation of NH$_4^+$ to NO$_3^-$ yields H$^+$ ions and is also sensitive to changes in soil pH being negligible below pH 4.5. An important recent development is the discovery that archaeal ammonium oxidisers – which perform the first step in the process of nitrification which was previously thought to be carried out exclusively by bacteria – are the most abundant ammonium oxidisers in soil and are likely to play a substantial role in the N cycle (Leininger et al., 2006).

Plants and microbes take up NO$_3^-$ produced by nitrification, but large amounts of NO$_3^-$ may also be lost from soil through leaching and denitrification. Compared to positively charged cations (e.g. NH$_4^+$) that are retained in soil by negatively charged soil colloids (e.g. clay minerals), NO$_3^-$ is highly mobile in soil solution. As a result, when not taken up by microbes or plants, or when soils have a low anion exchange capacity, it is often lost from
soil to drainage waters. This is especially common in coarse textured, sandy soils after rain events, in heavily fertilised soils and when soils are irrigated. Nitrate is also lost from soil by denitrification, which is the microbial reduction of NO$_3^-$ to gaseous NO, N$_2$O and N$_2$. Denitrification is carried out by a wide range of microorganisms under conditions of high NO$_3^-$ and C supply, and low oxygen, and is especially prevalent in wetlands that either receive inputs of NO$_3^-$ or when they have a surface aerobic zone where nitrification occurs, such as in paddy fields. Denitrification is also common in fertilised grasslands in high rainfall regions, such as western Britain, where soils may be wet for much of the year. High rates of denitrification are of particular concern since it is a major source of N$_2$O, a potent greenhouse gas that contributes to the depletion of ozone in the stratosphere.

### 5.3.3 Nitrogen fixation

Some plants gain N via the process of N fixation, which is driven by microbes that have the capacity to reduce molecular N$_2$ to NH$_3$ and incorporate it into amino acids for protein synthesis. These microbes either live freely in soil or they live in symbiotic association with plants, forming nodules in the root where they receive carbohydrate from the plant to meet their energy needs and in turn supply the plant with amino acids formed from reduced N. Legumes are the most widely known N fixers and are associated with bacterium of the genus *Rhizobium*; however, non-legumes such as alder that have the actinomycete *Frankia* as their endophyte are also common. These symbiotic associations can supply large amounts of N into soil systems in both natural and agricultural systems, and their benefits for soil fertility have been recognised since pre-Roman times. In agricultural grasslands, legumes can contribute up to 150 kg N ha$^{-1}$ an$^{-1}$, and the main route by which this N enters the soil is through the breakdown of plant litter inputs that are enriched in N but also via the exudation or leakage of N-rich exudates from roots. Another important pathway of N fixation is through free-living bacteria that can fix N while decomposing litter and soil organic matter (Vitousek and Hobbie, 2000). Non-symbiotic N fixers occur widely, but their contribution to the N cycle is of most importance in natural ecosystems where they fix relatively small, but significant, amounts of N (<3 kg N ha$^{-1}$ an$^{-1}$) (Cleveland et al., 1999).

### 5.3.4 Phosphorus mineralisation and the role of organic forms

Organic forms of P account for 30–60% of total P in most soils and up to 90% in high organic matter soils. Soil organic P is derived from a combination of plant, animal and microbial residues, and the turnover of P through the organic and microbial biomass pools is an important component of P cycling in soil-plant systems (Condron et al., 2005). The quantities and dynamics of organic P in soil are influenced by a combination of soil and environmental conditions, together with land-use and land management practices. Continued inputs of P in mineral fertilisers and organic manures result in elevated levels of plant-available P in soil, especially in surface horizons. Phosphorus inputs accumulate in soil in inorganic and organic forms, and the latter can make a significant contribution to plant P nutrition via mineralisation, which is governed by the production of extracellular phosphatase enzymes by plant roots and soil microbes. One such example comes from a Mediterranean Antipodean study where the application of P fertiliser to pastures in ancient soils leads to the accumulation of organic P relative to inorganic P, with a particular accumulation of labile bioavailable forms (George et al., 2007) (Figure 5.3).
Figure 5.3 Phosphorus accumulation and forms under ancient Australian soils (Mediterranean climate), after fertiliser application (three treatments from a long-term pasture trial with either (1) unfertilised, (2) initially fertilised or (3) continuously fertilised). Data presented are the mean of 6 replicates with l.s.d. \( (P = 0.05) \) for both inorganic and organic fractions for each extract shown as bars. Panels with different letters (uppercase, inorganic P; lower case, organic P) within the same extract are significantly different between P treatments averaged over the entire profile. Reprinted from George et al. (2007). With kind permission from CSIRO.
As with N, the mineralisation of organic P is partly regulated by the C:P ratio of substrates. In general, when the C:P ratio is greater than 100, P is immobilised by microbes, which have a relatively high P requirement (1.5–2.5% P by dry weight compared to 0.05–0.5% for plants). As such, microbes compete strongly with plants for available P in soil. The importance of microbial immobilisation of P is illustrated by the fact that they often contain as much as 20–30% of the total soil organic P pool, which is considerably higher than the proportion of C (~1–2%) or N (~2–10%) contained in soil microbes. Microbial P is therefore an important source of available P in soil. Mycorrhizal fungi, which are associated with the roots of some 80% of terrestrial plants (Smith and Read, 1997), also play an important role in plant uptake of P (and N). Many studies demonstrate the ability of mycorrhizal fungi to enhance nutrient supply to plants, as well as to provide resistance to pathogens, insectivorous herbivores and drought (Smith and Read, 1997). For example, studies in grasslands show that arbuscular mycorrhizal (AM) fungi can enhance plant productivity by up to twofold (Van der Heijden et al., 1998), which is mostly due to increased plant uptake of P, although acquisition of N through AM fungi might also play a role (Hodge et al., 2001). Ectomycorrhizal fungi (EM) play a key role in plant nutrient uptake in boreal and temperate forests (Read and Perez-Moreno, 2003) where extensive ectomycorrhizal hyphal networks forage for nutrients in soil and excrete extracellular enzymes that degrade complex organic compounds. Mycorrhizal fungi increase P acquisition through two mechanisms: (1) they produce phosphatase enzymes that cleave ester bonds that bind P to C in organic matter, thereby releasing phosphate (PO$_4^{3-}$) that can be taken up by the fungus and passed on to the plant; and (2) they produce low molecular weight organic acids, such as oxalic acid, which enhance the availability of soil P by increasing weathering rates of sparingly soluble Ca, Fe and Al P minerals. For example, a study of P utilisation in the rhizosphere of AM ryegrass (Lolium perenne) compared with EM radiata pine (Pinus radiata) in New Zealand soils revealed that P depletion extended further from the root surface (rhizoplane) of pine (5 mm) than ryegrass (3 mm), while enhanced mineralisation of soil organic P in the pine rhizosphere was attributed to a combination of higher phosphatase enzyme activity, microbial activity and root exudate production (Chen et al., 2002) (Figure 5.4).

In an African system, George et al. (2002) demonstrated that in extremely P-deficient soils the growth of plants that exude phosphatase into the rhizosphere (i.e. plants that have a capacity to try to use organic P) leads to a depletion of organic P (NaOH-Po; Figure 5.5). This research confirmed the role of organic P in contributing to plant nutrient uptake when inorganic P supply is low (NaOH-Po was previously thought of as being recalcitrant or non-plant available) under field conditions.

Clearly, organic forms of P represent the key reservoirs for P mineralisation and recent research from tropical rainforests studied in Panama has confirmed the key role of organic forms in these highly productive systems (Turner and Engelbrecht, 2011). Figure 5.6 shows scatter plots of the relationship between concentrations of soil P fractions (determined by solution $^3$P nuclear magnetic resonance (NMR) spectroscopy, a key means of separating and identifying organic P compounds) and soil chemical properties (total P, soil pH and total carbon).

### 5.3.5 Role of soil animals

The soil hosts a highly diverse community of invertebrates, including nematodes, collembolans, mites and enchytraeids, which gain most of their nutrition from roots and their exudates, organic matter and from feeding on soil microbes (Bardgett, 2005). Although most mineralisation of nutrients is governed by the activities of bacteria and fungi, their ability to do this is strongly affected by soil animals. Soil animals affect microbial communities either directly, through
selectively feeding on fungi and bacteria, or indirectly by fragmenting of organic matter, disseminating microbial propagules and by altering nutrient availability (Bardgett, 2005). These interactions between microbes and animals drive processes of energy flow, decomposition and nutrient cycling, and therefore contribute to plant nutrient acquisition and plant growth.

As previously mentioned, microbes can immobilise significant quantities of N and P, making them unavailable for plant use. However, these nutrients can be liberated through

Figure 5.4 This shows the effect of distance from the root surface on microbial biomass carbon, inorganic and organic P forms determined in the rhizosphere of two New Zealand soils. There is clear depletion of both inorganic and organic forms of P close to the root, which is related to the presence of microbes. From Chen et al. (2002). With kind permission of Elsevier.
the predation of microbes by protozoa, nematodes and microarthropods, which excrete nutrients that are in excess of their own requirements into the soil environment. This ‘re-mobilisation’ of nutrients is termed the ‘microbial loop’ (Clarholm, 1985), and via this mechanism, soil animals can increase the availability and uptake of nutrients by plants, thereby stimulating plant growth (Bardgett, 2005; Ingham et al., 1985; Setälä and Huhta, 1991) (Figure 5.7). In addition, this grazing of microorganisms can lead to changes in the structure, size and activity of the soil microbial community, which in turn can indirectly affect processes of nutrient cycling and plant nutrient uptake.

Soil fauna also affect nutrient cycling by physical alteration of decomposing material and of the soil environment. For example, some animals consume plant detritus and egest this material into soil as faecal pellets, which provide a favourable environment for microbial growth, leading to enhanced rates of decomposition and nutrient release. Also, earthworms consume large amounts of plant litter along with soil mineral particles, and these two fractions are mixed together in the earthworm gut and then egested as surface or subsurface casts. These casts contain significantly greater numbers of microbes and have higher enzyme activities than the surrounding soil, leading to enhanced rates of N mineralisation. Phosphorus availability is also much greater in casts than in the surrounding soil because of the stimulation of phosphatase activity (Sharpley and Syers, 1976). While these positive effects of casting on microbial activity are often transitory (Lavelle and Martin, 1992), their net effect, along with other earthworm activities, is the stimulation of total soil nutrient availability and enhanced plant nutrient uptake (Edwards and Bohlen, 1996).

5.3.6 Influence of soil biodiversity

In recent years, a number of studies have examined whether the diversity of soil organisms affects processes in the soil, including nutrient cycling. There is some evidence that decomposition of organic matter can be positively influenced by the species diversity of saprophytic fungi and soil arthropods (Setälä and McLean, 2004). However, these effects mostly occur
Soil Conditions and Plant Growth

at the low diversity end of the diversity spectrum and at levels of diversity that are likely to be well below that which is found in real ecosystems (Wardle, 2002). There is also evidence that the diversity of ectomycorrhizal fungi (Jonsson et al. 2001) and arbuscular mycorrhizal fungi (Maherali and Klironomos, 2007; Van der Heijden et al., 1998; Vogelsang et al., 2006) can affect plant growth, with responses ranging from positive to negative. In general, however, the prevailing message that emerges from studies is that changes in the abundance of particular species and the nature of trophic interactions are more important drivers of

Figure 5.6 Concentrations of soil organic phosphorus fractions determined by solution $^{31}{P}$ NMR spectroscopy and chemical properties (total phosphorus, soil pH, total carbon) of soils under lowland tropical rain forest in central Panama. All correlations are statistically significant at $p = 0.05$. From Turner and Engelbrecht (2011). With kind permission from Springer Science and Business Media.
ecosystem function than is diversity per se. This is consistent with the view that there is a high degree of redundancy within soil communities, as evidenced by the prevalence of omnivory in soil food webs (Scheu and Setälä, 2002), although it is clear that some species are more redundant than others (Bardgett, 2005). Very little is also known about the ecological effects of diversity of soil bacteria that perform specialised belowground functions, such as nitrogen fixation and nitrification. However, these types of functions are often performed by a physiologically and/or phylogenetically narrow range of taxa, when compared to ‘broad’ processes such as nitrogen mineralisation (Schimel et al., 2005). As a result, it is widely thought that processes such as N fixation and nitrification are more sensitive to changes in microbial community composition and diversity than is nitrogen mineralisation. For example, symbiotic nitrogen fixation in legumes is characteristically performed by a small number of strains of *Rhizobium*, and losses of a subset of these strains can greatly impair symbiotic nitrogen fixation and legume growth (Giller et al., 1998). It has been argued that ‘broad’ processes of mineralisation and immobilisation can be ‘aggregated’ into individual components, based on specific enzyme activities and spatial distribution of microbes in microsites that are sensitive to microbial community composition (Schimel et al., 2005). This emphasises the need for a better understanding of the spatial organisation of microbial communities within the soil and the need to simultaneously disentangle the specific functions of microbes involved in nutrient cycling.

### 5.4 Processes of nitrogen and phosphorus losses from soils

#### 5.4.1 Nutrient losses to water and air

Nitrogen and P cause environmental problems when they are transferred from productive agricultural systems, both in water bodies (water quality) and the atmosphere (air quality). A summary of the problems and impacts associated is provided in Table 5.2 (modified from...
Table 5.2  Impacts of phosphorus and nitrogen losses on the wider environment.

<table>
<thead>
<tr>
<th>Substance of concern</th>
<th>Environmental and other impacts</th>
<th>Scale of agricultural contribution</th>
<th>On-farm sources</th>
<th>Scale of impact</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nitrate (NO₃⁻)</td>
<td>Water quality</td>
<td>Major source</td>
<td>Intensively managed land (inputs from fertiliser, manures, slurry, legumes and feeds)</td>
<td>Local: on-farm surface waters Regional: surface waters; catchment; aquifers National/international: maritime waters</td>
</tr>
<tr>
<td></td>
<td>● Eutrophication</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>● Health</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Economic loss</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>● Loss to farmers</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nitrite (NO₂⁻)</td>
<td>Water quality</td>
<td>Major source</td>
<td>Managed land</td>
<td>Local: on-farm surface waters Regional: surface waters and wells</td>
</tr>
<tr>
<td></td>
<td>● Fish stocks and health</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ammonia (NH₃)</td>
<td>‘Acid rain’</td>
<td>Major source (&gt;85%)</td>
<td>Fertilisers (urea) Excreta Manures and slurry</td>
<td>Local: ‘on-farm’ deposition Regional: deposition into natural ecosystems National/international: cross-boundary transfer of NH₃ and deposition</td>
</tr>
<tr>
<td></td>
<td>● Acidification of soils</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>● Eutrophication of natural systems</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Direct toxicity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nitrous oxide (N₂O)</td>
<td>Greenhouse gas</td>
<td>Substantial (likely to increase in importance as other sources decrease)</td>
<td>N fertilisers Excreta</td>
<td>Global</td>
</tr>
<tr>
<td></td>
<td>● Global warming</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ozone interactions</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nitric oxide (NO)</td>
<td>Tropospheric ozone precursor</td>
<td>Probably minor</td>
<td>Combustion, fertilisers Manures and slurry</td>
<td>Global</td>
</tr>
<tr>
<td>Phosphorus (P)</td>
<td>Water quality</td>
<td>Substantial – increasing as industrial point sources decrease</td>
<td>Fertilisers Manures, slurry P-enriched soils</td>
<td>Local: on-farm surface waters Regional: surface waters, catchments National/international: maritime waters (cross-boundary transfer)</td>
</tr>
<tr>
<td></td>
<td>● Eutrophication</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Health</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>● Toxins from algal bloom</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Economic</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>● Cost of removal</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Source: Modified from Carton and Jarvis (2001).
When transferred into waterways, they play a role in the ecology of freshwater environments through acidification and eutrophication. Acidified waters are generally characterised by low pH, high SO$_4^{2-}$ and NO$_3^-$, and both acidification and eutrophication result in major changes in the species composition of biotic communities, with reductions in species diversity and changes in productivity being common (Foy and Withers, 1995; Hornung, 1999; Vollenweider, 1968; Vollenweider and Krekes, 1982). Redfield (1934) reported that marine biomass tends to have a constant atomic ratio of 105:15:1 C:N:P and this is also applicable to freshwater systems (Uhlmann and Albrecht, 1968). The implication of this is that adding P to a water body will have a much greater effect on aquatic ecology than does adding either N or C (Hudson et al., 2000; Redfield, 1934; Uhlmann and Albrecht, 1968). Eutrophication describes the process by which lakes accumulate nutrients, with a resulting eutrophic state. This can lead to accelerated algal growth, but may have other effects such as changes in species composition, high oxygen demand and a loss of fish species and biodiversity (Foy and Withers, 1995; Hornung, 1999; Hudson et al., 2000; OSPAR, 1998; Uhlmann and Albrecht, 1968; Vollenweider, 1968; Vollenweider and Krekes, 1982). In addition to problems with eutrophication, NO$_3^-$ has been suggested as the cause of methaemoglobinemia in young infants and also with stomach cancer, but there is some controversy over the extent of this (Addiscott, 1999), while NO$_2^-$ is known to contribute to stomach cancer and is toxic to fish, especially salmonids.

Ammonia (NH$_3$) is a reactive gas that has effects both on atmospheric chemistry and on terrestrial and aquatic ecosystems (Pain and Jarvis, 1999). After reaction with acidic compounds in the atmosphere, deposition of NH$_3$ on land as a gas or as ammonium salts can add N to nutrient-poor soils (e.g. heathlands and moorlands) and change the types of plants and thus the biodiversity. Ammonium salts can also be transported over long distances in the atmosphere and contribute to acid rain effects. It is estimated that the worldwide N$_2$O emitted directly from agricultural fields is 20–30% of the total N$_2$O emitted annually (Brown and Jarvis, 2001; Mosier, 1994). Greenhouse gases trap heat in the lower atmosphere, reflecting infrared energy back to the surface, partly responsible for global warming. Many soils of the world that are naturally low in nutrients are therefore particularly sensitive to enrichment by, for example, atmospheric inputs of N or inputs of fertiliser. Losses of P from soil to the atmosphere do occur as phosphine (PH$_3$) gas (Roels and Verstraete, 2001). Phosphine emissions are unlikely to have a major impact because transport is likely to be short range and localised, but there remains as yet a lack of substantial research evidence either way here.

5.4.2 Processes of nitrogen emissions to air and water

Under anaerobic soil conditions, denitrification occurs when microorganisms are capable of using the oxygen from NO$_2^-$, NO$_3^-$ and N$_2$O in the place of elemental oxygen with the emerging gaseous end products of N$_2$O, dinitrogen (N$_2$) and NO (Jarvis et al., 2001). In some soils, chemical denitrification can also occur (see also Section 5.3.2). Most denitrification takes place in the upper layer of soils and the gases produced are released into the atmosphere. The rate of denitrification depends on soil characteristics, soil moisture and aeration, the supply of carbon, temperature and available substrate (NO$_2^-$ and NO$_3^-$). It is most likely to occur in poorly drained, fine-textured soil and in situations with high water tables where anaerobic conditions are most likely to be present (De Clercq et al., 2001). A separate process by which N is released to the atmosphere is volatilisation of NH$_3$ from soils, added fertiliser, excreta and manure (Pain and Jarvis, 1999). In excreta and manure, the main source of NH$_3$ is urea, a major nitrogenous constituent of animal urine and some
inorganic fertilisers, especially prevalent in the aqueous phase. Therefore, the application of liquid-based animal manures or inorganic fertilisers can increase the potential for volatilisation depending on the land application method. High soil pH, dry conditions and high temperatures also increase the potential for the NH$_3$ volatilisation.

Nitrogen leaching can be a major means of N loss from the soil ecosystem (De Clercq et al., 2001; Jarvis, 2000; Wilson and Ball, 1999). Because it is negatively charged, NO$_3^-$ is not readily adsorbed to soil colloids and thus it is commonly present in soil solutions and easily leached away. Some NH$_4^+$ and organic N compounds are also easily leached, but not to the same extent as NO$_3^-$. Leaching potential is governed by soil type and structure (higher potential in easily drained soils), rainfall patterns and the supply of readily available N. Higher levels of mineralisation that occur in autumn after dry summer periods can result in substantial quantities of N for leaching, as the assimilative capacities of crops decline (Scholefield et al., 1993).

### 5.4.3 Processes of phosphorus emissions to air and water

The atmospheric component of the P cycle is small, with only a small amount of phosphine (PH$_3$) gas being generated (Stein et al., 1996). Possible sources of PH$_3$ include decaying organic matter and breakdown of metal compounds (Roels and Verstraete, 2001). In terms of mass balance, there is little available information to make any difference, but in the absence of further information it is thought that the soil–atmospheric loss pathway is generally insignificant. Conversely, processes of P transfer to water are more substantial, with three mechanisms being solubilisation, detachment and incidental – described by Haygarth et al. (2005) as a transfer continuum (Figure 5.8) (Haygarth et al., 2005; Withers and Haygarth,

![Figure 5.8](image_url)
The continuum concept was developed to describe P flows from agricultural land to water and helps promote the notion of connectivity from fine-scale soil processes and the wider river catchment. Although it was devised for P, it has generic value for N. ‘Solubilisation’ refers to the release of molecules or macromolecules of P from soil surfaces and soil biota into soil water. It reflects long-term management history, and there is strong evidence that solubilisation potential increases with increased soil P status. Heckrath et al. (1995) used Olsen-P measurements to define a ‘change point’ value of 60-mg kg\(^{-1}\) soil in an arable soil, above which there is a much-enhanced tendency to release P to land drains. Detachment is the movement of soil particles with P attached, often associated with soil erosion, and is a physical mechanism for mobilising P from soil to waters (Kronvang, 1990; Sharpley and Smith, 1990). Soil erosion (Burnham and Pitman, 1986; Elliot et al., 1991; Evans, 1990; Heathwaite and Burt, 1992; MAFF, 1997; Morgan, 1980; Quinton, 1997) plays the role of particle transfer in P loss (Catt et al., 1994; Kronvang, 1990; Sharpley and Smith, 1990; Zobisch et al., 1994). The role of small colloids in the detachment and transport of P has also been described (Haygarth et al., 1997), reflecting the point that P strongly adheres to surfaces. Incidental transfers are the circumstances by which applications of manure or fertiliser coincide with conditions favouring fast discharges of excess water in overland or preferential flow (Haygarth and Jarvis, 1999).

5.5 Nitrogen and phosphorus use in agricultural systems

The amounts and forms of N and P inputs in managed agricultural systems have changed dramatically over the past 150 years as a consequence of the development and use of N and P fertilisers (Vitousek et al., 2009). Global N fertiliser use is currently 90 million tonnes which represents 50% of total N inputs, while most of the 20 million tonnes of phosphate rock mined annually is used in fertilisers. However, it is important to note that benefits of these N and P fertiliser inputs have not been evenly distributed around the world. Increased fertiliser use has been concentrated in Europe and North America, with recent increases in emerging economies such as China, India and Brazil. In intensively managed arable systems, continued large inputs of fertilisers and manures have resulted in the accumulation of significant amounts of N and P in soil, often to levels that exceed agronomic optima. This N and P loading has in turn led to increases in N and P losses from soil in drainage and runoff, with consequent adverse impacts on water quality related to accelerated eutrophication. In stark contrast, the availability and use of N and P fertilisers in large areas of Africa, Asia and Central and South America have been and remain limited, to the extent that crop productivity is commonly constrained by the lack of adequate N and P inputs on often degraded soils.

As a consequence of the potential environmental impacts of losses from the soil and, in some systems, the limited availability, there is an increasing emphasis in developed intensive arable systems on improving the utilisation efficiency of N and P inputs with the objective of reducing environmentally damaging N and P losses (Christensen, 2004; Frossard et al., 2009). This desire to improve efficiency is also being driven by recent and predicted increases in the cost of N and P fertilisers. Somewhat ironically, improving the efficiency of N and P inputs in intensive systems will also potentially benefit low-input arable systems where the focus is to maximise the agronomic benefit of limited N and P fertiliser resources.

There are fundamental differences in the pathways and dynamics of N and P cycling in the soil-plant environment which will determine the nature and efficacy of any management
regime designed to improve nutrient-use efficiency and reduce losses. Biological processes primarily control the transformations and bioavailability of N in soil, and there are significant potential nitrate-N losses from soil during drainage as well as gaseous losses as dinitrogen, ammonia and nitrous oxide. Nitrogen losses from soil are important from both agronomic (nitrate) and environmental (nitrate, nitrous oxide) perspectives (Jenkinson, 2001; Liu et al., 2010). On the other hand, P dynamics are controlled by a combination of biological and chemical processes, and while drainage losses are limited there are important issues regarding restricted availability and utilisation associated primarily with the chemical reactivity of phosphate in the soil (Frossard et al., 2000).

The nature of N dynamics in the soil-plant system means that enhanced spatial and temporal coincidence between availability and plant uptake is necessary to optimise the utilisation of soil and applied N. As P is less mobile than N, optimisation efforts have focused on enhancing the solubility and potential bioavailability of soil and fertiliser P. Based on extensive research, a variety of best management practices have been developed to tighten N and P cycles in different agroecosystems and thereby increase nutrient-use efficiency, lower fertiliser inputs and reduce the risk of adverse environmental impacts. Best management practices for N include the selection of appropriate crop rotations and varieties to optimise productivity, maintenance of plant cover to minimise soil erosion and nutrient loss in drainage, use of soil analysis in combination with nutrient budget models to calculate appropriate levels of fertiliser and manure inputs (including timing and methods of application) and the use of appropriate pest and disease control strategies (Goulding et al., 2008). For P, the focus has been on improving the management of P inputs in intensive, integrated livestock-cropping systems to reduce diffuse P losses (Condron, 2004). This includes identifying critical P source areas within a farm or catchment based on a combination of soil type, P status and hydrology, and then reducing the risk of P loss by restricting P inputs to these areas and by installing reactive buffers such as riparian strips where appropriate. In most cases, the implementation of these practices can result in improved management of N and P resources and inputs in relation to crop requirements, which in turn reduces loss and enhances long-term sustainability.

The use of short-rotation cover crops and green manures also has the potential to significantly enhance the efficiency of N and P use in a wide range of temperate and tropical cropping systems (Cherr et al., 2006; Fageria, 2007; Thorup-Kristensen et al., 2003). Table 5.3 shows the impacts of catch crops and green manures in different input cropping systems. In temperate regions, cover crops such as grasses and brassicas are grown during autumn and winter between the main crops, and their primary function is to capture and retain mineral N released from soil organic matter and crop residues, which is susceptible to loss by leaching. Green manures are most commonly legume cover crops (e.g. clover and medic) that capture N but also add N to the system via biological fixation. Cover crops and green manures are not harvested and are incorporated prior to main crop establishment to provide N and other nutrients that are released by mineralisation. Green manures are used extensively in tropical cropping systems to increase N inputs, reduce soil erosion and increase soil organic matter. Above-ground accumulation of N in cover crops and green manures ranges from 20 to over 600 kg N ha\(^{-1}\), not including the N accumulated in the root biomass. The amount of N in cover crops and green manures depends on a combination of factors including species (legume and non-legume), climate (temperate and tropical), growth period and soil fertility. For example, temperate legumes such as vetch (\textit{Vicia}), clover (\textit{Trifolium}) and medic (\textit{Medicago}) can accumulate 150–250 kg N ha\(^{-1}\). The release of N during cover crop and green manure decomposition following incorporation can
Nitrogen and phosphorus cycles and their management

Table 5.3 Impacts of catch crops and green manures in different input cropping systems.

<table>
<thead>
<tr>
<th>Potential Positive Impacts</th>
<th>Potential Negative Impacts</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Catch crops in high input systems</strong></td>
<td>Increased N fertiliser requirement due to cover crop uptake and immobilisation during residue decomposition</td>
</tr>
<tr>
<td>Reduced N leaching loss</td>
<td>Increased N loss due to poor temporal coincidence with crop demand</td>
</tr>
<tr>
<td>Improved soil fertility</td>
<td>Loss of cash crop as a result of undersowing green manure</td>
</tr>
<tr>
<td>Erosion control</td>
<td></td>
</tr>
<tr>
<td>Improved soil physical conditions</td>
<td></td>
</tr>
<tr>
<td><strong>Green manures in low-input systems</strong></td>
<td></td>
</tr>
<tr>
<td>Consistent crop yields</td>
<td></td>
</tr>
<tr>
<td>Increased soil organic matter</td>
<td></td>
</tr>
<tr>
<td>Increased soil fertility, including N and P mineralisation</td>
<td></td>
</tr>
<tr>
<td>Erosion control</td>
<td></td>
</tr>
<tr>
<td>Improved soil physical conditions</td>
<td></td>
</tr>
<tr>
<td>Improved crop rooting depth</td>
<td></td>
</tr>
</tbody>
</table>

Source: Adapted from Thorup-Christensen et al. (2003).

Figure 5.9 The seasonal dynamics of potential nitrate leaching (percolation), mineralisation (availability) and crop uptake under northwest European conditions for a spring barley undersown with ryegrass as a nitrate catch crop. The vertical grey zones indicate periods susceptible to elevated nitrate leaching. Taken from Chrisensen (2004). With permission from CAB International.

significantly reduce or even eliminate the requirement for N fertiliser. The agronomic effectiveness of cover crop and green manure N depends on the spatial and temporal coincidence between N release and crop demand, which is determined mainly by a combination of plant maturity (C:N), the timing of incorporation and prevalent soil and environmental conditions (temperature and moisture) during residue decomposition (Figure 5.9). In this example, only a proportion of the total N accumulated in cover crops and green manures will be utilised by the subsequent crop. For example, it has been found that 15–20% of cover crop N is taken up by the succeeding crop when incorporated in spring, compared with only 5–15% for autumn incorporation. Most of the N taken up or fixed by a single cover or green manure crop is retained in the soil and thereby makes a continuing long-term contribution to crop N requirements. The inclusion of a green manure crop can also enhance soil P availability and subsequent crop P acquisition. A number of legumes
have been shown to increase the dissolution and bioavailability of sparingly soluble P minerals including phosphate rocks, which are poorly utilised by most crop species. This has been mainly attributed to the enhanced exudation of organic acids (e.g. citric acid). Furthermore, recycling of labile green manure biomass containing significant amounts of N and P may increase the rate of organic N and P mineralisation in soil, which may contribute to reducing fertiliser requirements. The use of green manures may be particularly useful in organic cropping systems, where inputs of soluble N and P fertilisers are not permitted and N and P requirements are met by biological fixation and by the application of sparingly soluble phosphate rock (Condron, 2004).

Various soil treatments may also be used to improve the efficiency of N and P cycling in arable cropping systems. In this context, soil treatment includes the addition of chemical compounds or microbial inoculants that are designed to improve nutrient retention and acquisition and thereby reduce fertiliser inputs. For example, the application of the nitrification inhibitor dicyandiamide to intensively managed grassland soils has been shown to significantly reduce nitrate leaching and nitrous oxide emission, which in turn reduced N fertiliser requirement (Moir et al., 2007). It is possible that nitrification inhibitors may be used in some cropping systems to tighten N cycling. Inoculation of the plant rhizosphere with microorganisms has the potential to enhance the bioavailability of mineral and organic forms of P in soil (Richardson, 2001). For example, under laboratory conditions, it has been shown that a range of soil bacteria and fungi (Bacillus, Pseudomonas, Penicillium, Aspergillus spp.) can increase dissolution of sparingly soluble P minerals found in soil, which was attributed to the enhanced production of low molecular weight organic acids such as oxalic and citric acids. However, while there is some evidence to show that plants responded positively to soil inoculation with some of these microorganisms, this may have been due to direct growth promotion rather than improved soil P acquisition.

Certain soil amendments may also play a role in tightening N cycling in soil. Biomass-derived black carbon (charcoal) in soil originates from a combination of natural and anthropogenic fires and is widely acknowledged to be the most stable form of organic carbon in soil, with an estimated turnover time and a half-life of thousands of years. Biochar is a specific form of charcoal produced by pyrolysis of organic material, which involves low temperature (700 °C) combustion in limited oxygen. Biochar can be made from a wide range of crop and wood residue materials at a variety of scales with up to 50% of the feedstock carbon being converted to biochar. Biochar amendment of soil at high rates of application can potentially be a means of increasing total soil carbon. There is an emerging debate that says that biochar production has the potential to be used to sequester atmospheric carbon dioxide and convert it to stable soil carbon and thereby contribute to the mitigation of ongoing and predicted global climate change (Woolf et al. 2010). It has been estimated that widespread adoption of biochar production from agricultural and forest waste biomass could reduce global greenhouse gas emissions by 1.8 gigatonnes per annum, which is equivalent to 12% of current anthropogenic emissions. In addition to the potential for carbon sequestration, there is the possibility that biochar amendment of soil could improve soil fertility and plant productivity enhancing nutrient retention and bioavailability, including biological N fixation. Until recently, most of the research on the benefits of charcoal amendment on soil fertility was based on observations made on Amazonian Dark Earths (Terra Preta), which were formed by continued additions of large quantities of charcoal to soil several thousand years ago. The beneficial effects of contemporary charcoal amendment of agricultural soils, especially under temperate conditions using biochar made by pyrolysis, remain to be established. However, with regard to the effect of biochar
Nitrogen and phosphorus cycles and their management

5.6 Future soil nutrient cycles and environmental change

5.6.1 Future of phosphorus fertilisers

Unlike N, P is a non-renewable resource and the productivity of managed ecosystems is heavily dependent on continued inputs of P in the form of mineral fertilisers manufactured from phosphate rock (Condron, 2004; Frossard et al., 2000). This fertiliser P is required to maintain the high levels of production in developed agroecosystems and to increase the productivity of underdeveloped and degraded soils. The future nature and use of P fertiliser will be largely determined by the availability and supply of phosphate rock. It is widely acknowledged that there are finite reserves of readily accessible phosphate rock and that these are being exhausted at an increasing rate to meet increased demand linked to a combination of population growth and accelerated economic development (Cordell et al., 2009; Smit et al., 2009). Figure 5.10 shows the P flows through the global food system. The current annual global production of phosphate rock is 20 million tonnes, and known phosphate rock reserves are estimated to be 2,400 million tonnes. The latter represents 120 years of supply at current production rates, but it has been predicted that continued increases in demand could mean that global production of phosphate rock may begin to decline from as early as 2030. While phosphate rock deposits are found in many locations around the world, the largest reserves are located in Morocco and China (60% of total), with significant quantities present in the USA and South Africa (20%). It is likely that increased demand and declining reserves will lead to pressure on supply and price, including increased volatility. For example, in the USA, the cost of superphosphate fertiliser was relatively stable over the 30 year period from 1975 at $140–300 t\(^{-1}\) (average $215 t\(^{-1}\) ), although by 2008, the cost had increased to $800 t\(^{-1}\) in response to increased demand and shortage of supply. The potential impact of declining supply and increasing the cost of P fertilisers in the future will be especially acute in areas with high population growth such as much of Africa, where the availability of P fertiliser is already limited by a combination of economic and logistic constraints.

At least part of the increased pressure on P resources could be alleviated by improved utilisation of fertiliser P and recycling of P contained in various agricultural and urban waste streams (Frossard et al., 2009). There is an urgent need to further investigate and quantify practical means of improving the effectiveness and utilisation of existing and alternative P fertilisers (e.g. phosphate rock and waste P). For example, the inclusion of selected legume green manure crops could be used to increase the dissolution and utilisation of sparingly soluble forms of soil and fertiliser P. In intensive animal production systems, there is considerable potential to develop new and innovative waste management regimen designed to enhance P reuse and recycling and thereby reduce excessive P loading onto soils from organic waste disposal. This may require waste treatment and transport to areas where the agronomic benefit of P addition can be realised. In the long-term, it would be advantageous if the feed production system and related animal production facilities could be located in the same area that would enable efficient P utilisation and recycling. The ongoing expansion in
Figure 5.10  Phosphorus flows through the global food system (redrawn from Cordell et al. (2009)). With kind permission from Elsevier.
urbanisation across the globe will further increase net transfer of P from farmland to cities on regional and transnational scales. Given the predicted increasing value of P, it should be possible to recover a significant proportion of the P present in liquid and solid municipal and industrial waste streams and return these to rural areas as fertiliser. However, this will require considerable investigation and investment in new waste management technologies related to the collection, processing and transport of waste P resources. This includes increasing the agronomic effectiveness and fertiliser value of sparingly soluble forms of P commonly found in many waste products such as biosolids.

5.6.2 Future of phosphorus and nitrogen cycles and land-use change

Soils perform the key function of harbouring and regulating the cycling of N and P and critically support the increasing global demand for food and fibre (crops, milk, meat, wool and hide production) from an expanding human population. Phosphorus and N cycles through soils and plants have always underpinned civilisation, as their role is fundamental to the provision of food. But there are future changes on the horizon. On the one hand, the global N cycle is increasing as a result of anthropogenic interference (Gruber and Galloway, 2008), but, on the other hand, there are projections for reductions in P supply to soil (Elser and Bennett, 2011). Confusingly, at the same time there are predictions for increasing demands on soil, not as agricultural land, but as land for providing other services to society, especially in densely populated countries where multiple demands by the society are made. There is an imperative for soil managers to protect productive land and the ability of the soil and the landscape to undertake their nutrient cycling functions (Haygarth and Ritz, 2009).

References


6 Properties and management of cationic elements for crop growth

Philip J. White¹ and Duncan J. Greenwood²

¹The James Hutton Institute, Dundee, UK
²University of Warwick, Warwick, UK

This chapter describes the properties and management of both the ‘macronutrient’ cationic elements (potassium, calcium and magnesium) and the ‘micronutrient’ cationic elements (zinc, iron, copper, manganese and nickel) for crop nutrition. It begins by noting that transpiration-driven mass flow of the soil solution is required to deliver many of these elements to the root surface for adequate plant nutrition. The amounts of each element required by plants and their essential functions are then reviewed, and the molecular mechanisms by which root cells take up essential cationic elements from the rhizosphere solution are described. Aspects of soil geochemistry that determine the phytoavailability of essential mineral cations in the soil solution are presented, including the interactions between cations in the soil and the effects of soil pH. Management practices for optimising the delivery of essential mineral cations for crop production and quality are discussed, and the chapter concludes by considering the prospects for harnessing crop genetics to improve the use of cationic elements in crop production.

6.1 The movement of essential cationic elements to the root surface

Plants acquire many of the mineral elements required for their growth and fecundity from the soil solution in their cationic forms (Table 6.1; White, 2003; Maathuis, 2009). These essential cationic elements include potassium (K), calcium (Ca), magnesium (Mg), iron (Fe), manganese (Mn), zinc (Zn), copper (Cu) and nickel (Ni). The phytoavailability of these elements in the soil solution is often determined by pH (Figure 6.1; Taiz and Zeigler, 2006).

The movement of cationic elements in the soil solution is determined primarily by the processes of mass flow and diffusion (Barber, 1995; Fageria et al., 2011). Mass flow, which is also known as mass transfer or bulk flow, is the term given to the movement of solutes that is proportional to the flow of the solution in which they are dissolved. The mass flow of cations to the root surface is driven by transpiration from the shoot and the consequent uptake of water by the plant. The delivery of a particular cation to the root surface is calculated as the product of the concentration of the cation in the soil solution and the rate of transpiration. Mass flow of the soil solution will deliver ample quantities of cations that are present at high
Table 6.1 The functions of mineral elements acquired as cations from the soil solution and the critical leaf concentrations for their sufficiency and toxicity in non-tolerant crop plants.

<table>
<thead>
<tr>
<th>Element</th>
<th>Form acquired</th>
<th>Physiological functions</th>
<th>Critical leaf concentrations (mg g⁻¹ DM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Potassium (K)</td>
<td>K⁺</td>
<td>Enzyme activation; cellular osmoticum; counter-cation for anion accumulation and electrogenic transport</td>
<td>5–40 &gt;50</td>
</tr>
<tr>
<td>Calcium (Ca)</td>
<td>Ca²⁺</td>
<td>Membrane and cell-wall structure; counter-cation for anion accumulation in vacuoles; cytosolic signalling</td>
<td>0.5–10 &gt;100</td>
</tr>
<tr>
<td>Magnesium (Mg)</td>
<td>Mg²⁺</td>
<td>Constituent of chlorophyll; photosynthetic charge separation; enzyme cofactor; nucleic acid stabilisation</td>
<td>1.5–3.5 &gt;15</td>
</tr>
<tr>
<td>Iron (Fe)</td>
<td>Fe²⁺, Fe³⁺ chelates</td>
<td>Photosynthesis; mitochondrial respiration; C and N metabolism; production and scavenging of reactive oxygen species; regulation of transcription and translation; hormone biosynthesis</td>
<td>50–150 × 10⁻³ &gt;0.5</td>
</tr>
<tr>
<td>Manganese (Mn)</td>
<td>Mn²⁺ Mn chelates</td>
<td>Photosystem II; enzyme activation in photosynthesis, C and N metabolism, RNA polymerase</td>
<td>10–20 × 10⁻³ 0.2–5.3</td>
</tr>
<tr>
<td>Copper (Cu)</td>
<td>Cu⁺, Cu²⁺, Cu chelates</td>
<td>Photosynthesis; mitochondrial respiration; C and N metabolism; protection against oxidative stress</td>
<td>1–5 × 10⁻³ 15–30 × 10⁻³</td>
</tr>
<tr>
<td>Zinc (Zn)</td>
<td>Zn²⁺, Zn chelates</td>
<td>Structural stability of proteins; regulation of transcription and translation; oxidoreductases and hydrolytic enzymes</td>
<td>15–30 × 10⁻³ 100–300 × 10⁻³</td>
</tr>
<tr>
<td>Nickel (Ni)</td>
<td>Ni²⁺, Ni chelates</td>
<td>Constituent of urease</td>
<td>0.1 × 10⁻³ 20–30 × 10⁻³</td>
</tr>
<tr>
<td>Sodium (Na)</td>
<td>Na⁺</td>
<td>Osmotic replacement of K⁺; counter-cation for anion accumulation and electrogenic transport; C4 and CAM metabolism</td>
<td>Beneficial 2–5</td>
</tr>
<tr>
<td>Aluminium (Al)</td>
<td>Al³⁺</td>
<td>Herbivore defence; prevention of Fe-toxicity</td>
<td>Beneficial 40–200 × 10⁻³</td>
</tr>
<tr>
<td>Cobalt (Co)</td>
<td>Co²⁺</td>
<td>Nitrogen fixation</td>
<td>Beneficial 10–20 × 10⁻³</td>
</tr>
<tr>
<td>Lead (Pb)</td>
<td>Pb²⁺</td>
<td>–</td>
<td>– 10–20 × 10⁻³</td>
</tr>
<tr>
<td>Cadmium (Cd)</td>
<td>Cd²⁺ Cd chelates</td>
<td>–</td>
<td>– 5–10 × 10⁻³</td>
</tr>
</tbody>
</table>

The critical concentration for sufficiency is defined as the concentration in a diagnostic tissue that allows a crop to achieve 90% of its maximum yield. The critical concentration for toxicity is defined as the concentration in a diagnostic tissue above which yield is decreased by more than 10%. It should be recognised that critical tissue concentrations depend upon the exact solute composition of the soil solution and can differ greatly both between and within plant species. The latter differences reflect both ancestral habitats and ecological strategies.

Source: Data compiled from Maathuis (2009), Hänsch and Mendel (2009), Pilon-Smits et al. (2009) and White and Brown (2010).
concentrations in the soil solution relative to their concentrations in plant tissues, but often insufficient quantities of cations that are present at low concentrations in the soil solution relative to their concentrations in plant tissues, for the mineral nutrition of crops (Table 6.2). Diffusion is defined as the movement of a solute from a region of high concentration to a region of low concentration. When the uptake of a particular cation by a plant exceeds the amounts delivered by mass flow, a concentration gradient develops for that cation in the rhizosphere. In principle, the diffusion of a particular cation is calculated as the product of the concentration gradient and the diffusion coefficient of the cation, as described by Fick’s First Law (White, 2003). However, since soil is neither homogeneous nor a solution, to calculate the diffusion of a solute to the root surface, this equation must be modified to include a tortuosity factor and a term for the soil volume occupied by water (Barber, 1995; Fageria et al., 2011). Diffusive movement of cations in the soil solution is generally slow and is effective in delivering cations to the root only from distances of 0.1–15 mm from the root surface (Barber, 1995). Nevertheless, the delivery of K, Fe and Mn to the root surface is often determined by local diffusion in the vicinity of the root (Table 6.2). In addition to mass flow and diffusion, cations can be intercepted during the growth of a root system. This can contribute significantly to the mineral nutrition of the plant for nutrients whose concentrations in the soil solution approach or exceed those found in plant tissues, such as Ca, Mg, Mn and Zn (Table 6.2).
6.2 Cationic elements in plants

6.2.1 Potassium

Potassium (K⁺) is the most abundant inorganic cation in plants, and shoot concentrations up to 100 mg g⁻¹ dry matter (DM) have been reported (Broadley et al., 2004; Watanabe et al., 2007). Potassium is present at high concentrations in most plant tissues but is particularly concentrated in growing tissues and reproductive organs. It is required by plants for various biochemical and biophysical functions (Table 6.1; Maathuis, 2009; White and Karley, 2010). In particular, K⁺ is required for the activation of many enzymes, as a cellular osmoticum for rapidly expanding cells and as a counter-cation for anion accumulation and electrogenic transport processes (Leigh and Wyn Jones, 1984; Maathuis, 2009; White and Karley, 2010). The K⁺ concentrations in metabolically active cellular compartments are maintained at approximately 100 mM, whilst vacuolar K⁺ concentrations vary considerably from several hundred millimolar to a minimum of about 20 mM (White and Karley, 2010). Apoplastic K⁺ concentrations fluctuate according to K⁺ availability and transpirational water flows (White and Karley, 2010). Crop demand for K is directly related to its N demand, and the K requirements of plants are linearly related to those of N as they grow. Potassium deficiency is often observed in rapidly growing plants and K-fertilisers are, therefore, frequently applied in crop production. Potassium is delivered to the roots through a combination of transpiration-driven mass flow of the soil solution and, since K concentrations in the soil solution are generally high, local diffusion in the rhizosphere (Barber, 1995).

6.2.2 Calcium

Shoot Ca concentrations generally range from 1 to 50 mg g⁻¹ DM, reflecting both Ca availability in the soil solution and the Ca requirements of different plant species (Broadley et al., 2003; Watanabe et al., 2007). Tissue Ca concentrations are lower in commelinoid monocots (cereals and grasses) than in non-commelinoid monocots or other angiosperm species growing in the same environment, which appears to be directly correlated with the cation exchange capacity of cell walls and the free carboxyl groups of cell-wall pectins (White and Broadley, 2003). Calcium is required as a divalent cation (Ca²⁺) for various structural roles in the cell wall and cellular membranes, as a counter-cation for organic and inorganic anions in the vacuole and for coordinating cellular responses to developmental cues and environmental challenges through changes in cytosolic Ca²⁺ concentration (White and Broadley, 2003). Since Ca²⁺ is cytotoxic, submicromolar concentrations of Ca²⁺ must be maintained in the cytosol, whereas supramillimolar Ca²⁺ concentrations can be present in the vacuole and apoplast (White and Broadley, 2003). In the cytosol, Ca²⁺ is buffered by chelation with soluble organic compounds, whereas in the vacuole Ca is often present as soluble or insoluble oxalate or phytate salts (White and Broadley, 2003). Calcium deficiency is rare in nature, but can occur on soils with low cation exchange capacity (CEC). By contrast, Ca toxicity can occur on calcareous soils, although the flora of these soils is most often determined by insensitivity to P and Fe deficiencies (Lee, 1999). Calcium concentrations in the soil solution are generally in the millimolar range, and sufficient Ca for plant nutrition can be provided by transpiration-driven mass flow of the soil solution (Barber, 1995).
6.2.3 Magnesium

Shoot Mg concentrations range from 1 to 10 mg g\(^{-1}\) DM in plants with an adequate Mg supply (Wilkinson et al., 1990; Broadley et al., 2004). When growing in the same environment, most angiosperm species have a similar shoot Mg/Ca quotient, with the exception of members of the Caryophyllales order, which often have a higher shoot Mg/Ca quotient than other angiosperm species (Broadley et al., 2004, 2008; Watanabe et al., 2007). Magnesium is required for the structure and conformation of nucleic acids, as a cofactor for numerous enzymes, as a counter-cation during photosynthetic charge separation and as a constituent of chlorophyll. About 75% of leaf Mg appears to be associated with protein synthesis, through its roles in ribosomal structure and function, and between 15% and 20% is associated with chlorophyll (Wilkinson et al., 1990; White and Broadley, 2009). Most cellular Mg is associated with organic acids and proteins (Mengel et al., 2001), and a cytosolic Mg\(^{2+}\) concentration of about 0.4 mM is generally maintained (White et al., 1990). Magnesium deficiencies can occur in plants growing on shallow or coarse-textured soils, where Mg\(^{2+}\) leaches through the soil profile; on strongly acidic soils, where Al\(^{3+}\) and Mn\(^{2+}\) can inhibit Mg\(^{2+}\) uptake; and on alkaline soils, where carbonate formation and excess Ca\(^{2+}\), K\(^{+}\) and Na\(^{+}\) can reduce Mg\(^{2+}\) availability and inhibit Mg\(^{2+}\) uptake (Wilkinson et al., 1990; Römheld and Kirkby, 2007; White and Broadley, 2009). The concentration of Mg\(^{2+}\) in the soil solution generally lies between 125 \(\mu\)M and 8.5 mM, which is sufficient for transpiration-driven mass flow to supply sufficient Mg to roots for adequate plant nutrition (Wilkinson et al., 1990; Barber, 1995).

6.2.4 Iron

Shoot Fe concentrations above 50–150 \(\mu\)g g\(^{-1}\) DM are required for adequate plant nutrition (Mengel et al., 2001; White and Brown, 2010). Iron is a redox-active metal that functions in photosynthesis, mitochondrial respiration, nitrogen assimilation, production and scavenging of reactive oxygen species, regulation of transcription and translation, and in the biosynthesis of plant signals such as ethylene, gibberellic acid and jasmonic acid (Hänsch and Mendel, 2009; Guerinot, 2010). It is an integral component of many proteins and enzymes either as a haem group, as occurs in cytochromes, peroxidases and catalases; as non-haem iron bound to sulphur atoms, as occurs in ferredoxin, nitrite reductase, sulphite reductase, glutamate synthase and iron-storage proteins; or as non-haem non-sulphur bound iron, as occurs in the plastidic Fe-storage protein ferritin (Hänsch and Mendel, 2009; Guerinot, 2010). Up to 90% of the Fe in leaf cells is contained in chloroplasts and is distributed equally between stroma and thylacoids (Hänsch and Mendel, 2009; Guerinot, 2010). Iron deficiency occurs on a variety of soils and is generally caused by low Fe phytoavailability rather than low abundance of Fe in the soil (Follett et al., 1981; Shuman, 1998; Schmidt, 1999; Frossard et al., 2000; Fageria, 2009; White and Broadley, 2009). The amount of Fe in the soil solution is decreased by an increase in redox potential and/or pH and is determined by the presence of Fe(OH)\(_3\) in well-oxidised soils, by Fe\(_5\)(OH)\(_8\) in moderately oxidised soils and by FeCO\(_3\) in highly reduced soils (Lindsay and Schwab, 1982). The Fe concentrations in solutions from most agricultural soils range from 10\(^{-8}\) to 10\(^{-6}\)M, but Fe concentrations in solutions from alkaline or calcareous soils rarely exceed 10\(^{-10}\)M (Frossard et al., 2000; Fageria, 2009). The low Fe concentration in the soil solution, together with its small diffusion coefficient, makes plants reliant on transpiration-driven mass flow of the soil solution for the delivery of sufficient Fe to their roots for nutrition (Barber, 1995). Iron deficiency associated with soil
alkalinisation is referred to as lime-induced chlorosis. It is estimated that up to one-third of the world's agricultural soils are calcareous and susceptible to lime-induced Fe chlorosis (Follett et al., 1981; FAO 2010). At the other extreme, Fe-toxicities can occur in waterlogged or flooded soils, in which Fe$^{3+}$ is reduced to Fe$^{2+}$, which increases Fe$^{2+}$ availability and Fe uptake by plants (Schmidt, 1999; Fageria, 2009).

### 6.2.5 Manganese

Shoot concentrations above 10–20 μg Mn g$^{-1}$ DM are required for plant growth and fecundity (Mengel et al., 2001; White and Brown, 2010). Manganese serves as a catalytically active metal in various enzymes, including the water-splitting reaction of photosystem II, superoxide dismutase and oxalate oxidase, and also regulates the activity of several enzymes catalysing reactions in intermediatory metabolism (malic enzyme, isocitrate dehydrogenase and PEP carboxykinase), fatty acid biosynthesis, nitrogen metabolism (glutamine synthetase and arginase), the shikimic acid pathway and subsequent pathways leading to the formation of aromatic amino acids, lignins, flavonoids and the phytohormone indole acetic acid, gibberellic acid biosynthesis and RNA polymerase (Hänsch and Mendel, 2009; Williams and Pittman, 2010). Manganese deficiency can occur in plants growing on organic soils, in which Mn$^{2+}$ is chelated by organic material; alkaline and calcareous soils, in which Mn is precipitated as manganese oxides (MnO$_2$, Mn$_3$O$_4$, Mn$_2$O$_3$); and in coarse-textured sandy soils, in which Mn$^{2+}$ is subject to leaching (He et al., 2005; Fageria, 2009). Since most Mn in soils is associated with organic matter, it is often concentrated in the topsoil (Shuman, 1985; Havlin et al., 2005; Fageria, 2009). Global crop production is restricted by Mn toxicity to a greater extent than Mn deficiency (Mengel et al., 2001; He et al., 2005). Tissue concentrations above 0.2 mg Mn g$^{-1}$ DM can be toxic to some plants, although the ability of an individual plant to tolerate Mn toxicity is influenced by plant genotype, plant growth rate and the concentration of silicic acid in the soil solution (Table 6.1; Horst, 1988; White and Brown, 2010). Several plants that hyperaccumulate Mn to concentrations greater than 10 mg Mn g$^{-1}$ DM are also known (Baker and Brooks, 1989; Whiting et al., 2005). Intriguingly, plants lacking sufficient Zn often accumulate excess Mn, suggesting a physiological interaction between these two elements (Brady and Weil, 2002). Manganese toxicity is prevalent in plants grown on acid mineral soils and can also occur in plants growing on waterlogged or flooded soils as redox potential becomes more negative (Mengel et al., 2001; Kochian et al., 2004; Fageria, 2009). On acid soils, Mn toxicity can be corrected by increasing soil pH through liming.

### 6.2.6 Zinc

Shoot Zn concentrations vary greatly, from 50 to 200 μg Zn g$^{-1}$ DM in agricultural plants supplied with adequate Zn to over 10 mg Zn g$^{-1}$ DM in Zn-hyperaccumulator plants growing in soils with high Zn concentrations (Broadley et al., 2001, 2007). Zinc occurs in plants as Zn$^{2+}$, which is associated with a multitude of proteins. For example, in Arabidopsis thaliana, Zn$^{2+}$ is associated with over 2,300 proteins in 181 gene families (Broadley et al., 2007). These proteins include a large number of zinc-finger-containing proteins and transcription factors, oxidoreductases and hydrolytic enzymes such as protein kinases, metalloproteases and phosphatases (Broadley et al., 2007; Hänsch and Mendel, 2009; Clemens, 2010). Most agricultural soils contain 10–300 μg Zn g$^{-1}$, which would be sufficient for adequate crop nutrition if it were phytoavailable, and Zn$^{2+}$ concentrations in the soil solution generally
range from $10^{-8}$ to $10^{-6}$ M (Alloway, 1995; Barber, 1995; Welch, 1995; Frossard et al., 2000; Broadley et al., 2007). Zinc has a small diffusion coefficient and sufficient Zn for plant nutrition must be provided to the roots by transpiration-driven mass flow of the soil solution (Barber 1995). Zinc deficiency is the most common micronutrient deficiency in crops (Cakmak, 2002, 2004; Alloway, 2004; White and Broadley, 2009). High pH often limits Zn phytoavailability, and Zn deficiencies occur in plants growing on eroded calcareous or alkaline soils, especially in arid and semiarid environments, where Zn$^{2+}$ concentrations in the soil solution can be as low as $10^{-11}$–$10^{-9}$ M (Hacisalihoglu and Kochian, 2003; Broadley et al., 2007; Cakmak, 2008; Clemens, 2010). Zinc toxicity is less common than Zn deficiency, but can occur in crops growing on acidic soils enriched by anthropogenic Zn inputs, such as agricultural soils treated with sewage sludge, soils of urban and peri-urban areas and land contaminated by mining or smelting activities (Chaney, 1993; Broadley et al., 2007).

### 6.2.7 Copper

Plants require concentrations above 1–5 μg Cu g$^{-1}$ DM for their growth and fecundity (Mengel et al., 2001; White and Brown, 2010). Copper is a redox-active metal and, consequently, has roles in many aspects of plant biochemistry including photosynthesis and mitochondrial respiration, carbon and nitrogen metabolism, protection against oxidative stress and cell-wall remodelling (Burkhead et al., 2009; Hänsch and Mendel, 2009). Copper is the catalytic metal in many oxidases because of its high affinity for dioxygen molecules. In Arabidopsis thaliana, Cu is associated with over 100 different proteins, including Cu-binding proteins/chaperones and various enzymes (Kramer and Clemens, 2005). Several of these enzymes are important for mounting plant defence responses to pests and pathogens (Fageria et al., 2011). About 50% of leaf Cu is associated with chloroplasts, where it participates in photosynthetic reactions (Hänsch and Mendel, 2009). For some essential functions, either Cu or Fe can be utilised and, depending upon their Cu and Fe nutrition, plants can switch their metabolism between enzymes catalysing the same biochemical reaction containing either Cu or Fe (Hänsch and Mendel, 2009). Examples include Cu-nitrite versus heme-nitrite reductase, Cu/Zn-superoxide dismutase versus Fe-superoxide dismutase and cytochrome oxidase versus di-iron oxidase. In legumes, Cu deficiency reduces nodulation and N$_2$ fixation. High pH is often the major factor restricting Cu phytoavailability, and Cu deficiency is often observed on calcareous or alkaline soils (Lindsay, 1979; He et al., 2005). Copper deficiency is also observed in plants growing on highly organic soils, since organic matter chelates Cu tightly (He et al., 2005). The affinity of organic matter for Cu is also the reason why, in many soils, Cu is concentrated in the topsoil (Fageria, 2009). Copper has a small diffusion coefficient and transpiration-driven mass flow of the soil solution must provide Cu for plant nutrition (Barber, 1995). Typical Cu$^{2+}$ concentrations in the soil solution range from $10^{-9}$ to $10^{-6}$ M (Barber, 1995; Welch, 1995). Excessive Cu is detrimental to plant growth and tissue concentrations above 15–30 μg g$^{-1}$ DM are toxic to most plants, with the exception of Cu-hyperaccumulator plants in which leaf Cu concentrations over 1 mg Cu g$^{-1}$ DM have been reported (Table 6.1; Baker and Brooks, 1989; Broadley et al., 2001; Mengel et al., 2001; He et al., 2005; Whiting et al., 2005; White and Brown, 2010). Copper rarely reaches toxic concentrations in natural soils, but anthropogenic activities, such as the mining and refining of metal ores, the repeated use of Cu-containing fungicides, pesticides or herbicides, and the application of Cu-rich manures from commercial pig or poultry farms, have contaminated both natural and agricultural land (Welch et al., 1991; He et al., 2005).
6.2.8 Nickel

Trace amounts of Ni are sufficient for adequate plant nutrition and tissue concentrations above 2–4 ng g\(^{-1}\) DM suffice for plant growth and fecundity (Brown et al., 1987; Tejada-Jiménez et al., 2009; White and Brown, 2010). However, crop plants generally contain tissue concentrations of 0.1–5 μg Ni g\(^{-1}\) DM (Mengel et al., 2001). Nickel is required by plants for the activation of the enzyme urease, which catalyses the hydrolysis of urea to carbon dioxide and ammonia. Tissue concentrations exceeding 10–50 μg Ni g\(^{-1}\) DM are toxic to most plants, with the exception of Ni-hyperaccumulator plants, which can tolerate in excess of 1 mg Ni g\(^{-1}\) DM (Broadley et al., 2001; Mengel et al., 2001; He et al., 2005; Whiting et al., 2005; Reeves, 2006; Fageria, 2009; Tejada-Jiménez et al., 2009; White and Brown, 2010). Nickel deficiency is rarely seen in nature and Ni concentrations in the soil solution from most agricultural soils are less than 10 μM Ni, which does not produce Ni toxicity (Welch, 1995; Mengel et al., 2001). However, Ni toxicity can occur worldwide on soils derived from Ni-rich serpentine or ultrabasic rocks or as a consequence of anthropogenic activities, such as the mining and refining of metal ores or the application of Ni-rich biosolids or municipal composts to agricultural land (He et al., 2005; Fageria, 2009). Like Fe, Zn, Mn and Cu, Ni has a small diffusion coefficient and transpiration-driven mass flow of the soil solution must provide Ni for plant nutrition (Barber, 1995).

6.2.9 Beneficial and phytotoxic elements

In addition to these essential mineral elements, several mineral elements that are considered beneficial to plants, which promote growth for various taxa under certain environmental conditions, are also acquired from the soil solution as cations (Table 6.1). These include Na, which can replace K in its biophysical roles and is required by C4 and CAM plants for the regeneration of PEP from pyruvate; Co, which is essential for nitrogen fixation in legumes and may enhance drought resistance and herbivore defence; and Al, which has also been implicated in herbivore defence and the prevention of Fe-toxicity (Mengel et al., 2001; Pilon-Smits et al., 2009). Plant roots are exposed to these and various other non-essential elements that are present as cations in the soil solution and can reach toxic concentrations in plants. Excessive concentrations of Na, Al, Co, cadmium (Cd) and lead (Pb) commonly reach toxic concentrations in the soil solution either as a consequence of natural or anthropogenic activities (see Chapter 7).

Sodium is present in high concentrations in saline and sodic soils, which occur worldwide but mostly in the arid subtopics where rainfall is low (Frossard et al., 2000; FAO, 2010). In these soils, excess soluble salts accumulate near the surface of the soil. They occur both naturally and anthropogenically when water containing high salt concentrations is used for irrigation. Saline soils are defined as having an electrical conductivity greater than 4 dS m\(^{-1}\), whereas sodic soils are defined simply as having a high ratio of Na to Ca and Mg in the soil solution (see Chapter 7). Sodic soils can be acid or alkaline, saline or non-saline, although most are alkaline soils with the dominant anion being bicarbonate (Qadir and Schubert, 2002). Sodium toxicity is thought to limit crop production on between 5% and 15% of potential agricultural land (Munns and Tester, 2008). Plants vary appreciably in their ability to tolerate high Na concentrations in the environment (see Chapter 7). At one extreme are halophytes and at the other extreme are sensitive plants, whose growth is impaired at leaf concentrations below 2–5 mg Na g\(^{-1}\) DM (Munns and Tester, 2008). Symptoms of Na toxicity in plants are caused both directly, through effects of Na\(^+\) on plant metabolism, and
indirectly, through the osmotic effects of Na⁺ or Na⁺ inhibition of transport of other cations, especially Ca²⁺ (Munns and Tester, 2008). Crop yields on saline soils can be improved by management practices that reduce Na⁺ concentrations in the soil solution or reduce the ratio of Na⁺ to other cations in the soil solution and/or the cultivation of Na-tolerant crops. Traditionally, saline soils are remediated by leaching soluble salts from the soil profile by irrigation with freshwater, whereas sodic soils are remediated through the application of Ca²⁺, often as gypsum, followed by flushing the soil with freshwater (see Chapter 7).

Aluminium limits crop production in acidic soils, which constitute over 40% of the world’s arable land (Von Uexküll and Mutert, 1995; Sumner and Noble, 2003). It is present in the soil solution as Al³⁺, Al(OH)²⁺ and Al(OH)₃⁺ (Kinraide, 1991; Mengel et al., 2001; Kochian et al., 2004). Aluminium toxicity is manifest primarily by the inhibition of root elongation and altered root architectures, which are both associated with the effects of toxic concentrations of Al³⁺ at the root tip (Kinraide, 1991; Mengel et al., 2001). Leaf concentrations greater than 40–200 μg Al g⁻¹ DM are toxic to most plants, although some plants can tolerate leaf concentrations in excess of 1 mg Al g⁻¹ DM (Table 6.1; Barceló and Poschenrieder, 2002; Watanabe et al., 2007). These Al-hyperaccumulator plants occur in about 45 families, belonging mostly to the basal branches of the rosid and asterid clades (Barceló and Poschenrieder, 2002; Jansen et al., 2002; Watanabe et al., 2007). Many native plants growing on acidic soils have developed Al-tolerance through extracellular or symplastic detoxification mechanisms (Barceló and Poschenrieder, 2002; Kochian et al., 2004; Pilon-Smits et al., 2009). Plant roots reduce Al phytoavailability by secreting organic acids or mucilage to chelate Al, raising rhizosphere pH and binding Al to cell-wall components (Ma et al., 2001; Kochian et al., 2004; Delhaize et al., 2007). Aluminium entering plant cells is rendered non-toxic by sequestration in the vacuole complexed with organic acids (Ma et al., 2001; Kochian et al., 2004). Liming, especially with dolomitic lime (CaMg(CO₃)₂), is an effective way to raise soil pH and avoid Al and Mn toxicities and Ca and Mg deficiencies in acid soils (Fageria, 2009). However, the cultivation of Al-excluding or Al-tolerant crops, or plants generally tolerant of acid soils, can also facilitate agricultural production on these soils. Such crops are generally native to acid soils and are constitutively adapted to overcome this constraint (Fageria, 2009).

Cobalt is not very abundant in most soils. Its concentration generally ranges from 1 to 40 mg kg⁻¹ soil and its concentration in the soil solution approximates 0.2–0.7 μM (Mengel et al., 2001). Leaf concentrations are typically in the range of 0.1–10 μg Co g⁻¹ DM, and concentrations greater than 10–20 μg Co g⁻¹ DM are toxic to most plants (Table 6.1; Palit et al., 1994; Mengel et al., 2001; Watanabe et al., 2007). Cobalt toxicity can be observed in plants growing in natural soils formed from serpentinite or other ultrabasic rocks. Nevertheless, several plants native to these areas can tolerate leaf concentrations in excess of 1 mg Co g⁻¹ DM (Baker et al., 2000). The majority of these Co-hyperaccumulator plants belong to the Lamiaceae, Scrophulariaceae, Asteraceae and Fabaceae (Baker et al., 2000).

High concentrations of Cd and Pb occur both naturally and through anthropogenic activities. Natural mineral outcrops can be enriched in Cd through the weathering of Cd-rich rocks and Cd pollution of the environment has occurred through the mining and refining of metal ores or through the application of Cd-containing phosphate fertilisers, sewage sludge and municipal composts to agricultural soils (He et al. 2005; Kirkby and Johnson, 2008). Leaf concentrations greater than 5–10 μg Cd g⁻¹ DM are toxic to most plants, although some plants have adapted to grow on sites with high soil Cd concentrations and can tolerate leaf concentrations in excess of 100 μg Cd g⁻¹ DM (Table 6.1; Baker and Brooks, 1989; Broadley et al., 2001). The majority of the plants that hyperaccumulate Cd occur in the order
Brassicales (Broadley et al., 2001). Soils with high Pb concentrations can occur naturally, but the mining and refining of metal ores, incineration of coal and waste, and the presence of Pb in automobile fuel have made a significant contribution of Pb accumulation in agricultural soils (Colburn and Thornton, 1978; Bacon et al., 1996). Leaf concentrations greater than 10–20μg Pb g\(^{-1}\) DM are toxic to most plants, with the exception of several Pb-hyperaccumulator plants (Table 6.1; Broadley et al., 2001).

6.3 The uptake of essential cationic elements from the rhizosphere

The concentration of a cation in the soil solution affects both the rate at which it can be delivered to the root system and the transport mechanism required for its uptake into a root cell. Some cations, such as K\(^{+}\), Ca\(^{2+}\) and Mg\(^{2+}\), are sufficiently available and mobile in well-fertilised agricultural soils to supply plant demand, but the root system must forage for less readily available cations, such as Fe\(^{2+}\), Fe\(^{3+}\), Mn\(^{2+}\), Zn\(^{2+}\), Cu\(^{+}\), Cu\(^{2+}\) and Ni\(^{2+}\), which have low concentrations in the soil solution and small diffusion coefficients (Barber, 1995). With the exception of K\(^{+}\), the delivery of essential mineral cations to the root surface is largely determined by transpiration-driven mass flow of the soil solution, rather than by local diffusion (Table 6.2; Barber, 1995; Fageria et al., 2011). Hence, the production of an extensive root system to provide a large surface area for cation uptake in unexplored volumes of soil and the modification of rhizosphere chemistry and biology by root exudates to increase the concentrations of mineral cations in the soil solution improve the acquisition of essential mineral cations (see Section 6.6).

Mineral cations entering roots from the soil solution can reach the xylem, for delivery to the shoot, via intracellular (symplastic) or extracellular (apoplastic) pathways. Most essential mineral cations enter the root symplast through transport proteins located in the plasma membrane of epidermal and/or cortical cells and are transported across the root through the cytoplasm of root cells connected by plasmodesmata (Karley and White, 2009; Puig and Peñarrubia, 2009; Tejada-Jiménez et al., 2009; White and Broadley, 2009). They are then loaded into the xylem by transport proteins located in the plasma membrane of stellar parenchyma cells. The formation of a Casparian band in the root endodermis at an early stage of development restricts the apoplastic movement of solutes to the xylem. This enforces entry of most essential mineral cations to the symplast and, therefore, the selectivity of cation transport to the shoot. Among the essential cationic mineral elements, only Ca is thought to reach the xylem via a predominantly apoplastic pathway, although a significant proportion of the Fe and Zn acquired by plant roots can reach the xylem by this route in particular plant species and/or when these elements occur at high concentrations in the soil solution (White, 2001; White et al., 2002; Taiz and Zeigler, 2006; Broadley et al., 2007). In addition, since the apoplastic pathway is relatively non-selective, toxic cations such as Na\(^{+}\) and Cd\(^{2+}\) can reach the xylem by this route (Lux et al., 2004; Plett and Møller, 2010). The acquisition of cationic elements reaching the xylem via an apoplastic pathway is generally restricted to the extreme root tip and to regions in which lateral roots are being initiated (White, 2001).

The uptake of cationic mineral elements that are delivered to the xylem via a symplastic pathway generally exhibits both saturatable and non-saturatable components (Figure 6.2; White, 2003). The concentration-dependence of cation uptake and relative magnitude of the saturatable and non-saturatable components reflect the properties, activity and abundance of transport proteins in the plasma membrane of root cells. A multitude of transport
proteins in the plasma membrane of root cells catalysing uptake from the soil solution and xylem loading of essential cationic minerals have been identified (Figure 6.3). Potassium can enter root cells through voltage-gated inward-rectifying K⁺-selective channels (KIRC), such as those encoded by members of the “shaker” gene family, non-selective cation channels (VICC), such as those encoded by members of the cyclic nucleotide-gated

Figure 6.2 General relationship between the uptake of a cation and its concentration in solution in (a) plants with sufficient cation and (b) plants lacking the cation. Total uptake (solid line) is represented as the sum of a ‘high-affinity’ saturable ( - - - ) and a non-saturatable ( - - - ) component. The capacity of the saturatable component is increased in plants lacking the cation.

Figure 6.3 Transport proteins in the plasma membrane of root cells. See text for details.
channel (CNGC) and glutamate receptor (GLR) gene families; or ‘high-affinity’ K⁺ transporters, such as K⁺/H⁺ symporters (HAK/KUP), K⁺/Na⁺ co-transporters (HKT/Trk) and cation-H⁺ exchangers (CHX) in the plasma membrane (Karley and White, 2009; White and Karley, 2010). Potassium is loaded into the xylem through voltage-gated, outward-rectifying K⁺-selective channels (KORC) present in the plasma membrane of root pericycle and stelar parenchyma cells. Calcium can enter root cells through Ca²⁺-permeable cation channels including hyperpolarisation-activated Ca²⁺ channels (HACC), voltage-independent non-selective cation channels (VICC) and depolarisation-activated Ca²⁺ channels (DACC; White and Broadley, 2003; Karley and White, 2009; McAinsh and Pittman, 2009; Dodd et al., 2010). Efflux of Ca²⁺ from root cells to the apoplast is catalysed by Ca²⁺-ATPases in the plasma membrane. Magnesium enters root cells either through non-selective cation channels or through Mg²⁺-transporters (MGT) in their plasma membrane (Karley and White, 2009; White and Broadley, 2009). It is speculated that ATPases catalyse Mg²⁺ efflux from root cells into the xylem. Since many proteins catalysing K⁺ and, especially, Ca²⁺ transport are non-selective cation channels, Na⁺ can also enter root cells through these transporters (White, 1999; Munns and Tester, 2008).

Plants exhibit two strategies for acquiring Fe from the soil (Figure 6.4; Schmidt, 1999; Mengel et al., 2001; Guerinot, 2010). Roots of non-graminaceous species (Strategy I Plants) acidify the rhizosphere and release organic acids and phenolic compounds to increase Fe³⁺ concentrations in the soil solution. These compounds chelate Fe³⁺, which is subsequently reduced to Fe²⁺ by ferric reductases in the plasma membrane of root epidermal cells and Fe²⁺ is taken up by Fe²⁺ transporters encoded by members of the ZIP gene family (Puig and Peñarrubia, 2009; White and Broadley, 2009; Guerinot, 2010). Cereals and grasses (Strategy II Plants) can take up Fe²⁺ through Fe²⁺ transporters, but they also release phytosiderophores (structural derivatives of mugineic acid) into the rhizosphere to chelate Fe³⁺ and take up Fe³⁺-phytosiderophore complexes through homologues of the maize yellow stripe protein (YSL) in the plasma membrane of root cells (Puig and Peñarrubia, 2009; White and Broadley, 2009; Guerinot, 2010). Zinc can be taken up across the plasma membrane of root cells as Zn²⁺, catalysed by Zn²⁺ transporters encoded by members of the ZIP gene family, or, in cereals and grasses, as a Zn-phytosiderophore complex, catalysed by YSL proteins (Grotz and Guerinot, 2006; Broadley et al., 2007; Puig and Peñarrubia, 2009; White and Broadley, 2009; Clemens, 2010). Members of the heavy metal P₁B-ATPase family load Zn²⁺ into the xylem. The uptake of Cu by root cells occurs through high-affinity Cu⁺ transporters of the copper transporter (CTR) family and/or by Cu²⁺ transporters encoded by members of the ZIP gene family, and Cu²⁺ is loaded into the xylem by heavy metal P₁B-ATPases (Grotz and Guerinot, 2006; Burkhead et al., 2009; Puig and Peñarrubia, 2009; White and Broadley, 2009; Clemens, 2010). In cereals and grasses, Cu can also be taken up as a Cu-phytosiderophore complex, catalysed by YSL proteins. Manganese uptake by root cells is catalysed by Mn²⁺ transporters encoded by members of the ZIP gene family (Puig and Peñarrubia, 2009; Williams and Pittman, 2010) and it has been speculated that Ni²⁺ also enters root cells via these transporters (Pilon-Smits et al., 2009; Tejada-Jiménez et al., 2009). Furthermore, since some ZIPs, such as IRT1, are relatively non-selective between cations, they can transport Co²⁺ and Cd²⁺ into root cells. It is also conceivable that Mn and Ni can be taken up by root cells as phytosiderophore complexes catalysed by YSL proteins (Puig and Peñarrubia, 2009; Tejada-Jiménez et al., 2009; Guerinot, 2010).

Both the activity of transport proteins and the expression of genes that encode them are regulated dynamically to ensure appropriate tissue concentrations of essential mineral
The abundance and/or activity of specific high-affinity transport proteins are increased in plants lacking K, Fe, Zn, Cu or Mn and enable the acquisition of these elements from soil solutions containing low concentrations of their cations (Figure 6.2). In addition, when Fe, Zn, Cu or Mn are in short supply, roots efflux more protons, organic acids, phytosiderophores and enzymes capable of degrading organic compounds to increase the concentrations of these elements in the soil solution (Broadley et al., 2007; Pilon et al., 2009; Puig and Peñarrubia, 2009; White and Broadley, 2009; Clemens, 2010; Guerinot, 2010; White and Karley, 2010). These root traits are often complimented by associations with mycorrhizal fungi and by fostering beneficial microbial communities in the rhizosphere (Rengel et al., 1999; Harrier and Watson, 2003; Barea et al., 2005; Morgan et al., 2005; Broadley et al., 2007; Smith and Read, 2007; Cavagnaro, 2008; White and Broadley, 2009).

Since the delivery of cationic elements to the root surface is generally determined by their concentrations in the soil solution, their mobilisation from recalcitrant soil constituents can have a major effect on their acquisition by plants, whereas increased root uptake capacity alone often has little effect on the acquisition of these elements by field crops (Barber, 1995).

**Figure 6.4** Root responses to iron deficiency in non-graminaceous (Strategy I) and graminaceous (Strategy II) plant species. In Strategy I plants, increased acidification of the rhizosphere by H+-ATPases and the induction of ferric reductase activity to reduce Fe(III) chelates in the rhizosphere releases Fe2+ for uptake across the plasma membrane of root cells by Fe-deficiency-inducible, high-affinity Fe2+ transporters. In Strategy II plants, enhanced synthesis and release of phytosiderophores (PS) into the rhizosphere chelate Fe2+, Zn2+, Cu2+ and Mn2+, and these chelates are transported across the plasma membrane by Fe-deficiency-inducible transport proteins. Adapted from White (2012). With kind permission of Academic Press London.
6.4 Cationic elements in the soil solution

Differences in the complement and concentrations of cationic elements in soil solutions reflect the enormous diversity of soils across the world. Base cations are found in rock fragments, in minerals, in aqueous solution and on the negatively charged surfaces of soil particles. The quantities of base cations on negatively charged surfaces vary from about 1 to 200 cmol kg\(^{-1}\) soil and are often used for estimating fertiliser requirements (Havlin et al., 2005; Fageria, 2009). Some of this variation results from differences in agronomic practice, but for the most part it results from processes within the soil and their dependence on the parent material and the extent of weathering. Weathering breaks down rocks and rock fragments by water- and temperature-dependent physical processes, such as freezing and thawing. Chemical changes in soil take place by hydrolysis and oxidation and by the effects of atmospheric deposition (e.g. of S and Na), carbon dioxide and organic acids and other compounds produced by the activities of animals, plants and microorganisms. Initially, weathering produces small pieces of rocks and mineral grains, which have similar composition as the parent rocks, but, as weathering continues, clays and amorphous materials are produced, which contain less mineral elements than the parent rocks. Even these products are further transformed by subsequent weathering. In addition, organic matter is accumulated in soils as a result of the formation and degradation of plant debris. However, despite the complexity of the soil-forming processes, widely applicable principles are emerging that govern the formation of the different components of soil and their cationic properties. In this section, these principles are introduced by first discussing the cationic properties of the main components of soil and the key stages of their formation. The quantitative relationships governing the dependence of base cation concentrations in solution on each other, on the CEC of the soil (the sum of all cation equivalents held on exchange sites) and on pH are then discussed.

6.4.1 The properties and formation of soils

The characteristics of mineral fragments vary enormously depending on the rocks from which they are derived. For example, fragments derived from some forms of magma that have cooled slowly contain few base-rich cations, whereas fragments derived from magma that has cooled rapidly are generally rich in base cations (Wild, 1993). There is similar diversity in mineral fragments derived from the two other major rock types, namely sedimentary rocks (such as limestone) which, as the name implies, are formed by sedimentation, and metamorphic rocks, which are formed by recrystallisation under high temperatures and pressures. Generally, fragments derived directly from parent rocks are large and occur within the sand and silt particle size fractions. Consequently, their surface area per unit mass is small and their CEC is less than those of clays of similar composition (Table 6.3). Nevertheless, if the parent material is base-rich, the CEC of the sand and silt particles can be substantial. Average values of 19 cmol kg\(^{-1}\) for the sand fraction and 33 cmol kg\(^{-1}\) for the silt fraction in the upper 20–30 cm layers of soil have been reported (McAleese and McConaghy, 1957). These values are significant even when compared with smectite clays that are derived from similar base-rich rocks.

Mineral particles less than 2 \(\mu\)m in diameter are classified as clays. First, it must be noted that, although the examples given here illustrate the main processes of clay formation, clays can vary widely in both their physical and chemical properties. In general, clays
Soil Conditions and Plant Growth

consist of two distinct units: silica–oxygen and aluminium hydroxide. The silica–oxygen unit consists of Si$^{4+}$ at the centre of a tetrahedron of three O$^{2-}$ ions (Mott, 1988a). These tetrahedra link with one another in different ways, but in all cases there is a Si–O–Si bond. In clays, these bonds form two-dimensional paper-like sheets that are only one Si layer thick. The aluminium hydroxide unit is arranged geometrically so that Al$^{3+}$ is at the centre of an octahedral arrangement of six OH$^{-}$ that are polymerised into sheets one Al atom thick. The formation of these sheets and the way they are combined affect the cationic properties of clays, as is illustrated by kaolinite and smectite. Kaolinite is formed from one Si sheet adjacent to one Al sheet. The bonds between the two sheets are strong, shrinking and swelling is small, and the clay is resistant to weathering. It is formed in acid environments (Wilson, 1999). It has a low CEC (Table 6.3), contains few base cations and has almost zero permanent charge. However, since it has exposed OH$^{-}$ groups, its surface charge is pH dependent (Tan, 1998). By contrast, smectite consists of one Al sheet sandwiched between two Si sheets held together by Si–O–Al links (Mott, 1988a). The bonding between sheets is weak, shrinking and swelling is considerable, and it is readily degraded by weathering into kaolinite (Wilson, 1999). It is formed in base-rich environments and cations such as Mg$^{2+}$ and Fe$^{3+}$ become entrapped within ‘holes’ of the Al sheet and also replace some of the Si$^{4+}$ and Al$^{3+}$ ions during its formation (Mott, 1988a). The mix of cations involved in this amorphous replacement varies in different forms of smectite, but, since these cations have lower valencies than Si$^{4+}$ or Al$^{3+}$, all these forms have a high negative charge and a high CEC of, typically, 100 cmol kg$^{-1}$ (Table 6.3). It is noteworthy that Cu$^{2+}$ never takes part in amorphous replacement because it is too large to fit within the available spaces. The weak bonding between layers also means that K$^{+}$ and NH$_4^{+}$, uniquely among cations, can become entrapped between layers and are then said to be fixed. Fixed K$^{+}$ and NH$_4^{+}$ are only slowly available to plants, and the significance of this is discussed later in Section 6.5. Another important group of clays is hydrated Fe and Al oxides. These clays have weak electronegative charges in acid conditions and an electropositive charge in alkaline conditions, which allows them to adsorb anions such as phosphate. They also form covalent bonds with heavy metal cations such as Cu$^{2+}$, Zn$^{2+}$ and Pb$^{2+}$. The adsorption capacity of these clays ranges from 3 to 30 cmol kg$^{-1}$ (Tan, 1998).

The elucidation of the effects of weathering, and the sequence of soil formation, is complicated by the difficulty in finding soil pairs with the same parent materials that have experienced different intensities of weathering. One difficulty is the huge disparity in age and weathering between soils developed after an ice cap and those where none has ever been present. The processes of soil formation have also been confounded by sufficient deposition of elements from the atmosphere to sustain vegetation on some highly weathered soils.

### Table 6.3  Typical values for the cation exchange capacity (CEC) of major components of soils.

<table>
<thead>
<tr>
<th>Soil colloid</th>
<th>CEC (cmol kg$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Soil organic matter</td>
<td>110–290</td>
</tr>
<tr>
<td>Smectite</td>
<td>60–150</td>
</tr>
<tr>
<td>Kaolinite</td>
<td>3–15</td>
</tr>
<tr>
<td>Sesquioxides</td>
<td>2–4</td>
</tr>
</tbody>
</table>

*CEC varies with pH.

(Jackson et al., 1971; Chadwick et al., 1999). Nevertheless, studies have been made on enough suitable pairs of soils to suggest that, as the duration or intensity of weathering increases, the order of events is as follows: (1) Hydrolysis of the primary rocks with a loss of base cations, which is initially rapid but then decreases sharply with time (Syers and Walker, 1969; Walker and Syers, 1976; Bain et al., 1993). (2) Formation of smectite- or kaolinite-type clays, depending on the pH of the surrounding media (Fisher and Ryan, 2006). (3) Breakdown of smectite clays into kaolinite and then into gibbsite, Al(OH)$_3$ (Wilson, 1999; Fisher and Ryan, 2006). (4) Leaching of silica, but not of coarse-grained quartz, decomposition of kaolinite and an increase in the proportions of Fe and Al oxides in soil (Aniku and Singer, 1990; Troeh and Thompson, 2005; Uehara, 2006). (5) Leaching of Fe and Al oxides, with the consequence that the remainder consists of nutrient-free coarse-grained quartz (Jackson et al., 1971; Troeh and Thompson, 2005). Thus, weathering results in a loss of base cations and the ability of soils to retain added cations. Weathering is accelerated by high temperatures and rainfall.

In addition to its inorganic constituents, soil also contains organic matter, which, because of its low bulk density, can constitute a substantial proportion of the total soil volume. Soil organic matter exists in many different forms (see Chapter 4). In the context of this chapter, and as a general rule, the CEC of soil organic matter varies with pH and is typically 200 cmol kg$^{-1}$ (Table 6.3). Some components of soil organic matter form strong complexes with cations such as Cu$^{2+}$, Fe$^{2+}$ and Mn$^{2+}$. These are generally insoluble, but they can be solubilised under some environmental conditions (Tan, 1998).

6.4.2 Relationships between cations on exchange sites and in the soil solution

The solid matrix of soils has both negatively and positively charged components. Negative charges adsorb base cations and dominate temperate soils. Two key parameters defining cation exchange between the solid matrix and the soil solution are the CEC of the soil and its percent base saturation. The CEC of a soil defines its capacity to adsorb and exchange all types of inorganic cations including protons and aluminium. The percent base saturation is the percentage of the total CEC that is saturated with base cations (i.e. cations other than protons and aluminium). Both are important as they influence the extent to which soils can meet plant demands for cations such as K$^+$ and Mg$^{2+}$ without frequent applications of these macronutrient elements. The percent base saturation is strongly correlated with soil pH (Figure 6.5) and, together with soil CEC, these parameters have a decisive influence on lime requirement. Provided the soil CEC is high, when the soil pH is over 7 it is buffered strongly by the reaction of H$^+$ on base carbonates, such as CaCO$_3$, and when soil pH is between 5 and 6 it is buffered by the replacement of Ca$^{2+}$, Mg$^{2+}$, K$^+$ and Na$^+$ on the exchange sites with H$^+$. Below pH 4, the soil is often strongly buffered by hydroxyl–aluminium complexes that have formed during the breakdown of clay minerals (Goulding and Annis, 1998). Subsoil acidity is a serious problem and very difficult to correct. Much emphasis is often placed on maintaining the pH of the ploughed layer by liming which, among other things, should prevent the spread of acidity to the subsoil.

In this section, for the sake of clarity, ideal relationships between cations on exchange sites and those in dilute soil solutions are presented initially to avoid the complications associated with the effects of solution ionic strength on the chemical activity of each cation. Adjustments that need to be made for the effects of solution ionic strength, and the consequences of the formation of uncharged ion pairs, are considered subsequently.
The same principles appear to govern exchange processes irrespective of the nature of the exchange sites. Within the micro-volume of solution adjacent to the negative charges on the solid matrix, there is a diffuse double layer in which the cation concentrations decrease and the anion concentrations increase with distance. Beyond this minute distance from the surface, the sum of the positive charges of the cations must always equal the sum of the negative charges of the anions in the soil solution to maintain electrical neutrality (Figure 6.6). Strict rules govern the interchange between cations on the exchange sites (denoted by the subscript ‘exchange’) and those in the soil solution (denoted by subscript ‘solution’). According to the law of mass action, when equilibrium is established for the monovalent cations $A^+$ and $B^+$:
\[
\frac{A^+_{\text{solution}}}{B^+_{\text{solution}}} = K_{ex} \left( \frac{A_{\text{exchange}}}{B_{\text{exchange}}} \right)
\]

where \( K_{ex} \) is a selectivity index. The larger the \( K_{ex} \), the greater the relative adsorption of \( B^+ \) compared with \( A^+ \).

A number of more empirical relationships have been advanced for the exchange of monovalent, divalent and trivalent cations with one another. Divalent cations, such as \( \text{Ca}^{2+} \) and \( \text{Mg}^{2+} \), are absorbed less strongly than monovalent cations, such as \( \text{H}^+ \) and \( \text{K}^+ \). Schofield developed a fundamental law governing the ratio of the concentrations of two cations \( M_1 \) and \( M_2 \) of valency \( Z_1 \) and \( Z_2 \) in soil solution (Schofield, 1947; Schofield and Taylor, 1955a, b). It states that the ratio \([M_1]^{1/Z_1}/[M_2]^{1/Z_2}\) remains constant irrespective of soil/water ratio provided that the cation contents on the exchange complex remain approximately constant, which is usually the case as the amounts of cations in the soil solution are usually small compared with those on the exchange sites. If the ionic concentrations are other than dilute, they must be replaced with ionic activities as described in Section 6.4.3 (Schofield and Taylor, 1955a, b). The predictions of this relationship have been confirmed experimentally for the following ratios: \( \text{K}:\text{Ca}, \text{Ca}:\text{Al}, \text{K}:\text{Al} \) and \( \text{Na}:\text{Al} \) (Bower, 1959; Salmon, 1964; Bache, 1970). Thus, the quotient \([K^+]/([\text{Ca}^{2+}]+[\text{Mg}^{2+}])^{0.5}\) has a fundamental basis as a measure of the K-status of a soil and has proved useful in elucidating the K dynamics of soils (Evangelou et al., 1994). The total cation concentrations in solutions from many soils are dominated by the concentrations of \( \text{Ca}^{2+} \) and \( \text{Mg}^{2+} \), and the sum of their charge equivalents is, therefore, approximately proportional to the sum of the charge equivalents of anions. The approximate constancy of \([K^+]/([\text{Ca}^{2+}]+[\text{Mg}^{2+}])^{0.5}\) over a wide range of soil water contents, together with the quantitative charge-equivalent relationship between \( \text{Ca}^{2+}+\text{Mg}^{2+} \) and anions, has proved useful in developing simulation models for the responses of crops to K-fertilisers (Greenwood and Karpinets, 1997a, b) and the way in which these are affected by the application of N-fertilisers (Zhang et al., 2007).

An important aspect of the K economy of soils concerns K⁺ fixed between some clay sheets and its release into the soil solution (Karpinets and Greenwood, 2003). This process buffers the available K⁺ in the soil solution. Typically, the soil solution contains about 0.1–0.2% of the total K in the soil, a further 1–2% is present as exchangeable K⁺ held in cation exchange sites and about 1–10% is present as fixed K⁺ associated with clay lattices (Mengel et al., 2001; Fageria, 2009). The remainder of the soil K is effectively unavailable over the growth period of a typical arable crop. Interchange between K⁺ in solution and exchangeable K⁺ is very rapid, but the exchange between K⁺ in solution and fixed K⁺ is slow. The amounts of each of these components, and the rates of interchange between them, are important for the self-organisation of soil K pools, in which the solution and exchangeable forms of K⁺ tend toward critical values characteristic of a soil (Karpinets and Greenwood, 2003). If a crop acquires K⁺, the concentration of K⁺ in solution will fall which, if the concentration falls below the critical value, prompts the slow release of fixed K⁺ into solution. This can replenish the exchangeable K⁺ pool. If K-fertiliser is added to the soil, it will rapidly increase K⁺ in solution and replenish the exchangeable K⁺ pool and, if these rise above the critical values, some K⁺ will enter the fixed K⁺ pool. Changes in the net balance between crop K⁺ acquisition and the addition of K-fertilisers will seldom equal the sum of the changes in solution and exchangeable K⁺ because of the buffering effect of fixed K⁺.

An important group of soils derived from sedimentary rocks consist largely of CaCO₃. Calcium carbonate has no cation exchange properties, but it reacts with carbonic and other
acids to form soluble Ca salts. Carbon dioxide dissolved in water forms carbonic acid which, though a weak acid, reacts with CaCO$_3$ to release Ca$^{2+}$ ions into the solution. It has long been recognised how the various processes can be defined in terms of classical chemical equilibria (Knox, 1912). The key equations and values of the corresponding coefficients are as follows (Hodgman, 1953):

<table>
<thead>
<tr>
<th>Equilibria</th>
<th>Coefficients at 25 °C</th>
</tr>
</thead>
<tbody>
<tr>
<td>$[Ca^{2+}][CO_3^{2-}] = S$</td>
<td>$4.95 \times 10^{-9}$</td>
</tr>
<tr>
<td>$\frac{[H_2CO_3]}{[P]} = K_1$</td>
<td>$3.29 \times 10^{-2}$</td>
</tr>
<tr>
<td>$\frac{[H^+][HCO_3^-]}{[H_2CO_3]} = K_2$</td>
<td>$4.3 \times 10^{-7}$</td>
</tr>
<tr>
<td>$\frac{[H^+][CO_3^{2-}]}{[HCO_3^-]} = K_3$</td>
<td>$5.61 \times 10^{-11}$</td>
</tr>
</tbody>
</table>

where P is the partial pressure of CO$_2$ expressed in atmospheres and concentrations are expressed in molarities. Division of Equation 6.3 by Equation 6.4 gives:

$$\frac{[HCO_3^-]^2}{[H_2CO_3][CO_3^{2-}]} = K_2 / K_3$$

(6.5)

Multiplication of Equation 6.5 by Equation 6.1 gives:

$$\frac{[HCO_3^-]^2[Ca^{2+}]}{[H_2CO_3]} = K_2S / K_3$$

(6.6)

Electrical neutrality must be maintained and thus:

$$2[Ca^{2+}] + [H^+] = [HCO_3^-] + 2[CO_3^{2-}]$$

As the concentrations of [H$^+$] and [CO$_3^{2-}$] are small compared with those of [Ca$^{2+}$] and [HCO$_3^-$], they may be neglected so that $2[Ca^{2+}] \approx [HCO_3^-]$.

Substitution of these values in Equation 6.6 gives:

$$\frac{4[Ca^{2+}]^3}{[H_2CO_3]} = K_2S / K_3$$

(6.7)

Elimination of [H$_2$CO$_3$] by Equation 6.2 gives:

$$[Ca^{2+}] \approx \frac{3}{8} \sqrt{S.K1.P.K2 / (4*K3)}$$

(6.8)

which after substitution of the above coefficients gives:

$$[Ca^{2+}] \approx 0.00713 \sqrt{P}$$

(6.9)
Multiplication of Equations (6.2), (6.3) and (6.4) and substitution of CO₃ by Equation 6.1 gives:

\[ P = \frac{[H^+]^2 \cdot S}{[Ca^{2+}]K1K2K3} \]

and substitution of Ca²⁺ in Equation 6.9 gives:

\[ P^{\text{HSP}} = \frac{[H^+]^2 \cdot S}{7.13 \times 10^{-3} K1K2K3} \]  \( (6.10) \)

Equations (6.8) and (6.10) are of crucial importance for understanding the cationic relationships in these soils. Equation (6.8) indicates that Ca²⁺ concentration increases approximately in proportion to the cube root of the partial pressure of CO₂. So, in the surface layers of soil where there is high metabolic production of CO₂, there will be higher concentrations of both CO₂ and Ca²⁺ than in the subsoil, where metabolic activity is lower. Leaching will result in the movement of Ca²⁺ from the surface soil to the subsoil, where it will be precipitated. Another consequence of low CO₂ concentrations and, thus, CO₂ partial pressures in the subsurface soil is that the H⁺ concentration will fall and the pH will rise as defined by Equation 6.10.

Soils often receive strong acids, such as nitric acid from the nitrification of ammonia in soil and sulphuric acid from rainfall that dissolve CaCO₃ to form soluble salts. Such acids are much stronger than carbonic acid and, in consequence, can give rise to higher concentrations of Ca²⁺ and greater Ca²⁺ leaching down the soil profile.

### 6.4.3 Caveats concerning high ionic strength and uncharged complexes

The concentration of an ion gives a value to the upper limit of its reactivity. If the ionic strength of a solution rises, then the reactivity of an ion is reduced. To correct for this, concentrations are multiplied by an activity coefficient for an appropriate ionic strength of solution. As an approximation to the Debye–Huckel equation, the activity coefficient can be calculated from:

\[ -\log_{10} f_i = \left( A Z_i^2 I^{1/2} / (1 + a_i B I^{1/2}) \right) \]

where \( f_i \) is the activity coefficient of ion i, of valency \( Z_i \), A and B are functions of temperature and \( a_i \) is the effective diameter of the ion. I is the ionic strength given by:

\[ I = \frac{1}{2} \sum C_i Z_i^2 \]

where \( C_i \) is the molar concentration of ion i. The activity coefficients of trivalent ions are depressed more than those of divalent ions which are depressed more than those of monovalent ions by increase in ionic strength (Tan, 1998).

Another complication in estimating the phytoavailability of cations in a soil solution arises because some pairs of cations and anions combine to form uncharged complexes or ion pairs. An example of this is calcium sulphate. In solutions containing a concentration less than that at which precipitation occurs, there exists in solution the following species:

\[ \text{CaSO}_4^{\text{s}} \leftrightarrow \text{Ca}^{2+} + \text{SO}_4^{2-} \]
Soil Conditions and Plant Growth

where CaSO₄° is uncharged. The equilibrium constant, Keq, is given by:

\[
Keq = \frac{\left(\text{Ca}^{2+}\right)\left(\text{SO}_4^{2-}\right)}{\left(\text{CaSO}_4°\right)}
\]

where the terms in parenthesis are activities.

The complicated calculations required to determine the concentrations and activities of cations in soil solutions have led to the development of several sophisticated computer programs. Prominent ones include GEOCHEM (Sposito and Mattigod, 1980; Parker et al., 1995; Shaff et al., 2009) and PHREEQC (Parkhurst and Appelo, 1999), both of which include many inorganic and organic species in widely different environments including those with solid cation exchange sites. Although, recent models and associated experimental work reveal some discrepancies between model predictions and actual measurements (e.g. Huber and Montserrat, 2002; Unsworth et al., 2006), these programmes have been fundamental to developing our understanding of the chemistry of soil solutions. Further development of these computer programmes, including improvements in the treatment of activity coefficients and ion pairs, will be especially valuable for studying the effects of aluminium and cationic micronutrients in the soil solution as they are prone to speciation (Mott, 1988b; Parker et al., 1988; Kinraide, 1991; Percival et al., 1999; Ge et al., 2000; Cancès et al., 2003; Almås et al., 2007).

6.5 Management practices increasing the acquisition of essential mineral cations

6.5.1 Potassium

Most soils contain K concentrations between about 0.3 and 25 g kg⁻¹ soil, and soil solutions often contain K⁺ concentrations between 0.1 and 1 mM (Mengel et al., 2001; Fageria, 2009). The K⁺ available for plant growth – which is present as K⁺ in the soil solution, as exchangeable K⁺ associated with cation exchange sites and as fixed K⁺ associated with clay lattices – amounts to less than 2–10% of the total K in the soil, a large proportion of which resides in the topsoil (Section 6.4.2). The requirements for K-fertiliser are determined primarily by the ability of the soil to provide sufficient K⁺ to satisfy the demand of a particular crop. Although crops differ in their K requirements (Section 6.6), and in some plant species Na⁺ can substitute for K⁺ in its biophysical roles (Section 6.2), their basic K requirements are ultimately established by their rate of growth, which is generally determined by the availability of N. Soils vary greatly in their ability to supply K⁺ to a crop. Some soils can release more K⁺ than is required by crops such that no additional K-fertiliser is necessary, whilst other soils have no capacity to retain or supply K⁺ and must receive sufficient K-fertiliser each year to meet the entire crop requirement. Some soils are so well buffered that the optimum strategy is to attempt to maintain a given level of exchangeable K⁺ through regular applications of K-fertilisers. To provide a more quantitative basis for K-fertiliser practice, the interrelationships between soil K⁺ pools and K⁺ concentration in the soil solution, the movement of K⁺ to the root and its acquisition by the crop and the growth-driven K demand of a crop have been incorporated into predictive models. These models have been tested against data from field experiments and are now embedded in decision support tools for K-fertiliser practice (Zhang et al., 2007).

The main additional sources of K for plant nutrition are crop residues and composts (0.30–0.45% DM; Defra, 2010), animal manures (0.65–1.04% DM; Defra, 2010),
K-bearing minerals and K-fertilisers. Most K-fertilisers are applied to field crops as KCl, KNO₃, K₂SO₄, K₂SO₄·MgSO₄ or KPO₃ (Table 6.4; Lægrid et al., 1999; Fageria, 2009). Strategies for optimising the use of K-fertilisers include: (1) Liming acid soils, which not only increases the retention and availability of applied K⁺ by replacing Al³⁺ on cation exchange sites with Ca²⁺ and K⁺ but also improves root growth and, thereby, the uptake of water and mineral nutrients. (2) Maintaining adequate phytoavailable K⁺ in the soil solution by applying K-fertilisers at an appropriate rate and time for a particular crop (Eckert, 1987; Kopittke and Menzies, 2007). (3) In wide-row crops, using a banded placement of K-fertiliser close to the developing root system in a K-deficient soil, rather than broadcasting K-fertiliser. (4) Supplying adequate moisture to allow K⁺ acquisition by roots in regions of the soil where K⁺ is available. (5) Incorporating crop residues after harvest, which can contain up to 80% of the K acquired by a cereal crop and up to 50% of the K acquired by a legume crop (Fageria, 2009). (6) Improving the organic matter content of soil, using crop residues, composts or animal manures, which, in addition to supplying K and other essential mineral elements, improve soil structure, water relations and mitigate against erosion (see Chapter 8). (7) Using crop species and genotypes within species that acquire and utilise K⁺ most efficiently to produce a commercial yield (see Section 6.6).

Table 6.4 Commonly applied inorganic fertilisers containing essential cationic elements (Fageria 2009; White and Broadley 2009).

<table>
<thead>
<tr>
<th>(A) Macronutrients</th>
<th>(B) Micronutrients</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Potassium</strong></td>
<td><strong>Iron</strong></td>
</tr>
<tr>
<td>KCl, KNO₃</td>
<td>FeSO₄·7H₂O</td>
</tr>
<tr>
<td>K₂SO₄; K₂SO₄·MgSO₄</td>
<td>Fe₂(SO₄)₂·4H₂O</td>
</tr>
<tr>
<td>KPO₃</td>
<td>(NH₄)₂SO₄·FeSO₄·6H₂O</td>
</tr>
<tr>
<td>Kainit [MgSO₄ + KCl + NaCl]</td>
<td>Fe[OH]₃PO₄·H₂O</td>
</tr>
<tr>
<td><strong>Calcium</strong></td>
<td><strong>Manganese</strong></td>
</tr>
<tr>
<td>Calcitic lime [CaCO₃]</td>
<td>MnCl₂</td>
</tr>
<tr>
<td>Dolomitic lime [CaMg(CO₃)₂]</td>
<td>CuCl₂</td>
</tr>
<tr>
<td>Epsom salts, kieserite [MgSO₄]</td>
<td>ZnSO₄·H₂O; ZnSO₄·7H₂O</td>
</tr>
<tr>
<td><strong>Magnesium</strong></td>
<td><strong>Copper</strong></td>
</tr>
<tr>
<td>Dolomite lime [MgCO₃]</td>
<td>CuSO₄·H₂O; CuSO₄·5H₂O</td>
</tr>
<tr>
<td><strong>Zinc</strong></td>
<td><strong>Zn-carbonates</strong></td>
</tr>
<tr>
<td>Basic slag [CaSiO₃]</td>
<td>CuO; Cu₂O</td>
</tr>
<tr>
<td>Slaked lime [Ca(OH)₂]</td>
<td>Zn₃(PO₄)₂</td>
</tr>
<tr>
<td>Struvite [(NH₄)Mg(PO₄)·6(H₂O)]</td>
<td>Zn-EDTA chelates</td>
</tr>
</tbody>
</table>

Fertilisers containing essential cationic elements include:

- KCl, KNO₃, K₂SO₄, K₂SO₄·MgSO₄ or KPO₃
- Calcitic and dolomitic limes
- Epsom salts and kieserite
- Basic and slaked slags
- Struvite
- Wood ash
- FeSO₄·7H₂O, Fe₂(SO₄)₂·4H₂O
- (NH₄)₂SO₄·FeSO₄·6H₂O
- Fe[OH]₃PO₄·H₂O
- MnCl₂, CuCl₂, ZnSO₄·H₂O; ZnSO₄·7H₂O
- CuSO₄·H₂O; CuSO₄·5H₂O
- CuO; Cu₂O
- Zn₃(PO₄)₂
- Cu-silicates, CuO; Cu₂O
- Na₂CuEDTA
- NaCuHEDTA
- Zn-silicates
- Zn-EDTA chelates
- Zn-EDTA chelates
- Zn-HEDTA chelates
- Zn-HEDTA chelates
- NaFeEDDHA
- Mn-methoxyphenylpropane
- NaFeHEDTA
- Mn-polyflavenoid
- NaFeHEDTA
- Mn-polyflavenoid
- NaFeEDDHA
- Mn-methoxyphenylpropane
- NaFeHEDTA
- Mn-polyflavenoid
- NaFeEDDHA
- Mn-methoxyphenylpropane
6.5.2 Calcium and magnesium

Most soils contain sufficient Ca and Mg concentrations for adequate crop nutrition (White and Broadley, 2003, 2009). Soil solutions often contain Ca$^{2+}$ concentrations in excess of 1 mM and Mg$^{2+}$ concentrations between 125 μM and 8.5 mM (Mengel et al., 2001; White and Broadley, 2003, 2009). A lack of Ca seldom restricts the growth of arable crops, but Mg deficiencies can occur on sandy soils in the early stages of crop growth (Wilkinson et al., 1990; White and Broadley, 2009). In a similar manner to K$^+$, the Mg$^{2+}$ immediately available for plant growth is present as Mg$^{2+}$ in the soil solution and as exchangeable Mg$^{2+}$ associated with cation exchange sites, which can constitute up to 5% of the Mg in the soil (Mengel et al., 2001). The most effective agronomic practice to increase Ca$^{2+}$ and Mg$^{2+}$ concentrations in the soil solution is liming. Calcite lime (CaCO$_3$), dolomite lime (MgCO$_3$) and dolomitic lime (CaMg(CO$_3$)$_2$), which increase soil pH, are the most appropriate Ca-fertilisers and Mg-fertilisers for acid soils, although burned lime (CaO), slaked lime (Ca(OH)$_2$), basic slag (CaSiO$_3$) and wood ash are also commonly used Ca-fertilisers (Table 6.4). Other Ca-fertilisers include gypsum (CaSO$_4$), Ca(NO$_3$)$_2$, superphosphates (14–20% Ca), rock phosphates and farmyard manures (0.5–2.3% Ca). Other common Mg-fertilisers include composites (~0.24% Mg), farmyard manures (0.2–0.3% Mg), MgSO$_4$ (Epsom salts or kieserite) and MgO (Draycott and Allison, 1998; Mengel et al., 2001; Fageria, 2009; White and Broadley, 2009; Defra, 2010). In addition, magnesium ammonium phosphate (struvite) has recently received attention as an alternative Mg-fertiliser, since it has potential as a sustainable P source for agriculture (Parsons and Smith, 2008). The application of lime increases Ca and/or Mg in the topsoil, whereas water-soluble fertilisers distribute these elements throughout the soil profile. Prophylactic foliar applications of soluble Ca-fertilisers are often made to horticultural crops to prevent Ca-deficiency disorders (Shear, 1975; Ho and White, 2005) and foliar applications of MgSO$_4$ are also common on some crops (Draycott and Allison, 1998). Applying Ca-fertilisers and Mg-fertilisers at an appropriate time and rate, as well as maintaining appropriate Ca and Mg concentrations and Ca/Mg quotients in the soil solution, are important factors for crop production (Eckert, 1987; Kopittke and Menzies, 2007). Supplying adequate moisture and improving the organic matter content of soil also benefit the acquisition of Ca$^{2+}$ and Mg$^{2+}$.

6.5.3 Micronutrients

Although the total amounts of Fe, Zn, Cu and Mn in most soils are sufficient for crop nutrition, the phytoavailability of these mineral elements restricts production in many regions of the world (Follett et al., 1981; Alloway, 1995; Barber, 1995; Welch, 1995; Loneragan, 1997; Shuman, 1998; Schmidt, 1999; Graham et al., 1999; Frossard et al., 2000; Rengel, 2001; Broadley et al., 2007; Fageria, 2009; White and Broadley, 2009). Iron rapidly becomes unavailable in alkaline and calcareous soils. For this reason, the application of many inorganic Fe-fertilisers (Table 6.4), including FeSO$_4$, Fe$_3$(SO$_4$)$_2$, FeO, Fe$_2$O$_3$, (NH$_4$)$_2$SO$_4$, FeSO$_4$, Fe(NH$_4$)HP$_2$O$_7$, Fe(NH$_4$)PO$_4$ and Fe-frits, to soils is usually ineffective in combating Fe chlorosis. However, the application of Fe chelates, such as NaFeEDTA, NaFeHEDTA and NaFeEDDHA, to soils is effective (Mortvedt, 1991; Shuman, 1998; Rengel et al., 1999). In addition, the availability of Fe in the rhizosphere can be increased by soil acidification with elemental S (Shuman, 1998). By contrast, the application of soluble fertilisers to the soil is often an effective way to correct Zn, Cu and Mn deficiencies (Shuman, 1998; Harris et al., 2007; Cakmak, 2008; Fageria, 2009; White and Broadley, 2009). Common Zn-fertilisers...
include $\text{ZnSO}_4$, ZnO, $\text{ZnCO}_3$, $\text{ZnS}$, $\text{Zn}_2(\text{PO}_4)_2$, zinc silicate frits, Na$_2$ZnEDTA, NaZnHEDTA and raplex zinc, ZnPF (Table 6.4). Copper sulphate (CuSO$_4$), CuO, Cu$_2$O, CuCl$_2$, Na$_2$CuEDTA, NaCuHEDTA and Cu-frits are common Cu-fertilisers (Table 6.4). Copper can also be supplied to crops in sewage sludges and animal manures. Manganese sulphate (MnSO$_4$), MnO, MnO$_2$, MnEDTA, MnCO$_3$, MnCl$_2$ and Mn-frits are used as Mn-fertilisers and are often applied in combination with N-fertilisers and/or P-fertilisers that induce a local acidification of the rhizosphere soil (Fageria, 2009). Foliar applications of soluble mineral fertilisers, usually FeSO$_4$, $\text{ZnSO}_4$, CuSO$_4$ or MnSO$_4$, are also used as a temporary or emergency measure to correct micronutrient deficiencies in crops (Gupta, 1979; Loneragan, 1997; Rengel et al., 1999; Cakmak, 2002, 2004, 2008; Fang et al., 2008; Fageria et al., 2009). The trace amounts of Ni required by plants can be supplied as NiCl$_2$, Ni(NO$_3$)$_2$ or NiO. Additional sources of Fe, Zn, Cu and Mn for crop nutrition include crop residues and organic manures, which not only provide phytoavailable sources of these elements but also improve soil structure and water relations.

As an additional or alternative strategy to the application of mineral fertilisers, the yields of crops sensitive to deficiencies in Fe, Zn, Cu or Mn can be increased either by intercropping with plants that are better able to access and mobilise these elements or by including such plants in crop rotations (Rengel et al., 1999; Jolley et al., 2004; Naeve, 2006; Graham et al., 2007; Inal et al., 2007). Soil microorganisms can also be exploited to increase the phytoavailability of these elements for crop production (see Chapter 3; Rengel et al., 1999; Barea et al., 2005; Morgan et al., 2005; Lynch, 2007; Smith and Read, 2007; White and Broadley, 2009). Several studies have found that both mycorrhizal associations (Kothari et al., 1991; Caris et al., 1998; Rengel et al., 1999; Harrier and Watson, 2003; Cavagnaro, 2008) and inoculants of growth-promoting bacteria (Rengel, 2001; Whiting et al., 2001; Barea et al., 2005) can increase tissue Fe, Zn and Cu concentrations and yields of crop plants. Manganese occurs in more than one oxidation state ($\text{Mn}^{4+}$, $\text{Mn}^{3+}$, $\text{Mn}^{2+}$) and the redox potential of the soil solution affects the solubility and availability of Mn markedly (Mengel et al., 2001). Thus, microbes that oxidise Mn can have a great influence on its phytoavailability (Nogueira et al., 2007).

6.6 Genetic strategies to improve the efficiency of acquisition and physiological use of essential cationic elements

6.6.1 Plant traits improving the acquisition of essential cationic elements

Plant growth can be restricted by the lack of any essential mineral element. Although a lack of Ca or Ni rarely limits crop production, deficiencies of K, Mg, Fe, Zn, Cu and Mn can occur on agricultural soils worldwide (Pretty and Stangel, 1985; Fageria, 2009; White and Broadley, 2009). Where there is an absolute lack of a mineral element in the soil, it must be supplied directly to crops to increase production and the objective of plant breeding is to maximise the uptake and physiological utilisation of the mineral element(s) supplied as fertiliser. In situations where sufficient mineral element is present in the soil but crop production is constrained by its phytoavailability, plant breeding seeks, in addition, to produce genotypes with a greater ability to mobilise the mineral element from its recalcitrant forms in the soil.
The delivery of most cationic mineral elements to the root surface is determined largely by transpiration-driven mass flow of the soil solution (see Sections 6.1 and 6.2). Nevertheless, their acquisition is enhanced by (1) the production of an extensive root system in unexplored volumes of soil, which reduces the path of water movement through the soil and increases the surface area available for cation uptake, and (2) the modification of soil chemistry and biology by root exudates, which increases the concentrations of cationic mineral elements in the soil solution. For cationic elements that are concentrated at the soil surface, such as Mn, Cu and Ni in some soils, a topsoil foraging phenotype can be beneficial. However, for cationic elements that are present at high concentrations in the soil solution, such as K, Ca and Mg, or elements that are distributed evenly within the soil profile, such as Fe and Zn, proliferation of roots throughout the soil volume is most appropriate. For elements that are taken up predominantly at the tip of roots, such as Ca and in some plant species Fe, the number of root apices in moist soil is of critical importance. For the latter elements, extensive root branching and the generation and maintenance of root tips in moist soil are advantageous.

6.6.2 Differences between plant species in their nutritional requirements

Plant species differ in their demand for essential mineral elements (see Section 6.2). For example, the nutritional demand for Ca is less in grasses and cereals than in other plant species and is generally higher in plants native to calcareous soils (Broadley et al., 2003; White and Broadley, 2003). Similarly, the nutritional demand for Mg is less in grasses and cereals than in other plant species, and members of the Caryophyllales appear to have a greater demand for Mg than other eudicot species (Broadley et al., 2004, 2008). The demand for K is high in all plant species, but leaf K concentrations expressed on a fresh weight basis are often highest in members of the Poales, Ranunculales and Rosales (Broadley et al., 2004; Watanabe et al., 2007). By contrast, the demand for essential micronutrients is low and, although some higher level phylogenetic trends have been reported in leaf concentrations of Ni, Zn and Cu (Broadley et al., 2001, 2004, 2007; Watanabe et al., 2007), and with the notable exception of the rare hyperaccumulator plants, these may have little practical consequence for plant mineral nutrition. There is also considerable variation both between and within plant species in their ability to acquire and utilise essential cationic elements for growth and seed production and, consequently, their ability to grow on soils with low phytoavailability of essential cationic elements (Broadley et al., 2007; Fageria, 2009; White and Broadley, 2009). There is also variation in the ability of plants to tolerate mineral toxicities arising in acidic, alkaline, saline, sodic and calcareous soils (Fageria, 2009). Variation in these characteristics can be used to develop novel crops for infertile and/or hostile soils.

6.6.3 Differences between and within plant species in the acquisition and utilisation of essential cationic elements

Several definitions are available in the literature relating to the efficiency of use of mineral elements in agriculture (White et al., 2005). Here, the efficiency of agronomic use of an essential mineral element (MUE) is defined as the dry matter yield of a crop per unit of a mineral element available to it (g DM g⁻¹ Mₐ). This is numerically equal to the product of crop content (M) per unit mineral element available (g M g⁻¹ Mₐ), which is referred to as a
crop’s uptake efficiency (MUpE), and crop yield per unit content (g DM g⁻¹ M), which is referred to as a plant’s utilisation efficiency (MUtE).

The ability to acquire and utilise K⁺ is fundamental to reducing the amount of K-fertiliser used in agriculture. In general, grasses and cereals often achieve their growth potential with a lower K⁺ supply than legumes, vegetable or tuber crops and also compete best in K-limited ecosystems (Pretty and Stangel, 1985; Mengel et al., 2001; Fageria, 2009). Some arable crops show marked responses to K-fertiliser. These include potatoes, sugar beet and some vegetables. Differences in response to K-fertiliser and agronomic KUE have also been demonstrated among genotypes of several crop plants (Trehan, 2005; Fageria, 2009). Variation in K uptake efficiency (KUpE) has been observed among genotypes of many plant species (Trehan, 2005; Rengel and Damon, 2008; Fageria, 2009; White et al., 2010). Greater KUpE has been attributed by various authors to (1) increased K⁺ uptake capacity of root cells, which reduces the K⁺ concentration in the rhizosphere solution and increases diffusional K⁺ fluxes to the root surface; (2) release of more non-exchangeable K⁺ by root exudates, which increases the K⁺ concentration in the soil solution; (3) proliferation of roots into the soil volume, which not only increases the area for K⁺ uptake but also reduces the distance required for K⁺ diffusion and water flow; and (4) higher transpiration rates, which accelerate the mass flow of the soil solution to the root surface (Jungk and Claassen, 1997; Baligar et al., 2001; Høgh-Jensen and Pedersen, 2003; Trehan, 2005; Rengel and Damon, 2008; Fageria, 2009; White et al., 2010). Within-species genetic variation in tissue K concentration and KUtE has also been reported (Baligar et al., 2001; Rengel and Damon, 2008; Fageria, 2009; White et al., 2010). Greater KUtE has been attributed to better K⁺ redistribution within the plant, from older tissues to younger tissues with immediate K requirements, and/or a better ability to maintain appropriate cytoplasmic K⁺ concentrations, either by anatomical adaptations or by the substitution of K⁺ for other solutes in the vacuole (Rengel and Damon, 2008; White et al., 2010).

In most environments, the concentrations of Ca²⁺ and Mg²⁺ in the soil solution are high enough for sufficient Ca and Mg to be delivered to the root system by transpiration-driven mass flow of the soil solution for adequate plant nutrition (Wilkinson et al., 1990; Barber, 1995; White and Broadley, 2003 and 2009). However, since human dietary deficiencies of both Ca and Mg are prevalent in both developing and developed countries (Thacher et al., 2006; Broadley and White, 2010), efforts to increase Ca and Mg concentrations in edible portions through plant breeding are currently being pursued (White and Broadley, 2009). Considerable variation has been observed among genotypes of several species in their growth responses to Ca²⁺ and Mg²⁺ in the soil solution and also in their tissue Ca and Mg concentrations, which often reflect differences in the acquisition and accumulation of these elements (White and Broadley 2003, 2009; Havlin et al., 2005; Broadley et al., 2008). For example, calcifuge ecotypes achieve their growth potential at lower Ca²⁺ concentrations in the soil solution than calcicole ecotypes and are also generally more tolerant of acid soils (White and Broadley, 2003). Similarly, crops that are adapted to acid soils generally have lower Ca requirements (Fageria, 2009).

The total concentrations of Fe, Zn, Cu and Mn in most soils would be sufficient for crop nutrition if these elements were available in the soil solution (see Section 6.5). Hence, there is considerable interest in breeding crops that produce higher yields on previously infertile soils through greater acquisition and/or better physiological utilisation of these elements (Graham et al., 2007; Lynch, 2007; White and Broadley, 2009). The acquisition of mineral elements with restricted mobility in the soil, such as Fe, Zn, Cu and Mn can be improved by (1) investing more biomass in the root system; (2) producing a greater number and more
Soil Conditions and Plant Growth

186

even spread of roots; (3) developing a more extensive root system, with longer, thinner roots
with more root hairs; and (4) proliferating lateral roots in mineral-rich patches (White et al.,
2005; Lynch, 2007; White and Broadley, 2009). In addition, the exudation of protons, phy-
tosiderophores and organic acids by roots, which displace cations from their binding sites in
the soil, and the secretion of enzymes capable of degrading organic compounds that chelate
mineral cations, such as phytases, phosphatases and proteases, can also improve their phy-
toavailability and acquisition by plants (Welch, 1995; Rengel, 2001; Abadía et al., 2002;
Alloway, 2004; Ismail et al., 2007; Broadley et al., 2007; Lynch, 2007; Uren, 2007; Degryse
et al., 2008; Fageria, 2009; White and Broadley, 2009). Increasing Fe(III) reductase activity
in tandem with Fe$^{2+}$ transport capacity can increase Fe acquisition by Strategy I Plants,
whilst greater production and exudation of phytosiderophores in tandem with an increased
capacity for Fe-phytosiderophore uptake can increase Fe acquisition by Strategy II Plants
(White and Broadley, 2009). There is considerable within-species genetic variation in root
architecture, root exudation and root uptake capacity that might be harnessed to improve the
acquisition of Fe, Zn, Cu and Mn from unfertilised soils (White et al., 2005; Lynch, 2007;
Fageria, 2009). Genotypic differences have been observed both between and within plant
species in their growth responses to Fe, Zn, Cu and Mn availability in the soil solution.
Crops that are highly susceptible to Fe deficiency and respond well to the addition of
Fe-fertilisers, include sorghum, soybean, common bean, proso millet and rice, whilst
Fe-efficient crops include wheat, sunflower and amaranth (Fageria, 2009). Among cereals,
maize responds more to the application of Zn-fertiliser than does millet, rice or wheat
(Rashid and Fox, 1992). Oats, spinach, wheat, alfalfa and barley generally respond well to
Cu-fertilisers, whilst beans, peas, brassica, rye, grass, potatoes and soybeans respond less
well (Mengel et al., 2001). Oats are also extremely susceptible to Mn deficiency and respond
well to Mn-fertiliser applications, whilst rye, which is efficient in Mn uptake and utilisation,
responds less dramatically to the application of Mn-fertilisers (Welch et al., 1991; Havlin
et al., 2005). There is also considerable within-species variation in the uptake of Fe, Zn, Cu
and Mn, in tissue concentrations of these elements and in the physiological utilisation of
these elements to produce yield (Graham, 1988; Chaney et al., 1989; Cakmak et al., 1998;
Rengel, 1999; Hacisalihoglu and Kochian, 2003; Grusak and Cakmak, 2005; Jolley et al.,
2006; Evans et al., 2007; Pfeiffer and McClafferty, 2007; Wu et al., 2007; Fageria, 2009;
White and Broadley, 2009; Broadley et al., 2010).

6.7 Concluding remarks

Plants acquire eight essential mineral elements from the soil solution in their cationic form
(Table 6.1). Three macronutrients, K, Ca and Mg, are required at tissue concentrations of
several mg g$^{-1}$ DM and five micronutrients, Zn, Fe, Cu, Mn and Ni, are required at lower
tissue concentrations. In addition, several other mineral elements present as cations in the
soil solution are considered beneficial for the growth of various taxa under certain environ-
mental conditions. These include Na, Co and Al. Under most agricultural conditions,
cationic elements are taken up by all regions of the root through transport proteins located
in the plasma membrane of epidermal or cortical cells and are delivered to the xylem via a
symplastic pathway. Notable exceptions to this generalisation are Ca and, in some plant
species, Fe, whose uptake is restricted to the root apex and whose movement to the xylem is
predominantly through the apoplast. In addition, when a high concentration of a particular
cation is present in the soil solution, a substantial proportion reaching the shoot might reach
the xylem though an apoplastic pathway. Many genes encoding transport proteins catalysing the influx of essential cationic elements to the root symplast and their delivery to the xylem have been identified (Figure 6.3).

Most cationic mineral elements are delivered to the root surface by transpiration-driven mass flow of the soil solution (Sections 6.1 and 6.2). Thus, their absolute concentrations in the soil solution are critical for adequate plant nutrition. The concentrations of cations in the soil solution are determined by the chemistry of the parent rocks, the processes of weathering, agronomic practices and vegetation processes (Sections 6.4, 6.5 and 6.6). Appropriate management strategies, such as the manipulation of soil pH or the application of mineral fertilisers, crop residues or animal manures, can increase the acquisition of cationic mineral elements by plants. In acid soils, which restrict crop production in many areas of the world, the application of dolomitic lime is recommended to raise soil pH, alleviate Al and Mn toxicities and prevent Ca and Mg deficiencies. The pH of alkaline calcareous soils, where crop production is limited by high pH, Ca\(^{2+}\) and bicarbonate concentrations, can be lowered by the addition of S, Fe sulphates or aluminium sulphate and fertilisers containing urea, ammonium or phosphate, although the effects of these amendments are often short lived. Saline soils can be remediated by leaching soluble salts from the soil profile by flushing with freshwater, and sodic soils can be remediated through the application of Ca\(^{2+}\), generally as gypsum, followed by flushing with freshwater (see Chapter 7).

Where crop production is restricted by low phytoavailability of Fe, Zn, Cu or Mn, these elements can be supplied as soil or foliar fertilisers (Section 6.5). Alternatively, genotypes that produce higher yields on infertile soils, whether through greater acquisition and/or better physiological utilisation of these elements, can be grown (Section 6.6). Genotypes with greater acquisition of Fe, Zn, Cu and Mn often possess an extensive root system that exploits the spatial and temporal variations in mineral phytoavailability, and their roots release protons, phytosiderophores, organic acids and diverse enzymes into the rhizosphere to increase the concentrations of Fe, Zn, Cu and Mn in the soil solution. The cultivation of companion plants or soil microorganisms that release protons, phytosiderophores, organic acids and hydrolytic enzymes into the rhizosphere can also be exploited to increase the phytoavailability of these elements for crop production.

References


Hänsch, R. and Mendel, R.R. (2009) Physiological functions of mineral micronutrients (Cu, Zn, Mn, Fe, Ni,


Properties and management of cationic elements for crop growth


7 Managing adverse soil chemical environments

Mark E. Hodson¹ and Erica Donner²

¹Environment Department, University of York, York, UK
²Centre for Environmental Risk Assessment and Remediation, University of South Australia, Mawson Lakes, South Australia, Australia

7.1 Potentially toxic metals (PTMs)

All metals are potentially toxic. As the fifteenth to sixteenth century physician Paracelsus said ‘the dose determines the response’. However, within studies of contaminated soils/adverse environments, there is a clear perception of typical potentially toxic metals. These are often erroneously referred to as heavy metals, an ill-defined term (Duffus, 2002; Hodson, 2004) that is used as a synonym for nasty metals that have a toxic effect. In many respects, the phrase ‘potentially toxic metal’ is as unsatisfactory as ‘heavy metal’ since, as Paracelsus tells us, all metals, indeed all substances, can be toxic. However, the phrase does at least make the point that the metals in question though present need not be having a toxic effect. The majority of studies concerned with potentially toxic metals are concerned with Al, Cd, Cr, Cu, Hg, Ni, Pb and Zn, and examples of case studies in this section will focus on these elements.

7.1.1 Sources of PTMs

PTMs are naturally occurring elements and are therefore ubiquitous in the soil and wider environment. They cause a problem when present at concentrations which are elevated above critical levels. These critical levels are the subject of much research and indeed legislation and depend on a number of factors such as soil properties but also proposed land use.

Natural sources

Metals are present in rocks and, therefore, through natural pedogenic processes end up in soils. Concentrations of PTMs in ‘typical’ uncontaminated soils are listed in Table 7.1. Fuller listings are given in, for example, Kabata-Pendias (2000), Bowen (1979) and Alloway (2013a).
Airborne dust and aerosols are also increasingly recognised as inputs to soil systems (e.g. Derry and Chadwick, 2007). Soils have naturally elevated concentrations of PTMs when developed on metal-enriched rocks, for example soils developed on ophiolite complexes and mineral deposits (Bowen, 1979; Alloway, 2013a; Kabata-Pendias, 2000) and when weathering processes have led to the concentration of specific elements. In the latter case, this is normally classically immobile elements such as Al and Fe in highly weathered tropical soils leading to the formation of laterites (e.g. Brimhall et al., 1991; Samouelian and Cornu, 2008; Fandeur et al., 2009).

**Anthropogenic sources**

Many textbooks have been written on the anthropogenic enrichment of metal concentrations in soils (e.g. Alloway, 2013a; Kabata-Pendias, 2000). In brief, metal enrichment can occur due to elevated concentrations of metals in soil amendments and/or accidental addition of metals to soils due to industrial activities and spillage. Classic examples of increased metal concentrations due to soil amendments are the occurrence of Cd in phosphate fertilisers (Smolders and Mertens, 2013) and elevated concentrations of PTMs in sewage sludge applied to land (Alloway, 2013b). This latter example, at least in Europe, is historic. Present-day sewage sludges that are licensed for land disposal have very low concentrations of metals.

### 7.1.2 Controls on mobility/availability of PTMs

A wide variety of factors control the mobility and availability of PTMs to plants in soils. Ultimately, mobility depends upon the PTMs being present in the soil solution and movement of that solution through the soil, whilst availability depends upon the PTMs being in solution and to a certain extent the speciation, that is the chemical form, of the PTM in the solution. In this section, we will briefly review the impacts of pH, adsorption, Eh, organic matter and speciation on mobility and availability.

pH is perhaps the most important variable governing mobility and availability. Metals, with the exceptions of those that form oxyanions such as Cr (e.g. CrO$_4^{3−}$), are more soluble under acidic than alkaline conditions. In a pure solution comprising dissolved metal M,
Managing adverse soil chemical environments

197

Water, H⁺ and OH⁻ as pH increases from acid to alkali, metals typically change their speciation from cations to neutrally charged hydroxide molecules to hydroxyanions, i.e.:

\[
M^{n+}_{(aq)} + nOH^-_{(aq)} \Rightarrow M(OH)_n^{0}_{(aq)} 
\]

\[
M^{m+}_{(aq)} + mOH^-_{(aq)} \Rightarrow M(OH)_{m+n}^{-m}_{(aq)} \quad (7.1)
\]

If the concentration of the neutrally charged hydroxide ion is sufficiently great, precipitation will occur. In reality, of course the situation is more complex. Firstly, the presence of other cations and anions in solution affects what complexes form and secondly ions interact with surfaces. However, the generalisation holds that PTMs are more mobile and available at lower pH. Speciation impacts availability and will be discussed at the end of this section. The impact of surfaces on metal mobility will now be briefly touched upon.

Surfaces provide sites where metals (or indeed any molecule in solution) may precipitate out of solution, provided concentrations are high enough. An energy barrier (activation energy) has to be overcome for precipitation to occur. This barrier is lower when precipitation occurs on a surface. However, the main way in which surfaces impact on PTM mobility and availability is through adsorption. Adsorption can be defined as the ‘accumulation of substance/material at an interface between the solid surface and the bathing solution’ (Sparks, 1995). There are two main varieties of adsorption: specific and non-specific (Figure 7.1).

Specific adsorption is considered to be the more permanent of the two. It involves covalent-type bonds between specific (hence the name) surfaces and molecules in solution and renders PTMs immobile and unavailable. Non-specific sorption involves ionic-type bonds between exchange sites on surfaces and charged molecules in solution. Non-specific sorption is reversible. It therefore immobilises PTMs, but only temporarily, and ions held on exchange sites are often thought to be available to plants (and other soil organisms). Protons compete with cationic PTMs for negatively charged exchange sites and typically at lower pHs such sites are dominated by H⁺. Although published about 30 years ago, the textbook of Lindsay (1979) is still an excellent reference that contains partition coefficients, phase diagrams and solubility data for a range of PTMs, which allow prediction of the impact of pH both on solubility and exchange site occupancy.

As well as competing for exchange sites, pH is also important in generating such sites. Soil particles often have either a negative or positive charge on their surface. This charge may be pH independent or pH dependent. pH independent charge arises from charge imbalances in the solid due to the isomorphous substitution of ions of differing charge, for example Al replacing Si or Mg replacing Al in mineral structures, resulting in a deficit of positive charge and therefore generating a negatively charged surface. Such pH independent charge is typical of 2:1 clays such as smectite. pH dependent charge is more typical of 1:1 clays such as kaolinite and oxyhydroxides of Al, Fe and Si. At low pH, the hydroxyl groups on the surface of such solids protonate, resulting in a positively charged surface, whilst at higher pH the hydroxyls deprotonate thereby generating negative surface charge. The pH at which such surfaces have a neutral charge is known as the point of zero charge (PZC) and this varies from solid to solid. Typically, the PZC of Fe-oxyhydroxides is 7–9, of Al-oxyhydroxides is 8–9.2, of Mn-oxyhydroxides is 1.5–4.6 and of clay minerals is 5–6. Thus, in a ‘typical’ slightly acidic soil Fe- and Al-oxyhydroxides have a positive charge and Mn-oxyhydroxides have a negative charge.

It is difficult to make generalisations about the role that organic matter plays in PTM mobility and availability. The dissociation of carboxylic and phenolic acid groups can provide exchange sites for PTMs. However, PTMs can also bond specifically to organic matter, often forming ring structures in which the PTM bonds to two or more hydroxyl groups (Figure 7.2).
These so-called chelate complexes are very stable and, when associated with solid organic matter, lead to the immobilisation of PTMs. When organic matter is present in a dissolved state, complexed PTMs can be held in solution at concentrations far higher than predicted in inorganic systems. Although the mobility of PTMs held in solution like this may be high, their availability is usually low (see later).

The redox state of soils (how reducing or oxidising they are) impacts on metal mobility and availability in a number of ways. Under reducing conditions, highly sorptive oxyhydroxides such as Fe- and Mn-oxyhydroxides will dissolve and any adsorbed PTMs will be...
released into solution. In addition, generally the lower valance state of multivalent cations is more soluble. These increases in solubility can be offset, to a certain extent by the precipitation of reduced species such as sulphides, though clearly this depends on S concentrations in the soil. Typically, reducing conditions develop when soils are poorly drained. In such cases the mobility of the PTMs will be relatively low although they will be available for uptake by plants and animals that can tolerate the reducing conditions.

Finally, in this section speciation must be mentioned. There is increasingly good evidence that, at least in the short term, the speciation of metals governs their availability for uptake (Di Toro et al., 2001; Van Gestel and Koolhaas, 2004; Steenbergen et al., 2005; Thakali et al., 2006a,b; Arnold et al., 2007; Lock et al., 2007; Allen et al., 2008). It is now widely accepted that metals are available for uptake when present as free metal ions and also as simple inorganic ions such as chlorides whilst metals present as larger organo-metal complexes are not available because the metals are held too tightly to the organic ion to sorb onto, and then pass through, biological interfaces. Thus, generally as ionic strength and dissolved organic carbon content of soil solution increases, a decrease in bioavailability would be expected. However, metal complexes will dissociate to release free metal ions and maintain steady-state conditions as free metal ions are taken up by organisms. Thus, although metals in organo-metal complexes may be unavailable in the short term, they do represent a reservoir of potentially available metals held in the soil.

7.1.3 Impacts of PTMs on plant growth: toxic effects including case studies

Potentially toxic metals can be subdivided into those which are essential for plant growth and those which are not. The impacts of deficiencies of those elements which are essential have been dealt with in Chapter 6. This section will therefore be concerned solely with the impacts of excesses of PTMs on plants. In broad terms, an excess of any PTM will have a toxic effect with severity of that effect increasing with the availability and concentration of the PTM. The actual value of an ‘excess’ of a PTM which results in harm will depend on soil properties (see Section 7.1.2) and also the plant. Plants that have evolved to tolerate unusually high concentrations of PTMs are discussed in Section 7.2.

Kabata-Pendias (2000) provides a useful table summarising the toxic effects of different PTMs on plants, which is reproduced in Table 7.2.

The causes of these toxic effects are summarised as:

- changes in permeability of the cell membrane (Ag, Au, Br, Cd, Cu, F, Hg, I, Pb, UO₂);
- reactions of thiol groups with cations (Ag, Hg, Pb);
- competition for sites with essential metabolites (As, Sb, Se, Te, W, F);
- affinity for reacting with phosphate groups and active groups of ADP or ATP (Al, Be, Sc, Y, Zr, lanthanides and possibly all other PTMs);
- replacement of essential ions (mainly major cations) (Cs, Li, Rb, Se, Sr);
- occupation of sites for essential groups such as phosphate and nitrate (arsenate, fluorate, borate, bromate, selenate, tellurate and tungstate).

Although toxicity varies from plant to plant and is dependent on soil conditions, in broad terms, the most toxic PTMs to plants are Hg, Cu, Ni, Pb, Co, Cd and possibly Ag, Be and Se (Kabat-Pendias, 2000). Due to variability in soil conditions and plant tolerance, it is not possible to state with any degree of accuracy the concentration of these elements which causes
<table>
<thead>
<tr>
<th>Element</th>
<th>Symptoms</th>
<th>Sensitive crop</th>
</tr>
</thead>
<tbody>
<tr>
<td>Al</td>
<td>Overall stunting, dark green leaves, purpling of stems, death of leaf tips and coralloid and damaged root system</td>
<td>Cereals</td>
</tr>
<tr>
<td>As</td>
<td>Red-brown necrotic spots on old leaves, yellowing or browning of roots, depressed and wilting of new leaves</td>
<td>Legumes, onions, spinach, cucumbers, bromgrass, apricots, peaches</td>
</tr>
<tr>
<td>B</td>
<td>Margin or leaf tip chlorosis, browning of leaf points, decaying growing points and wilting and dying-off of older leaves. In severely affected pine trees, necrosis occurs on needles near the ends of shoots and in the upper half of the tree</td>
<td>Cereals, potatoes, tomatoes, cucumbers, sunflowers, mustard, apple, apricots, citrus, walnut</td>
</tr>
<tr>
<td>Be</td>
<td>Inhibition of seed germination and reduced growth, degradation of protein enzymes</td>
<td>–</td>
</tr>
<tr>
<td>Cd</td>
<td>Brown margin of leaves, chlorosis, reddish veins and petioles, curled leaves and brown stunted roots. Severe reduction in growth of roots, tops and number of tillers (in rice). Reduced conductivity of stem, caused by deterioration of xylem tissues. Reduction of chlorophyll and carotenoids</td>
<td>Legumes, spinach, radish, carrots, oats</td>
</tr>
<tr>
<td>Co</td>
<td>Interverinal chlorosis in new leaves followed by induced Fe chlorosis and white leaf margins and tips and damaged root tips</td>
<td>–</td>
</tr>
<tr>
<td>Cr</td>
<td>Chlorosis of new leaves, necrotic spots and purpling tissues, injured root growth</td>
<td>–</td>
</tr>
<tr>
<td>Cu</td>
<td>Dark green leaves followed by induced Fe chlorosis, thick, short or barbed-wire roots, depressed tillering. Changes in lipid content and losses of polypeptides involved in photochemical activities</td>
<td>Cereals, legumes, spinach, citrus seedlings, gladiolus</td>
</tr>
<tr>
<td>F</td>
<td>Margin and leaf tip necrosis and chlorotic and red-brown points of leaves</td>
<td>Gladiolus, grapes, fruit trees, pine trees</td>
</tr>
<tr>
<td>Fe</td>
<td>Dark green foliage, stunted growth of tops and roots, dark brown to purple leaves of some plants (e.g. Bronzing disease of rice)</td>
<td>Rice, tobacco</td>
</tr>
<tr>
<td>Hg</td>
<td>Severe stunting of seedlings and roots, leaf chlorosis and browning of leaf points</td>
<td>Sugar beets, maize, roses</td>
</tr>
<tr>
<td>Li</td>
<td>Chlorotic and necrotic spots on leaves and injured root growth</td>
<td>Citrus</td>
</tr>
<tr>
<td>Mn</td>
<td>Chlorosis and necrotic lesions on old leaves, blackish-brown or red necrotic spots, accumulation of MnO₂ particles in epidermal cells, drying tips of leaves, and stunted roots and plant growth</td>
<td>Cereals, legumes, potatoes, cabbage</td>
</tr>
<tr>
<td>Mo</td>
<td>Yellowing or browning of leaves, depressed root growth, depressed tillering</td>
<td>Cereals</td>
</tr>
<tr>
<td>Ni</td>
<td>Interverinal chlorosis [caused by Fe-induced deficiency] in new leaves, grey-green leaves and brown and stunted roots and plant growth</td>
<td>Cereals</td>
</tr>
<tr>
<td>Pb</td>
<td>Dark green leaves, wilting of older leaves, stunted foliage and brown short roots</td>
<td>–</td>
</tr>
<tr>
<td>Rb</td>
<td>Dark green leaves, stunted foliage and increasing amount of shoots</td>
<td>–</td>
</tr>
<tr>
<td>Se</td>
<td>Interverinal chlorosis or black spots at Se concentration of about 4 ppm and complete bleaching or yellowing of younger leaves at higher Se content, pinkish spots on roots</td>
<td>–</td>
</tr>
<tr>
<td>Ti</td>
<td>Chlorosis and necrosis of leaves, stunted growth</td>
<td>Beans</td>
</tr>
<tr>
<td>Tl</td>
<td>Impairment of chlorophyll synthesis, mild chlorosis and slight cupping of leaves, reduced germination of seeds and growth of plants</td>
<td>Tobacco, cereals</td>
</tr>
<tr>
<td>Zn</td>
<td>Chlorotic and necrotic leaf tips, interveinal chlorosis and leaves, retarded growth of entire plant and injured roots resemble barbed wire</td>
<td>Cereals, spinach</td>
</tr>
</tbody>
</table>

Source: Reproduced from Kabata-Pendias (2000). Reproduced with permission from Taylor & Francis Group LLC.
Managing adverse soil chemical environments

toxicity in plants, but concentrations in the order of single to a few hundred ppm in plant leaf tissue have a toxic effect. Plants have evolved a variety of methods for tolerating excessive concentrations of PTMs (Figure 7.3) and these mechanisms include (Kabata-Pendias, 2000; Memon and Schroder, 2009):

- selective uptake of ions;
- reducing the permeability of cell walls or modifying membranes either structurally or functionally;
- isolating ions in an immobile form via the precipitation of minerals or fixation by ligands;
- modification of the impacted metabolic pathway, for example by increasing the enzyme system that is affected or bypassing the affected metabolic pathway;
- alteration of enzymes so that the PTM is incorporated without enzyme functionality being affected;
- enhanced excretion of the PTM via leaf shedding, guttation, root exudates, volatilisation, etc.

Thus, one strategy when confronted with soil containing elevated concentrations of PTMs is simply to grow plants that can tolerate these concentrations. However, there is then the potential of moving the PTMs through the food chain, particularly if the plants grown are used as a food source. Thus, the most usual route is to remediate the soil. There are many possible remedial treatments for contaminated land, which are dealt with in specialist texts (e.g. La Grega et al., 2001; Nathanail and Bardos, 2004). However, Section 7.1.4 will consider phytoremediation, the use of plants to remediate contamination. Phytoremediation may be applied to both PTMs and organic contaminants. The latter are outwith the scope of this chapter. Application of phytoremediation to organic contaminants is well described in, for example, Gerhardt et al. (2009), Cunningham et al. (1996) and Schnoor et al. (1995).

7.1.4 Phytoremediation

Phytoremediation is the use of plants to remediate contaminated soils. The concept has its origins in the measurement of high concentrations of metals (primarily Ni) in the tissues of plants growing in soils, which in turn contained elevated concentrations of PTMs (e.g. Minguzzi and Vergano, 1948; Doksopulo, 1961; Severne and Brooks, 1972; Cole,1973; Jaffré and Schmidt, 1974; Jaffré et al., 1976; Brooks et al., 1977). It was later realised that such plants could be transferred from their native habitats to contaminated environments where, by accumulating similar concentrations of metals in their tissues, the plants would gradually remove the contaminants from the soil (Yamada et al., 1975; Chaney, 1983; Baker and Brooks, 1989). Since the initial exploration of phytoremediation, several distinct concepts have evolved which are explored later. Useful recent reviews of phytoremediation include Zhao and McGrath (2009), Dickinson et al. (2009) and Robinson et al. (2009).

Phytoextraction

Phytoextraction is the term used for the basic method of phytoremediation. Plants which extract metals into their tissues, thereby removing the contaminants from the soil, are grown. The method relies on the identification of hyperaccumulators, the name given to plants that can accumulate high concentrations of metals in their tissues. Hyperaccumulator
Figure 7.3 Overview of our current understanding of adaptations to hyperaccumulate metals, in particular zinc. (I) Enhanced zinc uptake by root cells is thought to be driven by metal transport genes such as ZIP4. In some hyperaccumulators such as *Thalassia caerulea* a peri-endodermal thickening (apoplastic barrier along the root axis) is present. (II) Active xylem loading depends on reduced vacuolar root sequestration (main zinc storage is thought to be in the cortex) and enhanced activity of a P-type metal ATPase gene, *HMA4* in stellar parenchyma, leading to efficient efflux of zinc from the symplasm. Efflux of citrate seems to play a role in zinc translocation. Loading of zinc in the xylem can occur as Zn-nicotianamine complexes. (III) In leaves Zn is unloaded in vessel-associated cells. (IV) Detoxification is assumed to be operated by chelation of metals. $<>$ refers to chelation. Possible ligands of zinc in the cytoplasm are histidine and nicotianamine. (V) Vacuolar sequestration in the leaves is thought to be the main pathway of detoxification of metals. Zn is mainly stored in vacuoles or mesophyll and/or epidermal cells. In the vacuoles, a large pool of malate favours the formation of Zn-malate complexes. (VI) Other adaptive processes include homeostasis of other nutrients, in particular iron and phosphorus. Adapted from Verbruggen et al. (2009). Copyright John Wiley & Sons.
is a loosely defined term. Brooks et al. (1977) defined hyperaccumulators, with reference to Ni, as those plants in which tissue concentrations exceed 1000 mg kg\(^{-1}\). Subsequent authors have suggested that the tissue concentrations required to classify a plant as a hyperaccumulator depend on the element with concentrations of 100 (e.g. Cd, Baker et al., 1994), 1000 (e.g. Ni, Brooks et al., 1977) and 10 000 mg kg\(^{-1}\) (e.g. Zn, McGrath, 1998) being proposed. Kabata-Pendias (2000) describes hyperaccumulators as plants that accumulate metals to concentrations greater than background levels and more than other species grown in the same soil. A range of hyperaccumulators for a range of metals have been identified. Metals that can be hyperaccumulated include As, Cd, Cr, Cu, Mn, Ni, Pb, Se and Zn (e.g. Baker et al., 1994; Kumar et al., 1995; Reeves and Baker, 2000; LaCoste et al., 2001; Ma et al., 2001; Galeas et al., 2007; Casado et al., 2008; Fernando et al., 2008). Table 7.3 summarises numbers of hyperaccumulators identified by Reeves and Baker (2000). The high number of Ni hyperaccumulators reflects the large amount of research carried out on plants growing in soils derived from serpentinite and other ultramafic rocks. Table 7.4 gives selected examples of more recent studies that have identified hyperaccumulators. Additional plant genera that include hyperaccumulators include Arabidopsis, Astragalus, Brassica, Garcinia, Helianthus, Maytenus, Pelargonium, Pteris, Stanleya and Thlaspi (Cunningham and Ow, 1996; McGrath, 1998; Galeas et al., 2007; Fernando et al., 2008; Gonzaga et al., 2009; Liang et al., 2009).

Many laboratory studies have been published in the peer-reviewed literature in which plants have been shown to accumulate high concentrations of metals from aqueous solution or from soils amended with metal salts. In both cases, the metals have a high degree of availability. There are comparatively fewer studies in which plants have been shown to accumulate metals from contaminated sites (as opposed to naturally metal-enriched soils) in the field. In the authors’ opinion, this demonstrates that phytoremediation is not yet a fully developed technology (e.g. Van Nevel et al., 2007) though the high costs associated with field trials and potentially the commercial sensitivity of field trial results may also explain the relative paucity of published results from field trials compared to pot and greenhouse trials. A few examples of field studies are given in Table 7.5.

Phytoextraction has two clear advantages over many remedial methods – it is of low cost and low impact. Amongst its disadvantages are that it can take a long time and that most hyperaccumulators are element specific and do not address the range of contaminants found at a site. The issue of time is largely a result of the fact that the mass of metal removed from a contaminated site is a function of both the concentration of metal in the plant tissues and the biomass of the plant. The majority of hyperaccumulators have a low biomass and a low growth rate, a
result of the energy cost of accumulating metals. One phytoremediation method that attempts to get around the high concentration–low biomass problem is assisted phytoextraction.

**Assisted phytoextraction**

In this method, soil chemistry is manipulated to render the contaminants more available to plants so that normally non-accumulating, high biomass plants can be grown and used to remove metals from soil. This technique has its origins in attempts to render Pb, which is generally immobile in soils, available to plants so that it might be extracted rather than being left immobile in the soil. In this authors’ view, there is a certain irony in the amount of research effort that has been expended towards mobilising otherwise relatively inert contaminant lead in soils. In the majority of studies to date, soil chemistry is manipulated by the addition of a chelation agent, most usually EDTA (e.g. Huang and Cunningham, 1996; Lombi et al., 2001; Lopez et al., 2007; Shibata et al., 2007; Dourmet et al., 2008; Saifullah et al., 2009; Sun et al., 2009) but also (in recognition that EDTA is a significant pollutant and relatively recalcitrant) ammonium thiocyanate (Anderson et al., 1999), citric acid, NH₄-citrate, oxalic acid, SS-ethylenediamine discursinic acid (EDDS), nitrilotriacetic acid (NTA) (Duquene et al., 2009), EDTA, tartrate and glutamate (Doumett et al., 2008), the phytohormones indole-3-acetic acid, gibberellic acid and kinetin (Lopez et al., 2007), citric acid and NTA (Irtelli and Navari-Izzo, 2006). The purpose of the chelating agent is to move more of the contaminant into solution from which it may be taken into plant tissues. Whilst assisted phytoremediation can result in high metal concentrations in high biomass plants thereby reducing the time required for remediation, the fate of the soil amendment and the mobilised contaminants requires careful monitoring. There is the possibility that both components could leak from the system into groundwater (Nowack et al., 2006).

Table 7.4 A selection of plant species and PTMs that they hyperaccumulate.

<table>
<thead>
<tr>
<th>Plant species</th>
<th>PTMs</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Pteris vittata</em> L.</td>
<td>As</td>
<td>Bui et al. (2011)</td>
</tr>
<tr>
<td><em>Pityrogramma calomelanos</em> L.</td>
<td>As</td>
<td>Bui et al. (2011)</td>
</tr>
<tr>
<td><em>Thelypteris palustris</em></td>
<td>As</td>
<td>Anderson et al. (2011)</td>
</tr>
<tr>
<td><em>Pteris vittata</em></td>
<td>As,</td>
<td>Kalve et al. (2011)</td>
</tr>
<tr>
<td><em>Viola principis</em> H. de Boiss</td>
<td>As</td>
<td>Lei et al. (2008)</td>
</tr>
<tr>
<td><em>Baccharis sarothroides</em> Gray</td>
<td>As,</td>
<td>Hoque et al. (2008)</td>
</tr>
<tr>
<td><em>Solanum</em> photoacinocarpum</td>
<td>Cd</td>
<td>Zhang et al. (2011)</td>
</tr>
<tr>
<td><em>Cynodon dactylon</em> L.</td>
<td>Pb,</td>
<td>Bui et al. (2011)</td>
</tr>
<tr>
<td><em>Lonicera japonica</em> Thunb.</td>
<td>Cd</td>
<td>Liu et al. (2011a)</td>
</tr>
<tr>
<td><em>Potentilla griffithii</em> Hook.</td>
<td>Cd,</td>
<td>Qiu et al. (2011)</td>
</tr>
<tr>
<td><em>Sedum alfredii</em> Hance</td>
<td>Cd,</td>
<td>Wu et al. (2007)</td>
</tr>
<tr>
<td><em>Leersia hexandra</em> Swartz</td>
<td>Cr</td>
<td>Liu et al. (2011b)</td>
</tr>
<tr>
<td><em>Spartina argentinensis</em></td>
<td>Cr</td>
<td>Redondo-Gomez et al. (2011)</td>
</tr>
<tr>
<td><em>Phytolacca</em> acinosa</td>
<td>Mn</td>
<td>Xu et al. (2007)</td>
</tr>
<tr>
<td><em>Ricinus communis</em> L.</td>
<td>Pb</td>
<td>Liu et al. (2008)</td>
</tr>
<tr>
<td><em>Tephrosia</em> candida</td>
<td>Pb</td>
<td>Liu et al. (2008)</td>
</tr>
<tr>
<td><em>Debregeasia</em> orientalis</td>
<td>Pb</td>
<td>Liu et al. (2008)</td>
</tr>
<tr>
<td><em>Eleusine indica</em> L.</td>
<td>Pb,</td>
<td>Bui et al. (2011)</td>
</tr>
<tr>
<td><em>Cyperus</em> rotundus* L.</td>
<td>Pb,</td>
<td>Bui et al. (2011)</td>
</tr>
<tr>
<td><em>Equisetum ramosissimum</em> (Vauch)</td>
<td>Pb,</td>
<td>Bui et al. (2011)</td>
</tr>
<tr>
<td>Soil type</td>
<td>Metals present</td>
<td>Plant type</td>
</tr>
<tr>
<td>--------------------------------------------------------------------------</td>
<td>----------------</td>
<td>---------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Paddy field polluted by Zn-refining plant dust</td>
<td>Cd</td>
<td>Indica rice cultivars MORETSU and IR-8</td>
</tr>
<tr>
<td>Soil mixed with pyrite ore roasting wastes</td>
<td>As, Co, Cu, Pb, Zn</td>
<td>Populus alba L., Populus nigra L., Populus tremula L., Salix alba L.</td>
</tr>
<tr>
<td>Historically sewage sludge amended soils</td>
<td>Cd, Zn</td>
<td>Thlaspi caerulescens, Arabidopsis halleri</td>
</tr>
<tr>
<td>Thermal power plant fly ash mixed with soil and lime</td>
<td>Cr, Cu, Pb, Zn, Mn, Fe, Ni, Cd</td>
<td>Pongamia pinnata, Tectona grandis, Delbergia sissoo, Cassia siamea, Eucalyptus hybrida, Dendrocalamus strictus</td>
</tr>
<tr>
<td>Contaminated soil</td>
<td>Cd, Cu, Pb, Zn</td>
<td>Salix viminalis</td>
</tr>
<tr>
<td>Sewage sludge disposal site</td>
<td>Cd, Zn</td>
<td>Nine varieties of Salix</td>
</tr>
</tbody>
</table>
Phytovolatilisation

Phytovolatilisation has been used for the remediation of both inorganic (e.g. Se and Hg) and organic (e.g. BTEX, TCE and MTBE) contaminated soils. It depends on contaminants being taken up into the plant, remaining in solution and being carried to the leaves from where the contaminants volatilise (Chaney et al., 1997; Baeder-Bederski-Anteda, 2003). In the case of Hg phytovolatilisation, genetic modification of *Arabidopsis* and *Nicotinia tabacum* is carried out to produce plants which express merA and merB genes and which take up methylmercury from soil solution, convert it to Hg(II) which is then volatilised as elemental Hg (e.g. Raskin, 1996; Rugh et al., 1998; Ruiz and Daniell, 2009). The volatilised Hg is significantly less toxic than methylmercury. For Se, it is the inorganic form of the element that is taken up and then converted to a less toxic organic form (dimethyl selenite) which is volatilised. The Se uptake path is related to that of S and a wide variety of genetically engineered pathways have been exploited to enhance the uptake and volatilisation of Se (e.g. Kubachka et al., 2007; Pilon-Smits and Leduc, 2009).

7.2 Salt-affected soils (saline, sodic and saline-sodic soils)

There are more than 930 million hectares of salt-affected land worldwide, equating to more than 7% of the world’s total land area (FAO, 1998). The management of these soils has been challenging humans since the advent of irrigated agriculture many millennia ago (Umali, 1993). They have even been implicated in the decline of historic civilisations and societies, such as Babylon and Sumeria in Asia, Carthage in North Africa and the Hohokam Indians in present-day Arizona (Umali, 1993; Schofield et al., 2001). An indication of the global distribution of salt-affected soils can be obtained from Figure 7.4.

7.2.1 Definitions of saline and sodic soils

The US salinity laboratory has defined saline soils as those with a saturated paste extract electrical conductivity (ECe) of $> 4$ dS m$^{-1}$ (Richards, 1954). This value corresponds to a total cation concentration (TCC) of 40 mmol L$^{-1}$ or an osmotic potential of $-145$ kPa (Sumner et al., 1998) and approximates the level at which many plant species are negatively affected by salt accumulation (Figure 7.5).

Although this value is widely agreed upon in the literature, the measurement of ECe is relatively laborious, hence the EC of a 1:5 soil:water extract (EC$_{1:5}$) is often used instead of the ECe. In such cases, the saturation water content can be estimated on the basis of soil texture and a conversion factor applied to predict the ECe. The accuracy of this conversion varies according to soil type. Typically, the ions responsible for the high electrical conductivity are chlorides and sulphates of calcium, magnesium, sodium and potassium, i.e. soluble salts that can reach high concentrations without precipitation occurring.

Sodic soils are generally defined by the parameter of exchangeable sodium percentage (ESP), which expresses the concentration of Na$^+$ held on soil exchange sites as a percentage of the cation exchange capacity of the soils as shown in Equation 7.2 (concentrations in cmol$_c$ kg$^{-1}$):

$$ESP = 100 \times \frac{\text{Exchangeable Na}}{\text{Cation Exchange Capacity}}$$  \hspace{1cm} (7.2)
Figure 7.4 Global distribution of saline/sodic soils (these soils would fall within the category of salinity/alkalinity according to this map key). http://soils.usda.gov/use/worldsoils/mapindex/stresses.html. With permission from USDA-NRCS, Soil Survey Division, World Soil Resources. For a colour version of this figure, please see Plate 7.1.
Soil Conditions and Plant Growth

It is worth noting that the term ‘sodic soil’ has now effectively replaced the old term ‘alkali soil’ as examples of acid sodic soils were increasingly being reported in the literature (e.g. Ford et al., 1993; McKenzie et al., 1993). Other expressions that have been previously used in reference to sodic soils include the terms black alkali, solonetz, slick-spot, non-saline sodic, saline sodic and saline alkali (Letey, 1984; Sumner et al., 1998; Qadir and Schubert, 2002).

The sodicity of irrigation water and soil solutions is defined using the parameter of sodium adsorption ratio (SAR) which is a measure of the relative concentrations of Na\(^+\), Ca\(^{2+}\) and Mg\(^{2+}\) in solution and can be calculated as shown in Equation 7.3 (concentrations in mmol L\(^{-1}\)):

\[
SAR = \frac{[\text{Na}^+]}{\left[\text{Ca}^{2+} + \text{Mg}^{2+}\right]^{1/2}}
\]  

(7.3)

Rengasamy and Marchuk (2011) recently proposed the introduction of a ‘cation ratio of soil structural stability’ (CROSS). This new concept is analogous to the SAR but incorporates both the differential dispersive effects of Na\(^+\) and K\(^+\) and the differential flocculating powers of Ca\(^{2+}\) and Mg\(^{2+}\). The validity of this concept as an index of structural stability across a wide range of soils requires further testing. It is quite likely however that CROSS may be a more effective predictor of soil structural behaviour where levels of K and Mg are high (e.g. in soils irrigated with recycled water).

Unlike saline soils, total concentrations of ions in sodic soils can be relatively low with an EC of <4 dS m\(^{-1}\). However, the ratio of Na\(^+\) to Ca\(^{2+}\) and Mg\(^{2+}\) ions is high. The original US salinity laboratory definition of sodic soils as those with ESP > 15 (Richards, 1954) has proved controversial over the years, partly due to the large-scale variability in salinity and sodicity levels across continents and partly due to a lack of consideration of the effects of

![Diagram illustrating the classification of normal, saline, saline-sodic and sodic soils in relation to soil pH, electrical conductivity (EC), sodium adsorption ratio (SAR) and exchangeable sodium percentage (ESP). Also shown are the ranges for different degrees of sensitivity of plants to salinity.](image)

**Figure 7.5** Diagram illustrating the classification of normal, saline, saline-sodic and sodic soils in relation to soil pH, electrical conductivity (EC), sodium adsorption ratio (SAR) and exchangeable sodium percentage (ESP). Also shown are the ranges for different degrees of sensitivity of plants to salinity. a, sodic soil, pH < 8.5; b, saline-sodic soil, soil pH generally < 8.5; c, normal soils, pH < 8.5; d, sodic soils, pH > 8.5. Reprinted from Brady and Weil (2010) by permission of Pearson Education, Inc., Upper Saddle River, NJ.
electrolyte concentration on sodic soil behaviour, for example the onset of soil dispersion (Szabolcs, 1979, Halliwell et al., 2001). As a result, there is currently no universally accepted ESP threshold value for sodic soils. In the USA, the value of 15 is still generally accepted, whereas in Australia, an ESP of 6 is most commonly used to denote sodic soils (Northcote and Skene, 1972). Sumner (1993) emphasised that sodic behaviour (e.g. dispersion) can even be apparent in soils with ESP values <1, demonstrating that sodicity is not confined to the traditional definitions of ESP presented earlier. The problems associated with ESP- and SAR-based sodicity definitions have prompted several authors (e.g. Sumner 1993; Shaw et al. 1994; Sumner et al. 1998) to recommend that new classification systems based on soil behaviour (e.g. dispersibility) and its limitations for productivity to be devised.

Regardless of issues relating to classification, broad definitions of sodic soils indicate that their key defining factor is the presence of a high proportion of exchangeable and/or soluble sodium ions relative to other cations (Shaw et al., 1994). Furthermore, these excess sodium ions are often at levels that can adversely affect soil structure and crop growth and are characteristically present in the form of salts capable of causing alkaline hydrolysis on cation exchange complexes (Schofield et al., 2001). Sodic soils can be acid or alkaline, saline or non-saline (Qadir and Schubert, 2002); however, most are alkaline with the dominant anion being the bicarbonate ion.

Some soils demonstrate characteristics of both saline and sodic soils and are hence termed ‘saline sodic’. These soils not only have elevated levels of neutral soluble salts but also demonstrate an excess of sodium salts capable of alkaline hydrolysis. In such cases, crop growth may be adversely affected not only by osmotic stress and the effects of excess salts, but also by the high Na+ levels and associated structural degradation of the soil (Burrow et al., 2002). It depends on a combination of factors such as the SAR, overall salt content and mechanical energy inputs (e.g. tillage practices) as to whether the soil will be flocculated or dispersive and whether or not other poor physical properties such as crusting will be exhibited (Sumner et al., 1998).

### 7.2.2 Origins and properties of saline and sodic soils

The formation of saline and sodic soils is intimately related to the water balance and salt distribution of the soil system. Both the salinisation and sodification of soils are caused by the accumulation of excess soluble salts at or near the soil surface. The source of the salts has frequently been used to distinguish between primary and secondary salinity/sodicity (Schofield et al. 2001; Qadir and Schubert, 2002), where primary salinity/sodicity relates to natural sources of salts, and secondary salinity/sodicity relates to anthropogenically sourced salts (e.g. pumping saline groundwater or effluent onto land as irrigation water). Rengasamy (2006) provides an alternative classification, with three major types of salinity defined on the basis of soil and groundwater processes; these being groundwater associated salinity, non-groundwater associated salinity and irrigation associated salinity.

**Saline soils**

Naturally occurring saline soils are most common in arid and semi-arid regions of the world, where low rainfall and leaching rates can lead to the accumulation of excess salts. The source of these salts can be varied, including Quaternary weathering of geological parent materials (Ford et al., 1993), aeolian sources (Shaw et al., 1994), windborne deposition from sea water and saltwater ingress through underground aquifers (Abrol et al., 1988).
The term *dryland salinity* is used to denote the occurrence of saline soils in a non-irrigated environment. This does not mean that dryland salinity is not anthropogenically induced, however, as although saline soils do occur naturally in some areas, agricultural practices in arid and semi-arid regions have been linked to a substantial increase in dryland salinity (Rengasamy, 2002, 2006). One of the major causes of increasing salinisation has been large-scale land clearance and deforestation for agricultural production, with the replacement of deep-rooted perennial species by shallow-rooted annual species allowing increased infiltration of water which then leads to rising watertables (Abrol et al., 1988; Fitzpatrick et al., 1994). In such cases, it is common for the soluble salts present in the soil to also rise up through the profile, concentrating in the soil surface layers and affecting plant growth. The amount of additional deep drainage or recharge necessary to cause salinisation can vary from as little as 5 mm to over 130 mm depending on the original salt/water balance and the system dynamics (Fitzpatrick et al., 1994). Where an impermeable layer is present in the soil profile, saline seeps may also occur, with infiltrating water being intercepted and conducted laterally to discharge on low lying land or in landscape depressions (Abrol et al., 1988).

The introduction of new irrigation systems and the expansion of existing systems is an ongoing worldwide trend (FAO, 2002) and *irrigation-induced salinity* is increasing hand-in-hand with the global increase in irrigated land area. Abrol et al. (1988) reported that almost 50% of the irrigated land in arid and semi-arid regions of the world was affected to some degree by salinisation and/or sodification. One of the major causes of increasing irrigation-induced salinity has been inadequate attention to drainage systems design and maintenance (Umali, 1993) as inadequate drainage may result in rising groundwater levels (Endo et al., 2002). In such a case, the soil may not only become waterlogged following rainfall, but may also become salinisated through the movement and accumulation of salts from naturally saline soil strata, from saline irrigation water and/or from the capillary rise of groundwater and soluble salts (Abrol et al., 1988; Umali, 1993). Salt accumulation and translocation is also determined partly by irrigation method, water quality and quantity (Pessarakli, 1991; Schofield et al., 2001).

Even in areas where drainage is adequate, irrigation-induced salinity may occur as a result of inadequate leaching. If salt build-up has occurred and insufficient volumes of good quality irrigation waters are applied, leaching, which is dependent on soil properties and permeabilities (Endo et al., 2002) may be inadequate to remove the salts from the root zone. It is worth noting that the re-use of wastewaters and effluents for irrigation is increasing in many countries around the world (Balks et al., 1998; Bond, 1998). Effluents typically contain relatively high salt loadings, which often include high sodium concentrations (Balks et al., 1998), suggesting there is a risk of the soils becoming degraded by effluent irrigation (Jayawardane et al., 2001; Tillman and Surapaneni, 2002).

**Sodic soils**

In some ways, the formation of sodic soils closely parallels that of saline soils, with the major difference being that the accumulating salts are high in Na⁺. Sodic soils present particularly complex problems for management however, in that not only the chemical properties but also the physical properties of the soils may be seriously affected. Some of the environmental conditions that promote the genesis of sodic soils include the presence of shallow saline groundwater, textural discontinuities throughout the profile (i.e. duplex soils), the occurrence of perched watertables close to the surface, low slope gradients and
impeded drainage (Fitzpatrick et al., 1994; Bui et al., 1998). Increasing soil sodification can lead to serious degradation of soil structural and physical properties, with concomitant impacts on plant growth and productivity (Rengasamy and Olsson, 1991). This topic is dealt with in detail in Chapter 9, but to summarise, as the Na⁺ concentration in a soil increases, repulsion forces between clay particles also increase, and at a unique level of sodicity and salinity for each soil (depending on the interactions of factors such as particle size, clay mineralogy, hydration, pH, ESP, EC and organic matter content (Qadir and Schubert, 2002)), clay swelling and dispersion will occur (Oster and Shainberg, 2001). Slaking may also result, with soil macro-aggregates breaking down to micro-aggregates upon wetting (Levy et al., 1998). These processes can lead to reduced porosity, as clay swelling and slaking causes a reduction in the number and size of pores at the soil surface and thereby limits the infiltration of rain or irrigation waters (Qadir and Schubert, 2002). The dispersion of clay particles can further reduce hydraulic conductivity and soil permeability as clay platelets are released from soil aggregates, blocking pores and plugging soil-water flow paths (Endo et al., 2002). Dispersion can result in a very unstable soil structure, and hence conditions that may be highly undesirable for agricultural land uses. Other soil physical changes and effects that are known to occur in association with sodic soils include increased susceptibility to surface crusting and sealing, hardsetting, runoff, erosion, tunnelling and poor aeration (Rengasamy and Olsson, 1991; Ford et al., 1993; Shaw et al., 1994; Qadir and Schubert, 2002).

7.2.3 Plant growth in saline and sodic soils

Plants differ substantially in their tolerance of salt stress (e.g. Figure 7.6), ranging from sensitive species such as *Oryza sativa* (rice) to highly tolerant (halophytic) species such as *Atriplex amnicola* (saltbush) which can even maintain good growth at salinities greater than sea water (Munns and Tester, 2008).

Halophytes can be defined as plant species which can reproduce (i.e. complete the lifecycle) under natural conditions of salt concentration of at least 200 mM NaCl and as such are distinguished from those species which are relatively salt tolerant but do not usually live

![Figure 7.6](image-url) **Figure 7.6** Diversity in the salt tolerance of different plant species, shown as increases in shoot dry matter after growth in solution or sand culture containing NaCl for at least 3 weeks, relative to plant growth in the absence of NaCl. For a colour version of this figure, please see Plate 7.2.
under saline conditions (Flowers and Colmer, 2008). More information on salinity tolerance in natural halophytes, constituting about 1% of the world’s flora, can be found in Flowers and Colmer, (2008), whilst Maas and Grattan (1999), Shannon and Grieve (1999) and Steppuhn et al. (2005) present data showing the salinity tolerance of a very extensive range of agricultural crop species. With some exceptions (e.g. sugar beet), most crop plants are relatively salt tolerant at germination even though the germination process may be delayed (e.g. Maas and Poss, 1989). Sensitivity of plants typically increases during emergence and vegetative development, although research investigating this to date has frequently been difficult to interpret due to differences between controlled experimental conditions and realistic field conditions (Läuchli and Grattan, 2007; Tavakkoli et al., 2010). As soil salinity and sodicity are commonly accompanied by additional abiotic and biotic stresses, numerous interactive effects may be relevant, yet many of these (e.g. salinity and boron toxicity interactions) are still only poorly understood (Läuchli and Grattan, 2007). Figure 7.7 summarises key properties of the major types of salt-affected soils and some possible mechanisms of impact on plant growth occurring in these different types of soil.

Typically, saline and sodic soils impact on plant growth via one (or more) of three effects, resulting in reduced yields:

1. physiological drought due to a reduction in water availability caused by lowering of the osmotic potential (see Chapter 9 for details);
2. ionic imbalances in the plant cells; and
3. toxicity due to specific ions, most usually Na⁺, Cl⁻ or B (as a variety of oxyanions, predominantly B(OH)₄⁻ and at pH > 7 H₄BO₃⁻ and B₄O₇²⁻).

In many cases, high pH may also be a significant cause of plant stress in these environments, impacting on nutrient availability and plant metabolism. However, this is independent of the direct salt effects on plant growth, and as mentioned previously, not all salt-affected soils are alkaline.

These major effects on plant growth in salt-affected soils are summarised below, but the interested reader may also refer to several comprehensive reviews published in the literature, including those by Hasegawa et al. (2000), Zhu (2001), Sairam and Tyagi (2004), Bartels and Sunkar, (2005), Chinnusamy et al. (2005) and Munns and Tester (2008). Plant growth response to salt stress has been described as a two-phase process, with the initial response shortly after exposure attributed predominantly to osmotic stress and the second, much slower effect over the following days, weeks or months, attributed to salt accumulation in the plant (Munns, 2002, 2005). This salt-specific effect is particularly noticeable in the older leaves of the plant and may result in leaf death or premature leaf senescence with corresponding reductions in total photosynthetic leaf area and overall plant growth (Munns, 2002, 2005). It should be noted that this two-phase model is heavily based on results from glasshouse trials and hydroponic systems and therefore may be less representative for plants growing under field conditions.

**Physiological drought**

The initial phase of growth reduction following sudden exposure of a growing plant to salt stress (e.g. under experimental conditions) is primarily due to osmotic stress as the increased ionic strength (and hence osmotic potential) of the solution outside the root causes changes in cell–water relations. The initial response to this osmotic effect is cell dehydration and
Figure 7.7 Key properties of the major types of salt-affected soils and possible mechanisms of impact on plant growth. Soils are categorised based on sodium absorption ratio (SARe) and electrical conductivity (ECe) measured in soil saturation extracts and pH1.5 measured in soil:water suspension. The value of SARe (SARe ≈ ESP) of 6 and above to classify a soil as sodic is based on the Australian soil classification criteria (Isbell, 1996); however, the general categories shown are of relevance worldwide. Toxicity, deficiency or ionimbalance due to other elements (e.g. B, K, N, P) will depend on the ionic composition of the soil solution and the particular species of plant.
shrinkage as the ability of the plant to absorb water is compromised. This occurs within moments of salinisation, is very similar to the effects of water stress and differs both within and between genotypes (Munns, 2002, 2005; James et al., 2008; Rajendran et al., 2009; Tavakkoli et al., 2010). The effect is transient and recovery is rapid. By using a pressurisation technique to maintain experimental plants at maximum water status whilst the soil was progressively salinised, Passioura and Munns (2000) were able to show that the transient growth reduction following initial application of salt to growing plants was due solely to changes in cell–water relations. Under conditions of increasing salinity, plants must expend more and more energy to maintain turgor. For example, Figure 7.8 illustrates the energy input required for plants to access soil water in a sandy loam soil under different conditions of soil moisture and salinity. In a non-saline soil, plants were able to access sufficient moisture to maintain turgor as long as the soil contained as little as 5% soil-water content, whereas with increasing soil salinity the energy required by the plant to extract water from the soil increases. At an EC$_{1:5}$ of 1 dS m$^{-1}$ in the same sandy loam soil, plants required at least 18% soil-water content to take up sufficient moisture to prevent wilting.

![Figure 7.8: Energy (equivalent to soil matric plus osmotic potential) required by plants to take up water from a loamy soil as influenced by EC$_{1:5}$ and % soil-water content.](image)

**Impacts of high pH on nutrient availability and plant metabolism**

As discussed in Chapter 6, nutrient availability is linked to soil pH. The high pH of most sodic soils (>8.5) and many saline soils means that plant availability of essential cationic nutrients are often limited, thus resulting in nutrient deficiencies in plants growing in these soils.

**Ion toxicity and imbalance**

Given that the dominant salt in salt-affected soils is NaCl, high concentrations of both Na$^+$ and Cl$^-$ ions are likely to be present. For most plants to tolerate salinity, Na$^+$ and Cl$^-$ uptake must be restricted and the uptake of macronutrients such as K$^+$, NO$_3^-$ and Ca$^{2+}$ must be maintained. Chloride toxicity occurs in sensitive plants (many citrus and woody plant
species) when solution concentrations exceed 20 mM, whereas tolerant species (halophytes) are unaffected (e.g. *Atriplex hastata* or even stimulated (e.g. *Sueda maritime*) by concentrations in excess of 200 mM (Greenway and Munns, 1980; Marschner, 1995). Detailed information about Cl\(^{-}\) uptake and movement within plants and the mechanisms of Cl\(^{-}\) transport contributing to salt tolerance in plants can be found in White and Broadley (2001) and Teakle and Tyerman (2010), respectively, whilst Na\(^{+}\) tolerance and transport in higher plants has been reviewed by Tester and Davenport (2003). Considerable variation in plant tolerance is also exhibited for Na\(^{+}\), with plants ranging from Na\(^{+}\) tolerant or natrophilic to Na\(^{+}\) intolerant or natrophobic (Marschner, 1995). It is difficult to provide a value at which Na soil concentrations become toxic. This is because high Na\(^{+}\) concentrations in salt-affected soils typically coincide with high Cl\(^{-}\) and possibly also high HCO\(_3\)\(^{-}\) concentrations (i.e. high pH), and the toxic effect of Na is also strongly linked to the ratio of Na to Ca and its impact on Ca uptake (Marschner, 1995). Chloride and Na\(^{+}\) toxicity can be expressed as a decrease in yield, chlorosis and decreased photosynthesis (Sykes, 1992; Maas, 1993; Marschner, 1995; Xu et al., 2000). High NaCl concentrations can also inhibit Ca, Mn and NO\(_3\)\(^{-}\) uptake and therefore cause deficiency/imbalance in these nutrients. It can also lead to enhanced uptake of P and subsequent P toxicity (Marschner, 1995).

Boron (B) is also often found in high concentrations in saline/sodic soils, particularly in the sub-surface layers of these soils, where saline-sodic conditions prevail. Toxicity has been reported in sensitive species at soil solution concentrations of 1 mg L\(^{-1}\), whilst concentrations of 10–15 mg L\(^{-1}\) are toxic even to tolerant plants (Ryan et al., 1977; Chaudan and Powar, 1978; Bonilla et al., 1980; Johnson and Fixen, 1990). Visual symptoms of boron toxicity are initially dark green wilted leaves (Kabata-Pendias, 2000). However, as concentrations of B increase, chlorosis of leaf tips and/or margins and necrosis occurs. Toxicity appears to be due to metabolic impairment due to formation of complexes with NAD\(^{+}\) or the ribose of RNA as well as more specific metabolic pathway inhibitions (Loomis and Durst, 1992; Lukaszewski et al., 1992). Tolerance appears to depend partly on the ability to maintain low B shoot accumulation, and this has also been linked to mechanisms of active B efflux from roots (Sutton et al., 2007).

### 7.2.4 Amelioration and management of saline and sodic soils

Two major strategies are available for the amelioration of saline and sodic soils. The first strategy is based on the cultivation of salt-resistant crops. As already mentioned, the salinity tolerance of crop species varies significantly (Maas and Grattan, 1999, Shannon and Grieve, 1999; Steppuhn et al., 2005) and genetic manipulation may also lead to the development of tolerant cultivars (e.g. Zhang and Blumwald, 2001; Munns, 2009). The use of these salt tolerant species can significantly boost the productivity of agriculture in salt-affected soils. However, in areas where salt loads are continuing to rise (e.g. due to ongoing use of low quality irrigation waters) this must really be viewed as a short-term management option. Unless it is allied to other management actions, the soil will continue to become progressively more saline or sodic and will eventually be rendered infertile. A historic example of this can be seen in the Sumarian civilisation of Mesopotamia (Jacobsen, 1982; Wild, 2003). As their soils became increasingly saline/sodic, there was a change in agriculture from wheat to more tolerant barley but with no associated change in irrigation methods. As a result, the soil continued to degrade and ultimately became infertile to barley as well. Starved of cereal crops the civilisation collapsed.
The more successful long-term strategy for ameliorating salt-affected soils requires the salt concentration to be lowered by flushing out the problem salts. For saline soils, this is theoretically quite straightforward, simply requiring a sufficient volume of good quality irrigation water to wash the salts out of the soil. In practice, however, it is rarely so simple, as good quality irrigation water in salt-affected areas is often in very short supply. Moreover, for sodic soils, matters are frequently complicated by the degradation of soil structure (Chapter 9) and a source of Ca\(^{2+}\) ions to replace the dominant Na\(^{+}\) ions on the soil exchange sites is also generally required. The costs of sodic soil remediation may therefore be prohibitively high. Further information about salt-affected soil amelioration can be found in Qadir et al. (2000).

Salt tolerant crops

Crops are able to tolerate saline/sodic conditions through a variety of mechanisms (Flowers and Colmer, 2008; Munns and Tester, 2008; Türkan and Demiral, 2009; Munns, 2009), including:

1. selective exclusion of ions;
2. selective inclusion of ions:
   (a) to maintain cell turgor
   (b) to maintain K\(^{+}\) uptake
   (c) to partition Na\(^{+}\) and Cl\(^{-}\) into specific tissues or organelles (e.g. vacuoles);
3. induction and up-regulating of antioxidative enzymes;
4. accumulation of ‘compatible organic solutes’ (e.g. betaines, polyols, sugars and amino acids) to increase hyperosmotic tolerance; and
5. rapid and efficient salt excretion via salt glands or tissue shedding.

Generally, salt tolerant plants are known to have lower rates of Na\(^{+}\) and Cl\(^{-}\) transport to leaves compared with salt-sensitive plants, as well as greater ability to tolerate these ions and compartmentalise them into vacuoles, thereby preventing their build-up in the cytoplasm or cell walls (Munns, 2002; Munns and Tester, 2008).

Traditional methods of improving the salt tolerance of plants have focused on screening cultivars and using cross-breeding to produce cultivars with improved levels of both tolerance and yield. More recently, however, the genomics revolution has made genetic engineering of salt tolerant plants possible (e.g. Zhang and Blumwald, 2001) and has also facilitated more efficient molecular screening of cultivars for use in cross-breeding programmes (e.g. Munns and James, 2003; Lindsay et al., 2004; James et al., 2006). Nevertheless, genetic engineering for increased salt tolerance is not a simple procedure and agriculturally relevant progress in this direction has so far been quite limited. For instance, single gene modifications can be used to confer salt tolerance in transgenic plants, but they are also likely to incur significant fitness costs in terms of plant development and physiology (Denby and Gehring, 2005). Salt tolerance is a genetically complex trait, governed by multiple genes with additive, dominance and reciprocal effects (Genc et al., 2007), and, as noted earlier, numerous different mechanisms may potentially contribute to salt tolerance. It is also a developmentally regulated process, with tolerance of plants at one stage of development not always correlated with tolerance at other stages (Yamaguchi and Blumwald, 2005). Despite the challenges, efforts to confer favourable traits to improve crop salt tolerance are continuing (e.g. Choi et al., 2011; Gao et al., 2011) and further success in this direction is to
be anticipated (Mittler and Blumwald, 2010; Roy et al., 2011). By studying the naturally occurring variation of salt stress tolerance in varieties, landraces and wild relatives of crops, and studying the traits that contribute to tolerance, the genetic loci determining these traits can be determined. Once the molecular basis of traits contributing to tolerance have been identified, marker-assisted breeding and genetic modification technologies can be used to introduce these genes into current, high yielding elite cultivars.

**Reducing saline/sodic conditions**

For saline soils, good quality irrigation water (see ‘Ion toxicity and imbalance’) can be used to flush out salts from the soil profile. To achieve this, the volume of water applied needs to be sufficient to meet the water needs of the crops and to leach out the salts. The fraction of water required to pass through the root zone to leach salts to an acceptable level has traditionally been known as the *leaching requirement*, whilst the fraction of infiltrated water that actually percolates below the root zone is called the *leaching fraction*. This can be calculated from the depths of the irrigation and drainage waters and is also (under certain steady-state conditions) approximately equal to the ratio of the electrical conductivity of the irrigation water to the drainage water (Qadir et al., 2000). Details of the necessary calculations to determine the amount of irrigation water to apply are given by Ayers and Westcott (1985) and Rowell (1994). It is important to ensure that the flushed drainage water does not cause deterioration of the quality of drainage water downstream. To alleviate sodic and saline-sodic conditions, the key objective is to reduce the concentration of \( Na^+ \) in the soil; this is most usually done by additions of \( Ca^{2+} \), usually in the form of gypsum followed by flushing of the soil (Richards, 1954). Good drainage has to be maintained for this strategy to work and sometimes this requires addition of non-sodic soils or manures together with specialised ploughing regimes (e.g. deep ripping) to create a permeable upper soil layer.

**Preventing salinity/sodicity problems**

In order to avoid or ameliorate soil salinity/sodicity problems in the first place, good quality irrigation water should be used and the water table kept low. Irrigation water quality guidelines (e.g. Table 7.6) can be used to indicate water composition ranges suitable for restricted and/or unrestricted use in agriculture.

Maintenance of a low water table requires good drainage. For example, open or till drains or drainage wells may be used to maintain the water table at a depth of >2 m in arid and semi-arid regions to prevent groundwater rising by capillarity and evaporating, thereby increasing ion concentrations in the rooting zone of plants. As with flushing of soils, the drainage water must be disposed of with care in order not to create salinity/sodicity problems elsewhere (Ayers and Westcot, 1985).

Practices such as ‘partial root zone drying’ may also be practised in order to minimise the amount of salt added to the soil via irrigation (Kaman et al., 2006). Other relevant irrigation-related management approaches that have the potential to aid the management of sodicity/salinity problems such as optimal timing and duration of irrigation are also addressed in Chapter 9. In rain-fed agriculture, suitable crop rotation practices (e.g. rotation of shallow-rooted annual crops with deep-rooted perennial species) may help restore the balance between rainfall and water use, thus preventing/slowing the rise of salts to the surface (Munns, 2002).
7.3 Acid and alkaline soils

7.3.1 Designation of acid and alkaline soils

Soils may be designated as either acid or alkaline on the basis of soil pH, a parameter well recognised as a master variable affecting both soil chemistry and plant nutrition. Although pH 7.0 can technically be used to differentiate an acid soil from an alkaline soil, from a practical perspective soils with pH close to 7.0 (i.e. pH 6.5–7.5) are often considered as neutral. By contrast, soils with pH 5.5–6.5 (measured in 0.01 M CaCl₂) can be classed as slightly acidic, those with pH 4.5–5.5 as moderately acidic and those with pH < 4.5 as strongly acidic. Similarly, pH 7.5–8.5 can be considered slightly alkaline, pH 8.5–9.5 moderately alkaline and pH 9.5 and above strongly alkaline (Charman and Murphy, 2000). Delineating between these different degrees of acidity and alkalinity is of extensive practical significance in terms of land management, crop growth and related issues. For instance, key aspects of soil chemistry regulated by pH and relevant to plant growth include element availability and toxicity (Barber, 1995; Impellitteri et al., 2001) microbial activity (Entry et al., 2002) and root growth (Pavlovkin et al., 2009).

Soils tend to vary in terms of acidity/alkalinity both spatially and temporally due to differences in local soil forming processes (e.g. precipitation, temperature), parent materials (i.e. basic vs. siliceous rock), landscape characteristics (e.g. slope) and land management.
regimes (e.g. use of soil amendments, tillage practices, etc.), but most naturally occurring soils have a pH somewhere between pH 4 and pH 10. The actual soil pH in a given area is essentially a function of the soil composition (i.e. the nature and relative proportions of mineral and organic constituents) and related ion exchange and hydrolysis reactions (Ulrich and Sumner, 1991). Soil pH values <4 are relatively uncommon because the pH is buffered in this range by the dissolution of aluminosilicate and oxide minerals (Ulrich, 1991). Nevertheless, in special cases such as acid sulphate soils, soil pH values well below this level have been recorded (Melville and White, 2000; Fitzpatrick et al., 2009).

The major source of alkalinity in soils is the dissolution of carbonates and subsequent production of OH\(^{-}\) ions. Alkaline soils are 100% saturated with exchangeable base cations (e.g. Ca\(^{2+}\), Mg\(^{2+}\) and Na\(^{+}\)) and may have a very large acid-neutralising capacity depending on their carbonate content.

Soils containing free calcium carbonate (CaCO\(_3\)) are referred to as calcareous soils and are typically slightly to moderately alkaline. Under the FAO World Reference Base soil classification system, most calcareous soils fall within the group of Calcosols, comprising soils containing a calcic horizon with substantial secondary accumulation of carbonates (FAO, 1998). Calcareous soils may also contain free magnesium carbonate (MgCO\(_3\)). They frequently occur in arid/semi-arid regions (FAO 1998; Candy and Black, 2009) where low precipitation and biological activity equate to relatively low acid input and leaching, hence resulting in the net accumulation of alkalinity. However, alkaline soils may also occur in higher rainfall areas, where young soils are forming on alkaline parent materials (e.g. limestone). Calccic horizons may form via the leaching and precipitation of carbonates under natural precipitation regimes (e.g. seasonal leaching and precipitation cycles) and/or under irrigation scenarios whereby salts move upwards from the water table and precipitate near the top of the capillary fringe. They exhibit a wide range of morphologies and pedogenic origins (Candy and Black, 2009) and are sometimes densely cemented with poor drainage characteristics. These hardened calcic horizons are often referred to as calcretes and may present problems for plant root development and irrigation management.

Highly alkaline soils (pH = 8.5–10) are generally associated with the dissolution of sodium carbonate (NaCO\(_3\)) which is more soluble than CaCO\(_3\) and thus typically results in greater production of OH\(^{-}\) ions. These soils are also often sodic (see Section 7.2).

### 7.3.2 Distribution of acid soils

Acid soils occur naturally in many areas of the world (Figure 7.9) but acidification may also be anthropogenically induced.

Naturally occurring acid soils tend to be found in relatively humid areas where high rainfall leads to extensive leaching and mineral weathering such that base cations are depleted and the soil neutralising capacity exceeded. Acid soils with pH <5.5 in their upper horizons reportedly occupy around 30% (~3950 million ha) of the world’s ice-free land area (see Table 7.7) and are located predominantly in two main geographical belts, one in the humid northern temperate zone and the other in the humid tropics (von Uexküll and Mutert, 1995). Approximately 75% of these areas are also affected by subsoil acidity. The predominant natural vegetation communities occurring on these soils are coniferous forest in the northern belt and savannah and tropical rainforest in the southern belt (von Uexküll and Mutert, 1995).

Acid soils occur in all major regions of the world (see Table 7.8; FAO, 1991); however, the majority (~60%) are located in the humid tropics. From a global perspective, the largest
Figure 7.9  Global distribution of topsoil pH. Used with permission. FAO-UNESCO Soil Map of the World 2007. (http://www.fao.org/fileadmin/templates/nr/images/resources/images/Maps/geonetwork/ph_t.png). For a colour version of this figure, please see Plate 7.3.
areas of acid soils are located in North and South America (40.9% of the world’s acid soils by area) and Asia (26.4% of the world’s acid soils by area). From a regional perspective, the highest proportion of acid soils is found in south-east Asia where 63% of the region’s soils are classed as acid soils. By contrast, 35% of North American soils and 57% of South American soils are acidic (von Uexküll and Mutert, 1995).

Comparison of global land use statistics and acid soil distribution data indicate that a major portion of the world’s acid soils still remain under natural vegetation (von Uexküll and Mutert, 1995; Sumner and Noble, 2003). Nevertheless, with modern agricultural techniques for the amelioration and management of soil acidity (e.g. lime amendment and phosphate fertilisation), large areas of uncultivated acid soils are now considered potentially suitable for arable production. Indeed, these areas of undeveloped acid soils have even been referred to as agriculture’s last frontier (Borlaug and Dowswell, 1993), with recent land clearing and agricultural development of areas such as the Cerrado (savanna) region of South America used as an example of the latent potential for highly productive agricultural production on acid soils (e.g. Lopes, 1996; Yamada, 2005). Yet, although it is true that with the appropriate inputs of lime and fertiliser high agricultural yields may potentially be obtained on such soils, it should also be noted that these remaining areas of uncultivated

### Table 7.7 Global extent of acid soils classified according to major soil groups.

<table>
<thead>
<tr>
<th>FAO soil group</th>
<th>Soil taxonomy (approximate match)</th>
<th>Area (million ha)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fluvisols</td>
<td>Fluvents</td>
<td>50</td>
</tr>
<tr>
<td>Gleysols</td>
<td>Aquents, Aquepts</td>
<td>402</td>
</tr>
<tr>
<td>Regosols</td>
<td>Psammments</td>
<td>293</td>
</tr>
<tr>
<td>Arenosols</td>
<td>Psammments</td>
<td>280</td>
</tr>
<tr>
<td>Rankers</td>
<td>Haplumbrepts</td>
<td>61</td>
</tr>
<tr>
<td>Andosols</td>
<td>Andisols</td>
<td>34</td>
</tr>
<tr>
<td>Cambisols</td>
<td>Ochrepts, Tropepts</td>
<td>299</td>
</tr>
<tr>
<td>Podzoluvisols</td>
<td>Boralfs, Aqualfs</td>
<td>255</td>
</tr>
<tr>
<td>Podzols</td>
<td>Spodosols</td>
<td>415</td>
</tr>
<tr>
<td>Planosols</td>
<td>Aaquults</td>
<td>15</td>
</tr>
<tr>
<td>Acrisols</td>
<td>Ultisols</td>
<td>731</td>
</tr>
<tr>
<td>Nitosols</td>
<td>Ultisols</td>
<td>118</td>
</tr>
<tr>
<td>Ferralsols</td>
<td>Oxisols</td>
<td>727</td>
</tr>
<tr>
<td>Histosols</td>
<td>Histosols</td>
<td>270</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td><strong>3950</strong> (≈30% of global land area)</td>
</tr>
</tbody>
</table>
acid soils support important habitat for numerous flora and fauna species of notable conservation significance, whilst also delivering a variety of other ecosystem services (Klink and Machado, 2005). Moreover, large areas of acid soils currently under cultivation are still delivering suboptimal returns due to insufficient liming and soil management regimes, whilst potentially slipping further towards the development of subsoil acidity (AACM, 1995). Given these conditions, it is clear that the agricultural sector can still benefit from further efforts to improve the nutrient supply capacity of acid soils in existing agricultural areas in order to sustain and improve agricultural productivity (Fageria and Baligar, 2003). Further clearing of marginal lands requiring extensive inputs to facilitate efficient agriculture should thus be very carefully considered, as potential side effects include pollution of water resources and increased greenhouse gas emissions.

7.3.3 Anthropogenic drivers of soil acidification

Precipitation is naturally acidic when in equilibrium with atmospheric CO$_2$ and is thus a continuous source of acidity to soils. In addition, areas affected by industrial pollution (e.g. heightened atmospheric levels of oxides of sulphur and nitrogen) may also be subject to the precipitation of ‘acid rain’ which can contribute substantially to progressive soil acidification (Bini and Bresolin, 1998). Models and analyses of the acid-base status of global precipitation indicate that acid precipitation is of particular concern in North America, Europe and China (Greenfelt and Galloway, 1996; Galloway, 2001; Rodhe et al., 2002), where detrimental effects on soil quality have been observed (Tomlinson, 2003; Hicks et al., 2008; Lehmann et al., 2008). Much less is known about the extent of acid deposition in other regions of the world (WMO, 1996), although the results of Rodhe et al. (2002) indicate that potential problem areas include several regions with sensitive soils in southern, south-eastern and eastern Asia, and in central South America. These sensitive soils include those which have low base saturation and cation exchange capacity and hence limited buffering capacity to resist the effects of acid deposition (Kuylenstierna et al., 2001).

In practice, the study of acid rain induced soil acidification is complicated by the ongoing parallel process of natural acidification and the relatively long time scales involved in the acidification process (Stuanes et al., 1995; Matzner and Davis, 1996). Yet, although it is difficult to clearly discriminate between pollution-induced acidification and natural acidification on the basis of fundamental parameters such as low base saturation, low pH and low alkalinity, there are certain characteristics such as high soil solution concentrations of SO$_4^{2-}$ and NO$_3^-$, which specifically indicate the contribution of acid rain. Whereas these are amongst the key acidifying ions derived from acid rain, organic anions and HCO$_3^-$ are the primary drivers of natural acidification (Alewell, 2003). Agricultural practices can also accelerate the process of soil acidification. For example, the application of nitrogen-containing inorganic fertilisers is well recognised as a contributing factor to the development of acid soils (Stumpe and Vlek, 1991; Jensen and Hauggaard-Nielsen, 2003). This is because the nitrification of fertiliser ammonia effectively releases H$^+$ to the soil system as indicated in Equation 7.4:

$$2\text{NH}_4^+ + 4\text{O}_2 = 2\text{NO}_3^- + 4\text{H}^+ + 2\text{H}_2\text{O} \quad (7.4)$$

The continual loss of alkaline grain, pasture and animal products due to the removal of material during harvest is a further driver in the acidification of agricultural soils (von Uexkiill and Mutert, 1995) as is the increasing prevalence of legume-based pasture systems.
Managing adverse soil chemical environments

(Tang et al., 1997). Together these causes of increasing acidity on extensively managed agricultural lands can lead to ongoing land degradation and significant reductions in farm productivity. For example, acidification due to the prevalence of legume-based pasture systems is a growing problem in large parts of Australia whilst acidification of intensively cropped land is a serious problem in large areas of the USA and former USSR (von Uexküll and Mutert, 1995; Garvin and Carver, 2003).

7.3.4 Acid sulphate soils

One key group of acid soils which are particularly challenging to manage are the acid sulphate soils (ASS). These are soils which either contain sulphide minerals or are affected by the transformation of sulphide minerals (Fanning, 2002). As such, ASS either contain sulphuric acid or have the potential to form sulphuric acid (most typically through the oxidation of iron pyrite) in quantities greatly exceeding the acid-neutralising capacity of the soil (Dent and Pons, 1995).

Sulphidic soil materials are naturally occurring in many areas of the world (see Ljung et al., 2009), especially in low lying coastal areas (e.g. mangrove swamps), and over 17 million ha of ASS have been identified worldwide (Andriesse and van Mensvoort, 2006; Ljung et al., 2009). The oxidation of sulphidic soil materials can occur naturally as a result of soil drying during times of drought (Dent and Pons, 1995) but is also frequently instigated by human activities, such as the draining of floodplains and backswamps for agriculture, aquaculture or housing developments (Dent and Pons, 1995; Österholm and Åström, 2004; Ljung et al., 2009).

The course of pyrite oxidation is complex, involving both chemical and microbiological reactions. Complete oxidation of pyrite to form sulphuric acid involves up to 22 electron transfer reactions and proceeds via a number of steps (Nordstrom, 1982). However, the overall oxidation process can be represented by the following equation:

\[
\text{FeS}_2 + 15/4\text{O}_2 + 7/2\text{H}_2\text{O} \rightarrow \text{Fe(OH)}_3 + 2\text{SO}_4^{2-} + 4\text{H}^+
\]  

The pH of oxidised ASS can be as low as pH 2 (Dent and Pons, 1995) and acid leaching from these soils can have very serious consequences for soil nutrient availability (Golez and Kyuma, 1997), surface water quality (Sammut et al., 1996; Corfield, 2000), toxic metal availability (Åström, 2001), human health (Ljung et al., 2009), vegetation and aquatic life (Sammut et al., 1996; Hyne and Wilson, 1997; Lin et al., 2001) and even civil infrastructure (e.g. the corrosion of bridges, roads and pipelines) (Fitzpatrick et al., 2001). Careful management of these materials based on appropriate risk assessment when considering and/or undertaking their development is paramount to the prevention of serious environmental degradation, health effects and financial implications associated with unmitigated acid sulphate soil drainage.

7.3.5 Managing plant growth in acid soils

Although many plant species are well suited to growth in slightly acid soils, acid soil infertility and growth inhibition is a very common management issue in soils with pH <5.6 and is particularly problematic when pH is <5. Plants that are highly sensitive to acidity (e.g. lucerne and barley) tend to be adversely affected below pH 5, whilst more tolerant plants can grow normally until around pH 4.5. Below this, only very acid tolerant plants such as oats
Soil Conditions and Plant Growth

will grow without a significant reduction in production. Problems typically associated with acid soils include Al and Mn toxicities (Scott et al. 2000; Barcelo and Poschenrieder, 2002), nutrient deficiencies (e.g. Mo, Ca, Mg), legume nodulation failures (Bordeleau and Prevost, 1994) and inhibition of root growth (Marschner, 1991).

A key aspect of managing acid soils for plant growth is the pH amendment of the soil (Edmeades and Ridley, 2003). This is commonly practised in both agricultural and garden settings, generally via the application of acid-neutralising materials such as lime and dolomite and can be a simple and successful management strategy (e.g. Table 7.9).

Whilst some acid soils, particularly those occurring in tropical regions, are acid throughout the whole profile, others which have been exposed to more moderate weathering and leaching comprise acid upper layers underlain by alkaline layers with substantial acid-neutralising capacity at depth (Farina et al., 2000a, b). As liming is most effective in situations where only the surface soil layer is acidic, it is now recognised that liming to prevent acidification (particularly the development of subsoil acidification) is as important as liming for the purpose of acid soil amelioration, because this ensures that the land remains profitable in order to offset the cost of liming. By contrast, once acidification develops, the initial capital input required to return the soil to a suitably productive state may be preventative (Edmeades and Ridley, 2003).

Individual soils respond differently to liming and there is no optimal soil pH that will suit all crops or plant species (Edmeades and Ridley, 2003). Acid amelioration practices should thus be planned in accordance with the nature of the soil and required land use. Research in this area has seen a move away from the concept of soil lime requirement (laboratory-based determinations of the lime required to increase soil pH by a specific amount) towards the concept of biological lime requirement, meaning the amount of lime required to eliminate restrictions to plant growth (e.g. Slattery and Coventry, 1993). However, there are many factors at play in determining this level and crop responses to liming have repeatedly proved difficult to predict. Progress in the development of predictive relationships between relevant soil parameters and crop yields has been made (Edmeades et al., 1984; Dierolf et al., 1999), but further research efforts are required. In some cases, the expense of lime amendment can be managed by determining which of the factors that potentially limit plant growth on acid soils (e.g. Al and Mn toxicity, Ca deficiency, P deficiency, etc.) are of greatest relevance in a given setting and specifically addressing those issues. For example, Edmeades and Ridley (2003) cite cases of legume-based pastures on acid soils responding equally well to relatively small amounts of Mo fertilisation as to large dressings of lime. In such cases, an initial campaign of Mo fertilisation to improve productivity could help support the cost of ongoing pH management. Fageria and Baligar (2003) show increases in rice and bean yield in Cerrado and Varzea acid soils from Brazil in response to N fertiliser applications, N levels having previously been reduced in the soils due to net N removal due to harvesting and soil organic matter reductions (Table 7.10).

In situations where Al toxicity is identified as the primary concern, relatively small applications of lime to bring the pH to around 5.5 may effectively return plant growth to near-maximum productivity (Kamprath, 1970). It is clear that site-specific diagnosis and management of acid soil infertility can minimise the expense of amelioration.

Efforts should also be made to slow the rate of acidification through more efficient fertiliser use (to minimise nitrate leaching) and to deliberately choose less acidifying fertilisers (i.e. replacing fertilisers such as sulphate of ammonium and mono-ammonium phosphate with urea or legume fixed nitrogen). It may also be necessary to choose acid tolerant species and cultivars to grow, particularly in agricultural situations where cash
Table 7.9  Range and typical responses of crops and pastures to surface application of topsoil incorporated lime.

<table>
<thead>
<tr>
<th>Country</th>
<th>Crop</th>
<th>Soil acidity</th>
<th>Absolute range(^{a}) of response to lime (% relative to control)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>New Zealand</td>
<td>White clover-based pasture</td>
<td>pHW 5.0–6.0</td>
<td>0–10 (about 5)</td>
<td>Edmeade et al. (1984)</td>
</tr>
<tr>
<td></td>
<td>Clover-based pasture organic soils</td>
<td>pHW 4.0–5.0</td>
<td>10–140 (50–100)</td>
<td></td>
</tr>
<tr>
<td>Australia, NSW, Vic</td>
<td>Subterranean clover-based pasture</td>
<td>pHCa 4.2–4.7</td>
<td>−15 to 170 (20–60)</td>
<td>Scott et al. (2000)</td>
</tr>
<tr>
<td></td>
<td>Phalaris pastures</td>
<td>pHCa 4.1–5.0</td>
<td>0–900 (30–100)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cocksfoot pastures</td>
<td>pHCa 4.1–5.0</td>
<td>0–70 (about 20)</td>
<td></td>
</tr>
<tr>
<td>Australia, NSW, Vic</td>
<td>Wheat</td>
<td>pHW 5.0–6.0</td>
<td>0–38</td>
<td>Coventry et al. (1989)</td>
</tr>
<tr>
<td>USA, southern</td>
<td>Cotton</td>
<td>pHW 5.0–5.4</td>
<td>0–400 (10–30)</td>
<td>Adams (1984), Adams and Pearson (1967)</td>
</tr>
<tr>
<td></td>
<td>Corn</td>
<td>pHW 4.7–5.5</td>
<td>11–56 (10–30)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Soybean</td>
<td>pHW 4.7–5.8</td>
<td>0–56 (30–50)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Alfalfa</td>
<td>pHW 4.7–5.8</td>
<td>150 to &gt;1000 (100–200)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Other forage crops(^{b})</td>
<td>pHW 4.3–5.5</td>
<td>0–250 (10–30)</td>
<td></td>
</tr>
<tr>
<td>USA, Midwest</td>
<td>Corn</td>
<td>pHW 5.0</td>
<td>75–96(^{c})</td>
<td>McLean and Brown (1984)</td>
</tr>
<tr>
<td></td>
<td>Soybean</td>
<td>pHW 5.0</td>
<td>70–91(^{c})</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Alfalfa</td>
<td>pHW 5.0</td>
<td>20–84(^{c})</td>
<td></td>
</tr>
<tr>
<td>Various(^{d})</td>
<td>Soybean</td>
<td>A1 base saturation 75(^{\circ})</td>
<td>25–95(^{f})</td>
<td>Dierolf et al. (1999)</td>
</tr>
<tr>
<td></td>
<td>Corn</td>
<td>A1 base saturation 75(^{\circ})</td>
<td>66–95(^{f})</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Various(^{g})</td>
<td>A1 base saturation 75(^{\circ})</td>
<td>80–95(^{f})</td>
<td></td>
</tr>
</tbody>
</table>

\(^{a}\) Typical value by visual inspection of the data.

\(^{b}\) White clover, crimson clover, Bermuda grass, millet, sorghum, ryegrass.

\(^{c}\) Response at pH 5.0 relative to maximum yield.

\(^{d}\) All tropical soils (Oxisols and Ultisols) from ten countries.

\(^{e}\) Al base saturation % of ECEC.

\(^{f}\) Response at Al saturation of 75%, relative to maximum yield.

\(^{g}\) Groundnut, rice, cowpea, mungbean, pigeon pea, cassava, tanier, yams, sugarcane.

Source: Modified and used with permission from Edmeades and Ridley (2003).
flow is needed to facilitate further soil liming. In so doing, it should not be forgotten that specific toxicities and deficiencies are likely to be involved and the selected species/cultivars will need to be suitable from this perspective also. It will not do to select a cultivar that is relatively acid tolerant on some soils but which is highly sensitive to Mo deficiency if that is a key cause of the acid soil infertility for the field in question. In any case, given the increasing use of gene technology it is likely that additional advances in the production of acid tolerant crop species will continue to occur. Furthermore, isolation of key genes such as those underlying plant root Al tolerance are not only useful for genetic engineering applications but can also be used as markers to help select for tolerance traits in standard (i.e. traditional) plant breeding trials. Research has shown there to be a wide range of genetic diversity for acid soil and Al toxicity tolerance, which is not only conducive to crop improvement via hybridization and phenotypic selection but also provides a rich basis for the molecular manipulation of plant genotypes (Garvin and Carver, 2003).

In the special case of highly acidic ASS, management for plant and crop growth must be carefully balanced around the hydrological regime (Wilson et al., 1999). For example, shallow drains for agricultural land management are considered preferable to deep drains as they impart reduced risk of additional subsoil pyrite oxidation and acid production. In all cases, the further disturbance and exposure of unoxidised materials is to be avoided. Acid tolerant crops (e.g. sugar cane) can potentially be successfully farmed on ASS, with regular liming and good drainage management key to the success of such ventures. Where major acidification has resulted in scalded areas with visually distinctive sparse patches of ‘burnt’ or ‘scalded’ vegetation, rehabilitation is particularly difficult (Lin et al., 2001). Flushing with freshwater may be of some benefit although research in this area is still ongoing as there is considerable potential for the standing water after flooding to become highly acidic prior to discharge into local receiving waters (Melville and White, 2000).

### Table 7.10 Responses to fertilisation of rice and common bean grown in rotation in Cerrado and Varzea acid soils.

<table>
<thead>
<tr>
<th>Fertility level</th>
<th>Rice grain yield/t ha⁻¹</th>
<th>Common bean grain yield/t ha⁻¹</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cerrado soil</td>
<td>Varzea soil</td>
</tr>
<tr>
<td>Low</td>
<td>1.7ᵇ</td>
<td>4.3ᵇ</td>
</tr>
<tr>
<td>Medium</td>
<td>2.1ᵃ</td>
<td>5.5ᵃ</td>
</tr>
<tr>
<td>High</td>
<td>2.1ᵃ</td>
<td>5.5ᵃ</td>
</tr>
<tr>
<td>Medium + green manure</td>
<td>2.4ᵃ</td>
<td>6.3ᵃ</td>
</tr>
<tr>
<td>F test</td>
<td>*</td>
<td>**</td>
</tr>
</tbody>
</table>

*and ** indicate significance at the 0.05 and 0.01 probability levels. Within the same column, means followed by the same letter do not differ significantly at the 0.05 probability level by Tukey’s test.

Source: Modified and used with permission from Fageria and Baligar (2003).
7.3.6 Managing plant growth in alkaline soils

Moderately alkaline soils frequently provide the basis for highly productive agricultural land. However, at higher levels of alkalinity problems are increasingly encountered, and soils with pH 9.5 and above are generally considered unsuitable for agriculture (Charman and Murphy, 2000). Typical problems encountered in alkaline soils include plant nutrient deficiencies (e.g. Fe deficiency chlorosis), root growth inhibition and stunting (Gruber and Kosegarten, 2002). In addition, as highly alkaline soils are also often sodic, some of the problems associated with alkaline soils are in fact derived from their salt-affected nature and associated management issues (e.g. poor soil structure, poor drainage) as covered in Section 7.2, rather than from the alkalinity per se.

Respiring roots and microorganisms release CO$_2$ which hydrolyses in soil water to produce carbonic acid, thereby forming localised areas of acidity in alkaline soils which can aid in the plant uptake of key nutrients. Nevertheless, these and other biologically driven processes to solubilise key nutrients (such as the excretion of siderophores, protons and enzymes from roots) (Hedley et al., 1982; Grinsted et al., 1982; Dakora and Phillips, 2002; Hinsinger et al., 2003) are not always sufficient to overcome alkaline soil infertility, and active management is frequently required to mitigate the associated retardation of plant growth. The most common nutrient deficiency in plants growing on alkaline soils is iron (Fe) deficiency; however, numerous other plant nutrients such as P, Mn, Zn and Cu also have limited bioavailability in the alkaline pH range (Chapter 6).

Some arid and semi-arid zone plants are relatively well adapted to growing in alkaline soils (e.g. Abd El-Ghani, 1998; Buxbaum and Vanderbilt, 2007) because, as indicated earlier, the low precipitation and restricted leaching conditions common to arid areas are often a contributing factor to the formation of alkaline soils. By contrast, many important agricultural plant species have a low tolerance to this type of soil environment (e.g. Kerley et al., 2002). Despite this, it is far less common to practise wide-scale pH amendment of alkaline soils than it is to undertake liming and pH amendment of acid soils. Acidifying reactions can, however, be brought about through the application of elemental sulphur (S), which is subsequently oxidised to sulphuric acid by both chemical and microbially mediated reactions (Lindemann et al., 1991; Cifuentes and Lindemann, 1993; Neilsen et al., 1993). For the purposes of gardening, aluminium sulphate ($\text{Al}_2(\text{SO}_4)_3$) and iron sulphate ($\text{FeSO}_4$) are commonly used for alkaline soil amendment.

Given the limited broad-scale pH amendment of alkaline soils, it is often important for crops to be specially selected as suitable for alkaline soil cropping. Indeed, even within species some varieties are likely to be better adapted to alkaline soils than others (e.g. Brand et al., 1999). In addition to the careful selection of existing species and cultivars with useful adaptive characteristics, genetic engineering technology can also facilitate the identification, isolation and transfer of key genes from native species that are particularly well adapted to alkaline soil conditions. Micronutrient fertilisers may also be applied as a strategy for improving crop performance, either as a soil amendment or in foliar sprays or trunk implants (Wallace, 1991).

As calcareous soils are particularly common in regions of low rainfall, irrigation is frequently required in order to facilitate plant growth for agricultural production. This in itself may be linked to management problems if drainage through the calcic horizon is impeded and inorganic salts from the irrigation water accumulate within the soil profile, potentially leading to the development of salinity and/or sodicity.
References


8 Managing the soil physical environment for plants

Paul D. Hallett\textsuperscript{1} and A. Glyn Bengough\textsuperscript{1,2}

\textsuperscript{1}The James Hutton Institute, Dundee, UK
\textsuperscript{2}Division of Civil Engineering, University of Dundee, Dundee, UK

8.1 Introduction

The performance of plants is affected considerably by the soil physical environment. High-quality soil is characterised by a porous structure that drains well but stores water to support plants, breaks up easily and allows roots to elongate so that they can capture nutrients and water. If soil is damaged through compaction or drains poorly, the yield of crops, forest growth or quality of sports fields can decrease. The physical environment of soil drives competition between species and therefore vegetation ecology. People invest considerable time and money manipulating the physical conditions of soil to improve plant productivity. One of the greatest expenditures in agriculture is soil cultivation to improve seedbed conditions for seed germination, shoot emergence and root growth. Foresters also adopt soil cultivation if soil is damaged by harvesting machinery or has inherent physical properties that may hinder tree establishment and productivity. Land for amenity purposes, such as golf courses, sport fields and parkland, is sometimes modified with imported soil where the texture, packing and drainage are engineered to produce improved conditions for plant growth and recovery after mechanical stress.

This chapter describes the physical properties of soil that influence plants and how these properties can be managed. It also describes natural processes, induced by weathering or biological processes, that may improve soil physical conditions. Large areas of land have been disturbed by tillage, intensive grazing, vehicle damage and engineering projects with the long-term use of soils dependent on the natural physical resilience of soil. Careful soil management, through natural processes and sometimes mechanical intervention, makes it possible to enhance the soil physical environment for plants.

8.2 Basic physical properties of soil

This chapter begins with basic definitions of soil physical properties. Much of the information presented also requires an understanding of soil water physics (Chapters 2 and 9).
Information on processes underlying the formation of soil physical structure can be found in Chapters 3 (Concepts of soil fertility), 4 (Soil organic matter) and 10 (Plant-induced changes to soil processes).

### 8.2.1 Porosity, density and void ratio

Plant roots grow in the pores between soil particles. The abundance, interconnectivity, size distribution and tortuosity of the soil pore system therefore affect the capacity of soil to support plants. The most widely used soil property related to the abundance of pores is bulk density, \( \rho_d \), defined from the mass of soil, \( M_s \), in a given volume \( V_t \):

\[
\rho_d = \frac{M_s}{V_t} \quad (8.1)
\]

Bulk density is simple to measure from an extracted core of soil and is often used to describe compaction. As in equation 8.1, it is usually expressed in terms of dry mass of soil, but can also be expressed in terms of wet mass, which is more common for some potting mix manufacturers and geotechnical engineers. Mineral soils usually have bulk densities ranging from 1.0 to 1.9 Mg m\(^{-3}\) depending on texture and soil compaction (Keller and Hakansson, 2010). Organic soils have much smaller bulk densities because of the small density of organic compounds. Particle density, \( \rho_s \), is defined as \( \rho_s = M_s / V_s \), where \( V_s \) is the volume of soil particles alone (i.e. air space removed). For mineral soil, \( \rho_s \) varies from 2.6 to 2.7 Mg m\(^{-3}\) (Blanco-Canqui et al., 2006), but appreciable amounts of organic carbon can reduce this to about 1.5 Mg m\(^{-3}\) (Ruhlmann et al., 2006). Potting media used in horticulture have a much smaller density because of the greater proportion of organic compounds present. Wiberg et al. (2005) found a 76% coefficient of variation in the bulk density of 24 brands of potting media, mainly due to poor quality control. Moreover, the compressibility of potting media and temporal changes due to decomposition complicate the maintenance of prescribed bulk densities (Blok and Wever, 2008).

A simpler to interpret physical property is soil porosity, \( \phi \),

\[
\phi = \frac{V_p}{V_t} \quad (8.2)
\]

where \( V_p \) is the volume of pores. Unlike bulk density, particle density does not affect the measurement, so it is a better description of pore space in soil. Porosity can be determined from \( \rho_d \) and \( \rho_s \) by

\[
\phi = \frac{\rho_d - \rho_s}{\rho_s} \quad (8.3)
\]

Geotechnical engineers often describe the abundance of pores as void ratio, \( e \), defined as

\[
e = \frac{V_p}{V_t} \quad (8.4)
\]

As the denominator, \( V_t \), is fixed and thus remains constant regardless of soil density or porosity, it offers mathematical simplicity for soil-management studies (Dexter, 1997). \( \rho_d \) and \( \phi \), in contrast, are derived from \( V_t \) in the denominator, which changes with fluctuations in porosity. The conversion from \( \phi \) to \( e \) is

\[
e = \frac{\phi}{1 - \phi} \quad (8.5)
\]
8.2.2 Soil water

Water resides within soil pores, with bulk abundance defined by either gravimetric, \( w \), or volumetric, \( \theta \), water content,

\[
w = \frac{M_w}{M_i} \quad \text{and} \quad \theta = \frac{V_w}{V_i}
\]

where \( M_w \) and \( V_w \) are the mass and volume of water, respectively. Conversions between gravimetric and volumetric water content are a potential source of operator error, particularly if the property is poorly defined as ‘water content’ only and expressed as a percentage. Typically, manual measurements in the laboratory report gravimetric water content, whereas most soil water sensors report volumetric water content. If \( \rho_d \) is known, \( \theta \) can be derived from \( w \) by

\[
\theta = \frac{w \rho_d}{\rho_w},
\]

where \( \rho_w \) is the density of water.

Chapter 9 (Section 9.2.2) describes water retention in soil. Capillary forces between air, liquid and solid interfaces result in water held at negative matric potentials in unsaturated soils. There is an inverse relationship between the magnitude of matric potential (or suction) and pore radius, resulting in the gradual drainage of pores from largest to smallest as suction increases and the soil dries. The proportion of pores filled with water influences both aeration and the amount of stored water available for plants.

8.2.3 Soil strength

The rate of root growth is influenced considerably by soil strength, which is imparted by capillary forces and particle bonds in soil (Pärtel and Wilson, 2001; Bengough, 2003). Poor soil management that results in compaction (Glab, 2008; Alameda and Villar, 2009) or losses in organic matter (Emerson et al., 1994; Hati et al., 2006) generally increases soil strength, resulting in decreased root growth. A common field method for measuring soil strength is the use of a field penetrometer, which measures the resistance experienced by a metal cone being pushed into soil over a range of depths (Fritton, 2008; Sinnett et al., 2008). Another approach is the use of a shear vane, which has been advocated by some as a better measurement to describe mechanical resistance experienced by roots in the field (McKenzie and McBratney, 2001), but it is sometimes difficult to use for shallow measurements. Field measurements with either approach are influenced considerably by soil water content, because of changes in matric potential and degree of saturation (McKenzie and McBratney, 2001; Whalley et al., 2007). An alternative approach is to test soil cores equilibrated to prescribed matric potentials with a miniature penetrometer in the laboratory (Bengough and Mullins, 1990). This allows for testing at multiple matric potentials and removes the impact of water content on the variability of field results. Another limitation to penetrometer readings is the frictional resistance of the metal cone, so Bengough et al. (1991) developed a rotating-tip penetrometer to reduce its influence. In subsequent work, they found that using a rotating-tip penetrometer to reduce the frictional resistance reduced penetration resistance by more than 50% (Bengough et al., 1997).

Physical limitations to root growth are described in Section 8.5.
8.3 Soil structure

Soil structure is the physical framework in which all soil processes occur. The arrangement of solids, gases, liquids, voids, organic matter and organisms defines soil structure. It thus has a large influence on the capacity of plant roots to grow through soil, extract water and nutrients, exchange gases and interact with other plants, herbivores and pathogens (Angers and Caron, 1998). The range of sizes of different components of soil structure is enormous, varying by a factor of $>10^8$ between the diameter of small clay particles (typically $10^{-7}$ m) and soil hydrological features at field scale. Dexter (1988) described soil structure as the ‘spatial heterogeneity of the different components or properties of soil’, which embraces the complex environment created by the myriad of mineral particles, organic substances, organisms, weathering processes and anthropogenic impacts underlying soil formation. This definition also embraces the complexity of the resulting pore structure and the potential influence on the growth and functioning of plant roots. Young and Crawford (2004) described soil as ‘the most complicated biomaterial on the planet’.

A surge in research beginning in the 1990s showed conclusively that omitting soil structure from examinations of chemical, physical and biological processes in the natural soil environment often produces results that differ greatly from reality (Horn et al., 1994; Connolly, 1998). In the same decade, the rapid increase in available computer power combined with new mathematical approaches allowed soil structure to be defined far more accurately than was previously possible (Rieu and Sposito, 1991; Crawford and Matsui, 1996; Bird et al., 2000). Models of water and solute transport (Kosugi, 1999; Vogel et al., 2006; Zhang and Lv, 2009), biochemical transformations (Totsche et al., 2010) and plant root development (Pierret et al., 2007) among others have been developed where a major component is a detailed description of the soil pore structure.

Figure 8.1 illustrates the complex physical structure of soil, ranging in scale from the agglomeration of clay into domains to large-scale cracks. Textural porosity is derived from spaces left from the imperfect packing of soil minerals. Loam soils tend to have larger densities than clay soils because the greater array of particle sizes results in the in-filling of textural pores (Table 8.1). Structural pores form at small-scale (μm) from the aggregation of clay and organic matter complexes (Chapter 4). At larger scales, cracking and spaces left from decomposing roots or earthworm burrows form macropores. In general, textural pores provide water storage for plant roots, whereas structural pores provide rapid transport pathways.

The importance of structural pores is often underestimated. They are the large volume singularities in soil and consequently usually air filled. This spacious, aerobic environment is required by many micro-organisms (Emerson and Greenland, 1990). Structural pores are the weak point in soil that fractures under the action of tillage implements (Hallett et al., 1995), so their abundance and arrangement is important to the energy expended in creating seedbeds. Table 8.2 lists the size classification of soil pores and their relevance in terms of soil properties.

8.3.1 Soil structure formation

Soil structure formation is driven by various hydrological, mechanical and biological processes (Horn et al., 1994; Angers and Caron, 1998; Cosentino et al., 2006). The hydrological processes are the frequency and rate of wetting and drying of soil (Emerson
Soil Conditions and Plant Growth

Mechanical processes include the bonding of soil particles and external stresses from overburden soil, compaction or tillage. Hydrological and mechanical processes interact in the capillary cohesion of soil particles by water menisci in pores, shrinking and swelling dynamics, and the weight of overburden soil (Or and Ghezzehei, 2002). Biological processes include the mechanical deformation of soil by roots or fauna, exudates that alter

Table 8.1  Typical densities and porosities of soils with different textures and their parent material.

<table>
<thead>
<tr>
<th>Description</th>
<th>Bulk density (Mg m⁻³)</th>
<th>Porosity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clay texture</td>
<td>1.12</td>
<td>0.58</td>
</tr>
<tr>
<td>Loam texture</td>
<td>1.28</td>
<td>0.52</td>
</tr>
<tr>
<td>Sandy texture</td>
<td>1.61</td>
<td>0.39</td>
</tr>
<tr>
<td>Sandy loam (compacted)</td>
<td>1.90</td>
<td>0.29</td>
</tr>
<tr>
<td>Sandstone</td>
<td>2.12</td>
<td>0.20</td>
</tr>
<tr>
<td>Soil minerals</td>
<td>2.65</td>
<td>0.00</td>
</tr>
</tbody>
</table>

and Grundy, 1954). Mechanical processes include the bonding of soil particles and external stresses from overburden soil, compaction or tillage. Hydrological and mechanical processes interact in the capillary cohesion of soil particles by water menisci in pores, shrinking and swelling dynamics, and the weight of overburden soil (Or and Ghezzehei, 2002). Biological processes include the mechanical deformation of soil by roots or fauna, exudates that alter

Figure 8.1  Soil physical structure across a wide range of scales. Redrawn from Tisdall and Oades (1982). With kind permission from John Wiley & Sons. For a colour version of this figure, please see Plate 8.1.
Managing the soil physical environment for plants

Cycles of drying and wetting induce shrinking and swelling in non-rigid soils and can cause soils to slake or crack (Yoshida and Hallett, 2008). Cracks form macropores in soil that serve as rapid transmission pathways between upper and lower soil layers (Morris, 1992). They form if sufficient mechanical energy is transmitted to soil to break bonds between soil particles (Lima and Grismer, 1994). Plant roots exploit macropores as conduits through plough pans to allow access to subsoil water (White and Kirkegaard, 2010). In row-planted crops, greater water uptake close to plants can cause a hydromechanical gradient that drives the formation of deep inter-row cracks, particularly in heavy clays such as in paddy fields (Yoshida and Adachi, 2001). This can be important to field drainage in wet regions, but in dry regions, it can result in hydrologically isolated volumes of soil and the rapid transmission of surface water to groundwater (Lima and Grismer, 1994). Large surface cracks are a very visible soil structure feature, but cracks also occur at much smaller scales; they are the boundaries defining incipient soil aggregates that coalesce under the mechanical action of tillage to produce tilth (Hallett et al., 1995).

Tillage is the most abrupt form of mechanical disturbance that impacts upon soil structure. It has a desired outcome of fragmenting soil into an array of aggregate sizes that provide conditions to support plant growth from seed germination through to the longer term uptake of water, nutrients and oxygen by roots (Warkentin, 2008). Stresses imparted from wheels during tillage or other operations, however, may cause reductions in soil porosity through compaction (Horn et al., 1994) and decrease continuity of soil pores through shear damage (O’Sullivan et al., 1999). If pore continuity decreases, transmission pathways from upper to lower soil layers decrease, with implications for gas and water transport.

Just as the rate of water transport depends greatly on the soil pore structure, the ingress of water into soil pores can result in air pressure build-up that causes soils to slake (Figure 8.2). A well-structured soil will be able to dissipate air pressure build-up through connected air-filled pores (Henin and Santamari, 1975). Bonding between particles also controls the resistance of soil to slaking, similar to the resistance to cracking described earlier. Organic compounds in soil can increase the bond energy and may also decrease the wetting rate through hydrophobicity (Czarnes et al., 2000). Consequently, a depletion of organic matter can result in soils with a much greater susceptibility to slaking.

Slaking, cracking and mechanical deformation are influenced considerably by the type and amount of clay minerals in soils. The negative surface charge of many clays forms electrostatic forces of attraction that influence bond energy (Barzegar et al., 1995). Interactions with biological compounds, such as polysaccharides, can increase bond energy

### Table 8.2  Soil pore sizes and their relevance to soil properties.

<table>
<thead>
<tr>
<th>Size, m</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>$10^{-2}$</td>
<td>Macropores: cracks, clods</td>
</tr>
<tr>
<td>$10^{-3}$ (mm)</td>
<td>Tilled seedbed</td>
</tr>
<tr>
<td>$10^{-4}$</td>
<td>Smallest diameter for root penetration</td>
</tr>
<tr>
<td>$10^{-5}$</td>
<td>Largest water filled pores at field capacity</td>
</tr>
<tr>
<td>$10^{-6}$ (μm)</td>
<td>Smallest size accessible to bacteria</td>
</tr>
<tr>
<td>$10^{-7}$</td>
<td>Plants can access water held in these pores</td>
</tr>
<tr>
<td>$10^{-8}$</td>
<td>Space between clay plates (can be 50% of pore space)</td>
</tr>
<tr>
<td>$10^{-9}$ (nm)</td>
<td>Thickness of three layers of water molecules on clay surface</td>
</tr>
</tbody>
</table>
Soil Conditions and Plant Growth

significantly (Chenu, 1989; Chenu and Guérif, 1991). Clays with a 2:1 mineral structure (e.g. smectites) have a greater propensity to shrink and swell under cycles of drying and wetting than 1:1 minerals (e.g. kaolinite). The differential swelling of clays during wetting is another process that can cause slaking.

8.3.2 Aggregation

The aggregation of soil particles is a primary driver of soil structure formation (Emerson and Greenland, 1990). Chapter 4 described mineral–organic matter complexes that produce aggregated structures in soil. For plant roots, aggregation produces inter-aggregate pore spaces that ease penetration and improve aeration and intra-aggregate pore spaces that retain water and nutrients (Angers and Caron, 1998). These pore classes were defined in Section 8.3 as structural and textural, respectively. Aggregated soils generally have greater structural stability, resulting in less temporal fluctuation in pore structure over time (Le Bissonnais, 1996; Or et al., 2000). This is extremely important for the stability of cultivated seedbeds (Atkinson et al., 2007) as structural degradation can lead to decreased water infiltration and consequently greater soil erosion (Canton et al., 2009; Vermang et al., 2009). Moreover, physically degraded seedbeds are more prone to drought and waterlogging (Or and Ghezzehei, 2002).

The concept of ‘aggregates within clods’ was proposed by Currie (1966) and later refined to a proposed hierarchical order of soil aggregation in the highly cited paper of Tisdall and Oades (1979). It is now widely accepted that soils possess an aggregate hierarchy that underlies soil structure (Six et al., 2004). Figure 8.1 illustrates the hierarchical arrangement of soil particles. The lowest hierarchical order can be classified as individual mineral particles such as clay plates. This order combines to form flocules or domains of clay that represent the next hierarchical order. Subsequent combinations of orders form larger units in the aggregate hierarchy. Aggregates, as a topic, receive considerable attention in this chapter because the term has become analogous with soil structure and the concept underlies a large body of research related to root growth and rhizosphere development.

Figure 8.2 Build-up and dissipation of air pressure of soil is a major process leading to slaking of soil aggregates. For a colour version of this figure, please see Plate 8.2.
Aggregates are often demarcated into size classes, with agreement among many researchers that macro-aggregates are >250 μm, meso-aggregates are 53–250 μm and micro-aggregates are <53 μm (Chan et al., 1994; Six et al., 2000). Some researchers only use 250 μm as the demarcation between macro-aggregate and micro-aggregate, with meso-aggregates not defined (Six et al., 2004; De Gryze et al., 2006; Fristensky and Grismer, 2008; Wick et al., 2009). Stable aggregates are separated from bulk soil by imparting stresses either by mechanical agitation (Ashman et al., 2009), wet sieving (Le Bissonnais, 1996) or ultrasonic dispersion (Fristensky and Grismer, 2008). The concept is that unstable soil will fragment, whereas stable soil will remain aggregated and not pass through a sieve. Immersion in water and agitation simulates the action of rapid wetting (i.e. slaking) or raindrop impact in the field. As the amount of energy imparted to soil and its spatial distribution will influence the extent of fragmentation, aggregates collected through these approaches are thought by some to be biased considerably by the fractionation procedure (Ashman et al., 2003). This debate has been going on for some time (Beare and Bruce, 1993), but the fractionation approach described by Le Bissonnais (1996) is accepted by many to be the most robust sieving approach. There is a resurgence of interest in using ultrasonic dispersion because energy inputs are more controlled (Fristensky and Grismer, 2008; Zhu et al., 2009), although North (1979) found the approach to be insensitive for identifying changes in aggregation resulting from long-term soil-management practices.

Soil aggregate stability

The structure of soil can be a highly dynamic property, governed by a range of soil properties and external stresses. Freshly tilled seedbeds, for instance, can change over time as the tilth coalesces due to raindrop impact, fluctuations in soil water content (Dexter, 1997) and compaction from machinery (O’Sullivan and Simota, 1995). Dispersion and slaking of surface soil can lead to sealing with considerable effects on sediment loss through erosion (Skidmore and Powers, 1982). In a given soil, the quantity, quality and distribution of carbon are the major variables controlling stability (Zhang and Horn, 2001; Gerke and Kohn, 2002; Goebel et al., 2005; Park and Smucker, 2005), in addition to the amount of clay minerals and their mineralogy (Denef and Six, 2005). With the massive losses in carbon experienced in many managed soils, particularly those under arable production (Bellamy et al., 2005), declines in soil stability are posing a major threat to the environment and crop productivity (Diacono and Montemurro, 2010). The Great Dust Bowl that plagued the Midwest of the USA in the 1930s and dust storms that affect Beijing presently are examples where degradation of soil structural stability can have devastating effects.

There are numerous approaches for measuring soil stability, with considerable overlap with the techniques referred to in the Section 8.3.2 to fractionate soil into different sized ‘aggregates’. Many approaches have been developed since Yoder (1936) developed a standardized method of immersing soil in water on an oscillating nest of six sieves with decreasing aperture from top to bottom. He defined the lowering and raising of the sieves by 3.18 cm, at a rate of 30 cycles/min for a period of 30 min. This method forms the standard used by a large body of research on the stability of soil to wetting stresses (Kemper and Rosenau, 1986). After wet sieving, the size distribution of aggregates is assessed from the measured mass retained on each sieve size above 0.1 mm and uses Stokes Law to estimate smaller aggregate sizes. The size distribution of soil aggregates can be expressed as mean weight diameter (MWD) by
Soil Conditions and Plant Growth

\[ MWD = \sum_{i=1}^{n} \bar{x}_i w_i \]  

where \( \bar{x}_i \) is the mean diameter of each sieve fraction, \( w_i \) the mass of soil retained and \( n \) the size fraction. An increase in MWD indicates greater soil stability. Other methods to characterise aggregate size distribution have been proposed, such as the use of fractals (Perfect et al., 1992).

The wet sieving approach to assess aggregate stability has been modified by altering the capillary stress during wetting (Amezketa et al., 1996) or the use of different wetting liquids (e.g. ethanol) to isolate fragmentation mechanisms (Henin, 1948; Emerson, 1954; Le Bissonnais, 1996). It is often argued that the energy input from wetting soil is difficult to quantify, so mechanical agitation through shaking (Watts and Dexter, 1997) or ultrasonic dispersion (North, 1979; Fristensky and Grismer, 2008; Zhu et al., 2009) have also been used. Another test of wet aggregate stability developed by Emerson (1954) uses a visual examination of the dispersion of clay and slaking of small air-dry soil aggregates immersed in distilled water.

Dry soil stability is also measured, generally to evaluate the percentage of soil aggregates that are not erodible by wind (Chepil, 1952; Fryrear, 1985). The tests agitate samples of dry soil over one or more sieves with no exposure to water. Aggregates <0.42 mm are characterised as wind erodible, while aggregates >0.84 mm are non-erodible (Chepil, 1951). More recent research has examined the influence of the size distribution of aggregates on wind erodibility in greater detail, using measurements such as geometric mean diameter or fractals to characterise the distribution of aggregate size (Chepil, 1951; Popham et al., 2003). Dry soil stability is not only important for assessing the potential environmental impacts of poor soil management but also the abrasive impact of wind blown soil particles on plants (Michels et al., 1993; Baker, 2007). In extreme instances, damage of crops by blown soil during establishment can require complete replanting. Climate change is anticipated to increase periods of drought, prompting concern about the potential implications of wind erosion (Merrill et al., 1999). Wind breaks offer some protection, although this common feature of fields in the early twentieth century has declined as mechanised agriculture allowed for larger areas of land to be managed.

A plethora of research has been published on soil aggregate stability, with various critiques over the years that identify problems such as the lack of a unified fractionation technique and physical meaning of collected soil aggregates (Beare and Bruce, 1993; Amezketa, 1999; Ashman et al., 2003; Warkentin, 2008). However, aggregate stability tests do provide a useful measure of the propensity of soil to structurally collapse or erode, thereby allowing for comparison between different land-management practices. They also provide some insight into the structural arrangement of soil, such as the clay–organic matter complexes that form the building blocks of soil aggregates.

8.3.3 Approaches to measuring soil structure

Measures of soil aggregates and their stability are often used as a surrogate measurement of soil structure. This provides useful insight into mechanisms underlying the formation of soil structure, but the approach is destructive, so it is difficult to reconstruct the original structure of the soil. Using aggregates to describe soil structure has been compared to studying architecture by levelling a building with a wrecking ball. Other approaches include
state-of-the-art non-invasive three-dimensional (3-D) imaging and practical field-based assessments of soil structure.

**Non-invasive imaging**

Recent advances in non-invasive imaging using X-ray computer-aided tomography (CT) and nuclear magnetic resonance (NMR) allows for the 3-D physical structure of soil to be visualised on intact specimens (Taina et al., 2008). X-ray CT was first used with soil to estimate bulk density (Petrovic et al., 1982) and subsequently water content (Hainsworth and Aylmore, 1983). The fuzzy two-dimensional slices with poor resolution found in this early research have been replaced by 3-D images with a resolution as small as 3 μm in micro-focus scanners (Tippkotter et al., 2009) and 0.84 μm in nano-focus scanners (Sleutel et al., 2008). Dual energy source X-ray CT scanners are now available from various manufacturers as bench-top instruments that allow for the discrimination of soil, air and water (Rogasik et al., 1999). They are also beginning to allow for the discrimination of organic matter (Sleutel et al., 2008) and plant roots (Perret et al., 2007), although there is still considerable scope to improve the technique. A major limitation to the analysis of roots is the size of sample that is required and the time required for reconstruction of images assembled from huge data sets. This can limit sample scanning to <10/day and diminishes the resolution so that only macropores and large roots can be detected, although the technology is improving very rapidly (Gregory et al., 2009).

Synchrotron-source X-ray CT decreases scanning time while retaining resolution (Peth et al., 2008). Access time for users is limited as Synchrotrons are large-scale national facilities, but they have become a major research tool of some soil scientists (Feeney et al., 2006). Capture rates on small specimens can be rapid enough to monitor dynamic processes such as a water flow (Prunty et al., 2003).

NMR can discriminate the 3-D distribution of elements within soil but it is hindered by interference from ferromagnetic inclusions. The approach was used in the 1970s to characterise soil water content (Prebble and Currie, 1970) and later used to investigate pore size distribution, connectivity (Bird et al., 2005) and water retention characteristics (Jaeger et al., 2009). NMR is also used extensively to characterise organic matter in soil, including in studies of rhizosphere development (Chen et al., 2006) and root tissue decomposition (Mathers et al., 2007).

**Soil thin sections**

The 3D imaging tools discussed thus far are beginning to replace the older approach of thin sections (Kubiëna, 1938). In this approach, water is replaced with a solvent in intact soil specimens, the sample is then impregnated with a resin and finally cut and polished into a thin slice of soil set on a glass slide. Polishing of one surface can decrease the sample thickness to <20 μm. There are two major advantages to using thin sections over non-invasive techniques. First, microscopic techniques allow for the discrimination of the spatial distribution of clays, sand and organic matter; and second, the thin sections can also be sampled at the micro-scale to investigate the distribution of elements, either by removing material or laser ablation (Bruneau et al., 2002). Another advancement in thin sections is an ability to identify micro-organisms either through indiscriminate staining with calcofluor (Nunan et al., 2001) or the detection of specific organisms using molecular probes (Eickhorst and Tippkotter, 2008).
Intrusion and extrusion of liquids

Soil structure can also be inferred from the pore size distribution estimated from mercury porosimetry or water retention characteristics. Mercury porosimetry injects mercury into soil under pressure and allows for pores as small as 0.003 μm to be detected (Bartoli et al., 1999). The measurement of water retention characteristics is described in Chapter 9. Briefly, soils are desaturated on tension tables for smaller (less negative; i.e. 0 to −20 kPa) matric potentials or in pressure plate apparatus for larger (more negative; i.e. −20 kPa to −1500 kPa) matric potentials. Shrinkage of soil during drying influences the results, but methods have been developed to correct for this effect (Braudeau et al., 1999). Both water retention and mercury porosimetry are limited by the impacts of pore connectivity and shape so they do not provide a direct measurement of pore structure. However, equipment to measure water retention is commonplace in most soil physics laboratories and the data collected is also relevant to the accessibility of water by plants. Dexter and Bird (2001) proposed the use of water retention characteristics to provide a qualitative description of soil structure known as the $S$ factor. This is obtained from the slope at the inflection point of a water retention curve and reflects the spread of pore sizes.

Visual indicators of soil structure

More direct, qualitative measurements of soil structure examine distinct visual features (Mueller et al., 2009). In many national soil surveys, the shapes of aggregates prised from a soil profile were characterised. These data are simple to collect and provide valuable information on the physical condition of soil, thereby identifying soil types, regions or management practices where the soil physical conditions for plant growth may be limited. There are five major classes of soil structure: platy, prismatic, columnar, granular and blocky. Subclasses, such as subangular blocky, can also be defined. Soil with a granular structure, characteristic of grasslands high in organic matter, can change to a columnar or platy structure if carbon mineralisation and compaction from machinery damages aggregation. Platy structures found in plough pans limit penetration by plant roots and water. Readers should refer to general soil science textbooks such as Rowell (1994) and Marshall et al. (1996) for further information.

Soil physical quality indicators have been developed that characterise various features of a block of soil extracted with a spade. In the Visual Soil Structure Quality Assessment, a score is assigned based on the difficulty encountered when pushing the spade into the soil and extracting the block, the shape, strength and size of the largest aggregates, and the presence of gleyed anaerobic soil (Ball and Douglas, 2003). The block of soil is broken apart by hand to assess its friability and examine the porosity of larger aggregates. Scores from 1–5 are assigned, with 1 indicating a highly friable soil that provides an excellent environment for roots and 5 indicating a very compact soil, containing at best a few cracks. More detailed visual scoring of soil structure can be done with the Peerlkamp test or the French method ‘Le profil cultural’ (Mueller et al., 2010). The different approaches have been compared to each other, examined for operator variability and used on many soil types across the globe (Mueller et al., 2009). There is some need to refine the approach for particular regions or land uses. Water content and the time since tillage operations will also influence the results.
8.4 Soil tilth and seedbeds

In a review of research on soil structure, Warkentin (2008) quotes Cato in the ancient textbook ‘Roman Farm Management’:

What is the first principle of good agriculture? To plough well.
What is the second? To plough again; and then to manure.

Agriculturists like Cato, and his successors up to the nineteenth century, viewed soil structure as the tilth produced by soil cultivation. Although knowledge of soil structure produced by different cultivation implements or under different environmental conditions is plentiful, it involves mainly observational studies with minimal quantification of in situ properties and their variability with time (Atkinson et al., 2009). There has been abundant research investigating the distribution of clod and aggregate sizes following tillage (Braunack and Dexter, 1989a; Slowinska-jurkiewicz, 1994; Guerif et al., 2001). Earlier research sieved the tilth to separate size fractions, while other studies impregnated soil with resin and then studied macrostructure in large blocks or micro-structure to 20μm pore sizes in thin sections (Braunack and Dexter, 1989a). Much of this research was largely observational, although there are some examples of predicting physical properties, such as saturated hydraulic conductivity, from soil thin sections (Zhang et al., 2005). The importance of the size and geometry of soil structural units in relation to transport to roots was estimated by Passioura (1991). He predicted that the time taken for roots to extract water increased by up to two orders of magnitude for roots clumped within subsoil macropores as compared with uniformly distributed roots in a seedbed.

8.4.1 Soil tillage

Since Sir Edward John Russell published the first edition of this book in 1912, the practices used to cultivate soil have been revolutionised. On many soils around the globe, the plough has been replaced by reduced tillage systems; for example, over 20% of USA and almost half of Argentinian agricultural land is under zero tillage (Farage et al., 2007). Several drivers have prompted this change, including the ability to control weeds with herbicides, recognition of the deleterious impact of mechanical disturbance from cultivation on soil carbon mineralisation and a desire to save costs on fuel and labour (Triplett and Dick, 2008). The FAO and others have promoted reduced tillage for its capacity to decrease soil degradation, particularly in regions with less fertile, fragile or impoverished soils (Knowler and Bradshaw, 2007).

Figure 8.3 illustrates the various operations in cultivating soil using conventional, reduced and no-till tillage systems. Under conventional tillage, the soil is first inverted to a depth of about 20cm with a mouldboard plough. This buries residue and causes the initial fragmentation of the soil. A second pass with a disc breaks the clods produced from ploughing into smaller aggregates, which can be broken down further by a pass with tines or harrowing. Planting is done by drilling seeds into the cultivated seedbed. Section 8.4.2 describes the optimal seedbed conditions that farmers strive for with conventional tillage. Cultivation may be done after crop establishment for the mechanical control of weeds.

Conservation tillage disrupts soil to a shallower depth and is a generic term applied to any cultivation practice that is not conventional tillage. Essentially, it removes the use of a plough. The variety of practices defined as conservation tillage is enormous, varying from the direct drilling of seeds into uncultivated seedbeds in zero till to elaborate minimum tillage implements that subsoil, cultivate, plant and fertilise in one pass.
Soil Conditions and Plant Growth

Table 8.3 provides an overview of the advantages and disadvantages of soil tillage, particularly in relation to ploughing.

8.4.2 Desired physical properties of seedbeds

Considerable energy can be invested in creating seedbeds that optimise germination, plant growth and yield. Over 50% of the root mass of a developed plant are found in the tilled layer forming the seedbed, so its physical properties are very important to the production of food and biomass (Finney and Knight, 1973). Seedbed properties vary considerably depending on soil management, residue incorporation, tillage approaches, soil type and drainage (Guérif et al., 2001). In a good seedbed, almost all viable seeds will germinate and emerge (Nasr and Selles, 1995). Soil temperature is one of the most important physical properties of a seedbed that regulates germination (Addae and Pearson, 1992). Many models that simulate germination therefore use thermal time based on cumulative degree days. As water becomes scarcer, however, the significance of drought has been incorporated in hydrothermal time models (Rinaldi et al., 2005).

Loosening soil through cultivation decreases mechanical impedance to seedling emergence and root growth (Braunack and Dexter, 1989a). This has to be balanced against good seed–soil contact, which can be improved through rolling of the soil surface after sowing using pressures great enough to pack the soil but small enough to not impede plants (Passioura and Leeper, 1963; Adem et al., 1984). In general, an 'optimal' seedbed in cultivated soils consists of aggregates 1–5 mm in size (Russell, 1961). The packing of aggregates should form sufficient pores >75 μm in diameter that freely drain so that oxygen exchange maintains an aerobic seedbed (Adem et al., 1984). In addition, pores 0.2–30 μm in diameter are needed as a reservoir to store water that can be used by plants.
A great deal of research has investigated the optimal size distribution of aggregates for seedbeds (Braunack and Dexter, 1989b). It has been suggested that aggregates should be about one-fifth to one-tenth the size of the seed (Hadas and Russo, 1974), but for smaller seeds, this may not be practical and result in erosion. Whereas cereals perform well close to the 1 mm threshold described by Russell (1961), maize and sugar beet perform better at the larger 5 mm threshold. However, in the various studies reviewed by Braunack and Dexter (1989b), they noted that soil type and climate had a very large impact on the results, as shown in studies done in multiple years when conditions were either wet or dry. In general, a finer tilth results in better performance in drier conditions because the risk of anaerobic conditions is insignificant and seed–soil contact is greater.

Soil water potential regulates the rate that seeds imbibe water and this rate varies markedly between plants and crop cultivars. Al-Karaki (1998) demonstrated this impact for lentil seeds of different sizes (Figure 8.4). Smaller seeds (30 mg) imbibed water more rapidly than
larger seeds (58 mg), and a change in water potential from −0.03 MPa to −0.30 MPa decreased water uptake over 24 h by 40%. Research by Wuest (2007) suggests that vapour transport from soil to seed plays a substantial role in seed imbibition in unsaturated soils. Once germinated, root elongation is also affected by soil water potential, either by the potential gradient for water uptake or hypoxia if the soil is anaerobic or through mechanical impedance (Whitmore and Whalley, 2009).

8.4.3 Factors regulating soil fragmentation

The fragmentation of soil in the formation of seedbeds is controlled by a multitude of processes but the major drivers are: (1) internal pore structure, (2) water content, (3) mechanical and hydrological properties of the soil and (4) the mechanical action of the tillage implement. These drivers interact and generally result in decreased tillage efficiency in terms of energy input and fragmentation if soil is damaged through compaction or organic matter loss (Adem et al., 1984; Keller et al., 2007). In the soil thin sections shown in Figure 8.5, these impacts are visually apparent. Years of ploughing that depleted carbon and physically damaged the soil shown on the right has resulted in a narrower size distribution of pores and units of denser soil. A change in soil management, as shown for the soil on the left, can produce a broader range of pore sizes and improve aggregation of the soil. Fragmentation occurs through the coalescence of pre-existing pores and cracks in the soil, so the soil shown on the left should fragment into smaller aggregates with less energy input than the soil on the right (Grant et al., 1990).

Figure 8.4 Water uptake by small (filled symbols, 30 mg) and large (empty symbols, 58 mg) lentil seeds over time for different water potentials of PEG (circle −0.03 MPa, square −0.15 MPa and triangle −0.30 MPa). Data taken from Al-Karaki (1998). With kind permission from John Wiley & Sons.
Managing the soil physical environment for plants

The pore structure of the soil also affects water drainage and the presence of air-filled pores (Keller et al., 2007). Stresses imposed by tillage concentrate at the edges of air-filled pores, whereas the stresses will transmit through water-filled pores (Hallett et al., 1995). With increasing water content, fewer air-filled pores exist in the soil so fragmentation can decrease (Grant et al., 1990).

However, there is a trade-off between the influence of soil water content on the strength of the soil and the presence of air-filled pores (Dexter and Birkas, 2004; Keller et al., 2007). If soil is too wet, tillage deforms the soil plastically at the expense of the fragmentation of clods into smaller aggregates. When soil is dry, much greater energy needs to be expended in tillage and the resulting seedbed may require irrigation for germination. Considerable research has attempted to define the optimal water content for tillage. One approach is based on the transition between brittle, plastic and liquid behaviour, defined by Atterberg (1911):

- **Plastic limit, PL (or lower plastic limit)** – transition from brittle to plastic behaviour.
- **Liquid limit, LL (or upper plastic limit)** – transition from plastic to liquid behaviour.
- **Plasticity index** = liquid limit – plastic limit

Specific standards exists for measuring both LL and PL. For both tests the soil is initially sieved to either 425 μm (geotechnical standard) or sometimes 2 mm, remoulded and then tested at a range of water contents. PL is defined by the water content at which a thread of soil rolled to 3 mm begins to crumble. It is a difficult test to perform that requires considerable practice to avoid operator bias. LL can be determined using either a drop cone penetrometer or Casagrande apparatus. With a drop cone penetrometer, the sharp tip of a cone is placed in contact with the surface of the soil and then released. A penetration depth of 20 mm defines the LL. The Casagrande apparatus consists of a round brass cup that moves up and then drops 1 cm through the rotation of a cam. A paste of soil is placed in the cup and a groove is scratched down the centre. LL occurs at the water content when 25 drops closes the groove in the soil paste. The drop cone penetrometer is generally thought to provide more reliable results as it is simpler and induces less dynamic stress during testing.

When soil water content is at about 0.8 PL, the combined effects of limited plastic deformation, soil adhesion and the presence of air-filled pores can result in optimal tillage efficiency (Keller et al., 2007). The wettest water content for tillage is roughly the field capacity of the soil (water content following free drainage; see Chapter 9), whereas the

**Figure 8.5** Thin sections of topsoil from plots after 11 years of zero tillage (left) and conventional tillage by ploughing and discing (right). From VandenBygaart et al. (1999). With permission from Elsevier. For a colour version of this figure, please see Plate 8.3.
Soil Conditions and Plant Growth

The driest cut-off is more arbitrary and has been taken as the water content where the soil strength is double its value for the optimum water content (Dexter and Bird, 2001).

Another approach to identify the optimum water content for tillage uses water retention characteristics to characterise air-filled pore structure, soil composition and bulk density (Dexter et al., 2005). Figure 8.6 shows raw data and a fitted quadratic equation for the breakdown of soil clods over a range of water contents. Using pedotransfer functions, the optimum, lower limit and upper limit water content for tillage, $\theta_{\text{OPT}}$, $\theta_{\text{LTL}}$ and $\theta_{\text{UTL}}$ respectively, can be estimated. Due to the interdependency of soil properties (e.g. clay content and bulk density) and reliability of pedotransfer functions, caution is needed when adopting this type of approach (Dexter et al., 2005).

Good soil management strives to maintain or increase the range of water contents available for tillage. The major variables influencing the size of the range of water contents are organic matter (see Section 8.2.4 on soil aggregates) and pore structure. Both of these properties can be degraded through intensive soil cultivation that enhances carbon mineralisation, disrupts aggregates and decreases their stability (Six et al., 2004; Diacono and Montemurro, 2010) and mechanically deforms the soil through wheel damage or the action of an implement (Horn et al., 1994). Odell et al. (1960) found a positive correlation between PL and soil organic carbon, which suggests that soils depleted of carbon need to be tilled when drier. From a practical standpoint, land managers can be faced with a small window of opportunity to till soils efficiently, particularly in the case of hard-setting soils (Mullins et al., 1988).

8.4.4 Soil friability

Friability describes how soil breaks up into tilth, whether it is from the action of cultivation implements or simply crumbling soil in your hand when assessing physical structure in the

---

**Figure 8.6** Presence of soil clods following tillage at a range of water contents. Raw data are shown on the left, including the minimum value of clods corresponding to the optimum water content for tillage, $\theta_{\text{OPT}}$. The right-hand graph shows the fitted quadratic equation with the upper and lower tillage limits, $\theta_{\text{LTL}}$ and $\theta_{\text{UTL}}$, and the range of water contents for tillage. Reprinted from Dexter et al. (2005). With permission from Elsevier.
field. Christensen (1930) developed the first index to describe soil friability based on the variability of the strength of soil aggregates. This was refined by Braunack et al. (1979) who applied extreme statistics to describe the variability and scaling of strength of different-sized soil aggregates. This produced a highly flexible approach that produced fitting parameters that are related to the spread of strength between and within levels of the soil aggregate hierarchy.

The strength of soil aggregates can be assessed with a crushing test, where soil is compressed between flat parallel plates (Hallett et al., 2000). Failure force, $F$, relates to tensile failure stress, $\sigma$, by

$$\sigma = c \frac{F}{l}$$

(8.10)

where $l$ is the geometric mean of the three principal axis lengths of the aggregate and $c$ is a constant equal to 0.576 (Dexter and Kroesbergen, 1985). The distribution of failure stress is evaluated as,

$$\log \sigma = -k \log V + A$$

(8.11)

where

$$A = \log[\sigma_0 V_0^k \Gamma(1 + k)]$$

(8.12)

The parameters $\sigma$ and $V$ are the strength and volume of the aggregate being tested, while $\sigma_0$ and $V_0$ are the strength and volume of the basic elements of soil that bind to form aggregates. The Gamma function, $\Gamma$ can be obtained from tables and $k$ is a fitting parameter that is proportional to the scatter of the failure stress. $k$ was used by Utomo and Dexter (1981) as an index of soil friability. More complex models have been proposed that use energy rather than stress of fracture (Perfect et al., 1998) or relate more directly to the scaling of pore structure (Hallett et al., 2000). A simpler, more practical approach based on the coefficient of variation of aggregate strength has also been developed (Watts and Dexter, 1998) as an index of soil friability. Field methods based on drop shatter tests (Schjonning et al., 2007) and simpler laboratory methods for assessing aggregate breakdown (Ding et al., 2009) have also been developed to provide a measure of soil friability.

Utomo and Dexter (1981) proposed the following classification of soil friability based on observations from a range of soils:

$k < 0.05$ Not friable
$k = 0.050–0.10$ Slightly friable
$k = 0.10–0.25$ Friable
$k = 0.25–0.40$ Very friable
$k > 0.40$ Mechanically unstable

They found that cycles of wetting and drying or freezing and thawing increased $k$ because of the production of micro-cracks that decreased aggregate strength. Other studies have shown soil to be most friable as the water content approaches the plastic limit (Watts and Dexter, 1998), with more recent models of friability incorporating the impact of water content directly (Munkholm et al., 2007). Watts and Dexter (1998) made many observations of the impact of soil management on friability with the following major findings: (1) there was a positive correlation between friability and organic matter; (2) deep soil was less friable.
than shallow soil; (3) tillage of wet soil decreased the friability of soils following subsequent drying and (4) soil friability changed with seasons.

### 8.4.5 Post-tillage seedbed dynamics

At water contents generally much greater than $\theta_{\text{UTL}}$, soil can begin to deform under its own weight and the action of capillarity forces from pore water (Ghezzehei and Or, 2003). Weathering processes can also disrupt unstable soil aggregates through slaking and raindrop impact (Le Bissonnais, 1996). This produces a seedbed that varies over the season and can degrade rapidly if the soil is poorly managed or inherently unstable due to its composition. Consequently, the tilth produced by tillage can be short lived, giving roots a limited window of opportunity to benefit from the energy exerted in soil cultivation. Hard setting of the soil surface can occur rapidly following cultivation, resulting in a severe restriction to seedling emergence (Rengasamy et al., 2003). This process is characterised by an initial dispersion of the surface soil from wetting followed by the formation of continuous, strong crust on drying. Processes described previously that improve soil aggregation also decrease hard-setting in soils.

Changes in tilth over time for different wetting conditions of soil are shown in Figure 8.7 (Bresson and Moran, 2004). Under slow wetting, aggregates coalesce through their own weight, overburden and the action of capillary bonding with adjacent aggregates (Or, 1996).
Under more rapid wetting, aggregates disperse and break down, with rainfall inducing crusting. Arguably one of the greatest challenges to soil physics is the development of predictive models that describe seedbed dynamics from cultivation to harvest. Although commendable models have been developed to predict surface crusting and its influence on water infiltration (Assouline and Mualem, 2000), and to some extent the temporal change in pore structure over time (Or et al., 2000), research investment in this area has been minimal.

### 8.5 Critical soil physical conditions for plant growth

Soil physical conditions can limit both root and shoot growth if the soil is too hard, too dry, does not contain sufficient oxygen or is far from the optimal temperature (Kaspar and Bland, 1992; daSilva et al., 1994; Bengough et al., 2006a). There is much literature describing both root and shoot growth responses to these stresses, particularly in annual crops, and the main effects are summarised in Table 8.4. The main aspects of these responses will be considered here, and readers are referred to recent reviews on individual topics for more detailed study.

<table>
<thead>
<tr>
<th>Stress</th>
<th>Effects on plant growth</th>
<th>Soil physical measurement</th>
</tr>
</thead>
</table>
| **Mechanical impedance**  
(Bengough et al., 2011) | Slows root elongation (Figure 8.8)  
Increase in root diameter (up to two-fold)  
Slows leaf elongation/expansion at larger values of penetrometer resistance (Figure 8.9) | Penetrometer resistance  
Direct measurement of root penetration resistance |
| **Water stress**  
(Davies et al., 2005) | Slows root elongation (Figure 8.8)  
Small decrease in root diameter  
Slows leaf elongation/expansion (Figure 8.9) | Soil matric potential (tensiometer or psychrometer) |
| **Waterlogging** (Colmer and Voesenek, 2009):  
Hypoxia  
Anoxia | Slows or stops root elongation (Figure 8.8)  
Slows or stops leaf elongation/expansion in susceptible species, though shoot elongation may be substantially enhanced in prolonged submergence of tolerant species | Oxygen diffusion rate  
Oxygen concentration |
| **Non-optimal temperature:**  
Too hot  
Too cold  
(Kaspar and Bland, 1992) | Slows root elongation (Figure 8.8)  
Slows leaf elongation/expansion  
Changes to root orientation | Direct measurement of soil temperature (thermometer, thermocouples and thermistor)  
As described in this column in the cells above |
| **Multiple stresses:**  
Water stress and mechanical impedance  
Waterlogging and mechanical impedance  
Waterlogging and soil temperature | Interactions due to:  
Increase in soil strength as soil dries  
Decreased turgor in water-stressed roots restricts the maximum growth pressure that roots can exert  
Decrease in oxygen diffusion rate with increase in soil compaction  
Increased oxygen demand from roots waterlogged in warm soil | |
Figure 8.8  Root elongation rate expressed as a percentage of the unimpeded elongation rate in relation to (a) penetrometer resistance, (b) matric potential, (c) oxygen diffusion flux and (d) temperature.

Figure 8.9  (a) Leaf length and leaf area (hollow) of wheat seedlings as a function of penetrometer resistance at two time points (data from Masle and Passioura, 1987); (b) leaf elongation rate (% of unstressed) as a function of matric potential (data from Parent et al., 2010).
8.5.1 Local root response to soil conditions

Roots respond to the local soil environment due to the influence of the soil on cell division and expansion rates in the root meristem and elongation zones. Seedling root elongation is shown as a function of individual stresses in Figure 8.8. Here it can be seen that root elongation is slowed to approximately half of its unimpeded rate by soils with a penetrometer resistance of between 1 and 2 MPa; a matric potential of between −0.5 and −1 MPa for maize roots; an oxygen diffusion rate of between 45 and 60 ng/cm²/min for oat roots and a temperature diverging more than 10 °C from the optimum for pecan roots (this will be very species dependent). These represent substantial physical limitations to root elongation, where a seedling root system might only achieve half of the vertical penetration into the soil attained by an unimpeded root (though it should be noted that the angle of root growth may also be affected by stress [e.g. Sheppard and Miller, 1977]). In the field, plants often experience more than one physical stress. Soil strength often increases rapidly as the soil dries so that roots will be confronted with both water stress and mechanical impedance (Whitmore and Whalley, 2009; Bengough et al., 2011). Recent surveys of physical properties from a range of 19 soils indicated that mechanical impedance often poses a greater restriction to root elongation in drying soil than water stress (Bengough et al., 2011): the increase in soil strength on drying was such that more than half the penetrometer resistances measured were in excess of 2 MPa at a matric potential of −200 kPa.

Root elongation rate is essentially the product of the rate of adding new cells onto a cell file and the final cell length. The rate of cell wall yielding influences both the rate of cell partitioning in the meristem and also the final cell length and is thought to be primarily determined by the stiffness of the cell walls, as long as there is sufficient turgor to drive growth. The rate of root elongation (\(dl/dt\)) is then given by (Greacen, 1986; Bengough et al., 2006a)

\[
\left( \frac{dl}{dt} \right) = l.m(\sigma,\psi)(P - Y(\sigma,\psi) - \sigma(\psi))
\]  

(8.13)

where, \(l\) is the length of the zone of elongating root tissue, \(m\) is cell wall extensibility, \(P\) is turgor pressure, \(Y\) is cell wall yield threshold, \(\sigma\) is the root penetration resistance of the soil and \(\psi\) is the soil matric potential. \(P, Y, \sigma\) and \(\psi\) all have dimensions of pressure, while \(m\) is strain rate per unit pressure. Yield threshold and extensibility are both aspects of cell wall stiffness that are related to the rates of making and breaking tethers between cell wall microfibrils and are therefore probably interrelated (Passioura and Fry, 1992). By considering this equation in more detail, root responses can start to be understood at the cellular scale and the growth trajectories of individual cells can be deduced as they move from the meristem through the elongation zone to finally form mature root tissue (Chavarria-Krauser and Schurr, 2004).

Cell wall stiffness is strongly influenced by both soil strength and matric potential, as indicated in Equation 8.13. In the case of mechanical impedance, cell walls stiffen axially in hard soils, causing cells to expand radially and the root to thicken. Such stiffening is evident when roots grow through a hard soil layer into looser soil or are transferred into a non-impeding medium such as hydroponics (Bengough and Young, 1993; Croser et al., 1999). Root elongation rates may take a period of 2–5 days to recover from the stress, allowing time for cells formed in the absence of the stress to enter the elongation zone. For water-stressed roots, rapid local growth (strain) rates are maintained close to the root apex, despite a decrease in turgor caused by the negative soil matric potential (Sharp et al., 2004). This arises because of local cell wall loosening near the root apex, in contrast with stiffening of the cell wall that occurs in the basal region of the elongation zone.
Soil Conditions and Plant Growth

In the case of hypoxia, slower root growth is mainly due to decreased cell length, with cells becoming shorter and fatter (Eavis, 1972). As oxygen concentrations decrease further to approach anoxia, cell production in the root meristem slows too (e.g. $<0.3$ atm oxygen partial pressure for pea roots). In relatively waterlogging tolerant species, the formation of aerenchyma (continuous gas spaces formed by cell lysis or by cells separating) within the root cortex allows transport of oxygen to the root tip to be maintained, decreasing the likelihood of root tip death (Colmer and Voesenek, 2009). The formation of new, especially stem-borne, roots is a very common response to flooding and the new roots formed will generally contain a greater proportion of aerenchyma tissue.

In the field, roots may experience multiple stresses due to spatial and temporal variation in soil physical conditions, particularly soil water and soil density. Soil surface layers may become sufficiently dry that mechanical impedance and water stress may limit root growth and crop yield, while water is available deeper in the profile (Whalley et al., 2006). The least limiting water range is a method for determining the range of water contents for particular soil conditions where root growth is free from major physical limitation (daSilva et al., 1994). The water content at which penetrometer resistance exceeds $2$ MPa is often taken as the soil strength limit; if the soil is any drier than the limit, the roots are likely to be slowed by substantial mechanical impedance to growth. The water content corresponding to $-1.5$ MPa has been used for the water stress limit, while the water content corresponding to $10$ cm$^3$cm$^{-3}$ is taken as the aeration limit. Field capacity water content is often taken as an additional wet end limit, where root growth will be restricted if it is exceeded, although this is less obviously linked to a physical limitation to root growth. By combining this threshold approach with soil water data, it is then possible to characterise for how much of a growing season a root system has been exposed to particular physical stresses (da Silva and Kay, 2004; Bengough et al., 2006b), although it is clear that these threshold values may benefit from further refinement (Mohammadi et al., 2010). This type of approach is promising in relation to starting to scale up responses from the individual root tip in the laboratory to whole crop performance.

8.5.2 Root–shoot signalling and feedback responses

Soil physical conditions influence shoot growth directly via root–shoot signalling and may decrease growth of the whole plant if plant access to water and nutrients becomes limited (e.g. if the root system is too stunted to explore sufficient soil volume).

There has been substantial work on how soil water status influences shoot growth directly via root–shoot signalling. In loose growth media, seedling shoot growth is relatively more sensitive to drying matric potential than is root elongation (Sharp et al., 1988). Maize shoot elongation virtually ceased in vermiculite at $-0.81$ MPa, whereas root elongation was slowed by only $50\%$. Leaf elongation rate decreases approximately linearly with decreasing matric potential (Figure 8.9; Parent et al., 2010). In most drying soils, it is likely that a combination of mechanical impedance and water stress are involved in regulating leaf expansion rates, although this is a matter of ongoing debate, particularly because of the difficulty in separating the effects of soil strength and matric potential (Passioura, 2002). Classic experiments by Masle and Passioura (1987) showed that leaf area decreased linearly with increasing penetrometer resistance for wheat plants grown in compacted soil, while there was no evidence that water or nutrients were limiting leaf growth (Figure 8.9). In barley, leaf elongation rate decreased rapidly when the strength of the growing medium (a sand) was increased by applying an external compressive pressure, suggesting a direct role of soil strength (Young et al., 1997). The nature of root-sourced signals and their control of leaf
expansion is likely to involve multiple interactions between ABA and ethylene and to be modulated by factors such as xylem sap pH (Davies et al., 2005).

References


9 Soil water and its management

Anne Verhoef\textsuperscript{1} and Gregorio Egea\textsuperscript{2}

\textsuperscript{1}Department of Geography and Environmental Science, Soil Research Centre, University of Reading, Reading, UK
\textsuperscript{2}Area of Agro-forestry Engineering, School of Agricultural Engineering, University of Seville, Seville, Spain

9.1 Soil water management – a general overview

A plant requires water for photosynthesis, for tissue rigidity (turgidity) and to produce carbohydrates (Jones, 1992; Roberts, 2005). The leaves draw up water from the soil, via the roots, through the plant’s xylem. The driving force behind this process is transpiration: the loss of water from microscopic openings in the leaves (stomata). The main purpose of stomatal opening is to exchange CO\textsubscript{2} and O\textsubscript{2} with the atmosphere; CO\textsubscript{2} is needed to build carbohydrates, allowing plants to grow, while O\textsubscript{2} is expelled during photosynthesis and taken up during plant respiration (during which CO\textsubscript{2} is disposed of), a process that serves to maintain existing tissue. The water obtained from the soil also contains vital mineral nutrients which plants require for biochemical processes.

Transpiration is, in many ways, an undesirable trade-off of canopy gas exchange; it entails a significant loss of water, much more than is needed to maintain turgidity and for photosynthesis. A plant therefore needs a continuous supply of water, with adequate nutrients dissolved within it, to maintain healthy functioning. In areas where the vegetation’s water demand is frequently larger than the amount of water available in the soil, plants will suffer. This will result in reduced growth or even plant mortality. For natural vegetation, these periods of drought define the distinctive plant communities that have adapted to and can survive under these dry conditions.

When considering agricultural and horticultural field crops, supplementing water stored in soil via irrigation may be required to meet plant water requirements (see Sections 9.2.4 and 9.2.5), or water harvesting (Zhao et al., 2009) may need to be practised. In some cases (generally in regions that have an arid climate), virtually all water is supplied via irrigation, often from deep subsurface stores, because surface water stores will be rapidly depleted or are not available.

The irrigating farmer or horticulturalist ideally provides their crop with enough water to grow successfully, with the aim to optimise water use efficiency, WUE (see Section 9.3.2), while at the same time avoiding waste or creation of excess water around the plant roots. The latter will hamper gas exchange between roots and soil and between soil and atmosphere. Furthermore, it can cause toxic effects such as nitric oxide being formed during hypoxia of root tissues. Highly water-saturated soils cause oxygen starvation of plant roots (and soil microorganisms) due to the
fact that gases diffuse much more slowly through water than through air. Hence, removal of CO₂ and replenishment of O₂ are slowed down or cease entirely (see Visser et al., 2003).

When excess water interferes with respiration for prolonged periods, thereby hampering crop growth, removal of water is required. This can be through active removal of soil water, by installing artificial drainage systems or by enhancing the natural drainage capability of the land, using land-management practices.

Often, irrigation and drainage systems are installed together, in particular, when flushing of the soils is needed to remove excess salts in (semi-) arid regions (see Chapter 7). Some crops are purposely grown under ponded conditions, for example rice, so here removal of ‘excess’ water would not be desirable, apart from during the harvesting period.

Excess water can be the result of over-irrigation (whether accidentally or on purpose, i.e. to remove salts); it can also occur naturally (although is often exacerbated through adverse human intervention). It is caused by a combination of climate, soil, vegetation, hydrological and topographical factors (see Section 9.6.1).

The practices of irrigation and the improvement of natural drainage, as well as the conservation of soil water, through suppressing evaporative and drainage losses or through practices such as runoff control, are collectively referred to as soil water management. Water conservation can also be practised by introducing drought-resistance species or varieties that have a high WUE or via tillage or soil amendments.

Agricultural soil water management is crucial in the context of meeting food demand and improving food security (see Rosegrant et al., 2009 and Chapter 1). Experiences in the past have shown that soil water management needs to take place in an integrated and sustainable way. Soil water management forms an important part of ‘integrated water resources management’ (IWRM), which can be defined as a systematic process for the sustainable development, allocation and monitoring of water resource use in the context of social, economic and environmental objectives.

This chapter will focus on the biophysical aspects of soil water management only.

This chapter will address the issues related to soil water management by starting with a detailed description of the soil water balance. It continues with the definition of key soil physical concepts, properties and variables related to water in soil. The chapter also addresses how these variables can be measured or estimated and the ways in which they play a role in soil water management. This naturally requires a better understanding about the interaction between soil water and plants, so some important physiological concepts are briefly discussed as well as the physiological background to water shortage/excess water and indicators of water stress. Finally, there is a discussion of the two major branches of soil water management: irrigation and drainage, only briefly addressing soil water conservation techniques. This chapter should not be viewed as a summary of key irrigation and drainage handbooks. Rather it discusses the overarching concepts and refers to major reference works in the literature, as well as providing a brief summary of tools (e.g. software packages) available to aid the soil water management process.

9.2 Key concepts related to water in soil

9.2.1 Soil water balance

A key equation when discussing shortage or excess of water in soils is the water balance for the root zone, which is given by (see Hillel, 1998):

\[ \Delta \Theta = P + I + C - E - T - D - R \]  (9.1)
where $\Delta \Theta$ is the change in soil moisture content during a given period, $P$ is the precipitation, $I$ is the amount of irrigation, $C$ is the capillary rise, $E$ is the soil evaporation, $T$ is the transpiration, $D$ is the deep drainage and $R$ is the (surface and subsurface) runoff or runon (Figure 9.1). $P$, $I$ and $C$ are gain terms and $E$, $T$ and $D$ are loss terms, whereas $R$ can be either a gain (negative $R$, i.e. runon) or a loss (i.e. runoff). In Equation 9.1 the terms generally have units of depth of water per unit of time, typically mm H$_2$O day$^{-1}$. A soil profile will gain water when the combined gain terms are larger than the loss terms and vice versa. Hence, $\Delta \Theta$ will generally exhibit strong seasonal and interannual variations.

Equation 9.1 describes the soil water store, one of the stores of the terrestrial part of the hydrological cycle; other land stores are ice, groundwater and surface water (rivers, lakes and swamps). The soil water store is relatively small compared to, for example, the ice and groundwater stores, but it is this water that is used by the plants (including our crops) to sustain their growth, so it is of vital importance to mankind. Water is transferred between the land and the atmosphere via transpiration and evaporation. It also flows from the land to the oceans via rivers (Figure 9.1; Farr and Henderson, 1986; Price, 1996).

$P$ (and/or $I$) that is intercepted by the vegetation and does not drip off to reach the ground (throughfall), or reach the ground via the stems or tree-trunks (stemflow), can be evaporated back into the atmosphere; this is called interception loss, which can be significant for forests but is generally less so for crops (Herbst et al., 2008; Siles et al., 2010). The remainder of the water falling from the sky will directly reach the soil surface, where it can then infiltrate, soil conditions permitting.

The amount of water per unit area per unit time that enters the soil is called the infiltration rate and its value is determined by the precipitation/irrigation rate (the supply rate), the
initial soil moisture conditions and the water permeability of the soil surface layer (the hydraulic conductivity, see Section 9.2.5). The initial wetness and hydraulic conductivity determine the infiltration capacity, sometimes called the infiltrability. The soil infiltrability varies; this means that at certain times a soil can cope with larger supplies of water than at other times. Starting with a relatively dry soil profile, there will be a large gradient in wetness (and hence in soil water suction, see Section 9.2.2) between the thin surface layer where water is entering the soil and the underlying drier soil. This large driving force (see Section 9.2.5) will cause infiltrability to be relatively large initially. As the soil profile wets up, the suction gradient, and hence the flow rate, will decrease, although at the same time the hydraulic conductivity will increase, which will counteract the decrease in suction gradient to some extent. When the topsoil reaches saturation, infiltration rate will equal the saturated hydraulic conductivity, \( K_s \). Although the vertical suction gradient is now zero, gravity will still be acting on this water, which will ‘push’ the water in at a steady rate (see Section 9.2.5).

At this time, the soil has reached its final infiltration capacity or steady-state infiltrability. This means that from now on the soil can only accept precipitation or irrigation rates \( \leq K_s \). Water being supplied at rates higher than \( K_s \) will pond immediately when the surface is flat or will run off if the soil surface is sloping. Organic matter, slope, soil texture and antecedent soil moisture are important factors influencing infiltration capacity (see, e.g. Wilcock and Essery, 1984).

However, before runoff occurs, excess water will first collect in surface depressions, whose total volume is called the surface storage capacity. The rate of runoff will depend on the excess of rain/irrigation intensity over the infiltration rate, as well as on the slope and the roughness of the soil (Hillel, 1971). Runoff is undesirable from a water management point of view as it negatively affects the amount of water available in the rooting zone for plant growth. Furthermore, runoff can cause erosion, which can result in loss of fertile topsoil (Lal, 2003). Runoff is often called surface water flow, to distinguish it from below-ground lateral flow. Below-ground lateral flow is often split into groundwater, which occurs below the water table, and interflow, where water that has accumulated within the topsoil layers starts to move laterally (this will only happen when the surface is sloping). When this interflow reaches the surface, it may become ponded or surface water (Figure 9.1; Farr and Henderson, 1986).

The infiltration process results in downward water movement in the soil, causing vertical and temporal changes of soil water content in the root zone, which will influence the water available to plants (see Section 9.2.4). The surface layers, initially having gained water, will now lose this water to deeper layers. In the absence of groundwater, or when dealing with a deep groundwater table, this process is called redistribution. With a groundwater table present, this water movement is often called internal drainage (Hillel, 1971). The duration of this post-infiltration movement (whether redistribution or internal drainage) strongly depends on the soil type.

Water eventually flowing out of the root zone, and therefore out of reach of the roots, is frequently referred to as deep drainage, \( D \), or seepage. This water passes on to deeper substrate layers and eventually aquifers (e.g. sediments or rock formations), generally well beyond the reach of the plant roots. Drainage continues until the water reaches a vertical level below which all voids or pores are water-filled. The upper surface of this completely saturated zone is called the groundwater table (GWT; Farr and Henderson, 1986). When no, or little, lateral drainage is present (whether natural or through artificial drainage systems), continuing vertical inflow from above will fill the available pore spaces, causing
the groundwater table to rise progressively through the subsoil, then the topsoil, until it reaches the surface, when ponding occurs. The depth of the GWT will vary with time due to its interaction with the water balance terms (Equation 9.1 and Figure 9.1).

If an impermeable layer (e.g. a hardpan) is present within the subsoil, this layer will prevent internal drainage resulting in saturated soil conditions and eventual ponding of water on the surface, even though the horizons below the impermeable layer are not saturated. This is referred to as a perched water table.

Finally, water can also travel upwards from the fully saturated zone to the root zone, via a process called capillary rise, \( C \) in Equation 9.1. This is the result of the fact that soil pores act as capillaries that can transport water upwards, as will be explained in Section 9.2.2. Capillary rise can contribute significantly to root zone water content in fine-textured soils and will allow water lost via evaporation during the day to be replenished during the night.

Having discussed how water enters and drains through the profile and how the root profile interacts with groundwater or deep soil water stores (losing to or gaining water from it), the focus switches to the evaporative process, which returns water to the atmosphere. The process of evaporation involves the change of liquid water into water vapour, which requires energy (resulting from the net radiation reaching the land surface, see Section 9.4.3) to break the bonds of the water molecules (so-called latent heat of vaporisation, \( \lambda: \approx 2.45 \times 10^6 \text{ J kg}^{-1} \text{ water at 20°C} \)); this water vapour is then transferred to the atmosphere, either via soil pores (soil evaporation, \( E \)), stomata (transpiration, \( T \)) or via interception loss. The combined processes of soil evaporation and transpiration are often grouped together and termed evapotranspiration, although an increasing number of authors prefer the general term evaporation instead, because the physical processes behind both evaporation and transpiration are identical. During the evaporation process, water vapour will also have to negotiate a thin boundary layer close to the leaf and soil surfaces where air-flow is close to being laminar, and the main transfer process is diffusion, after which it can be transported, via turbulent transfer, to higher atmospheric layers.

The continuous mixing of the air, a result of the turbulent nature of air flow near the land surface, ensures a gradient of water vapour pressure between the evaporating surface and the near-surface atmosphere, where the atmospheric vapour pressure is lower than that at the surface. The combined effects of supply of energy and continual removal of water vapour (to maintain the water vapour pressure gradient) together determine the evaporative demand of the atmosphere. The evaporative demand, sometimes called evaporativity, therefore depends on the available energy, the relative humidity, wind speed and temperature. Together these variables determine the maximum possible rate of evapotranspiration from the evaporating surface. This maximum rate is often called the potential evaporation, \( E_p \) (this generally includes \( E \) and \( T \)). Evaporation can only occur at its potential rate if water is continually supplied from below to the site of evaporation, that being the soil pores or the stomata. If this supply falls below the evaporative demand, evaporation will occur at a rate below the potential rate: the actual evaporation, \( E_a \). Hence, \( E_a \leq E_p \). There are cases when \( E_a > E_p \), this occurs when additional energy is supplied by advection, for example, when well-irrigated cropland is situated next to fallow dryland.

Soil evaporation is an important soil water loss term when the soil surface is bare. Under annual field crops, this may occur during fallow periods and throughout periods when soil tillage, planting, sowing and germination take place. Soil evaporation also plays an important role during the initial crop growth stages, when the canopy only covers a small percentage of the field surface area. It can also be considerable for sparse canopies, e.g. orchards, where a lot of exposed bare soil exists below and between the individual trees. Covering the soil with
a mulch layer, or tilling the soil in such a way as to decrease evaporation (see Section 9.5.5),
will avoid unnecessary water loss and improve crop water use efficiency (see Section 9.3.2).

When considering soil evaporation two cases can be distinguished: steady-state evaporation
in the presence of a relatively shallow GWT and evaporation in the absence of a GWT.
The latter will be a transient-state process, resulting in a net loss of soil water, thereby
drying the soil profile. The first process may cause cumulative salinisation near the surface,
a problem often encountered in (semi-) arid regions (see Section 9.5.4 and Chapter 7).
The steady rate evaporation depends on depth of the water table and the dryness of the soil
at the surface as well as on the hydraulic properties of the soil (Hillel, 1971).

Soil drying (soil evaporation in the absence of a GWT) generally occurs in two stages.
During the first stage, the energy-limited stage, soil evaporation is entirely determined by
the evaporative demand; it assumes a saturated soil surface, $E_a = E_p$. During the second stage,
the water limited stage, $E_a < E_p$ and $E_a$ continues to decrease, until eventually $E_a \sim 0$. This
point will be reached relatively rapidly in coarse-textured soils, whereas soil evaporation in
medium- to fine-textured soil will persist much longer. The reasons for this will be elabo-
rated upon in Sections 9.2.2 and 9.2.5.

Transpiration is intimately linked to the process of photosynthesis and the extent of
stomatal opening, which in turn partly depends on how much water is available for plants in
the root zone; this process is discussed in Sections 9.2.4 and 9.3.2.

9.2.2 Water retention in soils

Not all of the soil water that has collected in the soil pores of the root zone is available for
transpiration. To explain the reasons behind this, the ways in which soil water is retained in
pores need to be understood. For plant roots to take up water from soil they need to over-
come the force with which the soil matrix holds on to the water. These so-called matric
forces do not just hold water against root water uptake, but also against evaporation and
drainage; this is called water retention.

There are three mechanisms for binding of water to the solid matrix (Koorevaar et al.,
1983): direct adhesion of water molecules to solid surfaces by London-van der Waals forces;
capillary binding of water, due to a combination of adhesive and cohesive forces; and
osmotic binding of water in diffuse electric double layers. Here, adhesion is the attraction
between water molecules and the solid soil particles. Water molecules also experience
strong intermolecular attractive forces with neighbouring water molecules; this is called
cohesion. At the water surface there is an enhancement of the intermolecular attractive
forces, which is called the surface tension, $\sigma$ (which has a value 0.0728 Nm$^{-1}$ at 20°C).
The combination of these adhesive forces, and cohesive forces between water molecules,
cause capillary binding in soils; the most important mechanism of water retention in coarse
to medium-textured soils. Only in fine-textured soils may osmotic binding exceed capillary
binding (Koorevaar et al., 1983).

The concept of capillary binding, caused by capillary action, deserves further explana-
tion, because pores in soils may be viewed as irregularly shaped non-cylindrical capillaries,
with radius $r$ (m). In soil, pores exist in different sizes, hence we are dealing with a ‘bunch’
of capillaries with a range of radii. The pore-size distribution (see Chapter 8), largely deter-
mined by texture and soil structure, therefore plays an important role in soil water retention.
Capillary action occurs when the adhesion of water to the walls of a capillary is stronger
than the cohesive forces between the water molecules. The height, $h_w$ (in m), to which capili-
ary action will take water (capillary rise) in a uniform circular tube with radius $r$ is given by
Soil water and its management

\[ h_w = \frac{2\sigma}{\rho_w g} \quad (9.2) \]

(see, e.g. Koorevaar et al., 1983, for a derivation of this formula) where \( g \) is the acceleration of gravity (N kg\(^{-1}\)) and \( \rho_w \) is the density of water (kg m\(^{-3}\)).

This equation determines the extent of the capillary fringe, i.e. the thickness of the soil layer above the groundwater table that remains largely saturated (very shallow for sandy soils and up to about 0.4 m thick for clay soils). This capillary fringe may lead to a root zone that experiences excess water, despite the fact that the water table is below the root zone. However, the occurrence of capillary rise also means that the soil will be able to transport water from deeper saturated layers towards the root zone, resulting in replenishment (C in Equation 9.1). Soil water is also transported as a result of gradients in soil matric potential; this will be discussed in Section 9.2.5.

The theory of capillary binding also determines when a soil pore of radius \( r \) (m) will empty when a root exerts a certain suction to overcome the soil water suction, \( S \) (in Pa, i.e. Nm\(^{-2}\)). This relationship is given by:

\[ S = \frac{2\sigma}{r} \quad (9.3) \]

Hence, the narrower the pore, the larger the root suction required to empty it, i.e. to enable air to enter the pore. Soil water suction, \( S \), is the amount by which the pressure in the soil water, \( P_w \), is less than atmospheric pressure, \( P_a \) (atmospheric pressure has been given a reference value of \( P_a = S = 0 \)); it is related to how tightly the water is held in the soil, as discussed in more detail in Section 9.2.3.

### 9.2.3 Key soil moisture contents and water retention curves

In a saturated soil, i.e. a soil where all pores are filled with water and moisture content, \( \theta \), equals \( \theta_s \), the saturated water content. This is the maximum soil moisture content possible in a soil, although in practice some trapped air will be present causing \( \theta \) to be slightly less than the total porosity. The suction required at which the first (i.e. the biggest) pores start to release their water, and hence air enters into these pores, is the so-called air-entry value. Once the largest pores have emptied, soil water will remain in progressively smaller pores and it will become harder for roots to extract water from these ever smaller water-filled pores, as they need to overcome larger and larger soil water suctions (see Equation 9.3). Water will also be held in the necks between larger air-filled pores and as thin films held on to the particle surfaces enclosing air-filled pores.

Uptake of soil water by plant roots becomes more difficult as the soil water suction increases (i.e. as the matric potential, \( \psi_m \), decreases, see Section 9.2.5). Most plant roots are generally unable to take up any more water when \( S \) has increased to about 1500 kPa or \(~150\) m, equivalent to a pore diameter of \(~2\) \(\mu\)m. Hence, this particular suction value, and its related soil moisture content, is called permanent wilting point, PWP, often defined as the point at which plants wilt but do not recover overnight. There are several factors that could affect PWP, including; crop species, soil type and crop stage of growth. Nevertheless, PWP is generally assumed to occur at \( S = 1500 \) kPa.
Another salient soil water suction-moisture content point in relation to root water uptake is that of field capacity, FC, although its usefulness, physical basis and definition are often disputed (see Hillel, 1971; Nachabe, 1998). Generally, this is defined as the water content in a soil profile at which internal drainage or redistribution has nearly stopped (is ‘very slow’); often 2–3 days after the soil was thoroughly wetted by heavy rain or irrigation. At this stage, all transmission pores (pores with diameter $>$ 50–100 μm) have been emptied. The suction at field capacity is often defined as $S = |\psi_m| = 3.3$ m (33 kPa), although 10 kPa is another widely used value to denote FC. Also, some scientists prefer to define it by using a hydraulic conductivity (e.g. the soil water suction when hydraulic conductivity $K = 0.05$ mm day$^{-1}$).

The relationship between soil water content ($\theta$) and soil water suction (or matric potential, $\psi_m$) is known as the water release curve, water release characteristic or water retention curve. The shape of the water retention curve is largely determined by a soil’s textural type, because of the predominance of different mechanisms causing binding of water to the solid matrix. Capillary binding of water can only take place up to suctions of 10 m. Suctions larger than 10 m are caused exclusively by adhesive and osmotic binding. Hence, the pore size distribution, as affected by soil texture and structure, will determine the unique shape of the water retention curve for each soil (see Figure 9.2). The different distribution of pore sizes in soils causes sandy soils to release most of their water at relatively low suctions, while clay soils lose very little water until field capacity and still have a lot of water left at PWP.

The determination of water retention curves is briefly discussed in Section 9.4.2. While the curve is a unique property of each soil, it demonstrates hysteresis (Koorevaar et al., 1983). This means that the water retention curve derived from progressive extraction of soil water starting from saturation (desorption) is different from that obtained by gradual addition of water to air-dry soil (absorption). Typically, soil water characteristics are obtained through desorption. Hysteresis can be caused by the irregularity of soil pores’ sizes and shapes, the entrapment of air during wetting, the lower contact angle of the meniscus during wetting than drying and different arrangements of clay particles during wetting and drying (Koorevaar et al., 1983).

**Figure 9.2** Typical water retention curves for a number of soils. Reprinted from Hugget et al. (2004). With permission from Hodder Arnold.
9.2.4 Available soil water

The amount of soil water available for plant growth, often called (plant) available soil water or available water capacity, is generally expressed as the difference between \( \theta \) at FC and \( \theta \) at PWP: \( \theta_{\text{FC}} - \theta_{\text{PWP}} \); this concerns water in the storage pores, with approximate diameters between 2–50\( \mu \text{m} \).

Table 9.1, derived from Rowell (1994), shows typical values of \( \theta_{\text{FC}} \) and \( \theta_{\text{PWP}} \) as well as available soil water, for ten different soil textural classes. Available soil water ranges between 0.10 for the sandy soil and 0.22 for loamy soils, although these minimum and maximum values are indicative only. Soils with the largest available water capacities are medium-textured. Although clay soils contain a lot of remaining pore water after initial drainage (large \( \theta_{\text{FC}} \)), the considerable number of pores of small diameter (<2\( \mu \text{m} \)) results in a large amount of soil moisture remaining once permanent wilting point is reached; this is not available to plant roots. These soils can only be dried out further through the process of evaporation.

Sandy soils mainly have pores in the transmission category (pore size >50\( \mu \text{m} \)) and hence will have lost most of their water already between saturation and field capacity. At PWP very little water will be left in the soil.

When used in crop water management (for irrigation purposes), the quantity \( \theta_{\text{FC}} - \theta_{\text{PWP}} \) is often multiplied by the thickness of the root zone and then referred to as total available water, \( TAW \) (see Allen et al., 1998). Next, to indicate that soil water availability actually decreases with decreasing soil moisture content (e.g. due to a reduced rate of water transfer from the bulk soil to the root zone, largely as a result of decreased hydraulic conductivity values, see Section 9.2.5), it is sometimes split into readily available water (\( RAW \)) and a decreasingly available range, the latter starting when \( \theta \) has reached a threshold value, \( \theta_t \). In crop models these concepts are often incorporated via a water stress coefficient, \( \beta \), (called \( K_j \) in Allen et al., 1998) a dimensionless transpiration reduction factor (ranging from 0 to 1):

\[
\beta = \frac{TAW - D_r}{TAW - RAW} = \frac{TAW - D_r}{(1 - p)TAW}
\]

(9.4)

where \( D_r \) is the root zone depletion (total water content in the root zone relative to field capacity; \( D_r = 0 \) at \( \theta = \theta_{\text{FC}} \) ) and \( p \) is the fraction of \( TAW \) that a crop can extract from the root zone, without suffering water stress, i.e. before crop evapotranspiration gets reduced. The factor \( p \) differs between crops. Equation 9.4 indicates a linear decrease in \( \beta \) starting at \( \theta_t \), although some authors assume a curvi-linear decline. Similar to \( \beta \), Sinclair (2005) and Metselaar and Sinclair (2007), introduced the concept of a relative daily transpiration rate, \( RT \), given by:

\[
RT = 1 - \psi_{m,s} \left( \theta / \theta_s \right)^{-b} / \psi_l
\]

(9.5)

Here, \( \theta \) and \( \theta_s \) are as defined before and \( \psi_{m,s} \) is the soil matric potential at saturation (or rather, at air-entry point). Parameters \( b \) and \( \psi_{m,s} \) are texture-dependent parameters for which standard values can be found (based on Clapp and Hornberger, 1978). Parameter \( \psi_l \) is the hydrostatic pressure in leaves during periods of active transpiration. \( RT \) is defined as the water flux of a plant growing on drying soil relative to a plant growing on well-watered soil, and it is calculated between the upper, \( \theta_u \), and lower, \( \theta_l \), limits of soil water available to support plant transpiration. \( RT \) is a function of the fraction of transpirable soil water, \( FTSW: (\theta - \theta)/(\theta_s - \theta) \). Sinclair (2005) defined this lower limit as \( \theta \) were \( RT < 0.1 \); it was
Soil Conditions and Plant Growth

Soil moisture contents (θ) at field capacity (FC) and permanent wilting point (PWP) for ten typical soils ranging in texture from clay to sand.

<table>
<thead>
<tr>
<th>Texture</th>
<th>PWP</th>
<th>FC</th>
<th>FC-PWP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clay</td>
<td>0.28</td>
<td>0.44</td>
<td>0.16</td>
</tr>
<tr>
<td>Silty clay</td>
<td>0.28</td>
<td>0.44</td>
<td>0.16</td>
</tr>
<tr>
<td>Clay loam</td>
<td>0.23</td>
<td>0.44</td>
<td>0.21</td>
</tr>
<tr>
<td>Silty clay loam</td>
<td>0.20</td>
<td>0.42</td>
<td>0.22</td>
</tr>
<tr>
<td>Sandy clay loam</td>
<td>0.16</td>
<td>0.36</td>
<td>0.20</td>
</tr>
<tr>
<td>Loam</td>
<td>0.14</td>
<td>0.36</td>
<td>0.22</td>
</tr>
<tr>
<td>Silt loam</td>
<td>0.14</td>
<td>0.36</td>
<td>0.22</td>
</tr>
<tr>
<td>Sandy loam</td>
<td>0.08</td>
<td>0.22</td>
<td>0.14</td>
</tr>
<tr>
<td>Loamy sand</td>
<td>0.06</td>
<td>0.18</td>
<td>0.12</td>
</tr>
<tr>
<td>Sand</td>
<td>0.05</td>
<td>0.15</td>
<td>0.10</td>
</tr>
</tbody>
</table>

The last column shows the plant available soil water.

Source: Data taken from Rowell (1994).

Equation 9.5 is an approximation of Sinclair’s original formula, which involved depth of root water extraction, a variable for the geometry of soil water extraction around roots, the hydraulic conductivity and the daily transpiration rate on well-watered soil. It appeared (Sinclair, 2005), from sensitivity analyses, that all of these factors played a relatively minor role and that $RT$ was largely determined by soil texture and $\psi_l$, as summarised in Equation 9.5.

Figure 9.3 The relative transpiration rate as measured for well-watered and water-stressed dwarf sunflowers, using load-cell lysimeters (see Section 9.4.3). Also indicated is the result of Equation 9.5 with $\psi_l = −1.2$ MPa. The good fit between Equation 9.5 and the measurements illustrate the validity of Sinclair’s (2005) theory, at least for this particular experiment.

found to range between 0.070 and 0.289, for sandy soil and clay soil, respectively, with the soil types as defined by Clapp and Hornberger (1978); see also Section 9.2.5. The upper limit was found to be 0.13 greater than the lower limit, for most soils, but these values were 0.08 and 0.15, respectively, for sand and silt soils. In this context, FC and PWP are not defined, although the lower and upper limits calculated in this way are very similar to those given in Table 9.1.
Figure 9.3 gives an example of the relative transpiration rate as measured for well-watered and water-stressed dwarf sunflowers. The experimental set-up involved mini-lysimeters (see Section 9.4.3 and Diaz-Espejo et al. (2005)).

9.2.5 Water flow in soils

As previously mentioned, apart from water retention, the rate of water flow in soils is also important in the context of crop water uptake. When roots extract soil water for plant growth and maintenance, it needs to be replenished again, to sustain this continual process and to avoid (permanent) wilting. Furthermore, water flow in soils is important when discussing infiltration and drainage rates. In considering horizontal water movement, soil water will move in the direction of increasing suction, i.e. from places where it is less tightly held by the soil matrix to where it is more strongly retained. However, during vertical redistribution of water, both gravity and suction difference will affect movement and it will depend on the vertical suction profile and vertical location within the soil profile whether water will move and in what direction it will move. The effects of gravity and soil water suction are often discussed in terms of potential energy; water will flow from where its potential energy is higher to where its potential energy is lower. It is assumed that high elevation and low suction equate to high potential energy.

In this framework of potential energy, the major factors influencing movement are the gravitational potential, \( \psi_g \), and the suction potential, more often called the matric potential, \( \psi_m \). The sum of these two is called the hydraulic potential:

\[
\psi_h = \psi_m + \psi_g \tag{9.6}
\]

Here \( \psi_h \) is a measure of the overall potential energy status of the water. Soil water moves to decreasing values of \( \psi_h \) (more negative values). The potentials in Equation 9.6 can be expressed in different units: J kg\(^{-1}\), J m\(^{-3}\) (= Pa) or in m, i.e. expressed as an equivalent head of water (see e.g. Koorevaar et al., 1983; Rowell, 1994).

The potential calculations require a reference point, for both potentials. For the gravitational potential, this is usually the soil surface (\( \psi_g = 0 \) at \( z = 0 \)). The rule is then, when working with potential expressed as an equivalent head of water, which is often the easiest, that \( \psi_g = z \) (with \( z \), the depth in the soil, taken as negative). The reference for \( \psi_m \) is free water (an open dish of water under atmospheric pressure). This means that \( \psi_m = 0 \) at the water table or when the soil is saturated. As soon as \( \theta < \theta_s \), \( \psi_m \) becomes negative. **In situ** values of \( \psi_m \) can be determined using tensiometers (see Section 9.4.2) or derived from soil moisture measurement, assuming the water retention curve is known. Note that below the groundwater table, soil water pressure becomes positive and hence \( \psi_m \) (now referred to as \( \psi_p \), the pressure potential, because the matric forces play no role anymore) increases with depth below the GWT, as a result of the increasing weight of the water column above the point under consideration.

Plotting hydraulic potential, \( \psi_h \), against soil depth is helpful to determine which parts of the soil profile experience upward flow (as a result of evaporation and transpiration) or downward flow (drainage); this is called a potential diagram. It is the gradient of \( \psi_h \) with depth, \( (\Delta \psi_h/\Delta z) \), that will determine the rate of soil water flow, \( F_w \) (ms\(^{-1}\)) together with the hydraulic conductivity, \( K(\theta) \) (ms\(^{-1}\)), as a function of \( \theta \); this is given by Darcy’s law:

\[
F_w = -K(\theta) \frac{\Delta \psi_h}{\Delta z} \tag{9.7}
\]
When the soil is saturated, hydraulic conductivity, \( K \), is at its largest: the saturated hydraulic conductivity, \( K_s \). As soon as \( q \) decreases, \( K(q) \) will decrease, for the following reasons. Firstly, as a soil dries there is much less cross-sectional area of water-filled pores through which water can flow. Secondly, water now has to take a more convoluted path; this drier soil is said to have a higher tortuosity, \( \tau \). Thirdly, the size of the water-filled pores plays an important role because the volume flow rate of water (the discharge rate, \( Q \), in \( \text{m}^3 \text{s}^{-1} \)) is more rapid through a large pore; Poiseuille’s law dictates that \( Q \) is proportional to \( r^4 \). Finally, the internal friction in the water, as characterised by the viscosity, \( \eta \), increases considerably with decreasing temperature, although this effect is often ignored in soil physics. Taking all these factors into account, \( K(q) \) can be approximated by (Koo revaar et al., 1983):

\[
K(q) = \frac{1}{8\eta\tau} \sum (\Delta \theta)_i r_i^2
\]

where \( i \) represents the \( i \)th pore size class. Hence, the hydraulic conductivity of a porous medium can be estimated from the size-distribution of the water-filled pores. The pore size distribution can be estimated from the soil water characteristic, if we consider the volumes of water (i.e. \( \theta \)) released for increments of suction are measured (the latter can be directly related to \( r \), see Rowell, 1994). This method works best in sandy soils, where soil water is largely present in capillary pores and the water retention curve exhibits a clear ‘chair-shape’ (see Figure 9.2). For medium- to fine-textured soils this approach works less well. Generally, \( K \) is either measured (see Section 9.4.4) or calculated using mathematical expressions. Equation 9.8 explains why \( K(q) \) decreases so sharply with drying and swelling/shrinking. It also explains why \( K_s \) and \( K(q) \) vary so much between soil types.

One phenomenon worth mentioning here is preferential flow. Gerke (2006) states that ‘preferential flow comprises all phenomena where water and solutes move along certain pathways, while bypassing other volume fractions of the porous soil matrix’. Coppola et al. (2009) distinguish four types of preferential flow: (1) preferential flow in non-capillary pores (macropores); (2) preferential flow in interaggregate pores; (3) ‘fingering’ at the interface of a less permeable layer above a more permeable one and (4) preferential flow due to spatial irregularities or temporal dynamics in soil water repellency.

Preferential flow may cause rain or irrigation water to bypass large regions of the root profile, so that water is lost from it via deep drainage, and not used to decrease the soil water deficit. However, some types of preferential flow (e.g. of type 4) may cause wet bulbs to occur in the root-zone, which in fact improve water availability. It is thought that some trees (e.g. needle-leaf trees or shrubs in dry climates) release water-repellent resin below their crown to induce this wet bulb (see Robinson et al., 2010).

The soil water characteristic and \( K(q) \) functions define a soil’s hydraulic properties. When water retention data are not available to calculate \( \theta_{FC} - \theta_{PWP} \) the equations proposed by Clapp and Hornberger (1978) and Cosby et al. (1984) can be used to calculate \( \psi_m(\theta) \) and \( K(\theta) \). Parameters required are \( \psi_m \), the matric potential at saturation (m), parameter \( b \) (unit-less) denoting the slope of the water retention curve, \( K_s \) and \( \theta_s \). By setting \( \psi_m \) to FC (generally taken to be 3.3 m) and PWP (150 m) values, the corresponding \( \theta \)-values (and hence available water) can be calculated if \( b \) and \( \theta_s \) are known. Standard parameter values can be taken from the literature (Clapp and Hornberger, 1978; Cosby et al., 1984) or values can be calculated from pedotransfer functions (e.g. Cosby et al., 1984). In this case the parameters are calculated as a function of percentage sand, silt and clay content, using equations based on a large numbers of soil samples. Although the Clapp and Hornberger equations are widely used, an alternative parameterisation, the Van Genuchten equations (Van Genuchten, 1980), generally
provides a better fit to the water retention and hydraulic conductivity data as a result of their mathematical form and the fact that they have three extra tunable parameters (residual soil moisture content, \( \theta_r \), and fitting parameters \( l \) and \( m \)). Typical parameter sets for the Van Genuchten approach, and related pedotransfer functions, can be found in Schaap and Leij (1998) and Wösten et al. (1999), for example.

### 9.3 Soil water and plants

#### 9.3.1 Introduction to soil–plant water relations

Water and air both reside in the soil pores and their relative proportion is crucially important for healthy plant growth. The plant root system provides the aerial parts (i.e. stem and leaves) with water and dissolved minerals. Under optimum soil moisture conditions, plant roots will extract soil water at the potential rate to meet the evaporative demand of the atmosphere. Evaporation of water from vegetation takes place through the leaves’ stomata, in a process known as transpiration, \( T \), which can be considered as a necessary ‘cost’ associated with the inward diffusion of \( \text{CO}_2 \) gas from the air and the process of photosynthesis or \( \text{CO}_2 \) assimilation. The extent of opening of the stomata is called stomatal conductance, \( g_s \).

A limited water supply in the soil is sensed by roots and communicated to the shoot to mediate (partial) stomatal closure (causing a decrease in \( g_s \)) and permit timely adjustment of the plant’s water losses by transpiration. The inter-dependence between photosynthesis and \( T \) means that as \( g_s \) decreases there is a decrease in biomass production, although this adaptation to water stress greatly depends on how efficiently plants use water under limited soil water availabilities. This plant behavioural trait is termed water use efficiency (WUE) and a detailed description of this concept is provided in Section 9.3.2. Due to the close link between photosynthesis and \( T \), a drop in \( T \) will cause a decrease in growth. Consequently, there is the need for irrigation to safeguard crop productivity and farmer’s profitability. Irrigation represents a key feature in the optimization process of on-farm water resources in water-limited areas.

Although there are still some gaps in our knowledge regarding stomatal regulation (Lawson, 2009), it is now widely accepted that stomatal aperture is regulated by both internal physiological factors (e.g. leaf water status, plant hormones such as abscisic acid, ABA) and external environmental factors, such as light (in the photosynthetically active wave-band), air \( \text{CO}_2 \) concentration and air humidity (Farquhar and Sharkey, 1982; Buckley, 2005). The physiological mechanisms controlling the response of stomata to environmental stresses, such as reduced soil water availability, are complex and still not fully understood which led to \( g_s \) initially being calculated from multiplicative factors depending on environmental variables (Jarvis, 1976; Stewart, 1988). However, more recently, various authors have attempted to model stomatal behaviour in a more physiologically explicit fashion (Dewar, 2002; Gao et al., 2002; Buckley et al., 2003).

Many plant physiological processes (e.g. cell expansion and leaf photosynthesis) are significantly affected by leaf water status. As with soils, leaf water status is commonly expressed as leaf water potential, the thermodynamic expression of water status of leaves. Leaf water potential at a given time of the day is, therefore, not only the result of soil water status but is also directly linked to transpiration rate. The classical equation of water flux (Van den Honert, 1948) synthesizes the role that both factors play in leaf water potential (\( \psi_l \)):

\[
\psi_l = \psi_s - (R_s + R_p) J_w
\]

(9.9)
where $\psi_s$ is the soil water potential (negative) given by the sum of $\psi_p$ (Equation 9.6) and the osmotic potential ($\psi_o$). $R_s$ and $R_p$ are the hydraulic resistances from bulk soil to roots, a function of $K(\theta)$, and from roots to shoot, respectively, and $J_w$ is the water flux through the plant (i.e. transpiration rate, later on in this chapter referred to as $T$), which is greatly determined by stomatal regulation. Water transport through the soil–plant–atmosphere continuum requires the osmotic component to be considered. In soils there are no semi-permeable membranes (such as those found in plant tissues) so the osmotic potential does not play a role in water transport and hence hydraulic potential is the driving force. However, in the case of water flow from soils to plants, water needs to pass through semi-permeable membranes and hence the osmotic effect plays a role (this also explains why the transpiration stream is reduced in saline soils).

The major factors influencing the leaf water potential are ion concentration, pressure and gravity. Consequently, $\psi_l$ may be expressed in terms of its individual components, usually written as the following sum:

$$\psi_l = \psi_o + \psi_p + \psi_g$$  \hspace{1cm} (9.10)

where $\psi_o$ (negative), $\psi_p$ (positive) and $\psi_g$ (positive) denote the effects of solutes (osmotic potential), hydrostatic pressure (turgor) and gravity (gravitational potential), respectively, on the free energy of water (Taiz and Zeiger, 2002). The gravitational component is generally omitted because it is negligible compared to the osmotic potential and the hydrostatic pressure of plant tissues. These different components of leaf potential can be used when monitoring plant water stress (see Section 9.3.3).

Soil salinity also plays a major role in canopy gas exchange and plant water relations in general. Dissolved solutes in the rooting zone generate a low (more negative) osmotic potential ($\psi_o$) that lowers the soil water potential. Therefore, the effect of dissolved solutes is similar (as a first approximation) to that of a soil water deficit, and most plants respond to excessive levels of soil salinity in the same way as for water stress. A major difference between low water potential environments caused by salinity and water shortage is that the total amount of soil water may be much higher in a saline soil, albeit it is not available for the plant. Further details on soil salinity and its management are provided in Chapter 7 and Section 9.5.3.

Soil drying may affect plant growth through physiological effects on plants, as described earlier, but also by direct effects on the soil physical environment. As reviewed by Whitmore and Whalley (2009), soil water shortage may affect vegetation in a number of interacting ways, highlighting that ‘soil drought’ is not a single and simple stress (see Chapter 8 for an account of how physical properties affect plant growth).

Excess soil water also adversely affects plant functioning, in this case through anaerobic stress. For detailed background information of the processes involved and plant mechanisms that underlie short- and long-term tolerance to external anaerobic conditions refer to the reviews by Vartapetian and Jackson (1997) and Visser et al. (2003).

### 9.3.2 Concepts of water use efficiency

The term ‘Water Use Efficiency (WUE)’ has traditionally been used as an indicator of plant performance. WUE can be defined in two different ways: from an eco-physiological or agronomical point of view (Gregory, 2004; Katerji et al., 2008).
The eco-physiological approach is based on the analysis, at a given instant, of the relationship between net CO₂ assimilation and transpiration (T), which can be used at the field-, plant- or leaf-scales (Steduto et al., 1997; Chen and Coughenour, 2004; Katerji et al., 2008; Morison et al., 2008).

For agronomists, WUE or ‘crop water productivity’ (Katerji et al., 2008; Geerts and Raes, 2009) is usually a value defined by yield per unit area, Y, divided by water used to produce yield. The denominator may refer to water lost by evapotranspiration (ET) or just T by excluding direct evaporation (E) from the soil and plant surfaces (via evaporation of intercepted rainfall). When this is the case, the expression is then normally termed ‘transpiration efficiency’.

The amount of water used can also be written in terms of all of the loss components in the soil water balance equation (see also Equation 9.1 in Section 9.2.1):

\[
WUE = \left( \frac{Y}{T} \right) \left( \frac{1}{1 + \left( \frac{E + D + R}{T} \right)} \right)
\]

(9.11)

This equation has been normalised with regards to transpiration (Gregory, 2004). Interpretation of Equation 9.11 makes it clear that any enhancement in WUE must come about by exploiting three key processes: (1) agronomic management practices that maximize T by reducing the other losses; (2) acquiring more carbon (biomass) in exchange for the water transpired by the crop (i.e. improving crop transpiration efficiency) and (3) partitioning more of the achieved biomass into the harvested product (Gregory et al., 1997; Condon et al., 2004; Gregory, 2004).

The pressing need to improve WUE of agricultural systems is not only focused on irrigated cropping systems, but also on rainfed systems. The sharp increase in wheat yields and rainfall-use efficiency observed in Australia during the last two decades is considered to be one third attributable to new (more efficient) cultivars and two thirds attributable to the adoption of improved agronomic practices (Angus et al., 2001; Stephens, 2002; Turner, 2004).

Figure 9.4  Schematic representation of the main determinants of WUE. Adapted from Katerji et al. (2008). With permission from Elsevier.
Soil Conditions and Plant Growth

One agronomic practice that has proven effective to reduce evaporation from the soil surface is ‘mulching’ with crop residues, both irrigated and rainfed systems (Gregory, 2004; Deng et al., 2006; see also Section 9.5.4). As Hsiao et al. (2007) noted, there is still much room for improving WUE in irrigated cropping systems, in part owing to the low water storage and conveyance (≈ 70%) and irrigation (≈ 37%) efficiencies globally (Wallace and Gregory, 2002; Gregory, 2004). As reviewed by Katerji et al. (2008), the sources of variability of WUE are numerous and sometimes of a very distinct nature and may be grouped into three categories related to (1) plant, (2) agro-techniques and and (3) environment (see Figure 9.4).

Among the latter, soil texture has proven to play a significant role in determining WUE. Katerji and Mastrorilli (2009) in a long-term experiment, reported that WUE of potato, corn, sunflower and sugar beet was reduced by 22–25% when crops were grown on clay soil as compared to those grown on loam soils. This was largely related to the available water, as explained in Section 9.2.4.

9.3.3 Plant indicators of water stress

Irrigation techniques for which the decision to irrigate is based on measured levels of soil moisture content will be presented in Sections 9.5.1 and 9.5.2. As outlined by Jones (2004), a potential problem with all soil-water-based approaches is that many plant physiological attributes respond directly to changes in plant tissue water status rather than to changes in the bulk soil water content (or water potential). This is because leaf water potential at a given time of the day is not only the result of soil water status but also is directly linked to evaporative demand through its influence on transpiration rate (see Equation 9.8).

Another reported drawback of soil-based irrigation management systems relates to the high soil heterogeneity combined with the limited volume of soil that some soil moisture devices sample (Jones, 2004; Verhoef and Campbell, 2005), causing inaccurate estimates of soil water stress. This fact causes many researchers to consider ‘plant stress sensing’ as the most reliable approach for precise irrigation scheduling.

As reviewed by Jones (2004), plant-based water stress indicators can be grouped into two categories: first, direct measurements of some aspect of plant water status; and second, other plant processes that respond sensitively to water stress. The first group comprises all the available techniques to measure tissue (leaf, stem, etc.) water status (e.g. pressure chamber, or psychrometer). The principles underlying the quantification and measurement of plant water status have been described elsewhere (Boyer, 1995; Nobel, 1999; Kirkham, 2004) and these sources should be consulted for more detail and background theory. The second group comprise gas exchange indicators such as stomatal conductance, which can be directly measured with specialized equipment (e.g. porometers) or indirectly derived through thermal sensing, sap flow measurements (to estimate transpiration rate), micromorphometric variations of plant organs (i.e. leaf, branch, trunk or fruit) or other physiologically based but remotely sensed indicators.

Direct estimation of plant water status is usually done via measurements of tissue water potential by means of psychrometric methods (Jones, 1992) or by means of the pressure chamber (Figure 9.5, top left) (Scholander et al., 1965; Hsiao, 1990). Psychrometer-based measurements of water potential, although valuable and with the possibility of automation, require sophisticated equipment and high level of technical skill, limiting their usefulness for irrigation scheduling purposes. Leaf water potential, \( \psi_r \), has been recommended for many years as a sensitive index for irrigation control in both field crops and fruit trees.
Soil water and its management

(Grimes and Yamada, 1982; Peretz et al., 1984). However, the interactions between \( \psi_l \) and the temporal fluctuations of the surrounding environment (Fereres and Goldhamer, 1990; Jones, 1990) make it difficult to identify a threshold \( \psi_l \) for use in irrigation scheduling.

As outlined by Jones (2004), a further limitation of plant water status (and \( \psi_l \) in particular) as an indicator of water stress results from the fact that it is subject to some physiological control. In extreme cases, plants termed as ‘isohydric’ (e.g. maize or poplar plants; Tardieu and Simonneau, 1998) present good endogenous control systems and may maintain a relatively stable leaf water status over a wide range of evaporative demand or soil water supplies. Under these circumstances, and especially in strongly isohydric species, the usefulness of plant water status as an indicator of water stress is limited. This is in contrast with those species such as sunflower or barley, which appear to have less effective control of leaf water status and have been termed ‘anisohydric’ (Tardieu and Simmoneau, 1998).

None of the aforementioned plant-based methods, although suitable as plant stress indicators in research studies, are well adapted for automation of irrigation scheduling or control and all are labour-intensive (therefore expensive, especially for predawn measurements). Several indirect methods for continuously measuring or monitoring water status are currently available as alternatives to direct measurement. Jones (2004) has made an extensive and comparative review of such methods.

Among these plant-based indicators, a substantial amount of effort and research has been devoted to develop precise irrigation scheduling based on changes in the stem diameter, although variations in diameter may also occur in other parts of the plant (Ueda and Shibata, 2001; Cermak et al., 2007; Carminati et al., 2009). The fundamentals and usefulness of this technique for irrigation scheduling have been extensively reviewed by Fernández and Cuevas (2010) and Ortuño et al. (2010). The trunk of woody plants (e.g. fruit trees or vines) represents a major plant water store that may contribute substantially to its transpiration stream even under well-watered conditions (Zweifel et al., 2007). Diurnal and seasonal changes in the water content of the extensible tissues of the stem which are driven by changes in xylem water potential cause diurnal variations in the trunk diameter (stem diameter variations: SDV) that may be used to derive some stress indices, for example in a number of fruit.

Figure 9.5 (top left) Scholander-type pressure bomb; (top right) LVDT sensor installed on a fruit tree trunk; (bottom left) Sap flow probe installed on a fruit tree trunk; (bottom right) infrared thermometer installed above a fruit tree canopy. Pictures taken at the experimental station of the University of Cartagena (Spain).
Soil Conditions and Plant Growth

Most SDV-records for irrigation scheduling during the last two decades have used linear variable differential transformers (LVDT) (Figure 9.5, top right), because of their robustness and high precision (Fernández and Cuevas, 2010), although other devices like strain-gauges have also proven successful for continuous measurement of stem, branch or even fruit diameters (Ueda et al., 1996; Link et al., 1998). One of the limitations of the method lies in the requirement of expert knowledge to interpret SDV records. This limitation results from the range of factors, besides water stress, that affect SDV, such as variations in growth pattern due to the tree’s age (Goldhamer and Fereres, 2001), crop load (Intrigliolo and Castel, 2007), tree size (Intrigliolo and Castel, 2006) or crop management (Fernández and Cuevas, 2010). Nonetheless, when the method is properly used, it presents great potential for automatically scheduling irrigation in commercial fruit tree orchards (Vélez et al., 2007; Egea et al., 2009).

Another plant-based indicator with great potential for irrigation scheduling is sap flow (SF). Various methods are available to calculate mass flow of water in the transpiration stream by using heat as a tracer [Figure 9.5, bottom left; for reviews (see Swanson, 1994; Smith and Allen, 1996; Kostner et al., 1998)]. Sap flow systems are easily automated, robust and reliable enough for operation in the field over extended periods of time. Fernández et al. (2008) demonstrated that the ‘transpiration ratio’, as defined by the ratio between the actual daily water use (measured with SF sensors) divided by the potential daily water use of well-watered plants, appeared to be a reliable tool to schedule irrigation in commercial fruit tree orchards and grapevines. However, keeping plants ‘well-watered’ can present problems related to over-irrigation (e.g. hypoxia), which can make the use of these plants as representatives of the orchard unrealistic.

There has also been an increasing interest in using infrared measurements of canopy temperature as an indicator of stomatal closure, transpiration and ‘crop stress’, since the closure of stomata leads to increases in leaf temperature through the well-reported transpirational cooling effect (Ehrler, 1973; Jones et al., 2002). Thermal methods have the advantage

Figure 9.6  Stem diameter variations recorded during two consecutive days in well-watered and water stressed almond trees in SE Spain (Egea, 2008). MDS, maximum daily shrinkage; SGR, stem growth rate.
of being more rapid than conventional gas-exchange measurements and of not interfering with the stomatal responses through physical contact with the leaves (Guilioni et al., 2008). With the development and availability of infrared thermometers (Monteith and Szeicz, 1962), there have been many attempts to relate canopy temperature to plant water stress conditions to provide guidance for irrigation scheduling. Most of these methods use stress index approaches, such as the crop water stress index (CWSI, Idso et al., 1981; Jones, 2004), that are indirectly or empirically related to stomatal conductance.

CWSI is based on measured canopy-air temperature difference ($\Delta T = T_c - T_a$). After obtaining the two limits of $\Delta T$ for potential (non-water stressed, $\Delta T_{nws}$) and null transpiration of a crop ($\Delta T_{dry}$), or similarly, for maximum and minimum canopy conductance, under given environmental conditions, the CWSI can be calculated as follows (Idso et al., 1981):

$$\text{CWSI} = \frac{\Delta T - \Delta T_{nws}}{\Delta T_{dry} - \Delta T_{nws}}$$  \hspace{1cm} (9.12)

The $\Delta T_{nws}$ and $\Delta T_{dry}$ limits used in Equation 9.12 can be obtained using different methods. A widely used one is the method by Idso et al. (1981), where two baselines are obtained empirically by relating ($T_c - T_a$) to air vapour pressure deficit values (see Figure 9.7).

An alternative approach makes use of direct infrared measurements over wet or dry reference surfaces, either natural or artificial (Jones, 1999a,b; Jones et al., 2002; Wang et al., 2005; Moller et al., 2007). Although this method has great potential, as long as caution is taken in the design and placement of the reference surfaces, presently it is not suitable to be used directly and continuously for irrigation scheduling. Testi et al. (2008), by using continuously recorded Infra-Red Thermometers (IRT) installed above the tree canopy (Figure 9.5, bottom right), demonstrated the potential of CWSI as a plant-based water stress indicator in irrigation scheduling of pistachio trees. Besides CWSI, other thermal indices, such as the one proposed by Jones (1999a) and successfully tested, among others, by Leinonen and Jones (2004) and Grant et al. (2007), have also demonstrated its accuracy as stress indicators.

The viewing angle and the sample of foliage targeted by the IRT devices that Testi et al. (2008) used can also be obtained with high-resolution satellite or air-borne remote sensing equipment. This highlights the possibility of remote thermal detection of water stress for irrigation applications, as demonstrated by Sepulcre-Cantó et al. (2006) with high-resolution airborne thermal images flown over non-homogenous orchard canopies. Traditionally, critical issues such as the optimum spatial and spectral resolutions, the turnaround time and repeat cycle have been major limiting factors of the usefulness of remote sensing products (e.g. infrared thermal images) for precision crop management. Airborne sensors can deliver higher spatial and spectral resolutions and are more flexible in terms of revisit time than satellite-based products (Suárez et al., 2008; Sobrino et al., 2009), but the high operating costs and the long turnaround times due to high volume of data processing acquired in large airborne campaigns have limited their use. Nonetheless, developments conducted on unmanned aerial vehicles (Berni et al., 2009) have opened the gate to provide useful remote sensing products in agriculture. In addition to canopy temperature, these systems may also provide other physiological and structural indicators for remote sensing detection of water stress, such as the physiological reflectance index (PRI) which is sensitive to xanthophyll pigment detection under water stress conditions (Suárez et al., 2008) or solar-induced chlorophyll fluorescence emission (Zarco-Tejada et al., 2009), due to the strong correlation demonstrated between steady-state chlorophyll fluorescence and stomatal conductance (Flexas et al., 2002).
9.4 Determination of soil water balance and related variables

A wide range of soil-based, as opposed to plant-based, irrigation scheduling methods is currently available. Despite the fact that plant-based water-stress indicators promise great potential, soil-based techniques to inform irrigation scheduling are currently still much more common. The soil-based irrigation-scheduling approaches broadly fall under two categories: either direct measurement techniques or water balance methods. Direct measurement techniques measure soil water content or matric potential *in situ*. Irrigation is then scheduled according to a target soil water content (or deficit) or matric potential. Water balance methods keep a running balance of the water in the soil profile from measurements of rainfall and irrigation amounts and estimates of evapotranspiration (Hess and Knox, 2002). Hence, this section will describe measurement techniques to determine $\theta$ and $\psi_m$, as well as key components of the water balance.

Monitoring of soil moisture content (and groundwater levels, using piezometers, see Section 9.4.2) is also essential once drainage systems have been installed, to ensure the artificially lowered water levels and lower moisture contents in the root zone are within the intended range and not adversely affecting crop growth.

### 9.4.1 Determination of soil water content

Techniques for measuring soil water content are often classified as either direct or indirect. The direct methods refer to those methods that involve removal of water from the soil matrix to directly measure the amount of water present in the sample. This has traditionally been achieved by recording the change in mass of a soil sample after it has been dried out to a constant mass at 105°C. This method requires the use of convective oven drying and is
commonly referred to as the gravimetric method (see also Chapter 8). Water content can be expressed either on a mass basis (gravimetric soil water content, $\theta_g$, in kg water kg$^{-1}$ dry soil) or on a volumetric basis (volumetric soil water content, $\theta$, water volume to soil bulk volume, m$^3$ water m$^{-3}$ soil). They are related through

$$\theta = \left( \frac{\rho_d}{\rho_w} \right) \theta_g$$

(9.13)

where $\rho_d$ is the dry bulk density of the soil (kg m$^{-3}$) and $\rho_w$ the water density (kg m$^{-3}$).

The gravimetric method has the advantage of being accurate and inexpensive, but it is a destructive, time-consuming method, that does not allow for repetitive measurements in the exact same location. Hence, it is less suitable when (near-) continuous measurements are required, e.g. for scientific purposes or for precision irrigation.

Indirect methods are based on the measurement of some $\theta$-dependent physical or chemical properties. These properties include the dielectric constant (relative permittivity), electrical conductivity, heat capacity, hydrogen content or magnetic susceptibility (see Durner and Lipsius, 2005 for a detailed overview). Some of the most relevant available methods for indirectly estimating $\theta$ are summarized in the rest of the section.

**Neutron moderation or neutron thermalization**

The neutron moderation technique is a long-standing method which represented, for many years, the benchmark field method to determine $\theta$. The neutron probe has a radioactive source of high-energy (fast neutrons) and a detector for slow, thermalized neutrons and is lowered into the soil via a preinstalled access tube. The neutrons that are emitted from the radioactive source lose energy as they collide with soil particles; these neutrons slow down dramatically when they collide with hydrogen nuclei, which are part of H$_2$O molecules and become thermalized. Hence, a measure of the quantity of thermalized neutrons returning to the detector within a given time period gives a good measure of $\theta$. The method requires a linear calibration between the count rate of thermalized neutrons and $\theta$ obtained from soil samples taken at the site. Unlike other devices, the neutron probe is not affected by salinity or air gaps. This method offers the opportunity of frequent and undisturbed revisiting of an experimental field, thereby offering the opportunity of long-term monitoring over a (number of) growing season(s). Unfortunately, the measurements cannot be automated. The soil sensing volume is large as compared to other methods, with a sphere of influence of variable radius: about 0.15 m in wet soil and up to 0.5 m in dry soil, thus limiting measurements to >0.3 m depth below the soil surface without special calibration for surface layers (Topp and Ferre, 2005).

Due to regulations and hazards concerning the use of radioactive sources, the method requires certified personnel and in recent years regulations and controls have become more and more stringent. This has led to many contractors withdrawing technical support, often causing owners to decommission their neutron probes. This issue, together with the fact that the instrument is expensive and somewhat cumbersome, has stimulated the development and use of other methods, largely electromagnetic.

**Electromagnetic methods**

These techniques are becoming widely adopted because they provide instantaneous measurements, do not require maintenance, and can provide continuous readings through automation (Muñoz-Carpena, 2004). The fundamental principle of the electromagnetic (EM) methods
relies on the strong relationship between soil bulk dielectric permittivity ($\varepsilon$) and $\theta$ found for many mineral soils (Topp et al., 1980; Malicki et al., 1996). This strong dependence stems from the high permittivity of free water ($\varepsilon \approx 80$) as compared to that of mineral soil solids ($\varepsilon \approx 2–9$) and air ($\varepsilon \approx 1$). These methods estimate $\theta$ from a property (e.g. wave travel time, impedance, capacitor charge time, or frequency shift) related to the response of an EM signal (travelling through the soil either on a transmission line or in the soil medium directly) to changes in the apparent permittivity of the soil, $\varepsilon_a$.

Electromagnetic methods, whether time domain reflectometry (TDR), frequency domain reflectrometry (FDR), or capacitance methods, all estimate water content based on the permittivity of the soil medium. The accuracy of these estimates depends on a two-stage calibration. The first being the measurement of $\varepsilon_a$ and the second a calibration between $\varepsilon_a$ and estimated $\theta$ (Blonquist et al., 2005). A common approach to establish the relationship between $\varepsilon_a$ and $\theta$ has been the empirical equation derived by Topp et al. (1980), among others (Topp and Reynolds, 1998), although most sensors have their own calibration curve provided by the manufacturer. This equation has proven to be accurate for many soils, but does not account for factors such as a temperature effect on determined water contents in some soils (see Verhoef et al., 2006; Fernandez-Galvez et al., 2007) or the loss of accuracy in soils containing high ion-exchange-capacity clays (Topp et al., 2000; Ferré and Topp, 2002). At low frequencies, EM sensors are salinity-dependent or more correctly dependent on the temperature- and water content-dependent bulk electrical conductivity that results from salinity (Robinson et al., 1998; Kelleners et al., 2004). In those cases, a soil-specific $\varepsilon_a-\theta$ relationship is required (Muñoz-Carpena, 2004).

### 9.4.2 Soil water pressure sensors

Positive pressures (i.e. below the GWT) can be determined using piezometers, tubes of a few cm diameter (to avoid capillary rise), of known length, open at both ends, installed in a soil profile where a water table is present. The water level in the piezometer is used to calculate the positive pressure potential at the bottom of the tube, $z_b$ ($\psi_p = 0$ at the water level, $z_w$, and hence $\psi_p = |z_b - z_w|$) (Kooistra et al., 1983). As discussed in Section 9.2.1, water in unsaturated soils is retained at suction, $S = -\psi_m$, which may be determined either by direct measurement of soil water pressure (tensiometers), by equilibrium water vapour pressure in soil air (thermocouple psychrometers), or by inference from $\theta$ of a reference porous medium in equilibrium with the surrounding soil (heat dissipation sensors, gypsum blocks, granular matrix sensors, filter paper).

Tensiometry is the most widely used technique to measure the soil matric potential under relatively moist soil conditions. A tensiometer is a device that consists of a porous cup connected to a pressure-sensing element (pressure gauge or transducer) through a water-filled rigid body tube. A Bourdon-type vacuum gauge is commonly employed for water potential measurements, although mercury manometers, if a greater accuracy is required, or pressure transducers, if automatic and continuous readings need to be recorded, can also be used. The porous cup is normally made of ceramic material because of its structural strength as well as permeability to water flow.

Tensiometer length may be variable and the use of a specific length will depend on the user’s needs according to factors such as crop type (i.e. soil depth explored by the root system) or the irrigation criteria to be adopted by the irrigator. It is important for tensiometers to be installed properly, with the ceramic cup firmly in contact with the surrounding soil. Then, the soil solution can flow into or out of the tensiometer through the small pores in the
ceramic cup until the liquid in the cup comes into equilibrium with the matric potential of the soil water around the cup. As the soil dries, soil matric potential decreases (suction/tension increases) and the tensiometer vacuum gauge reading increases. The opposite happens when an increase in soil water content (e.g. from irrigation or rainfall events) takes place.

The suction generated within the cup is transmitted through the liquid column to the gauge. The vacuum gauge does not provide direct measurements of the tension in the cup when there is a difference in height between the two, since the liquid between the cup and the gauge is in static equilibrium and, thus, the pressure in the liquid column increases linearly with depth (Koorevaar et al., 1983). Therefore, \( \psi_m \) in the tensiometer cup (here expressed in pressure units) will be given by the expression (Koorevaar et al., 1983):

\[
\psi_m = P_{\text{gauge}} + \rho g \Delta z
\]

where \( P_{\text{gauge}} \) is the pressure, relative to atmospheric pressure, of the liquid in the vacuum gauge and \( \Delta z \) is the difference in height between the cup and the gauge.

Tensiometers are instruments that have been extensively used for irrigation scheduling purposes because they provide direct measurements of soil moisture status, are easily managed and relatively inexpensive. However, they also present some drawbacks that might limit their use. In this regard, the typical measurement range (\( \psi_m \) down to \(-80\,\text{kPa} \) or \(-8\,\text{m}\)) remains a severe limitation for soil studies and for some irrigation schemes such as deficit irrigation scheduling (see Section 9.5.1). Furthermore, the intimate soil contact they require around the ceramic cup for consistent readings is not always guaranteed in swelling or coarse soils.

An alternative to tensiometers is thermocouple psychrometers (TCPs), reliable and accurate devices (provided that appropriate precautions are taken) for determination of water potential (sum of matric and osmotic potential). Separation of matric and osmotic water potential components can be achieved with one of the methods described in Andraski and Scanlon (2002). TCPs infer the water potential of the liquid phase of a soil sample from measurements within the vapour phase that is in equilibrium with the liquid soil water. As soil dries, the relative humidity of the soil air decreases. Water potential in the air phase in equilibrium with the soil sample can thus be determined from the existing theoretical relationship between relative humidity and water potential (see, e.g. Andraski and Scanlon, 2002; Durner and Or, 2005, for a detailed description of the technique). As pointed out by Andraski and Scanlon (2002), the main difficulties of this technique are first, the range of variation of relative humidity in the soil gas phase is very narrow as compared to the typical range of interest of soil water potential (= \(-0.01 \) to \(-1.5\,\text{MPa}\)); and second, temperature differences in the sensor-sample system may lead to large errors in the determination of water potential.

For accurate determinations, the method requires calibration of TCPs, consistency in measurement technique and careful cleaning and handling. A wide range of instruments has been developed and are commercially available for use in laboratory, glasshouse and field experiments (Andraski and Scanlon, 2002). However, notwithstanding the reliability and accuracy of these devices, the degree of complexity of the technique as well as the number of precautions that need to be taken has limited its broad application in field studies.

The last group of methodologies indirectly infers soil matric potential from water content of a reference porous medium in equilibrium with the surrounding soil. Techniques used are the heat dissipation method, electrical resistance method and filter paper method. In general, heat dissipation sensors consist of a heating element and an accurate thermocouple embedded within a porous ceramic matrix (Flint et al., 2002). The working principle is based on the fact
that thermal conductivity of a porous material is proportional to its water content. These sensors are relatively inexpensive and easy to read, but require individual calibration and temperature correction (Flint et al., 2002).

The electrical resistance method consists of electrodes embedded in a porous medium (block). Variations in electrical resistance of the porous media are due to variations in $\theta$ and hence to variations in the matric potential of the surrounding soil. These sensors have low resolution, are temperature dependent and require calibration. However, they are simple, inexpensive and are suited to schedule deficit irrigation (Muñoz-Carpena, 2004). Gypsum blocks are an example of electrical resistance sensors, which use gypsum as porous material. Gypsum serves as a buffer (to a specific level) against soil salinity changes, since variations in soil electrical conductivity would affect electrical resistance measurements. Major limitations of gypsum blocks are the degradation they undergo with time and the fact that they are not suitable for measurements under high soil moisture conditions (matric potential $> -30$ kPa). Some of the disadvantages of gypsum blocks have been overcome by Granular Matrix Sensors (Irmak and Haman, 2001).

### 9.4.3 Estimation of the soil water balance

The soil water balance, SWB, was given in Equation 9.1. The way in which this equation is generally employed in irrigation scheduling is to assume a root zone soil moisture depletion, $D_r$ (see Equation 9.4), of zero when the soil is at field capacity. Thereafter, evaporative losses and deep drainage increase $D_r$ and rainfall and irrigation decrease it. Irrigation will take place when $D_r$ has reached a certain level; this level mainly depends on crop type and crop stage, rooting depth, plant spacing and irrigation water availability.

Consequently, this kind of irrigation scheduling method (see Section 9.5.1) requires measurements/estimates of gain terms $P$ and $I$ (through the use of rain gauges or knowledge of sprinkler/trickle/dripper rates) and of loss terms $D$ and $ET$ (ideally of $E$ and $T$ separately). For a relatively flat surface, $R$ can be ignored. Often, capillary rise, $C$, is assumed to be negligible, which is a reasonable assumption if the soil is coarse-textured and groundwater levels are relatively deep. Drainage ($D$) below the root zone is usually ignored for irrigation scheduling purposes because $\theta < \theta_{fc}$; water flow will have diminished considerably, see Section 9.2.3 (but not stopped entirely, especially in fine-textured soils in which deep drainage and redistribution will continue for a considerable length of time).

Hence, the key water balance component to estimate is the evapotranspiration, $ET$. The most direct way to obtain some measure of evaporation is by using a pan evaporimeter or open-pan: a container, filled with water to a few centimetres below the rim. Some pans are traditionally mounted above the ground surface, others are sunk in the ground. Their shape (generally cylindrical), depth, diameter, material (usually metal) and colour vary as well. A fixed-point gauge in a stilling well serves to indicate the level of water in the pan. The new water level is found (e.g. daily) and the change from the previous reading is the pan evaporation, after which the water level will be brought to the fixed point again. The reservoir is often covered with wire-mesh netting to protect the pan from wild-life and wind-blown debris. The results from identical pans vary according to their position and exposure. A pan evaporimeter determines potential evaporation, not the actual evapotranspiration, and can, depending on location, overestimate potential evaporation because of advection effects. Hence, it is not ideal for irrigation or drainage scheduling.

Another method is weighing lysimetry, which involves monitoring weight changes of an extracted, undisturbed soil column (although in some cases repacked soil columns have
Soil water and its management

Evaporation is calculated from differences between regularly measured masses. Weighing can be automated and continuous, using balance beam, hydraulic or mechanical load-cell based lysimeters and can be corrected for precipitation, irrigation and often drainage. Alternatively manually weighed microlysimeters can be used (Boast and Robertson, 1982). These smaller lysimeters, consisting of cores, which weigh typically 1–3 kg, with diameters ranging between 5 and 20 cm and depths of between 5 and 30 cm, are used for measurements of soil evaporation below canopies.

Lysimeters vary in size and depth, mainly depending on the type of vegetation and economic considerations (see Klocke et al., 1985; Dugas and Bland, 1989; Allen and Fisher, 1990; Daamen et al., 1993). Large lysimeters, the ones in theory most useful for irrigation and drainage planning purposes, are relatively expensive and difficult to install and maintain. They are therefore mainly used for research, although their numbers have dwindled in recent years, to make way for micrometeorological methods (see Verhoef and Campbell 2005, for more information on lysimeters and meteorological methods for estimation of ET). Figure 9.8 shows the installation procedure and arrangement of eight lysimeters (0.55 m × 0.55 m × 0.25 m, uniformly packed with the sandy loam soil obtained in situ).
installed in an orchard planted with 35-year-old olive trees (ground cover ~50%), located at the CSIC experimental farm La Hampa (Coria del Rio, Sevilla, Spain). These lysimeters were installed at different orientations (N, E, S, W) around and distances from (1 m: the inner set of lysimeters, (ILS) and 2 m: the outer set of lysimeters (OLS)) the trunk of a single olive tree. Each lysimeter was made of perspex, a material with a very low thermal conductivity to avoid lateral heat conduction through the lysimeter casing.

Figure 9.9 shows the soil evaporation results for a 9-day drying period after an irrigation event on DOY 248 (5 September) in 2005, for the outer (OLS) and inner (ILS) lysimeters. For both sets of lysimeters, evaporation decreased sharply after irrigation, reaching a near-constant level of about 1 mm day\(^{-1}\). Because this was a sandy soil located in a hot and dry climate, first stage drying (see Section 9.2.1) would only have lasted for a few hours after irrigation; the daily sum of \(E\) was below \(E_p\) for all days, so that Figure 9.9 illustrates the second stage of drying.

The inner lysimeters were more or less continuously shaded and hence very little difference between their \(E\)-values was observed. For the OLS, radiation load was much more variable and depended strongly on orientation of the lysimeter around the tree. This is reflected in the larger variability in \(E\) for the OLS. Figure 9.9 indicates that it is worth keeping irrigated areas as small as possible because large water losses by \(E\) can occur, even for shaded soil; this can be achieved by micro-irrigation (Section 9.5.1).

In the soil physical approaches described, \(\Delta \Theta\) is measured using the type of soil moisture sensors described in Section 9.4.1 to indirectly estimate \(ET\), from Equation 9.1, assuming \(P\), \(I\), \(D\) and \(C\) are known or can be neglected (see Fernández-Galvez et al., 2007). This is not the same as scheduling according to a target soil water content, where the absolute values of \(\theta\) are used to guide irrigation scheduling, rather than the changes in the entire water balance.

A final way to obtain estimates of evapotranspiration is by employing meteorological theory and equipment. Two main classes of micrometeorological methods can be defined: the first focuses on obtaining experimentally \(in situ\) evaporative flux measurements (and those can be split into direct and indirect methods), whereas the second class concerns calculation of \(ET\) from well-known equations, for example the Penman-Monteith equation, driven by easily available weather variables. These meteorological methods rely on measurement of the surface energy balance (\(SEB\)), which is given by:
where $\lambda$ is the latent heat of vaporisation (~ 2.45 MJ kg$^{-1}$). Equation 9.15 is an important boundary condition for energy exchange processes at the land surface, determining how available energy (net radiation flux density, $R_n$, minus soil heat flux density, $G$) is partitioned between sensible heat flux density, $H$ (the flux of heat from the earth’s surface to the atmosphere that is not associated with phase changes of water) and latent heat flux density, i.e. the evapotranspiration, $\lambda ET$. The $SEB$ is directly related to the $SWB$ through the evaporation term, $\lambda ET$ (see Brutsaert, 1982; Hillel, 1998). In the $SEB$ energy flux is expressed in units of Wm$^{-2}$. $ET$ in mm day$^{-1}$ can be converted to $\lambda ET$ in Wm$^{-2}$, with knowledge of $\lambda$ and given that 1 kg H$_2$O m$^{-2}$ = 1 mm H$_2$O.

The first group of meteorological methods to determine $ET$, or rather $\lambda ET$, experimentally, is the direct method using Eddy Covariance (EC). This method is based on turbulence theory. The EC method relies on the fact that turbulence is random in nature and hence an atmospheric variable $A$ may be separated into its turbulent part, $A'$, and its non-turbulent, average part, $\overline{A}$, ($A = \overline{A} + A'$). This method calculates $\lambda ET$ from the covariance between the vertical wind speed, $w$ (ms$^{-1}$) and the specific humidity, $q$ (kg water kg$^{-1}$ air); $w'q'$. By obtaining fast-response measurements of state variables $w$ (using a three-axis sonic anemometer) and $q$ (from an adjacent Infra Red Gas Analyser, IRGA, see Figure 9.10) time series can be generated from which the perturbation values of the data points can be calculated (i.e. the fluctuations around the mean, as indicated by the primes).

Figure 9.10  Detail of eddy covariance (EC) equipment, comprising a three-axis sonic anemometer and the adjacent Infra Red Gas Analyser (IRGA). Picture taken at the experimental station of the University of Cartagena (Spain). For a colour version of this figure, please see Plate 9.3.
IRGAs are available in either a closed path (e.g. LI-COR’s LI-7000) or an open path (e.g. LI-7500) set up and have associated advantages and disadvantages (e.g. Leuning and Judd, 1996). Acquisition and post-processing software are required to obtain and record the raw data (about 20 MB a day), to generate the covariances and to apply the corrections (see, for example, www.geos.ed.ac.uk/abs/research/micromet/edisol/edidownl.html).

EC instruments are expensive, require post-processing software and some insight into meteorological theory. Consequently, the EC method is generally only used for research purposes. It is sometimes used to derive or test crop coefficients (e.g. Paco et al., 2006), which will inform crop models.

An indirect method is the Bowen ratio energy balance (BREB) method, which infers the evaporation from combining measurements of air temperature, $T_a$, and actual vapour pressure ($e$ in mbar), measured at two levels, with the SEB. This yields the Bowen Ratio, $\beta'$, which is defined by $H/\lambda ET$:

$$\beta' = \frac{\Delta T}{\Delta e}$$

(9.16)

where $\gamma$ is the psychrometric constant (~ 0.66 mbar K$^{-1}$). $\lambda ET$ is then given by

$$\lambda ET = \frac{R_n - G}{1 + \beta'}$$

(9.17)

The BREB technique requires two high-precision atmospheric humidity and temperature sensors (installed at different heights), a net radiometer and an estimate of soil heat flux (see Sauer and Horton, 2005, for example). Its robustness and simplicity result in relatively reliable estimates of evaporation, and the comparatively low cost of the sensors make it a useful method for estimates of crop water use (e.g. Todd et al., 2000).

An alternative, indirect method of estimating evaporation is by modelling. Models range from those calculating actual evaporation by multiplication of the potential evaporation by a crop coefficient; a factor depending on crop type, development stage and soil water status (Allen et al., 1998), to detailed mechanistic Soil Vegetation Atmosphere Transfer models (Shuttleworth and Wallace, 1988; Verhoef and Allen, 2000).

A reliable equation used widely for estimation of evapotranspiration for the purpose of irrigation and drainage scheduling (see Allen et al., 1998) is the Big-Leaf model or Penman-Monteith (Monteith, 1965) equation given by:

$$\lambda ET = \frac{\Delta(R_n - G) + \rho_a c_p VPD / r_a}{\Delta + \gamma(1 + r_s / r_a)}$$

(9.18)

where $r_a$ is the aerodynamic resistance (sm$^{-1}$) and $r_s$ (sm$^{-1}$) the canopy surface resistance. VPD (mbar) is the vapour pressure deficit ($e_s - e$; with $e_s$ being the saturation vapour pressure in mbar) at reference level, $z_r$ (m). Furthermore, $\Delta$ (mbar K$^{-1}$) is the slope of the vapour pressure temperature curve, $\gamma$ is the psychrometric constant (mbar K$^{-1}$), $\rho_a$ (kg m$^{-3}$) is the density of air and $c_p$ (J kg$^{-1}$ K$^{-1}$) is the specific heat of air at constant pressure.

The meteorological variables: $R_n$, $G$, $T_s$ (to get $\Delta$ and $\gamma$), $u$ (to calculate $r_a$) and VPD are relatively easy to obtain using standard meteorological instrumentation. However, the aerodynamic and surface resistance is less straightforward to determine. The aerodynamic resistance, $r_a$, is a function of the rate of turbulent mixing and can be calculated from measurements of wind speed, estimates of atmospheric stability and knowledge of surface roughness.
parameters (roughness length and displacement height; see Brutsaert (1982), Raupach (1994) or Verhoef et al. (1997) on how to derive these variables from standard estimates of canopy structure). Standard meteorological textbooks can be referred to for calculation of $r_a$ (e.g. Garratt, 1992).

If evaporation from the soil is assumed to be negligible, then $r_s$ can be equated to the bulk stomatal resistance of the leaves in the canopy: often calculated with $r_s = 1/(g_s \cdot LAI)$ where $g_s$ is the average leaf stomatal conductance (see Section 9.3.1) and $LAI$ is the leaf area index. This formula assumes that the air within the canopy is well-mixed, with all the leaves of the canopy being exposed to exactly the same atmospheric properties. Refer to Jones (1992) and Jones (2004) for overviews of methods to estimate stomatal conductance, $g_s$ (required to obtain $r_s$).

The exchange between the atmosphere and a dense (i.e. where the canopy covers the entire ground surface) sole crop or canopy is relatively well understood, and Equation 9.18 can yield very precise estimates of evapotranspiration. However, when the surface is sparsely vegetated, for example, before crop canopy closure has taken place, a significant percentage of the surface consists of exposed undergrowth or bare soil, so ideally a two-component or sparse canopy model is used. Such models introduce more explicit descriptions of aerodynamic transfer through the leaf boundary layer, and they have proved useful in describing a variety of sparse canopies, including crops and natural vegetation (Wallace et al., 1990; Van den Hurk et al., 1995; Allen et al., 1998; Verhoef and Allen, 2000). In some of these two-or multi-component models, there is no interaction between the component latent or sensible fluxes. However, there are many situations where fluxes from the substrate can significantly affect the fluxes from the upperstorey canopy. For example, Wallace et al. (1990) showed that, in a sparse millet canopy, transpiration could vary by up to 40%, depending on the fluxes of heat and water vapour from the soil below.

9.4.4 Estimation of hydraulic conductivity

Hydraulic conductivity is an important variable when describing water flow within soils (Equation 9.7). Knowledge of this variable is also essential in drainage studies. The design and functioning of subsurface drainage systems depends to a great extent on the soil’s saturated hydraulic conductivity ($K_s$); all drain-spacing equations make use of this parameter. To design or evaluate a drainage project, $K_s$ has to be determined as accurately as possible (Oosterbaan and Nijland, 1994).

Methods to determine (un)saturated hydraulic conductivity generally employ formulae based on the Darcy equation and the boundary conditions of the flow; these flow conditions are generally artificially imposed. The $K$-value is calculated from the formula using the values of hydraulic head ($\Delta \psi$) and discharge observed ($F_w$) under the imposed conditions (Equation 9.7). These methods can be grouped into in situ field methods and laboratory methods. In the latter case an undisturbed soil sample is extracted from the field and analysed under lab conditions. A variety of methods can be found in Klute (1972), Oosterbaan and Nijland (1994), Rowell (1994) and Bagarello et al. (2004).

In the last two decades, the Multistep Outflow method (see Van Dam et al., 1994; Vereecken et al., 1997), based on inverse parameter optimization for soil hydraulic functions [e.g. those presented in Van Genuchten (1980)], combined with transient outflow measurements, has been developed to make the measurement of laboratory hydraulic properties faster, more reliable and less laborious.
9.5 Water management to alleviate soil water shortage

Agricultural activity dominates the use of freshwater and accounts for some 70% of withdrawals from water resources globally. As most of the water withdrawn by agriculture is lost by evapotranspiration, in marked contrast to domestic and industrial withdrawals, agriculture accounts for 80–90% of freshwater used by humans (FAO, 2002; Morison et al., 2008). This highlights the importance of agriculture in the challenge of making the Earth’s available water serve the needs of its growing number of users. Furthermore, there is widespread agreement that climate change will exacerbate the present shortages of water in certain areas and may lead to increased drought (IPCC, 2007). Thus, reducing agricultural water use and making water resources more sustainable are increasingly urgent matters. This challenge requires combined efforts from multiple disciplines, addressing agronomic, physiological, biotechnological/genetic and engineering issues (Morison et al., 2008; Farooq et al., 2009). The utilization of monitoring techniques (such as those described in Section 9.4.3) to assess the level of soil water deficit, and the use of such measurements to control irrigation systems, will assist us to move towards precision agriculture. In this context, the aim is to satisfy plant water (and nutritional) requirements or alternatively achieve a desired and controlled level of water stress – so called deficit irrigation (see Section 9.5.1), while minimizing environmental impacts caused by excess applied water and subsequent leaching, and/or avoiding plant diseases related to excess water. Moreover, improving irrigation efficiency can contribute greatly to reducing production costs of vegetables and crops, making the industry more competitive and sustainable.

It is notable that irrigated agriculture is not sustainable in many (semi-)arid regions unless drainage is used simultaneously to provide salinity control (see Chapter 7 and Section 9.5.3) as well as for optimal root respiration.

9.5.1 Types of irrigation and irrigation scheduling

The irrigation system employed is a major determinant of on-farm irrigation performances and may be characterised by one of three categories: (1) surface irrigation, (2) sprinkler irrigation and (3) micro or localized irrigation. Surface irrigation systems are the most widespread worldwide, representing about 84.5% of the total irrigated area globally, followed by sprinkler irrigation (≈ 13.5%) and localized irrigation (≈ 2%) (AQUASTAT, 2010). The main indicator of farm irrigation efficiency is the application efficiency, which can be defined as the depth or volume of water added to the root zone store expressed as a ratio of the depth or volume of water applied to the field. In general, surface irrigation systems (e.g. furrow-, border- or basin-irrigation) apply large irrigation depths, and as these systems have traditionally been operated manually, over-irrigation is often practised (Smith et al., 2005). Consequently, on a global basis, on-farm application efficiency is estimated to be ~37%, which means that 63% of the water delivered to fields is lost as runoff and/or drainage (Wallace, 2000). Nonetheless, if the rate of water application is correctly matched to the infiltration rate of the soil and slope of the land, surface irrigation systems may significantly improve the crop’s performance.

The other two types of irrigation systems (i.e. sprinkler and localized) belong to the group of pressurized systems. Sprinkler systems, which include set, travelling rain-guns and continuously moving lateral systems (Pereira et al., 2002), are able to achieve better irrigation performances than surface systems.
Localized or micro-irrigation systems (i.e. micro-sprinkler, drip or subsurface irrigation systems) are designed to apply small and frequent irrigations. The working pressure of these systems is lower than in sprinkler systems. A common characteristic of these ‘high-tech’ schemes is that there is less wet soil surface exposed to the atmosphere and hence less soil evaporation (as well as less interception loss, which would occur after sprinkler irrigation). Localised irrigation systems also allow maintenance of a required/desired moisture level in the root zone, enable targeted fertiliser application (‘fertigation’), reduce the risk of disease by keeping the foliage dry and facilitate farm operations due to the fact that the soil surface remains dry. As with sprinkler irrigation, the application efficiency depends mainly on a number of system and management variables, the latter related to the duration of the irrigation and irrigation frequency (Pereira et al., 2002).

Subsurface drip irrigation systems are particularly advantageous as they can increase crop WUE because less water is lost as evaporation from the soil surface (Ayars et al., 1999). These systems involve permanent drip tapes below the soil surface. Depending on factors such as soil texture, crop type, the permanence of the system and whether or not the irrigation system is going to be used for plant germination, the installation depth may normally vary between 0.1 and 0.5 m (Ayars et al., 1999), although deeper installation depths have also been reported (Hutmacher et al., 1996). These systems present some additional advantages, such as a higher uniformity of application, less operation cost if well maintained, less weeds and unrestricted field access (e.g. for spraying or harvesting). In contrast, the main limitation of these systems is the susceptibility of the emitters to clogging by root penetration, rust, microorganisms or chemical precipitates, although with appropriate design and preventive practices this limitation might be overcome. The joint use of deficit irrigation programs (see later) and subsurface irrigation systems has proven to be a reliable strategy to significantly increase WUE in dry areas (Romero et al., 2004).

The advent of these localised irrigation systems has helped to reduce the water required in agri-horticultural crops but, at the same time, has highlighted the need for new methods of accurate irrigation scheduling. The choice of irrigation scheduling method depends to a large extent on the irrigator’s objectives and the irrigation system available. When water supplies are limited, as is the case in arid and semi-arid regions, farmers should place their emphasis on maximising WUE (see Section 9.3.2); these scheduling principles clearly deviate from those aimed at maximising yield by fully meeting crop water requirements.

In this regard, deficit irrigation (DI) strategies have demonstrated that, under water scarcity scenarios and with sufficient knowledge to manage DI optimally, the objective of maintaining or even increasing farmers’ profit may be attainable while reducing irrigation water use (Fereres and Soriano, 2007). The reasons why DI can improve crop water productivity are diverse and have been explained elsewhere (English and Raja, 1996; Fereres and Soriano, 2007). DI can be understood as any irrigation strategy applying water below the crop ET requirements (English, 1990). Until the development of precise irrigation methods (e.g. drip irrigation), there had been little room to manage irrigation with the degree of precision needed in optimal DI programmes, where controlling the timing of application is essential for avoiding the detrimental effects of stress.

Depending on factors such as the pattern of stress imposed, the intensity of water stress or the way irrigation water is delivered to the root-zone, various types of ‘controlled’ DI strategies can be practised to increase WUE: sustained deficit irrigation (SDI), regulated deficit irrigation (RDI), partial root zone drying (PRD) and supplemental irrigation (SI). The term SDI refers to those DI strategies that impose a uniform degree of water shortage throughout the growing season to deliberately seek a slow development of water stress...
Soil Conditions and Plant Growth

(as soil water reserves deplete) and a progressive plant adaptation to water deficits (Fereres and Soriano, 2007).

RDI strategies, as defined by Chalmers et al. (1981), are based on applying a certain degree of water deficit during specific periods of the growth cycle when the crops are less sensitive to water stress (non-critical periods), while full water needs are met during the remaining periods (critical periods). Experiments with RDI have been successful in many fruit and nut tree species, including almond (Goldhamer and Viveros, 2000; Egea et al., 2010), apricot (Pérez-Pastor et al., 2009) and citrus (Domingo et al., 1996) trees.

PRD is an irrigation technique that has been extensively studied in recent years (Dry et al., 1996). This approach is based on the generation of wet-dry cycles in the root-zone by alternating irrigation between different parts of the root system. This irrigation strategy seeks to promote the production of chemical signals (e.g. the plant hormone abscisic acid, ABA) from roots in dry soil and their transmission to the shoots to restrict water use (through triggering partial stomatal closure) while maintaining crop water supply from roots in the wet soil fraction (Stoll et al., 2000; Dodd, 2005). Considering the cost and management complexity of implementing PRD, adoption of this technique (in contrast to conventional DI, where the whole root system is irrigated) requires evidence that it stimulates beneficial physiological and agronomic responses, as compared to DI strategies receiving the same volumes of water. Dodd (2009) recently performed a comparative study between PRD and

![Figure 9.11](image_url)

**Figure 9.11** Crop yield ratio of PRD to DI at similar irrigation volumes (a ratio of 1 indicates that yield with both techniques is equivalent). The fraction of full irrigation is given as a thick line in each column. Columns are means ± SE of the number of experiments/seasons given in parentheses at the base of each column. Shaded columns denote pot experiments where the root system was confined. Significant differences between PRD and DI (P < 0.05) are indicated with an asterisk (*). From Dodd (2009). By permission of Oxford University Press.
conventional DI based upon 15 experiments comprising 10 different crop species. His results indicated that in no case did PRD significantly decrease yield and that, in six studies, PRD significantly increased yield by more than 15% (Figure 9.11). Sadras (2009) showed that for a broad range of horticultural crops and environments, yield per unit of applied water was, on average, significantly, but modestly (5%) higher in PRD than with conventional deficit irrigation.

Supplemental irrigation (SI), is used in some humid to temperate areas as a tactical measure to complement reasonably ample rainfall and stabilise production (Hsiao et al., 2007). In dry zones, where water supply is very limited, it is used as a form of DI by applying irrigation water in certain critical periods of the growing season to significantly improve crop performance relative to the rainfed situation. Examples of how SI may benefit yield and water use efficiency of crops grown in water-limited environments are given in Oweis et al. (2000) and Turner (2004).

To reduce uncertainty and risk associated with the employment of irrigation strategies, computer models that simulate crop performance can assist water managers in optimising a limited supply of irrigation water; Equation 9.18, the Penman-Monteith equation, is invariably at the heart of this type of models. The new FAO crop model AquaCrop represents an example of these useful tools (Raes et al., 2009; Steduto et al., 2009). AquaCrop simulates attainable yields of major herbaceous crops as a function of water consumption under rainfed, supplemental, deficit and full irrigation conditions. Another existing FAO tool is CROPWAT 8.0, a computer program for the calculation of crop water and irrigation requirements (including schedules for different management conditions and varying crop patterns) based on soil, climate and crop data. CROPWAT 8.0 can also be used to evaluate farmers’ irrigation practices and to estimate crop performance under both rainfed and irrigated conditions (see http://www.fao.org/nr/water/infores_databases_cropwat.html).

Two other well-established models are CropSyst (Stockle et al., 1997) and WOFOST (Diepen et al., 1989). Todorovic et al. (2009) compared the performance of AquaCrop with that of CropSyst and WOFOST in simulating sunflower (Helianthus annuus L.) growth under different water regimes in a Mediterranean environment. Their results indicated that despite the fact that AquaCrop requires less input information than CropSyst and WOFOST, it performed similarly in simulating both biomass and yield at harvesting.

### 9.5.2 Automated irrigation scheduling

The degree of automation of the irrigation systems described in Section 9.5.1, as well as their related advantages, is variable. An automated irrigation system saves time and effort for the irrigation manager and also allows higher flexibility in scheduling farm work. The irrigation control systems of these modern systems also offer the possibility to store information on control variables, which represent a valuable feedback to identify problems (e.g. water leakages, undesirable pH or electrical conductivity levels in the irrigation water, etc.).

There are three different ways to perform automated irrigation: time-based, volume-based or based on other (soil, plant or climate related) variables. In the former two groups the irrigation scheduler plays a fundamental role by setting either the time and duration of irrigation (time-based) or the volume of water to be supplied to crops (volume-based). In the latter, however, the system is designed to work more autonomously.

In time-based automation, the irrigation interval is determined from the amount of water to be supplied, application rate of the emitters and number of emitters. Volume-based automation represents another approach for automating irrigation and consists of measuring the
Soil Conditions and Plant Growth

water supplied during each irrigation event to automatically interrupt the water flow once the applied water volume has met the target volume. Both time-based and volume-based irrigation systems require an early estimate of crop evapotranspiration, ideally from one of the methodologies described in Section 9.4.3. An effective procedure to adjust estimated to real crop water needs is to monitor one or more of the available soil and plant water status indicators (reviewed in Section 9.3.3).

Irrigation automation based on measurements of one or more variables of the soil–plant–atmosphere continuum emerged from the need to increase precision in irrigation control to achieve specific objectives in crop management and to maximize irrigation efficiency. Effective operation of such systems requires a sensing system that determines irrigation need in real time; this prerequisite rules out large-scale manual monitoring programmes for such purposes and indicates a need for automated monitoring systems (Jones, 2004). These systems set the irrigation calendar by assessing one or more control variables that are continuously recorded. Edaphic factors, such as soil moisture content, meteorological factors, such as atmospheric evaporative demand, relative humidity, air temperature or solar radiation, or plant physiology-based factors, such as leaf temperature, micromorphometric variations of some plant organs or sap flow have been proposed as indicators to develop self-controlled automatic irrigation systems (Section 9.3.3).

As indicated by Fernández and Cuevas (2010), the characteristics of a good indicator for irrigation scheduling can be summarised as follows. The variable sensed by the indicator must be closely related to production parameters of economic importance (e.g. crop yield and fruit quality). The indicator response to water stress must be quick (earliness) and marked (intensity). It must also be highly sensitive, provide high signal intensity (i.e. very responsive to water stress) and a relatively low coefficient of variation (reduced ‘noise’ between sampling points). The sensors must allow automated and continuous data collection and transmission and must be reliable, robust, inexpensive, easy to install, operate and maintain. Finally, the sensor outputs must be easy to interpret and suitable for use in automatic irrigation control.

Irrigation scheduling based on soil water determination has been widely reported. A soil water-based irrigation control system uses real-time information on the soil water status to bypass a time-based preprogrammed schedule or to maintain soil water content in the root zone within a specified moisture range that is optimal for plant growth. These two approaches are bypass and on-demand, respectively (Muñoz-Carpena and Dukes, 2005). Effectiveness of these soil sensor-based irrigation technologies has been successfully assessed on field crops (Meron et al., 2001; Wang et al., 2007; Muñoz-Carpena et al., 2008), on residential green areas (McCready et al., 2009) or on ornamental potted plants (Nemali and van Iersel, 2006; Burnett and van Iersel, 2008).

Occasionally, automated irrigation is not devoted only to schedule irrigation under standard conditions related to crop water requirements, but also to deal with sporadically occurring situations, such as irrigation to prevent frost damage, or cutting off sprinkler irrigation under severe wind.

As pointed out by Greenwood et al. (2010), the use of sensor-based procedures for soil water management varies with the type of crop and environment. Accordingly, shallow-rooted crops (e.g. potted plants, lawns or urban landscapes) could be managed with soil water sensors installed at one depth only whereas deep-rooted agricultural crops would require a vertical profile of sensors to allow irrigation according to the distribution of soil water down the profile. This ensures optimal use of the stored water in the soil profile and hence substantial saving of irrigation water. Figure 9.12 summarizes a sequential methodology described...
Nemali and van Iersel (2006) devised an irrigation system that used calibrated, dielectric moisture sensors, interfaced with a datalogger and solenoid valves, to irrigate a soil substrate to a set-point (volumetric water content, $\theta$) and maintain it close to that target for prolonged periods. The irrigation system was able to maintain $\theta$ for a long period within an acceptable range of the set-point, despite large variations in environmental conditions and plant size, which resulted in little or no wastage of water. Likewise, Blonquist et al. (2006) used a dielectric moisture content sensor to schedule irrigation for turfgrass. To accomplish this, they defined a preset threshold $\theta$ value ($\theta_c$, see also Section 9.2.4 and Equation 9.4) determined from both soil physical properties and crops to be grown:

$$\theta_i = \theta_{fc} - \text{MAD}(\theta_{fc} - \theta_{wp})$$  \hspace{1cm} (9.19)

where $MAD$ (equivalent to $p$ in Section 9.2.4) is the management allowed depletion value (Cuenca, 1989). The $MAD$ is the fraction of $TAW$ that is allowed to be removed from the plant root zone before irrigation is required. Cuenca (1989) reported typical $MAD$ values as 0.33 for shallow-rooted, high value crops; 0.5 for medium-rooted, moderate value crops and 0.67 for deep-rooted, low value crops.

The irrigation controller was set to irrigate each day for a specific time and duration, but with the only possibility to open the solenoid valve when $\theta$ measured by the sensor was below $\theta$. Relative to irrigation recommendations based on evaporation estimates, the $\theta$
sensor-based system applied approximately 16% less water, while relative to a fixed irrigation rate of 50 mm week\(^{-1}\) the system applied approximately 53% less water. Many research papers have also documented the enhanced irrigation efficiency that can be attained in many field crops by means of soil water sensors without any noticeable detrimental effects on crop productivity. A few examples are those published by Zotarelli et al. (2009) and Muñoz-Carpena et al. (2008), who found that soil moisture sensor-based irrigation systems significantly reduced crop water requirements, water percolation and nitrate leaching from tomato plants.

When deficit irrigation strategies (Section 9.5.1) are used, monitoring the soil (or the plant) water status becomes even more critical for minimisation of risks. These irrigation strategies normally exploit some plant physiological traits or sensitivity differences to water stress of specific phenological stages to enhance water use efficiency (Fereres and Soriano, 2007). However, the fulfilment of the desired crop response under these strategies usually requires very precise irrigation water management, since the crop response may dramatically deviate from the expected response in a narrow range of soil moisture availabilities.

### 9.5.3 Irrigation and salinity

A major consequence of irrigated agriculture is the risk of increased soil salinity and its impact on the sustainability of the irrigation. Although salinization may be driven by different factors (Rengasamy, 2006), irrigation-associated salinity represents a major limitation to agriculture in all semi-arid regions. All irrigation waters contain salts and, as water evaporates, salts concentrate in the soil profile and must be displaced below the root zone before they reach a concentration that limits crop production. The salinization process on irrigated land may be accelerated under conditions of poor quality irrigation water, low hydraulic conductivity of soil layers, as found in heavy clay and sodic soils, and high evaporative conditions (Rengasamy, 2006). Under these circumstances, some of the irrigation water losses are unavoidable and are needed to maintain the salt balance, although they can be minimized with efficient irrigation methods and by appropriate management (see Chapter 7).

Soil salinization may be a major drawback in deficit irrigation (DI) programs (Section 9.5.1), where irrigation water use is reduced and the hazard of increased salinity due to decreased leaching is amplified. Therefore, the question as to whether DI can be used as a tactical measure to reduce irrigation water use over long time periods when supplies are limited by droughts or other factors is currently unresolved (Fereres and Soriano, 2007). More research in this field is required to investigate the sustainability of DI (and especially when low quality irrigation water, such as reclaimed water, is used) and to what extent it can contribute to the permanent reduction of water use in irrigated systems.

The mechanisms by which saline soil water affect plant growth can be divided into two phases (Munns, 2002). The first phase of the plant growth response is due to the osmotic effect of the salt in the soil solution, which reduces the ability of the plant to take up water and generates in the plant a similar response to that under water stress. The second phase takes place if excessive amounts of salt enter the plant tissues, since they will eventually rise to toxic levels in the older transpiring leaves, causing premature senescence and hence a reduction in the assimilate synthesis that may further limit growth (Munns, 2002). The ability of plants to tolerate salt stress and to maintain (or to improve) WUE when irrigated with saline water or grown in salinized soils is highly variable. Katerji et al. (2003) studied the
plant response to increasing soil salinity in ten species. They identified two groups of species; those which are salt-tolerant and that maintain, or slightly improve, their WUE with an increase in salinity (wheat, sunflower, potato, maize and sugar beet) and a second group of species (tomato, lentil, broad bean and chickpea) which are salt-sensitive and reduce their WUE with the deterioration of water quality (Figure 9.13). These results emphasise the importance of carefully choosing the best species to be grown in saline soils or with saline irrigation water. Indeed, breeding programs for salt tolerance are deemed to be an important way to cope with salinity in dry-land agriculture (Munns et al., 2006). Another potential risk of using poor quality irrigation water is the ‘sodium hazard’, leading to sodic soils which cause a decline in the structure of fine-textured soils, leading to poor infiltration of water and thus to salinization (through evaporation of irrigation water at the soil surface) and waterlogging (Munns et al., 2006). This underlines the close link existing between the processes of soil salinization, waterlogging and evaporation.

Figure 9.13  Relative (with respect to the non-saline treatment) water use efficiency (WUE) of different crops as a function of soil salinity during the crop season [ECe: electrical conductivity of the saturated paste]. Redrawn from Katerji et al. (2008) who derived the relationships from data obtained by Katerji et al. (2003). With kind permission from Elsevier.
9.5.4 Managing soils to improve agronomic water use efficiency

As explained in Section 9.3.1, an increase in WUE can be accomplished by managing plant-related traits such as improvement of crop transpiration efficiency or by partitioning more of the biomass into the harvested product. Changes in WUE achieved in this fashion are invariably related to choice of plant variety or manipulation of a plant’s phenological stage. This section will focus on agronomic (soil) management practices that maximise transpiration, \( T \), by reducing the other losses (\( E \), \( D \) and \( R \), see Equation 9.1) to increase WUE, as illustrated by Equation 9.11, together with techniques that aim to improve soil water availability for crop transpiration (either by affecting the water regime or the soil’s water holding capacity).

Evapotranspiration is a component of the energy balance (Equation 9.15), and a separate balance can be drawn up for the crop and soil surface component. Hence, any alteration in the available energy (e.g. through shading) will affect evapotranspiration (see also Equation 9.18). The aerodynamic resistance, \( r_a \), also determines evaporation (again, see Equation 9.18) and its value is affected by the roughness length: the smoother the soil surface the lower the evaporation.

Hatfield et al. (2001) review how soil-management practices can improve WUE. They conclude that soil management affects the process of soil evapotranspiration by modifying the available energy, the available water in the soil profile, or the exchange rate between the soil and the atmosphere. Their literature survey revealed a large variation in measured WUE across a range of climates, crops and soil management practices. It is possible to increase WUE by 25 to 40% through soil-management practices that involve tillage. Overall, precipitation use efficiency can be enhanced through adoption of more intensive cropping systems in semi-arid environments and increased plant populations in more temperate and humid environments.

Soil tillage is a powerful tool to increase plant water use by alleviating soil physical constraints that hamper root growth. For example, ‘deep ripping’ can increase the depth of rooting and hence improve the amount of water available to the crop, when poor soil structure physically impedes root growth. This technique will only be useful if water infiltrating the soil can actually reach these deeper layers. Rotations, alternating deep-and shallow-rooting crops, can lead to ‘biopores’ for a subsequent crop (Turner, 2004). The use of minimum tillage or conservation tillage involves leaving crop residues on the surface while controlling weeds by herbicides rather than tillage. This approach also allows earlier planting as delays resulting from using tillage to remove weeds are reduced. Next, the seed for the follow-on crop is sown with minimum disturbance of the soil surface. This practice has led to reduced losses of water by soil evaporation and increased yields (Unger, 1978; Cornish and Pratley, 1991; Stewart and Robinson, 1997; Turner, 2004). However, the greater retention of incoming rainfall through minimum tillage may also increase water losses through deep drainage that are detrimental in a landscape in which secondary salinity can develop (Turner, 2004).

Total available soil water (\( \theta_{pc} - \theta_{pwp} \)) in the shallow subsoil can be improved by adding soil organic matter (SOM). Adding SOM enhances water retention and hence will raise \( \theta_{pc} \) because of its hydrophilic nature and its beneficial effects on soil structure through increased soil aggregate formation and aggregate stability (Oades, 1984; Huntington, 2006). This leads to a larger amount of storage pores compared to transmission pores (> 50\( \mu \)m) and residual pores (<2\( \mu \)m). These changes in the physical properties also result in higher infiltration rates (Franzluebbers, 2002) and lower susceptibility to erosion (Wilson et al., 2009).

As already discussed in Section 9.3.2, increased water use by the crop itself via transpiration at the expense of water loss by weeds or from loss terms in the water balance equation
Soil water and its management

(soil evaporation, deep drainage, surface runoff, or lateral subsurface flow) will improve water use efficiency. Ways to achieve this include earlier planting to more closely match rainfall and the use of fertilizers to maximise yield and hence decrease soil evaporation losses. Similarly, high plant density and more even planting increases crop-water use and reduces soil evaporation, but only where ample soil water is available from rainfall or supplement irrigation. When crops rely on water stored in the root zone, low planting densities are preferred to ensure a larger availability of water per plant (Turner, 2004).

Fallowing land to conserve moisture is another widely adopted practice to improve yields of subsequent crops in water-limited environments. The benefits of this practice vary with season, soil type and management of the fallow land (O’Leary and Connor, 1997; Stewart and Robinson, 1997).

Another way to optimise soil water availability is by practising intercropping (the agricultural practice of concurrently cultivating two or more crops in the same field) or agroforestry (the combination of trees or shrubs with crops). Agroforestry has been promoted as a possible means of increasing the productive use of rainfall in water-limited environments, by using the water that is usually inaccessible to conventional cropping systems (i.e. soil water reserves that are below the crop rooting zone and/or rainfall occurring outside the normal cropping seasons). However, any potential increase in rainfall utilisation by agroforestry systems, when compared to either woodlots or open crop fields, must be offset against the evaporation of rainfall intercepted by the likely greater canopy size (tree and crop) and not therefore available for crop growth (Jackson, 2000). On the other hand, increased shading of the soil by the trees will decrease soil evaporative losses (Wallace et al., 1999).

Another practice aimed at reducing the amount of soil water lost is the covering of the soil surface by different types of materials; this is called mulching. The main aim is to prevent soil evaporation, either by blocking the soil pores entirely (e.g. by using plastic) or by creating a hydraulic discontinuity through using a porous medium with large pores. Many materials have been used as mulch, inorganic ones such as plastic film, sand and gravel, and organic ones such as crop residue, straw and compost. Mulch at the soil surface affects hydrological processes involving rainfall interception, infiltration, evaporation and dew deposition (Li, 2003). A mulch layer therefore changes the soil’s water balance. Mulching also affects the surface radiation budget (Lei et al., 2004) an effect largely related to the colour and nature of the mulch, which affects albedo (Kemper et al., 1994; Jonathan et al., 1998; Jalota et al., 2001; Anikwe et al., 2007). Hence, a mulch affects the energy available for evaporation (Novaka et al., 2000; Lei et al., 2004), thereby indirectly affecting the soil’s water balance. Through these combined effects a mulch layer will also affect a soil’s temperature regime. The influence of mulching on evaporation depends on the type of mulch, its texture (Xie et al., 2006) and layer thickness (Kemper et al., 1994). Its effectivity also depends on the stage of evaporation (Li, 2003).

Figure 9.14 shows that the evaporation from bare soil was higher compared to that from mulched soil during the first 2 days after irrigation (first stage of drying), after each irrigation event. This was caused by the mulch layer creating a hydraulic discontinuity, forcing the water to move in the vapour phase only. However, from day 3 (after irrigation) onwards, the evaporation from mulched soil was equal to or greater than that for bare soil. The reason for this is that the mulch-covered soil lost less water by evaporation during the first stage so that the moisture content in the mulched soil was higher than that in bare soil, thereby increasing hydraulic conductivity. Also, the bare soil forms a thin dry layer (a ‘self-mulching’ layer) hindering evaporation through its poor hydraulic properties. Similar observations were reported by other researchers when they used various mulch materials to reduce the
evaporation from bare soil (e.g. Modaihsh et al., 1985; Groenevelt et al., 1989; Van Wesemael et al., 1996; Mellouli et al., 2000).

Mulching can also change aerodynamic resistance and cause an enhanced turbulent transport of water vapour (and sensible heat) due to the relatively large roughness length of mulch layers such as gravel and plant residues (Nachtergaele et al., 1998; Xie et al., 2006). Additionally, mulching also avoids structural soil damage through water drop impacts, which can cause slaking and crusting of soil (and hence affect infiltration rate), as well as erosion (see Section 9.6.1).

A final point needs to be made here. Most of this section has focussed on how to increase and/or conserve the soil moisture available in the pore space of the root zone, for situations when the GWT is situated so far below the root zone that it does not influence the available soil water. In many cases, however, shallow water tables can be a valuable resource to provide crops with water (Kahlown et al., 2005; Qureshi et al., 2010). This will result in reductions in irrigation and hence improvements in (irrigation) water use efficiency. Ideally, this practice should be used only in areas where shallow water tables are a low salinity risk or where the local hydrology results in net recharge (e.g. Hurst et al., 2004).

9.6 Water management to alleviate excess water

9.6.1 Causes of excess water and remediation solutions

Root zone saturation can be caused by a variety of natural and anthropogenic factors, often occurring in combination. Natural factors comprise the climate, soil factors, plant factors, topography, geomorphological and landscape factors (Section 9.2.1). Examples of
anthropogenic influences that can cause suboptimal values of root-zone aeration, which often adversely affects crop growth, are over-irrigation and drainage of uplands resulting in flooding of lower-situated lands (D’haeze et al., 2003). Also, clearing of native vegetation for annual crops and pastures is recognised as a major cause of waterlogging (and secondary salinity), for example in southern Australia (Turner and Ward, 2002).

Focussing on soil factors, excess water in the root environment can be the result of a number of causes that are often attributable to poor physical soil quality. Firstly, below-optimal final infiltration capacity, combined with high-intensity rainstorms or high irrigation rates, may result in temporary ponding. Infiltration capacity largely depends on the soil texture type and initial water distributions (Giraldez and Sposito, 1985; Chu and Mariño, 2005). Low infiltration capacities can be the result of destruction of the surface tilth, generally caused by the mechanical impact of rain-drops on structurally unstable soils. The resulting finer particles fill up the larger pores or cause a crust to form (see, for example, Stroosnijder and Hoogmoed, 1984). Alternatively, runoff or runon leads to water erosion, thereby clogging up larger pores due to the deposition of finer particles. Yimer et al. (2008) found that changes in soil structure caused by surface soil compaction because of tillage and animal trampling, coupled with a smaller soil organic carbon content, caused a decline in infiltration capacity and soil moisture content after conversion of forest to cultivation and grazing in the highlands of Ethiopia.

Even if the infiltration rate is adequate, low hydraulic conductivity in the upper soil layers may cause poor internal drainage or slow redistribution, causing waterlogging within the soil profile, resulting in poor aeration around roots. Water can move through cracks or discontinuities in the soil structure, but there is reduced movement through the pores within the soil matrix. Drainage can only take place around the structural elements, not through them. In extreme cases, impermeable layers such as a plough pan, may lead to a perched water table, as discussed in Section 9.2.1.

Excess water can be alleviated by preventing or correcting low infiltration capacity by, for example, adding organic matter to the soil or covering the soil with crop residue to prevent damage to the soil structure. If insufficient coarse pores are available in the surface layers (i.e. the soil has a low hydraulic conductivity), cultivation can remedy this. A compacted subsoil can be ameliorated by deep tillage (using deep ploughs, deep cultivators or subsoilers), thereby destroying compacted layers or subsurface pans, although results with respect to crop yield are variable (e.g. Carter et al., 1996).

Alternatively, a rising groundwater table may give rise to anaerobic conditions. The extreme case is when the actual groundwater or the capillary fringe can reach the surface. A related problem here is secondary salinity; in southern Australia, for example, the replacement of deep-rooted perennial native vegetation with shallow-rooted annual crops has resulted in a rising water table and the development of a major secondary salinity problem (Barrett-Lennard, 2002; Turner and Ward, 2002). Establishment of tree belts, drains and perennial pasture, or a combination of these, will positively affect the water balance and ground water level and thereby alleviate water logging and secondary salinity in the rejuvenated landscape of southwestern Australia, while maintaining crop production at near-current levels (Turner and Ward, 2002). Rundle and Rundle (2002) presented a case study of farm-based solutions to water logging and secondary salinity in southwestern Australia. The plan included the grading of interceptor drains and the locating of dams along the drains to hold water for the dry summer months. Belts of three or four species of trees were planted on the down-slope side of the drains after leveling the spoil from the drain and deep-ripping the soil.
Noorduijn et al. (2010) assessed the positive effects of alley-farming agroforestry on groundwater levels through a 13-year alley farming trial incorporating different combinations of belt width, alley width and revegetation density. However, despite the reintegration of perennial plants back into farming systems in many regions the salinisation processes continues because the available perennial options are generally less profitable than annual crops. A possible solution is the partial restoration of productivity of saltland through revegetation with salt tolerant halophytes (Barrett-Lennard, 2002).

The highest permissible levels of the GWT below the surface depend on the crop and the soil type (as the latter will determine capillary rise). For agricultural crops, the depth of the GWT is best below 90 cm, although pastures and meadows can cope with higher water tables. The depth of the GWT becomes particularly important in the case of saline groundwater: the capillary fringe should then be well below bottom of the crop’s root system. For certain ecologically important plant species, like those growing on SSSIs (Sites of Special Scientific Interest), high water tables are essential, and drainage is undesirable (Kalusová et al., 2009; Araya et al., 2010).

Overbank flooding may also give rise to anaerobic conditions and hence adversely affect survival and growth of riparian species in floodplain ecosystems, for example during floodplain restoration. Natural resource practitioners should therefore take into account flood frequency and duration in their planting plans to increase the likelihood of restoration success (McCurry et al., 2010).

If the deeper subsoil is too impermeable to allow water to percolate below the crop root zone quickly enough, the field must be drained artificially. The efficiency of the drainage system can often be improved by deep tillage, as this increases the number of vertical waterways and the water can pass through from the surface to the drains. The success of the drainage system depends on the stability of the soil structure reflected in the temporal stability of the cracks and fissures.

### 9.6.2 Agricultural drainage systems

Excess water that cannot be alleviated by tillage or other soil-management practices can be remediated by artificially improving the natural drainage of the land. This is achieved by installing ‘Agricultural drainage systems (ADS)’. These are artificial systems that enhance water flow from the land, so that agriculture can benefit from the subsequently reduced water levels, via increased crop yields, improved diversification and increased intensification (where required), effective use of inputs, such as fertilisers, and improved mechanisation (Oosterbaan, 1994). ADS have the purpose of either reducing the flow of water over the soil surface (surface flow) or through the underground (inter- and groundflow), which leads to a distinction between ‘surface drainage systems’ and ‘subsurface drainage systems’.

A surface drainage system is applied when ponding frequently occurs on the soil surface, whereas a subsurface drainage system is applied when the water logging occurs in the soil (i.e. shallow GWTs or perched water tables; Oosterbaan, 1994). When both surface and subsurface water logging occur, a combined surface/subsurface drainage system is required.

This section briefly considers the configuration of sub-surface drainage systems, as surface systems are outside of the scope of this chapter. Subsurface field drainage systems consist of horizontal or slightly sloping channels made in the soil. They can be open ditches, buried pipe drains, or mole drains into and within which groundwater flows because of hydraulic gradients existing in the soil. The flow rate is affected by a number of environmental and technical factors (TANJI and KIELEN, 2002).
From a soil physical/hydrological point of view, the hydraulic conductivity of the soil as well as the configuration of the GWT and the hydraulic pressure of the groundwater are important. Furthermore, the rate of groundwater recharge by internal drainage or by horizontal groundflow from an external source needs to be considered. Physical characteristics of the drainage system that play an important role are the location of the drains (in relation to the groundwater level and the soil surface), the diameter of the drains, the slope of the drains and the spacing between the drains (Hillel, 1971; Van Schilfgaarde, 1974; Farr and Henderson, 1986; Oosterbaan, 1994).

**Direct and indirect effects of agricultural land drainage**

ADS will have an effect on the physical, chemical, biological and hydrological soil properties and variables, as well as on the soil water and energy balance components. These will, in turn, affect soil temperature and moisture content.

The direct effects of installing a drainage system in waterlogged land are: a reduction in the average amount of water stored on or in the soil (inducing drier soil conditions and reducing water logging) and an increased flow (discharge) of water through the system (Oosterbaan, 1994). These direct effects trigger a series of indirect effects, which can be either positive or negative. Some examples are given in the rest of this section (largely based on Oosterbaan, 1994).

Positive effects include increased aeration of the soil so that crops can develop a deeper root system. This can give winter-sown crops an advantage in spring as it will allow them to access a larger available soil volume for nutrient and water uptake. On heavy soils, the increased transpiration, due to increased soil water, can dry out the soil more and increase the amount of cracking, so that better drainage occurs during following rainstorms. In this way, vigorous crops on heavy land can help create and stabilise the natural drainage system. It will also improve anchorage of trees and crops (see Chapter 8).

Having drier top soils will also reduce the risk of frost heaving (which causes crop damage) because soils need to have a considerable amount of soil water to heave. Furthermore, drier top soils will generally result in a more stable soil structure, which will improve infiltrability and hence decrease surface runoff.

Agronomic advantages are higher availability of nitrogen in the soil due to increased nitrification (Feddes and van Wijk, 1977), higher and more diversified crop production, better workability of the land (Wind and Buitendijk, 1979), earlier planting dates due to increased spring temperatures, reduction of peak discharges by an increased temporary storage of water in the soil and removal of salts or other harmful substances from the soil due to an increased discharge (Section 9.5.3). This drainage water can be used for various purposes including irrigation (Sharma et al., 1990).

Negative effects comprise: (1) increased decomposition of organic matter, because of increased temperatures and reduced soil moisture contents affecting the root and microbial respiration; (2) soil subsidence (Todd, 1980); (3) acidification of potential acid sulphate soils (AARD and LAWOO, 1992); (4) increased risk of drought through reduced capillary rise from the ground water table (Section 9.6.1); (5) ecosystem damage if lowering of the water table adversely affects plant species and (6) problems caused downstream by the discharge of salty or otherwise polluted drainage water.

The indirect effects of drier soil conditions on weeds, pests and plant disease can be both positive and negative; the net result depends on the environmental conditions.
9.6.3 Design and implementation of drainage systems

Drainage should only be undertaken with clear agricultural aims (so-called agricultural object factors) in mind. In the context of this chapter: increasing crop yield (e.g. by extending the growth season or by being able to introduce new crops) or enabling soil tillage.

To achieve the agricultural targets requires a priori knowledge of criterion factors such as the extent of the saturation conditions and the soil salinity. Criterion factors are factors that are affected by the drainage system and at the same time influence the object factors. A more specific description of a criterion factor is called a criterion index, such as (1) the average depth of the watertable during the cropping season; (2) crop-season average salinity levels of the root zone and (3) the minimum number of days that the soil is workable during a critical period.

Decisions on the design and implementation of ADS are also influenced by environmental factors such as irrigation, precipitation, soil hydraulic properties, topography and aquifer conditions. These factors must be specified as environmental indices, such as the average seasonal rainfall, the maximum daily rainfall, the mean and vertical distribution of hydraulic conductivity and the variation in hydraulic conductivity with depth in the soil.

For further reading on land drainage for agriculture see Hillel (1971), Van Schilfaarde (1974), Wither and Vipond (1974), Farr and Henderson (1986) and Oosterbaan (1994). Reliable software, such as EnDrain to perform calculations on horizontal subsurface drainage systems in agriculture can be downloaded from http://www.waterlog.info/software.htm. SaltMod is a mathematical, numerical simulation model that describes the relations between crop rotation, water balance components, groundwater levels, soil salinity and subsurface drainage and SahysMod combines Saltmod with a model of ground water flow and hydraulics to account for large spatial variation through a network of polygons.

References


10 Plant-induced changes in soil processes and properties

Philippe Hinsinger

Institut National de la Recherche Agronomique, UMR Eco&Sols, Montpellier, France

10.1 Introduction

Since the early beginning of pedology with the work of Dokuchaev in the nineteenth century, vegetation has been recognised as one of the five major factors explaining the spatial distribution of soil types on planet Earth, together with the climate, parent material, topography and time. However, at such a scale, the biogeography of the vegetation is largely correlated with climatic constraints, so one may wonder whether these two factors are equally needed to explain pedogenesis. While temperature and the amount of water circulating through the system have long been recognised as key drivers of soil formation and biogeochemical cycles, the direct implication of plants has long been overlooked by both soil scientists and geochemists (Hinsinger et al., 2001; Lambers et al., 2009). There is little doubt that soil formation results primarily from a combination of physical and chemical processes leading to the breakdown and weathering of rock particles. The etymology of the word ‘weathering’ is explicit in this respect, and numerous works in experimental geochemistry (e.g. Bockheim and Gennadiyev, 2009) have demonstrated that purely chemical processes could explain the observed extent of weathering and the formation of typical secondary minerals present in the soil matrix such as clay minerals and metal oxides. However, their nature and amount depend on the combined effects of time, temperature and flux of solute transfer; the latter two being obviously related to the climate. Without plants, though, and especially higher plants with deep root systems, would there be any deep, differentiated soil at all or merely a thin regolith (Lambers et al., 2009)? One feature that distinguishes soil from the loose layer of rock material called regolith is the occurrence of a small amount of organic matter, which primarily originates from plants.

The first section of this chapter describes how higher plants have played a key role as the primary source of organic matter in soils and more generally in the biogeochemical cycle of carbon (C) on planet Earth. Compared with earlier life forms and primary producers such as photosynthetic bacteria, algae, lichens and bryophytes, higher plants have this special feature of deep root systems which have direct interactions with deeper soil horizons and
Soil Conditions and Plant Growth

Soil Conditions and Plant Growth can be responsible for a direct input of large amounts of C-rich compounds deep below the soil surface. A special focus of this first section will be on the role of the below-ground processes involved in the biogeochemical cycle of C.

The soil volume surrounding the roots of higher plants is distinct from the bulk soil as a consequence of a number of root activities. It is indeed unique in the intensity and nature of biophysical, biogeochemical/biochemical and biological/ecological processes that occur in the so-called rhizosphere (Darrah, 1993; Hinsinger, 1998; Hiltner, 1904 in Hartmann et al., 2008; Hinsinger et al., 2009). Plant roots are major drivers of such processes but these also largely rely on the indirect effects of microbial communities, which occur and are stimulated in the rhizosphere as a consequence of the rhizodeposition of C-rich (and sometimes N-rich) compounds (Jones et al., 2009). These combined effects of roots and rhizosphere microorganisms ultimately play a key role in the biogeochemical cycles of protons, nutrients and other elements. This will be addressed in two subsequent sections of this chapter, which concentrate on two major processes involved in soil formation and evolution, that is, soil acidification and weathering. While rhizosphere processes are often considered to occur at the microscopic scale of the soil–root interface, various parts of this chapter will attempt to demonstrate that, over time, they have a major impact at the global scale. How these processes affect plant growth either positively or negatively, and their ultimate consequences for ecosystem functions and services, will also be briefly addressed.

10.2 Roles of plants in carbon cycling: deposition, burial and emission

Higher plants are major contributors to photosynthesis among autotrophic organisms on planet Earth, especially in terrestrial ecosystems. These primary producers assimilate atmospheric CO₂ which is thereby incorporated in a whole range of organic compounds that are ultimately supplied to the soil directly as litter or via the loop of primary and secondary consumers, before being recycled back to the atmosphere by decomposers, a process which largely occurs in the soil. This is demonstrated, for example, by the comparison conducted by Johnston et al. (2009) and Hirsch et al. (2009) in the long-term Highfield ley-arable experiment at Rothamsted, United Kingdom. This long-standing grass ley was partly ploughed in 1949 and subsequently either managed by arable rotation or maintained as a bare fallow by regular tillage since 1959. Johnson et al. (2009) reported a significant decline in soil organic C over time, when grass had been converted to arable cultivation, and even more so in the bare fallow. The C content in the topsoil of the grass-treatment soil (54 g C kg⁻¹) was much higher than the arable soil (16 g C kg⁻¹) and the bare-fallow soil (10 g C kg⁻¹). After almost 50 years without significant plant C input, the bare-fallow soil contained ~30 t C ha⁻¹, a value <8% of that in the original grassland soil (Goulding et al., 2010). Hirsch et al. (2009) also showed that the bare-fallow soil supported smaller and less diverse microbial and mesofaunal communities than the planted soil treatments, as a consequence of the absence of an annual input of fresh organic residues of plant origin, compared with the arable and grass treatments which provided a richer C substrate. The changes that occurred were large in a short period of time (50 years) relative to the time needed for pedogenetic development of such soils (thousands or tens of thousands years). It shows that while soil evolution occurs over long time spans, it can also be subject to dramatic changes on a short time scale. The most extreme case of short-term change is that induced by topsoil erosion or landslides, that plant cover tends to restrict (Stokes et al., 2009). Lessons drawn from the
analyses of soil chronosequences are of special interest for understanding the long-term patterns of soil development. Richardson et al. (2004) have clearly illustrated with the Franz Joseph chronosequence in New Zealand that the organic C content of the soil increases from close to zero to a few percent in the first hundred of years, then builds further to reach values of up to 15% in the first millennium of soil development and then finally steadily decreases to values between 5% and 10% over the period 5000 to 120 000 years (Figure 10.1).

10.2.1 Photosynthesis, respiration, litter deposition and C sequestration in plants and soils

Global estimates of components of the C cycle are currently the following (Schlesinger, 1997): photosynthesis in terrestrial ecosystems (gross primary production, GPP) amounts to $120 \times 10^{15}$ g C year$^{-1}$ while C fixation by primary producers in the ocean amounts to $48 \times 10^{15}$ g C year$^{-1}$. In comparison, the sizes of the major reservoirs of terrestrial C are as follows: $550 \times 10^{15}$ g C in the standing vegetation, $300 \times 10^{15}$ g C in the litter and $1200 \times 10^{15}$ g C in the soils, while atmospheric C represents $780 \times 10^{15}$ g C. This suggests that one-sixth of the atmospheric CO$_2$ is recycled yearly through photosynthesis by terrestrial vegetation (GPP) (Trumbore, 2006). The net primary production (NPP) of terrestrial ecosystems amounts to $58 \times 10^{15}$ g C year$^{-1}$ (Field et al., 1998), out of which crops amount to a rather small proportion ($8 \times 10^{15}$ g C year$^{-1}$; Table 10.1), compared with savannas ($17 \times 10^{15}$ g C year$^{-1}$) and tropical rainforests ($18 \times 10^{15}$ g C year$^{-1}$). Such fluxes are worth comparing to that reported for soil respiration, $75 \times 10^{15}$ g C year$^{-1}$, which is actually larger than both terrestrial NPP and litterfall (Schlesinger and Andrews, 2000). Soil respiration is greater than NPP because a substantial part of soil respiration is related to respiration of roots and mycorrhizae (deriving from GPP). While above-ground C pools and fluxes are well documented, estimates of below-ground pools and fluxes exhibit greater variability (De Deyn et al., 2008). Distinguishing between root respiration and heterotrophic respiration (due to soil microbial and faunal activity) is especially difficult.

Looking back to geological times, the rise of land plants and particularly the emergence of vascular plants with greater photosynthesis and transpiration capacities (related to their innovative and extensively developed architecture) during the Devonian period is now recognised for its central role in the geochemical history of the Earth and its atmosphere (Beerling and Berner, 2005). The early forms of land plants at the beginning of the Devonian period looked like bryophytes and other plant life forms which had rhizomes but no roots and no large planate leaves (Beerling et al., 2001). Fossil records suggest that the vegetation formed a short chlorophyllous carpet made of short shoots rising from other stems running along the surface (Kenrick and Crane, 1997). In contrast, tall vegetation with megaphyll planate leaves became abundant during the late Devonian/early Carboniferous (Beerling et al., 2001). Interestingly, such evolution coincided with a major drop in atmospheric pCO$_2$, which was then much above present-day levels. It dropped from 3–4000 ppmv to less than 1000 ppmv, possibly as a result of increased photosynthesis and removal of atmospheric CO$_2$ by the megaphyll vegetation (Beerling and Berner, 2005). This fixed C was progressively buried as organic matter in soils. Most of the biogeochemical models account for organic C input as related to above-ground parts of the vegetation and implicitly assume that the below-ground portion mostly contributes to an additional input of C that is mostly concentrated in the topsoil. This is derived from the decay of above-ground parts which fall on top of the soil (trunks, branches, shoots of herbaceous plants, leaves and fruits).
Figure 10.1 Temporal evolution of soil properties along the Franz Josef soil chronosequence. Data are means of five plots at each site with the exception of site 7 where \( n = 4 \). Box plot symbols: horizontal lines are the median; shaded bars give 25% and 75% percentiles. From Richardson et al. (2004). Reprinted with kind permission from Springer Science and Business Media (Springer-Verlag)/International Association for Ecology.
10.2.2 Allocation of C to below-ground parts of the plants

A substantial proportion of the GPP is allocated to below-ground parts of the vegetation. According to Jones (1993), a median value of about 60% of the fixed C is used in the shoots (for shoot biomass production and respiration), while 40% is allocated to below-ground parts (Figure 10.2). For cereals, it has been estimated that over the whole growing season, the total amount of C allocated to below-ground parts ranges from 150 to 220 g C m\(^{-2}\) year\(^{-1}\) (Kuzyakov and Domanski, 2000). In perennial species, especially when exposed to limitation of below-ground resources (water or nutrient shortage), the allocation of fixed C to the root compartment can exceed 50% of the total C assimilated. In addition, root litter is usually a more recalcitrant C pool than above-ground tissues (shoots), thereby contributing to C sequestration in soils (De Deyn et al., 2008), as will be further discussed in this section. In addition, in contrast to the above-ground litter which is predominantly recycled back to the topsoil (although faunal activities can result in burial of such organic matter at depth), C allocation to roots contributes to a direct C input at depth, that is related to the vertical distribution of roots in the soil profile.

Meta-analyses of the vertical distribution of roots in soils (database of 250 root studies) showed that in most biomes more than 80–90% of the root biomass was found in the top (0–30 cm) soil (Jackson et al., 1996). This was further substantiated by a broader dataset of 475 soil profiles from 209 different geographical locations (Schenk and Jackson, 2002). However, Schenk and Jackson (2002) stressed that many of these studies underestimated root biomass as most authors did not investigate deep soil horizons. They stressed that the proportion of roots found in the topsoil decreased as the depth of soil investigated increased. Canadell et al. (1996) compiled published information about maximum rooting depths and showed that roots commonly grew as deep as 1–2 m in most crops. The biome average value for croplands and boreal forests was similar at about 2 m, while it extended to 2.6 m in temperate grassland, 2.9 m in temperate deciduous forests, 3.7 m in tropical deciduous forest, 7.3 m in tropical evergreen forest and 15 m in tropical grassland/savanna (Canadell et al., 1996). Some detailed studies in deep soil profiles have recently shown that a substantial proportion of root biomass occurs at depth and may thereby contribute a very significant contribution to deep soil C sequestration. Quantification of fine root dynamics down to a

<table>
<thead>
<tr>
<th>Table 10.1 Net primary production (NPP) in various terrestrial ecosystems.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tropical rainforests</td>
</tr>
<tr>
<td>Broadleaf and needleleaf forests</td>
</tr>
<tr>
<td>Broadleaf deciduous forests</td>
</tr>
<tr>
<td>Needleleaf evergreen forests</td>
</tr>
<tr>
<td>Needleleaf deciduous forest</td>
</tr>
<tr>
<td>Savannas</td>
</tr>
<tr>
<td>Perennial grasslands</td>
</tr>
<tr>
<td>Broadleaf shrubs with bare soil</td>
</tr>
<tr>
<td>Tundra</td>
</tr>
<tr>
<td>Desert</td>
</tr>
<tr>
<td>Cultivated agroecosystems</td>
</tr>
</tbody>
</table>

All values are expressed as 10\(^{15}\)g C, the total amounting to 56.4 10\(^{15}\)g C.
Source: Data compiled from Field et al. (1998). Reprinted with permission from AAAS.
depth of about 6 m in various Amazonian ecosystems showed a large input of C at depth (Trumbore et al., 2006): out of a total of 91 C m\(^{-2}\) year\(^{-1}\) in a degraded pasture and 78 g C m\(^{-2}\) year\(^{-1}\) in a mature forest, about 26% and 35%, respectively, of fine roots were found at 1–6 m depth while 40% and 30%, respectively, were found in the top 0.1 m. Even more impressive are the data obtained in fast-growing eucalypt plantations: in the first year of growth in the Congo, 16% of fine roots were found in the top 0.25 m, while 84% colonised the soil profile almost homogeneously to 3 m depth (Bouillet et al., 2002). In tropical crops such as sugarcane grown on poor, deeply weathered soils in Brazil, Battie Laclau and Laclau (2009) also reported large densities of fine roots to several metres; this was more marked in rainfed than irrigated conditions (Figure 10.3). In temperate crops, much larger root length densities are found in the topsoil for most species, and root length density often decreases exponentially with soil depth, as obtained in cauliflower and wheat in Figure 10.4 (Gregory, 2006a). Some crop species, however, exhibit a different pattern of root distribution with almost constant root length density at depth, as shown for oilseed rape and sugarbeet in Figure 10.4a. There is, thus, a need for more quantitative data about root distribution (both in terms of length and biomass) as this is crucial to (1) the exploitation of soil water and nutrient resources by plants and (2) the C sequestration occurring in the subsoil. Indeed, at a global level, only half of the total C sequestered in soils is found in the topsoil. In a temperate pasture, it was shown that this may be linked to greater stability against mineralisation in the subsoil, possibly due to a smaller priming effect occurring at depth because of the smaller supply of fresh material than in the topsoil (Fontaine et al., 2007). This phenomenon might well be related to root litter being more recalcitrant than shoot litter which makes a large proportion of topsoil organic C (De Deyn et al., 2008).

Besides litter deposition (including root turnover), another major process has to be accounted for when considering below-ground processes involved in C cycling. Indeed, a
substantial portion of the C allocated by plants to their below-ground parts is released during the life of roots via the so-called rhizodeposition process (Nguyen, 2003; Jones et al., 2009). Using 14C labelling experiments, data compiled for a broad range of species (271 datasets comprising 250 datasets of 17 herbaceous species largely dominated by Poaceae, and 21 datasets for 4 woody species) by Jones et al. (2009) showed a median value for C recovered in soil residues of about 5% of fixed C, while the median value for rhizosphere respiration was about 12%. Given that about half of rhizosphere respiration is due to microbial respiration, that is, CO2 emission as a consequence of the breakdown of rhizodeposits by rhizosphere microorganisms (Kuzyakov, 2006), it can be estimated that rhizodeposition is in about 11% of the total amount of C fixed by plants. This represents about a quarter of the amount of C allocated to below-ground parts of the plants. For cereals, over the whole growing season,
rhizodeposition would thus amount to 40–60 g C m\(^{-2}\) year\(^{-1}\) (Jones et al., 2009). This is a significant amount as it represents about 1% of the stock of C stored in the top 30 cm of an arable soil, that is, 4380–5990 g C m\(^{-2}\), assuming a soil organic matter content of 2–3% and a bulk density of 1.2 Mg m\(^{-3}\) (Smith et al., 1997). Besides its contribution to delivering C to the soil, rhizodeposition has many functional implications, as described in the following section.

### 10.2.3 Ecological significance of the rhizodeposition of C compounds

As pointed by Hiltner in 1904 when he first coined the term ‘Rhizosphäre’, which has a Greek origin (see Hartmann et al. (2008) for an English translation of Hiltner’s seminal paper), root exudates have major implications for both plant nutrition and health, via the stimulation of microbial activities (Hartmann et al., 2009). For reviews of these issues refer to Dessaux et al. (2010) and Richardson et al. (2009) for plant nutrition issues and to Raaijmakers et al. (2009) for plant health issues. Instead of root exudates, this chapter will refer to rhizodeposits because this term is broader and includes root secretions, root cap cells and root exudates (Nguyen, 2003; Jones et al., 2009). Rhizodeposits are sources of energy for soil microorganisms whose growth and activities are considerably enhanced in the vicinity of living roots, resulting in the so-called rhizosphere effect (Curl and Truelove, 1986; Lynch, 1990; Hartmann et al., 2009). Soils are a major reservoir of biodiversity and host impressive densities of microbial species and cells, but many of these are only active provided that they are close enough to sources of fresh C substrates. This typically occurs in the rhizosphere as a consequence of rhizodeposition (Hinsinger et al., 2009), as well as in the drilosphere as a consequence of earthworm activities (mucus release for instance) or in

![Figure 10.4](image_url)
the detritusphere, as a consequence of the burial of plant (or animal) residues in the topsoil (Brown et al., 2000). The delivery of fresh C compounds such as rhizodeposits into the soil can enhance the decay of older organic C compounds (i.e. soil organic matter) via the so-called priming effect which occurs as a consequence of the stimulation of microbial activity (Kuzyakov, 2002). Rhizodeposition thus has complex effects on the fate of soil organic matter and the global C cycle (Jones et al., 2009).

Many rhizosphere microbial activities that are stimulated by rhizodeposits have positive consequences on delivering nutrients to plants. This is especially true for the production of enzymes by rhizosphere microorganisms, which contribute to the hydrolysis of organic molecules containing nutrients (e.g. N, P or S) that would otherwise not be bioavailable to higher plants, although these can to some extent also release such enzymes (Richardson et al., 2009). In addition, some microbial activities are also implicated in the fate of inorganic compounds, as illustrated by the well-documented case of P-solubilising bacteria or fungi. The release of inorganic and organic acids by rhizosphere microbes can result in a dissolution of inorganic P compounds such as Ca phosphates, a capacity which seems to be distributed among many microbial genera in the rhizosphere (Rengel and Marschner, 2005; Richardson et al., 2009; Jones and Oburger, 2011). Another important effect is that of plant growth-promoting rhizobacteria (PGPR), which can also improve plant nutrition by indirect effects. For example, the stimulation of PGPR activities as a result of rhizodeposition can have positive feedbacks on root growth via hormonal signals (Faure et al., 2009), which ultimately enables the plant to access a greater pool of soil nutrients (Richardson et al., 2009). In contrast, there is not much evidence of competition for nutrients between plant roots and rhizosphere microorganisms, which would result in negative consequences of rhizodeposition on plant nutrition, although this may well occur when nutrient availability in soils is low (e.g. for iron especially in calcareous environments (Robin et al., 2008) and for ammonium or phosphate). Other major groups of rhizosphere microorganisms that are responsible for improving nutrient acquisition are microsymbionts (e.g. N₂-fixing rhizobia and mycorrhizal fungi). However, instead of competing for rhizodeposits with the many other rhizosphere microorganisms, these directly tap C compounds in the root, in the intimacy of symbiotic organs, root nodules and mycorrhizae. This long-lived coevolution of higher plants and soil microorganisms has resulted in highly efficient strategies to acquire nutrients (Brundrett, 2002; Sprent and James, 2007; Franche et al., 2009; Lambers et al., 2009), which is especially documented for N and P. Symbiotic N₂-fixation by Rhizobia legumes is a major component of the terrestrial N cycle (Vitousek et al., 2002; Gruber and Galloway, 2008; Herridge et al., 2008) and an underestimated process contributing to soil formation, which is a unique source of N input as N is almost absent in most rocks and thus in parent materials of soils (Richardson et al., 2004; Parfitt et al., 2005; Figure 10.1). The long soil chronosequences such as the Laupahoehoe site in Hawaii and the Franz Josef chronosequence in New Zealand clearly show that in their early stages of soil development (young soils), terrestrial ecosystems are primarily N limited, while at later stages of development (old soils) they become P limited (Vitousek and Farrington, 1997; Richardson et al., 2004; Parfitt et al., 2005). The role of mycorrhiza in P nutrition is also crucial as most plant species are mycorrhizal and recent reports have shown that even plants which do not positively respond to mycorrhizal infection (i.e. which do not show any improved growth when mycorrhizal) largely rely on mycorrhizal hyphae to acquire soil P (Smith et al., 2009). Indeed, once the mycorrhiza is established, the fungus turns off the P transporters that are located mainly in root epidermal cells and root hairs and turns on specific plant P transporters located in cortical cells where arbuscules are present, thereby forcing the plant to use the
mycorrhizal pathway of P uptake, in exchange for C from the host plant (Rausch et al., 2001; Harrison et al., 2002; Smith et al., 2003; Bucher, 2006; Smith et al., 2009). Bypassing rhizodeposition, root symbionts nonetheless represent a significant C cost for the host plants (e.g. mycorrhiza in Figure 10.2 and Jones et al., 2004), and there are several reports of reduced growth as a consequence of mycorrhizal infection (Ryan et al., 2005; Smith et al., 2009), although such symbiosis results in positive effects in many cases.

For plant health, there is an extensive literature on both positive and negative effects of the stimulated microbial populations in the rhizosphere (reviewed by Raaijmakers et al., 2009). Indeed, rhizodeposition can attract both the ‘good’ (beneficial) and the ‘bad’ (detrimental, e.g. pathogenic) microorganisms, as they are competing for the same C sources, especially when the availability of fresh C compounds is limited. Understanding how rhizodeposition can alter or even direct the structure and composition of microbial communities has made considerable steps forward with the advent of molecular ecology, and there are many more clues on how plants can select some microbial populations at the expense of others (Faure et al., 2009; Hartmann et al., 2009). There is a sophisticated molecular dialogue operating between microbes and between plants and microbes which suggests that, beyond the amount of C contained in the rhizodeposits or their stoichiometric ratios (e.g. C/N, C/P), minute concentrations of very specific compounds (e.g. antibiotics and signalling molecules) may well be the key driver of rhizosphere microbial communities (Faure et al., 2009). Some of these compounds are volatile and thus may affect a large soil volume, as evidenced by the long-distance recruitment of entomopathogenic nematodes in maize (Rasmann et al., 2005; Turlings and Ton, 2006). Understanding and making use of such processes offers considerable promise for innovative applications in the biocontrol of below-ground pests (Turlings and Ton, 2006) and diseases (Raaijmakers et al., 2009), although the management of their complex ecology remains challenging in agroecosystem conditions.

10.2.4 Root and rhizosphere respiration

Besides organic C, plants also play an important role in the fate of inorganic C in soils, with potential impacts on the global C cycle. The production of CO₂ either directly via root respiration or indirectly via microbial respiration in the rhizosphere as a consequence of rhizodeposition has already been addressed in Section 10.2.3 in terms of contribution to the C budget (Figure 10.2). It will also form part of the next section for its contribution to soil acidification. Plant-induced soil respiration is a major contributor to the build-up of soil pCO₂ and concomitant decrease in pO₂, which have major impacts on soil biology and soil biogeochemistry (Hinsinger et al., 2009). Because of respiration of below-ground organisms, bulk soil pCO₂ is normally 10- to 100-fold higher than that of the atmosphere (currently 360 cm⁻³ m⁻³). Values ranging from 7 000 to 24 000 (up to 32 000 under elevated atmospheric pCO₂) cm⁻³ m⁻³ have been reported in a forest soil (Karberg et al., 2005). It is, however, difficult to quantify how much of this was actually related to higher plant activities. Nevertheless, it has been estimated by various methods that the contribution of root respiration in eucalypt forest amounted to 48–70% of total soil respiration (Epron et al., 2006; Marsden et al., 2008), while the meta-analysis conducted by Hanson et al. (2000) showed a broader range (roughly 10–90%), with mean values consistently high in forest (46%) and non-forest ecosystems including crop-based agroecosystems (60%). Given that roots and rhizosphere microbes are major contributors to soil respiration, it would be expected that the rhizosphere should be a hot spot of elevated pCO₂ and decreased pO₂. However, this is little documented. Gollany et al. (1993) measured pCO₂ values of about 100 000 cm⁻³ m⁻³ at
1–3 mm from roots but also found only a gentle gradient of rhizosphere pCO₂ consistent with rapid diffusion of respired CO₂ towards the bulk soil. The same should also hold for pO₂ values, which might be expected to decrease in the rhizosphere, at least in aerobic conditions. Again, there are few measurements of pO₂ in the rhizosphere. Bidel et al. (2000) showed that pO₂ values in the rhizosphere of Prunus persica (L.) Batsch seedlings were much smaller in the meristematic region of the root as a consequence of intense metabolic activity and respiration. However, these data were obtained in an agar medium, which is known to be hypoxic, and thus hardly mimic the situation of well-aerated soils. In contrast, the case of wetland plants (and, among crops, lowland rice) growing in hypoxic soils has been extensively studied. To ensure adequate respiration of their root cells, those plants have evolved aerenchyma which conducts O₂ from the shoots to the roots (Armstrong, 1979). Rhizosphere pO₂ of such plant species thus increases as a consequence of O₂ leaking from roots (Flessa and Fischer, 1992; Revsbech et al., 1999; Armstrong et al., 2000; Blossfeld and Gansert, 2007). For lowland rice, Revsbech et al. (1999) reported that pO₂ was up to 20% of atmospheric pO₂ at the root surface, while being close to zero at distances greater than 0.4 mm from the root surface. This sharp gradient, which contrasts with the flat gradient of pCO₂, is the consequence of rapid consumption of O₂ in redox reactions occurring at the root surface, especially Fe oxyhydroxide precipitation in the so-called Fe plaque (Bacha and Hossner, 1977; Revsbech et al., 1999; Bravin et al., 2008). In situ, non-invasive sensors such as the optodes developed by Blossfeld and Gansert (2007) provide opportunities to acquire quantitative data on rhizosphere pO₂ and possibly pCO₂ as well (and monitor their temporal fate), in both anaerobic and aerobic environments (Blossfeld et al., 2010).

10.2.5 Role of plants in biogenic calcrete formation

Higher plants can leave long-lasting and visible imprints in soil profiles via the precipitation of biogenic CaCO₃, predominantly calcite, which ultimately forms calcretes (Freytet et al., 1997; Verboom and Pate, 2006a, b). Plants can thereby make a major and visible contribution to pedogenesis. A surprising case of pedogenic concretions induced by plants was that of calcrites observed by Braissant et al. (2003) and Cailleau et al. (2004, 2005) in tropical, acidic soils of Ivory Coast, that is, under environmental conditions which do not favour abiotic precipitation of Ca carbonate. The iroko (Milicia excelsa) trees and soil microorganisms in their vicinity are responsible for the formation of this mineral that would otherwise not occur in such soils (Cailleau et al., 2005). This process results from the large-scale formation of oxalate crystals in the wood tissues in both aerial parts and roots of the iroko trees (Cailleau et al., 2004, 2005), which are thereafter transformed into Ca carbonate by soil bacteria, which make use of only a small proportion of the oxalate for their C requirements (Braissant et al., 2002, 2003). Cailleau et al. (2004) calculated that such a process could be considered significant in the long term as a sink for atmospheric C. Braissant et al. (2002) even concluded that in the plant kingdom, this phenomenon of biologically induced mineralisation was a more efficient and longer term C sink than C sequestration in soil organic matter. Verboom and Pate (2006a, b) also report on the role of higher plants in the genesis of calcrites in acidic soil environments of Western Australia. In contrast with the previous examples, in calcareous environments biogenic precipitation of Ca carbonate as calcite in the rhizosphere has been frequently reported. It may occur first of all during the build-up of Ca which results from the large amount of Ca transferred towards the root surface as a consequence of plant-induced transpiration and subsequent mass flow of water and solutes (Jaillard, 1982; Callot et al., 1983; Hinsinger, 1998; Lambers et al., 2009). In
calcareous soils, this process ultimately leads to the precipitation of calcite crystals around roots. If root density is large and the ambient conditions are conducive to such biogenic precipitation in the rhizosphere, the coalescence of the calcite crystals around neighbouring roots can result in continuous calcrete formation (Jaillard, 1982; Callot et al., 1983; Lambers et al., 2009). Another root-mediated process leading to biogenic precipitation of calcite in calcareous soils does not occur in the rhizosphere but inside the root itself (Jaillard, 1982, 1983, 1984, 1985; Jaillard et al., 1991). The resulting, so-called, calcified roots appear as calcite crystals shaping the root cortex cells, while the stele remains free of precipitates (Figure 10.5). This process occurs as a consequence of root activity, which can dissolve abundant amounts of calcium carbonate in the outer rhizosphere as a consequence of respiration and proton release (see Section 10.3), with the Ca ultimately precipitating as calcite in the vacuoles of root cortical cells (Jaillard, 1987a, b). Jaillard (1984) calculated that this biogenic process is responsible for a considerable shift in the particle size of calcareous soils as a consequence of the isometric size of the calcified root cells (about 80 μm diameter). Jaillard (1984) calculated that within about 1000 years, a third of all the Ca carbonate contained in the highly calcareous soil of the studied site (a natural ecosystem with typical Mediterranean ‘garrigues’ vegetation cover close to Montpellier, South France) could have been dissolved and re-precipitated by the roots. Such an estimate is in line with that of Kuzyakov et al. (2006) who, based on the measurement of $^{14}$C incorporation in Ca carbonate derived from $^{14}$C-labelled root exudates, calculated that full re-crystallisation of loess...
carbonate would take 400–2000 years. Contrary to the previous case of calcrete formation in a tropical soil, such dissolution/precipitation of Ca carbonate occurring in calcareous soils would not contribute any net significant C sequestration. They are nonetheless one of the few visible pieces of evidence of the large impact of plants on soil composition and properties (Lambers et al., 2009). Such rhizoconcretions have been observed in some of the oldest paleosols described on Earth by Retallack (1997). Roots of higher plants can thus leave long-lasting morphological features in soil profiles and be a major contributor to pedogenesis via their direct role in the formation of biogenic Ca carbonates and calcrites (Freytet et al., 1997; Verboom and Pate, 2006a, b). To what extent such processes feedback onto plant growth remains to be studied.

### 10.3 Roles of plants in soil acidification

Respired CO₂ rapidly forms carbonic acid, H₂CO₃, which is a weak acid with a pK of 6.36. Plant-induced increase in pCO₂ thus results in a decrease in pH, except for those soils that are too acidic (pH < 6.36), that is, those soils for which H₂CO₃ remains essentially undisassociated. In contrast, in calcareous soils, the actual pH can be quite different from the expected value of 8.3, usually determined in soil-testing procedures which are conducted at ambient atmospheric pCO₂, instead of being conducted at the actual in situ pCO₂ of the soil in field conditions. Indeed, soil pH in calcareous soils is constrained by the dissolution/precipitation equilibrium of CaCO₃, and thus pCO₂, according to the following equation (Gras, 1974; Hinsinger et al., 2003):

\[
\text{pH} = -0.67\log(p\text{CO}_2) + 0.06 \text{(where pCO}_2 \text{ is expressed in atm)}
\]  

(10.1)

This suggests that for every tenfold increase in pCO₂, the corresponding equilibrium soil pH is expected to decrease by 0.67 pH units in calcareous soils. Based on reported in situ values of pCO₂ (Gollany et al., 1993), rhizosphere pH values are expected to be about 6.7–6.8 (Hinsinger et al., 2003), which is lower than the expected value of 8.3 (which would occur if soil pCO₂ was about the same as ambient pCO₂ in the atmosphere). At a global level, this relation explains why the current fast increase in atmospheric pCO₂ is resulting in the acidification of oceans with accompanying detrimental effects on ocean biogeochemistry, global geochemical cycles and aquatic ecosystems such as coral reefs (e.g. Rockström et al., 2009). Equation 10.1 shows that doubling the preindustrial atmospheric pCO₂ should induce a 0.2 pH unit decrease in sea water pH. In spite of the current attention on atmospheric pCO₂, it is surprising that there is so little data available about rhizosphere pCO₂ and its impact on below-ground organisms and biogeochemical cycles. Nevertheless, at the broad scale of ecosystems or watersheds, the contribution of (soil) respiration to acidification and proton budget is quite well documented and consistently shows that acidification rates are greater in soils of neutral to alkaline pH than in acidic soils (Van Breemen et al., 1983, 1984). Van Breemen et al. (1983) reported acidification rates in two oak forests in the Netherlands of up to 14.3 kmol H⁺ ha⁻¹ year⁻¹ in a calcareous environment (soil pH 7–8) but only about 4.8 kmol H⁺ ha⁻¹ year⁻¹ in an acidic environment (soil pH 3.7–3.9). However, the large buffering capacity of the calcareous soil masked this large acidification rate as a consequence of the consumption of protons by the dissolution of CaCO₃. As long as all the CaCO₃ is not completely dissolved, the soil pH hardly decreases, in spite of a large acidification rate (Van Breemen et al., 1983, 1984).
10.3.1 Role of plants in the proton budget and cycle

Van Breemen et al. (1983, 1984), Lelong et al. (1990) and Bourrié and Lelong (1994) showed that in addition to respiration, plants are largely responsible for the proton budget of terrestrial ecosystems via the control of the cation/anion balance (Table 10.2). In acidic soils, this is by far the major process controlling the acidification rate, with considerable variations according to the composition and productivity of the plant community. In forest ecosystems with acidic soils, cation uptake by the vegetation contributed 37–57% of the overall proton budget for beech, 25–64% for spruce and up to 81% for pine trees (Table 10.2). Lelong et al. (1990) and Bourrié and Lelong (1994) found that compared to tree-based ecosystems, alpine grassland ecosystems yielded smaller acidification rates, as a result of a smaller cation/anion balance (Table 10.2). This is also well documented in agroecosystems. For instance, in tropical pastures of varying composition and relying on various sources of N, Noble et al. (2008) reported over short-term periods (about 3 years) that the maximum acidification rate was obtained when the grass *Andropogon gayanus* was fertilised with ammonium-N, while the minimum rate was obtained when this grass was fertilised with nitrate-N. They also showed that in the absence of any fertiliser supply, the legume *Stylosanthes guianensis* fixing N\textsubscript{2} symbiotically had a much greater acidification rate than the grass (more than twofold larger), although not quite as great as when the grass was fertilised with ammonium-N. It is worth noting that in the latter case, the acidification rate was slightly greater than 16 kmol H\textsuperscript{+} ha\textsuperscript{-1} yr\textsuperscript{-1}, which is about 1–3-fold greater than the values reported by Van Breemen et al. (1983) for temperate forest ecosystems and much above those reported in Table 10.2. The ammonium-fertilised grass decreased the bulk soil pH from 4.10 to 3.64 in only about 3 years of growth, while the N\textsubscript{2}-fixing legume resulted in a bulk soil pH of 3.82 and the non-fertilised grass produced no significant pH change (Noble et al., 2008). These results of a short-term field experiment in Thailand are in agreement with a number of ‘over-the-fence’ comparisons made in Australia, which consistently showed that legume-based pastures resulted in greater acidification rates than grass-dominated pastures (Coventry and Slattery, 1991; Noble et al., 1997). This pH decrease in the acidic, poorly buffered soils of Australia caused subsequent problems of Al rhizotoxicity and nutrient deficiencies in crops. Acidification rates ranging from 0.2 to 10.6 kmol H\textsuperscript{+} ha\textsuperscript{-1} year\textsuperscript{-1} have been reported to occur (Ridley et al., 1990; Coventry and Slattery, 1991; Noble et al., 1997). Jarvis and Robson (1983a, b) clearly demonstrated that such acidification rates were largely explained by the cation/anion uptake imbalance as a result of N being largely taken up as an uncharged species (N\textsubscript{2}) in the N\textsubscript{2}-fixing legumes (see Section 10.3.2).

10.3.2 Root-induced changes of pH in the rhizosphere and their ecological relevance

At the rhizosphere scale, it is well established that protons are released by roots (hence rhizosphere acidification occurs) to compensate for an excess of cations over anions being taken up by plants (Nye, 1981; Haynes, 1990; Hinsinger et al., 2003). This is typically what occurs when the dominant form of N taken up is positively charged (ammonium), or to a lesser extent, when a major proportion of the N is taken up as the uncharged species N\textsubscript{2} in N\textsubscript{2}-fixing legumes (see Section 10.3.1). In contrast, if the plant takes up more anions than cations, which would typically occur when the dominant form of N taken up is negatively charged (nitrate), rhizosphere alkalinisation occurs. However, all of the major nutrient cations and anions actually taken up by the plant must be accounted for in order to predict the resulting
Table 10.2 Proton budget in watersheds corresponding to various ecosystems and dominant vegetation cover.

<table>
<thead>
<tr>
<th></th>
<th>Atmospheric deposition of H⁺</th>
<th>CO₂ input = respiration</th>
<th>Nitrogen cycle</th>
<th>Cation uptake</th>
<th>Anion uptake</th>
<th>Cation–anion balance</th>
<th>Cation release (weathering)</th>
<th>Anion fixation</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grass</td>
<td>324</td>
<td>609</td>
<td>25</td>
<td>198</td>
<td>-31</td>
<td>167</td>
<td>-805</td>
<td>-320</td>
<td>1156</td>
</tr>
<tr>
<td>Beech</td>
<td>324</td>
<td>250</td>
<td>-48</td>
<td>748</td>
<td>-59</td>
<td>689</td>
<td>-922</td>
<td>-293</td>
<td>1322</td>
</tr>
<tr>
<td>Spruce</td>
<td>324</td>
<td>472</td>
<td>62</td>
<td>280</td>
<td>-42</td>
<td>238</td>
<td>-1075</td>
<td>-21</td>
<td>1138</td>
</tr>
</tbody>
</table>

Positive values are proton sources, while negative values represent proton sinks, all values being expressed as mol\(\text{H}^+\) ha\(^{-1}\) year\(^{-1}\).

proton efflux from roots. Moreover, for those plant species responsible for large effluxes of charged exudates such as carboxylates, these should also be accounted for in the balance of ions crossing the plasma membrane of root cells (Hinsinger et al., 2003). Since the early work of Weisenseel et al. (1979) and Marschner and Römheld (1983), the use of pH-sensitive dyes combined with agar gel techniques has led to a large number of reports showing, in addition to the role of N source, (1) the plasticity of roots and their response to both Fe (e.g. Römheld, 1987; Vansuyt et al., 2003) and P deficiencies (e.g. Tang et al., 2004) and (2) the heterogeneity of such responses along the root system. The visual aspect of this technique is appealing, but it is essentially qualitative unless used in soil-less media such as agarose gel, when an adequate treatment of the signal can sometimes be achieved (Jaillard et al., 1996; Rao et al., 2002; Vansuyt et al., 2003). Recently, pH-sensitive optical sensors (the so-called optodes) suitable for noninvasive pH monitoring in soils have been successfully applied to map pH changes over time in plants grown in flat rhizotrons under laboratory conditions (Blossfeld and Gansert, 2007; Blossfeld et al., 2009). This promising technique has revealed the speed with which plant roots respond to changes in photosynthetic activity of the shoots as rhizosphere acidification developed very quickly after the onset of light in the wetland plant _Juncus effusus_ (Figure 10.6). In such plants that grow their roots in anaerobic conditions, it has been well established that the root-induced release of O₂ in the rhizosphere (as a consequence of leakage from the aerenchyma), and the subsequent oxidation of the soil, can make a substantial, and sometimes major contribution to the observed rhizosphere acidification, as also shown for lowland rice (Begg et al., 1994; Kirk and Le Van Du, 1997).

Locally, root-induced proton efflux or influx can result in large decreases or increases in rhizosphere pH, respectively, reaching commonly ±0.1 to 1 pH units and in some instances up to ±2 to 3 pH units (Hinsinger et al., 2003). There has been an abundant literature referring to rhizosphere pH changes since the seminal paper by Riley and Barber (1971), which reported on substantial pH variations in the rhizosphere of pot-grown soybean supplied with either nitrate or ammonium, and on their consequences for P bioavailability. Many of the published reports measured such pH changes in rhizobox-grown plants (i.e. more or less artificial settings) with various approaches and more or less spatially resolved and invasive techniques that have been discussed in detail by Jaillard et al. (2003) and more recently by Luster et al. (2009) and Neumann et al. (2009). Several of these works have produced quantitative data on pH gradients occurring in the rhizosphere, and their consequences for the availability of soil P (Gahoonia et al., 1992; Hinsinger and Gilkes, 1995, 1996, 1997; Zoysa et al., 1997). Most of these have reported on root-induced acidification patterns, although Gahoonia et al. (1992) also obtained an upward gradient in rhizosphere pH in ryegrass supplied with nitrate, which interestingly led to less P being depleted from the rhizosphere in the studied luvisol and to more P being depleted in the studied oxisol. The luvisol had a neutral bulk soil pH of 7.4, and the maximum pH increase recorded in the rhizosphere was +0.6 unit, while a 1.5 unit decrease was found when ammonium was supplied. In the oxisol, which had an acidic bulk soil pH of 5.5, similar pH changes were found as the maximum pH increase recorded in the rhizosphere was +0.5 unit for nitrate-supplied ryegrass, while a 1.4 unit decrease was found when ammonium was supplied.

As predicted by the model of Nye (1983), due to the faster diffusion rate of protons compared with phosphate ions, complex patterns of available P can occur in the rhizosphere, with some accumulation of dissolved P occurring at some distance from the roots instead of a continuous depletion gradient (e.g. Hübel and Beck, 1993; Hinsinger and Gilkes, 1996). The rapid diffusion of protons in the rhizosphere can thus lead to more extensive effects than one would first think of for pH-sensitive, poorly mobile nutrients such as P and some
micronutrients. For example, the work of Bravin et al. (2009a) showed that while Cu uptake models predicted sharp and narrow depletion zones with a computed gradient extending to <0.5 mm in the rhizosphere of durum wheat, the observed Cu gradient extended to about 6 mm from the root mat surface. Bravin et al. (2009a) demonstrated that this decrease in soil solution Cu concentration (and even more so for Cu^{2+}) was a consequence of the root-induced alkalisation of up to 2.8 pH units close to roots and which extended up to 6 mm. In this acidic Cu-contaminated soil, such root-induced pH change enabled the plant to reduce the toxic concentration of Cu found in the bulk soil (Bravin et al., 2009a) and ultimately the bioavailability of Cu (Bravin et al., 2009b; Figure 10.7). This process is likely to be of ecological relevance for the adaptation of plants to acidic soils where Al toxicity may otherwise considerably reduce root and hence plant growth (Bagayoko et al., 2000; Hinsinger et al., 2003, 2009). The artificial setting of rhizobox experiments used to obtain these results may

Figure 10.6 Temporal development of root-induced pH changes in the rhizosphere of growing roots of *Juncus effusus*, half an hour before and after the start of daylight, as measured noninvasively with a planar pH optode. The digital photograph (bottom) shows the two investigated roots. Reprinted from Blossfeld and Gansert (2007). With kind permission from Blackwell Publishing Ltd. For a colour version of this figure, please see Plate 10.3.
Soil Conditions and Plant Growth

Soil conditions can influence plant growth, and understanding these relationships is crucial for optimizing crop yield and health. However, underestimating the impact of rhizosphere conditions may lead to significant errors in predicting plant response to nutrients and stressors. For instance, rhizosphere gradients near dense root mats can be underestimated, leading to large overestimates of nutrient availability (Bravin et al., 2009a).

In addition, pH increases of about one pH unit have been reported in the rhizosphere of field-grown durum wheat plants (1–5 mm aggregates adhering to roots were sampled for this purpose) in an acidic soil (Michaud et al., 2007). Sampling field-grown plants, Bagayoko et al. (2000) also reported a substantial pH increase in the rhizosphere of maize in acidic oxisols and suggested that this would have enabled the plants to alleviate Al toxicity and nutrient deficiencies.

Figure 10.7 (a) Copper bioavailability (influx of Cu into the plant) as a function of rhizosphere pH in durum wheat supplied with either nitrate only (black symbols) or ammonium/nitrate (grey symbol) and (b) rhizosphere pH as a function of bulk soil pH (the range of bulk soil pH values was obtained by lime addition at various rates in an acidic soil for plants grown with nitrate only). The figures stand for OH\textsuperscript{−} efflux (for points above the y = x axis, corresponding to rhizosphere alkalisation) or H\textsuperscript{+} efflux (for points below the y = x axis, corresponding to acidification). Based on unpublished data and data from Bravin et al. (2009b). Reprinted with kind permission from Springer-Verlag.

Figure 10.7

(a) Copper bioavailability (influx of Cu into the plant) as a function of rhizosphere pH in durum wheat supplied with either nitrate only (black symbols) or ammonium/nitrate (grey symbol) and (b) rhizosphere pH as a function of bulk soil pH (the range of bulk soil pH values was obtained by lime addition at various rates in an acidic soil for plants grown with nitrate only). The figures stand for OH\textsuperscript{−} efflux (for points above the y = x axis, corresponding to rhizosphere alkalisation) or H\textsuperscript{+} efflux (for points below the y = x axis, corresponding to acidification). Based on unpublished data and data from Bravin et al. (2009b). Reprinted with kind permission from Springer-Verlag.

be thought to lead to large overestimates of the rhizosphere gradients in the vicinity of dense root mats instead of single roots. However, modelling has shown that this is not the case (Bravin et al., 2009a). In addition, pH increases of about one pH unit have also been reported to occur in the rhizosphere of field-grown durum wheat plants (1–5 mm aggregates adhering to roots were sampled for this purpose) in the same acidic soil (Michaud et al., 2007).

Sampling field-grown plants, Bagayoko et al. (2000) also reported a substantial pH increase in the rhizosphere of maize in acidic oxisols and suggested that this would have enabled the plants to alleviate Al toxicity and nutrient deficiencies.
Compared with rhizobox experiments in artificial laboratory conditions, there are fewer reports of pH changes in the rhizosphere of field-grown plants, with most such reports on natural or forest ecosystems (Häussling et al., 1985; Marschner et al., 1991; Clegg et al., 1997; Schöttelndreier and Falkengren-Grerup, 1999; Séguin et al., 2004; Turpault et al., 2007). Many of these report slight rhizosphere alkalisation in acidic (forest) soils. In arable soils, there are few reports on changes of pH in the rhizosphere of crops. Yang et al. (1996) reported systematic acidification of the rhizosphere of field-grown soybean, with more acidification associated with lateral roots than primary roots. They showed that such root-induced acidification was positively correlated with the pH of the bulk soil, so that while the pH change was almost nil for a bulk soil pH of 4.0, the pH change increased (in absolute value) with increasing bulk soil pH. This result is consistent with theory which shows that the buffering capacity of soil solution is lower near neutrality (Nye, 1981). It may also reflect the plasticity of roots which enables plants to cope with adverse soil (pH) conditions. For example, Youssef and Chino (1989) showed in rhizobox experiments that barley and soybean were able to increase their rhizosphere pH when grown in an acidic soil and conversely acidify their rhizosphere when grown at higher pH. More convincingly, in a single soil and a broad range of bulk soil pH obtained by various rates of lime addition, Chaignon et al. (2009) showed that oilseed rape was capable of slightly alkalising its rhizosphere when grown at bulk soil pH lower than 4.7 and of acidifying its rhizosphere at greater pH values. Similarly, Bravin et al. (2009b) demonstrated that durum wheat systematically increased its rhizosphere pH to a constant value of about 7.3 whatever the initial pH, unless supplied with ammonium (Figure 10.7). In this case, the plasticity resides in the ability of roots to adjust the proton influx according to the initial pH, as increasing the pH from 4.8 to 7.3 in this soil resulted in a fourfold greater proton influx (or hydroxyl efflux) than for increasing the pH from 6.9 to 7.3. Remarkably, this behaviour of durum wheat roots, which results in increasing alkalisation of the rhizosphere with decreasing pH, was also found in field-grown plants (Michaud et al., 2007).

### 10.3.3 Role of organic acids in rhizosphere pH changes

There has been a considerable amount of work on the so-called organic acids in the rhizosphere, and excellent reviews have been published on the subject including Jones (1998), Ryan et al. (2001) and Jones et al. (2003). These reviews clearly show that these should more correctly be referred to as organic anions or carboxylates (as most of these are derived from carboxylic acids) because they are exuded as anions and thereby do not result in specific pH changes per se (Hinsinger et al., 2003; Jones et al., 2003). They should, however, be accounted for in the anion/cation balance, at least when released in substantial amounts. This release varies considerably among plant species and is generally low in crop species with the exception of white lupin for which citrate and malate (among others) exudation from cluster roots is considerable (as reviewed by Neumann et al., 1999, 2000). The loss of negative charge which results from the release of carboxylates is counterbalanced by the efflux of an equivalent amount of either protons or K-ions (in the latter case, however, this means that protons have been released at a preliminary step to counterbalance for K uptake), hence contributing to rhizosphere acidification unless there is a net excess of anions taken up when accounting for the overall balance of charges crossing the root cell membranes. Indeed, it is well documented for cluster roots of white lupin that they result in a considerable release of both carboxylates (notably citrate) and protons and hence in distinct rhizosphere acidification (e.g. Neumann and Römheld, 1999). This may not be the case for other
Soil Conditions and Plant Growth

plants, especially crops which release much smaller amounts of carboxylates, and even rhizosphere alkalisation has sometimes been reported to occur in the rhizosphere of other cluster-rooted plants such as members of the Proteaceae (e.g. Roelofs et al., 2001). These carboxylates play many roles in nutrient acquisition, especially P (and to a lesser extent Fe), and in the alleviation of metal toxicities, as especially documented for the case of Al (Pellet et al., 1995; Ryan et al., 1995a, b, 2001, 2009; Kochian et al., 2004). This is due to their capacity to complex and for some of them chelate metal cations such as Ca-, Fe- or Al-ions (Gerke et al., 2000a, b) and to be involved in ligand exchange processes and thus in the desorption of P-ions (Geelhoed et al., 1999) in the rhizosphere (e.g. Jones, 1998; Hinsinger, 2001). Carboxylates released by roots, such as citrate, or produced by rhizosphere microorganisms and fungi such as some ectomycorrhizal fungal species have thus major implications for plant nutrition and in the weathering of soil minerals, but they play little, and only an indirect, role in root-induced soil pH changes compared with ion uptake and respiration (Hinsinger et al., 2003).

10.4 Roles of plants in the weathering of soil minerals

At the beginning of pedology, Dokuchaev and later Jenny pointed out that, besides climate and rock, vegetation is a major factor involved in soil formation and the diversity of soil types that can be found worldwide. With the development of experimental pedology in the 1950s (Bockheim and Gennadiyev, 2009), several studies showed the role of biochemical compounds and microbially mediated processes in soil formation. Biogeochemistry nevertheless remained essentially confined to geochemistry for a long time, although it was obvious to anyone dealing with the C cycle that higher plants were key players (see Section 10.2.1). At geological temporal scales, the emergence of deep-rooted plants during the Devonian coincided with deep soil profile development and the acceleration of rock weathering and soil formation rates in terrestrial ecosystems (Berner, 1992, 1997; Retallack, 1997; Beerling et al., 2001; Beerling and Berner, 2005; Taylor et al., 2009). Interestingly, Taylor et al. (2009) stressed that the five major plant-induced processes of biological weathering (Table 10.3) were ‘driven by plant photosynthesis fuelling plant growth and organic C release into the regolith and soils’. These five key processes, according to Berner et al. (2003) and Taylor et al. (2009) are (1) exudation, (2) respiration, (3) decomposition of plant litter, (4) evapotranspiration and (5) physical stabilisation of soils by root systems and plant cover. Although these authors stress that evapotranspiration plays a role via its impact on the flow of cations to plants, the uptake and translocation of cations and other elements by plants must be considered as another major plant-induced process of rock weathering and soil formation. This will be emphasised in what follows.

The physical effects induced by plants are beyond the scope of the present chapter which will concentrate on plant-mediated chemical weathering. Briefly, the plant-driven transpiration of water by plants directly affects the soil water residence times in numerous ways. Plants favour infiltration at the expense of surface runoff because plant roots create macropore-sized biopores and plant debris creates roughness at the soil surface and decreases the formation of surface crusts. At a local scale, the uptake of water by roots decreases the water content of the soil, which may ultimately reduce the weathering of minerals, given the necessary presence of water for such chemical reactions to occur to any significant extent. At a regional scale, plant communities can significantly affect the amount of rainfall (e.g. Pielke, 2001; Jonko et al., 2010). In addition to and/or combined with water uptake, during
their growth, and especially their radial growth, roots may contribute to the physical breakdown of rocks via (1) the formation of cracks or expansion of pre-existing cracks where roots proliferate and (2) the promotion of mechanical fractures in mineral grains (April and Keller, 1990; see Chapter 8). At the landscape or slope scale, root systems are key to the stabilisation of thick soil horizons by mechanically preventing processes such as landslides and erosion by surface or subsurface runoff water, as reviewed by Stokes et al. (2009). Such physical processes are linked with the physical properties of the root systems, by which plants contribute to deep soil profile formation and stabilisation. Plant-mediated chemical weathering of rocks and minerals can be subdivided into three major biogeochemical processes, which ultimately shift the chemical equilibria underlying chemical weathering processes, or alter the kinetics of such reactions (Table 10.3), by (1) changes in ionic concentrations as a consequence of water and ion uptake by roots or deposition of cations by throughfall and the decay of organic debris; (2) pH changes as a result of respiration, root-induced efflux or influx of protons or microbially mediated production of organic acids in the rhizosphere and (3) complexation or chelation of metal cations by root exudates or microbial metabolites in the rhizosphere.

### Table 10.3 Major processes of weathering of rocks and minerals induced by plants.

<table>
<thead>
<tr>
<th>Physical processes</th>
<th>Ion concentration changes</th>
<th>pH changes</th>
<th>Complexation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exudation of protons/low molecular weight organic acids</td>
<td>+++</td>
<td>+++</td>
<td></td>
</tr>
<tr>
<td>Respiration-induced increase in pCO₂</td>
<td>+++</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Deposition of litter comprising organic matter and cations</td>
<td>+++</td>
<td>+++</td>
<td></td>
</tr>
<tr>
<td>Evapotranspiration</td>
<td>+++</td>
<td>+++</td>
<td></td>
</tr>
<tr>
<td>Fragmentation and physical binding of particles/aggregates</td>
<td>+++</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Source: Adapted from Berner et al. (2003) and Taylor et al. (2009).

#### 10.4.1 Cation uptake by vegetation

Plants are major drivers of changes in ionic concentrations in soil solution which ultimately shift the dissolution/precipitation equilibria and either increase or decrease the kinetics of such processes. The uptake of water and inorganic ions (mineral nutrients, and also non-essential elements such as Si) at high fluxes by plant roots is responsible for important mass exchanges in the rhizosphere, which can drastically modify ionic concentrations in the soil solution and shift the chemical equilibria of reactions involving these ions (Hinsinger, 1998; Hinsinger et al., 2009; Lambers et al., 2009). Hinsinger and Jaillard (1993) demonstrated that the considerable decrease in K concentration occurring in the rhizosphere of ryegrass as a consequence of the sink effect of absorbing roots was directly responsible for the weathering of K-bearing phyllosilicates such as trioctahedral micas. This has been reported by several authors (Mortland et al., 1956; Spyridakis et al., 1967; Hinsinger et al., 1992; Kodama et al., 1994). Hinsinger and Jaillard (1993) and Hinsinger et al. (1992) showed that the root-induced depletion of soil solution K was the driving force for a rapid release of interlayer K
in the mica and the concurrent formation of the clay mineral vermiculite. Hinsinger and co-workers showed that as soon as the K concentration in the rhizosphere decreased below a threshold value of about 70–80μM (Figure 10.8), the weathering of the phlogopite mica became detectable by X-ray diffraction (XRD) through the appearance of a typical vermiculite peak forming at the expanse of the mica-characteristic peak. This process took only about two days for ryegrass (Hinsinger and Jaillard, 1993) and three days for oilseed rape (Hinsinger et al., 1993). Hinsinger et al. (2006) later reported that the uptake of K by plant roots led to a five and sixfold increase in the rate of release of interlayer K (Table 10.4) and thus in the rate of weathering of micas such as phlogopite and biotite, respectively, compared to unplanted systems where leaching resulted in only limited weathering over the 14-day experimental period. In the absence of plants, no significant vermiculite formation was detected by XRD, whatever the mica. No significant, root-induced transformation of a dioctahedral mica, muscovite, was found (Table 10.4), although the K budget suggested that some slight weathering had occurred. In contrast, a large proportion of the trioctahedral micas (phlogopite and even more for biotite) had been transformed into vermiculite in the rhizosphere (Hinsinger et al., 2006).

These results were obtained in the laboratory under rather artificial, controlled conditions. They are, however, in agreement with those computed by Taylor and Velbel (1991) and Velbel and Price (2007) who showed that taking account of the uptake of K by the vegetation of forested watersheds led to biotite weathering rates (432 mol K ha⁻¹ year⁻¹) which were up to 3.5 times greater than when neglecting this term in the K budget (124 mol K ha⁻¹ year⁻¹). In agroecosystems, accounting for such a rhizosphere process, which is the direct consequence of the uptake activity of plant roots, proved essential to understanding the observed K budgets in long-term fertiliser trials conducted by agronomists (Hinsinger, 2002). Indeed, most of the field trials conducted in Europe have shown that, in the absence of fertiliser K application, the cumulative K offtake by successive crops in the rotation was much greater than the observed decrease in exchangeable K, although only this pool was considered as bioavailable to crops (Hinsinger, 2002). Based on the difference, it can be calculated that the rates of release of nonexchangeable K amounted from 11 to 94 kg K ha⁻¹ year⁻¹ (Table 10.5; i.e. 281–2404 mol K ha⁻¹ year⁻¹) as an average value computed over the whole duration of these long-term (several decades) fertiliser trials (Hinsinger, 2002). This
means that, under such low input conditions, the release of nonexchangeable K contributed a major proportion (up to 90%) of K uptake by crops (Hinsinger, 2002), which is similar to results obtained with short-term experiments (4–8 days) conducted in pots in which K budgets were measured at the rhizosphere scale (Kuchenbuch and Jungk, 1982; Niebes et al., 1993). This major contribution of the release of nonexchangeable K was surprising at first sight as it has often been considered that this process should be negligible in such arable soils, because of bulk soil solution K concentrations typically ranging from 100 to 1000 μM.

The release of nonexchangeable K is mostly due to the release of interlayer K and concomitant weathering of soil micas and micaceous clay (illite-like) minerals and is governed by soil solution K concentration (Springob and Richter, 1998). These authors showed that while the rate of K release was considerably enhanced at soil solution K concentrations below 3 μM, it was almost nil at those concentrations commonly found in most soils of arable land in Europe. Claassen and Jungk (1982) estimated that while bulk soil solution concentrations of K were in the order of several hundreds of μM, in the immediate vicinity (about 1 mm) of maize roots, K concentrations could be as low as 2–3 μM; this explains the large rates of release of nonexchangeable K occurring in the rhizosphere (Kuchenbuch and Jungk, 1982; Niebes et al., 1993) in the short term and at the field scale in the long term (Bosc, 1988; Blake et al., 1999; Hinsinger, 2002).

The sink effect of plant roots and the subsequent depletion of soil solution K in the rhizosphere is thus the driving process for the release of interlayer K and concomitant weathering of micaceous minerals in soils (Hinsinger, 2002). Barré et al. (2007a) further proved the causal link between these two processes, based on a quantitative method of assessment of the mineralogical composition of the soil clay fraction derived from a deconvolution of X-ray diffraction patterns. These authors showed that the uptake of K by ryegrass in a pot experiment matched the increased amount of interstratified illite-smectite minerals in the rhizosphere, which formed at the expense of the illite-like clay minerals. These findings confirmed the earlier report of an increased amount of interstratified illite-vermiculite minerals in the rhizosphere of field-grown maize plants (Kodama et al., 1994). Barré et al. (2008) applied

<table>
<thead>
<tr>
<th>Table 10.4</th>
<th>Rates of weathering of three micas in an experimental weathering experiment in which the micas were leached with a K-free nutrient solution for 14 days in the presence or absence of ryegrass.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>K release (±weathering rate) [Log (molK (g mica)−1 s−1)]</td>
</tr>
<tr>
<td>Muscovite mica</td>
<td>Control (no plants)</td>
</tr>
<tr>
<td></td>
<td>Rhizosphere (ryegrass)</td>
</tr>
<tr>
<td>Phlogopite mica</td>
<td>Control (no plants)</td>
</tr>
<tr>
<td>Biotite mica</td>
<td>Control (no plants)</td>
</tr>
<tr>
<td></td>
<td>Rhizosphere (ryegrass)</td>
</tr>
</tbody>
</table>

The reactor was designed such that the roots had developed a planar root mat at the upper surface of a dialysis bag containing the particles of mica. The leachate was continuously removed by gravity, so that this device shall be considered as an open-flow system.

Source: From Hinsinger (1990) and Hinsinger et al. (2006).
Table 10.5  Potassium budgets and derived estimates of average, annual K release rates in 12 long-term fertiliser trials throughout Europe.

<table>
<thead>
<tr>
<th>Field trial location (Country)</th>
<th>Duration of trial (years)</th>
<th>Annual K budget = $K_{\text{fert.}} - K_{\text{offtake}}$ (kg K ha$^{-1}$ year$^{-1}$)</th>
<th>Annual change in exchangeable K (kg K ha$^{-1}$ year$^{-1}$)</th>
<th>Annual K release rate (kg K ha$^{-1}$ year$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clermont (France)</td>
<td>30</td>
<td>-105</td>
<td>-11</td>
<td>94</td>
</tr>
<tr>
<td>Versailles 2 (France)</td>
<td>18</td>
<td>-87</td>
<td>1</td>
<td>88</td>
</tr>
<tr>
<td>Bad Lauchstädten 2 (Germany)</td>
<td>33</td>
<td>-74</td>
<td>7</td>
<td>81</td>
</tr>
<tr>
<td>Bad Lauchstädten 1 (Germany)</td>
<td>33</td>
<td>-70</td>
<td>4</td>
<td>74</td>
</tr>
<tr>
<td>Versailles 1 (France)</td>
<td>32</td>
<td>-66</td>
<td>-4</td>
<td>62</td>
</tr>
<tr>
<td>Fjärningslöv (Sweden)</td>
<td>31</td>
<td>-68</td>
<td>-3</td>
<td>65</td>
</tr>
<tr>
<td>Amiens 1 (France)</td>
<td>22</td>
<td>-69</td>
<td>-17</td>
<td>52</td>
</tr>
<tr>
<td>Ekebo (Sweden)</td>
<td>31</td>
<td>-54</td>
<td>-3</td>
<td>51</td>
</tr>
<tr>
<td>Skierniewicze 1 (Poland)</td>
<td>32</td>
<td>-48</td>
<td>0</td>
<td>48</td>
</tr>
<tr>
<td>Kungsängen (Sweden)</td>
<td>36</td>
<td>-48</td>
<td>-3</td>
<td>45</td>
</tr>
<tr>
<td>Amiens 2 (France)</td>
<td>21</td>
<td>-49</td>
<td>-7</td>
<td>42</td>
</tr>
<tr>
<td>Vreta Kloster (Sweden)</td>
<td>30</td>
<td>-58</td>
<td>-18</td>
<td>40</td>
</tr>
<tr>
<td>Skierniewicze 2 (Poland)</td>
<td>32</td>
<td>-38</td>
<td>0</td>
<td>38</td>
</tr>
<tr>
<td>Rothamsted Broadbalk 1 (UK)</td>
<td>26</td>
<td>-28</td>
<td>0</td>
<td>28</td>
</tr>
<tr>
<td>Rothamsted Broadbalk 2 (UK)</td>
<td>30</td>
<td>-23</td>
<td>0</td>
<td>23</td>
</tr>
<tr>
<td>Toulouse (France)</td>
<td>11</td>
<td>-27</td>
<td>-16</td>
<td>11</td>
</tr>
<tr>
<td>Högåsa (Sweden)</td>
<td>30</td>
<td>-25</td>
<td>-18</td>
<td>8</td>
</tr>
</tbody>
</table>

The rate was deduced from the difference between K budget (amount of K applied as fertilisers minus amount of K exported in harvested products) and the change in exchangeable K (negative values indicate a decrease since the start of the trial) cumulated over the whole duration of the trial, divided by this duration in order to get comparable results (average annual rate).

the same methodology to soil clay fractions collected in a long-term K fertiliser trial and confirmed that the uptake of K by plants quantitatively matched the release of interlayer nonexchangeable K that was evident by the formation of interstratified illite-smectite minerals at the expense of the illite-like clay minerals. Barré et al. (2007b, 2008) concluded that higher plants have a major role in the biogeochemical cycle of K and in the formation and fate of expandable clay minerals (vermiculites or smectites) in topsoils, with illite-like clay minerals playing the role of a large K reservoir in soils of temperate regions of Europe. As stressed by Hinsinger (2002) and Hinsinger et al. (2006), no such release of interlayer K contained in micaceous minerals would be expected to occur in the absence of plants and the accompanying sink effect of their roots resulting in considerable depletion of soil solution K in the rhizosphere. The corresponding rates of release of about 10–100 kg K ha⁻¹ year⁻¹ found in long-term fertiliser trials throughout Europe (Table 10.5), either in intensive cropping systems (Bosc, 1988; Blake et al., 1999; Hinsinger, 2002) or in pasture systems (Holmqvist et al., 2003), were sometimes well above those reported by geochemical models based on biotite weathering in forested ecosystems (Taylor and Velbel, 1991; Velbel and Price, 2007). The latter authors computed K release rates of about 4–5 kg K ha⁻¹ year⁻¹ when neglecting the offtake by the vegetation and up to 17 kg K ha⁻¹ year⁻¹ when accounting for the offtake by forest trees. In addition to playing a key role in the weathering of micaceous minerals, such root-induced release of nonexchangeable K is of major importance for plant nutrition, covering the requirements of many crops in many soils, except in the most weathered soils of the tropical regions. This should be taken into account when designing K fertilisation strategies for the future, considering that K ores are finite resources (Manning, 2010).

10.4.2 Cation uplift in the soil profile as a consequence of plant activities

The concept of nutrient uplift was first put forward by Jobbagy and Jackson (2001, 2004). One of the most dramatic effect of plant-induced uplift is that reported for Si (Lucas et al., 1993; Lucas, 2001), which is a non-essential element for most plant species, although it is sometimes taken up at rates exceeding those of major nutrients (Epstein, 1999). Lucas et al. (1993) reported that in tropical forest ecosystems of Amazonian Brazil, Si was taken up from depth in soil profiles and subsequently deposited in the topsoil as a consequence of litterfall. This process explained the vertical distribution of Si and clay minerals in soil profiles. While geochemical models normally predict that the most weathered minerals should be found in the topsoil, a number of tropical soils typically exhibit a kaolinite-rich topsoil overlying gibbsite-rich horizons (as, e.g. in the study of Lucas et al., 1993). These authors demonstrated that the annual flux of Si to the topsoil by litterfall amounted to 41 kg Si ha⁻¹ year⁻¹ while Si leaching at depth was about fourfold smaller, amounting to 11 kg K ha⁻¹ year⁻¹. This shows that plant-promoted uplift of Si played a prominent role in Si biogeochemical cycling. The accumulated Si in the topsoil most probably occurred as phytoliths which later evolved in Si-rich clay minerals such as kaolinite via dissolution/precipitation processes. Alexandre et al. (1997) observed a considerable increase in Si content in topsoil horizons occurring as phytoliths, thus clearly demonstrating the role of higher plants in such biogeochemical uplift of Si. Higher plants are thus now considered to play a key role in the terrestrial biochemical cycle of Si (Basile-Doelsch et al., 2005; Derry et al., 2005) largely via uptake and sometimes via active Si uptake (Gérard et al., 2008).

While the work described in Section 10.4.1 showed that higher plants are major drivers of the weathering of K-bearing minerals and rocks as a consequence of root-induced processes
occurring in the rhizosphere, they are also responsible for considerable redistribution of K in soil profiles. Indeed, while K uptake occurs throughout the profile, with substantial contribution of the subsoil horizons (Kuhlmann, 1990; Witter and Johansson, 2001; Buxbaum et al., 2005), a substantial part of K accumulated in the vegetation (all of the K which is not removed with the harvested products) is recovered in the topsoil as a consequence of throughfall and litterfall, as quantified by Barré et al. (2007b, 2009) based on clay mineralogy and K budgets. Barré et al. (2009) and Velde and Barré (2010) stressed that such nutrient uplift can counteract cation leaching from surface soils (i.e. what is often considered as a major driving force of meteoric weathering and soil formation processes). Their measurements and calculations in temperate grassland soils suggest that K uplift by the vegetation may ultimately affect clay mineral stability in upper soil horizons (Barré et al., 2007b, 2009). While weathering is often considered to result in a decrease in soil fertility in the long term, due to nutrient cation leaching from the most weathered horizons, and especially the topsoil, biological activities of plants can conversely improve topsoil fertility through this uplift of nutrients such as K (Barré et al., 2009; Velde and Barré, 2010).

10.4.3 Plant-driven pH shift

Because the pH of interstitial water in rocks and soil solutions is a major driver of the weathering of minerals and rocks, root-induced pH changes in the rhizosphere can have big effects on such soil-forming processes. As reported by Dehéran (1892) in his Treatise of Agricultural Chemistry, an experiment conducted by Sachs in the nineteenth century had already demonstrated such a phenomenon. In this experiment, roots of beans were grown over the surface of a polished marble plate, which ultimately left clearly visible imprints in the rock (see also Trolldenier, 1987). The authors concluded that roots secreted an acid strong enough to dissolve Ca carbonate. More remarkably, such imprints of root systems have also been observed in situ (Figure 10.9) for plant roots growing between cracks in dense limestone plates (Jaillard and Hinsinger, 1993; Hinsinger et al., 2003). In the first half of the twentieth century, such acidic root secretions were attributed to carbonic and organic acids produced by rhizosphere microorganisms and roots through respiration and exudation. As explained in Section 10.3.2, it is now well known that such processes are primarily the consequence of excess uptake of cations over anions (possibly accompanied by an exudation of carboxylates), which is counterbalanced by proton efflux and rhizosphere acidification (Hinsinger et al., 2003). In calcareous environments such as in the case depicted in Figure 10.9, the contribution of rhizosphere respiration is another potentially significant cause of such dissolution of Ca carbonates around living roots (Jaillard, 1987a, b). In such soils where the alkaline bulk soil pH would otherwise be a major constraint for plant growth via its negative impact on the bioavailability of a number of nutrients, especially phosphorus (P) and metal micronutrients such as iron and zinc, rhizosphere respiration as well as local release of protons at root tips and the consequent decrease in rhizosphere pH are crucial survival strategies for plants (Hinsinger et al., 2003, 2009).

Poorly soluble minerals and rocks have also been reported to be dissolved as a consequence of rhizosphere acidification (Figure 10.10). The literature is particularly abundant for Ca phosphates and apatitic phosphate rocks because of the significance of their dissolution for plant nutrition and the biogeochemical cycle of P. The root-induced efflux of protons shifts the dissolution/precipitation equilibrium of such minerals and rocks according to the mass action law, thereby enhancing their dissolution (Khasawneh and Doll, 1978; Kirk and Nye, 1986) and ultimately the bioavailability of P (Figure 10.10). The ability of some plant
Plant-induced changes in soil processes and properties

species such as buckwheat, oilseed rape and various legumes to utilise P when supplied as a phosphate rock (i.e. francolite-like apatites) has been attributed to their capacity to release protons in their rhizosphere (Aguilar and van Diest, 1981; Bekele et al., 1983; Ruiz, 1992; Hinsinger and Gilkes, 1995, 1997). The impact of the form of N (nitrate vs ammonium) supplied on the extent of phosphate rock dissolution in the rhizosphere of ryegrass was measured by Hinsinger and Gilkes (1996), further substantiating such a causal relationship. In flooded soils, where adapted plants such as lowland rice are expected to rely solely on ammonium because of the ambient reducing conditions, root-induced dissolution of the acid-soluble pool of soil P (presumably Ca phosphates) has been shown to contribute a substantial proportion of the P acquired by rice crops (Kirk and Saleque, 1995; Saleque and Kirk, 1995). These authors showed that, in this particular case, the root-induced dissolution of soil P was only partly due to proton efflux to compensate for excess cation over anion uptake. In addition, the root-induced precipitation of the so-called Fe plaque (which is made of Fe oxyhydroxides) in the rhizosphere of rice and other wetland plants promotes a significant part of the observed rhizosphere acidification (Begg et al., 1994; Kirk and Le Van

Figure 10.9 Imprints of root systems observed in situ for plant roots growing in between cracks in dense limestone plates (Hinsinger et al., 2003). Reproduced with permission from Springer-Verlag. Photograph taken in a sample collected in the topsoil of a natural ecosystem with typical Mediterranean ‘garrigues’ vegetation cover at St Gély-du-Fesc, close to Montpellier (South of France). The photographed zone is about 5 cm × 7.5 cm. Photo credit: Philippe Hinsinger. From Jaillard and Hinsinger (1993). Reproduced with permission from John Wiley & Sons. For a colour version of this figure, please see Plate 10.4.
Soil Conditions and Plant Growth

Du, 1997; Hinsinger et al., 2003, 2009). Besides Ca phosphates, rhizosphere acidification also determines the solubility of silicates and the bioavailability of the cations contained in such minerals and rocks. Compared with ryegrass which did not result in any rhizosphere acidification, oilseed rape led to a substantial root-induced acidification of its rhizosphere and thus to a quite different weathering of a Mg-bearing silicate such as phlogopite mica (Hinsinger et al., 1993). Oilseed rape was responsible for the formation of hydroxy-interlayered vermiculite and for the release of metal cations such as Mg which are part of the octahedral sheet of trioctahedral micas such as phlogopite. These two processes implied that some acid-promoted dissolution of the crystal lattice of phlogopite occurred, contrary to the case of the root-induced formation of vermiculite reported for the rhizosphere of ryegrass (Hinsinger and Jaillard, 1993; Hinsinger et al., 1993).

The impact of pH changes on the weathering of micas and feldspars in the soil around ectomycorrhizae was shown by Arocena et al. (1999) and Arocena and Glowa (2000). Few reports have provided evidence for the in situ co-occurrence of such rhizosphere acidification and accelerated weathering rates (Courchesne and Gobran, 1997; Turpault et al., 2005). These works referred to forest stands but there has not been any equivalent report in agroecosystems. None of these studies demonstrated the causal relationship between these two processes but, nevertheless, suggest that rhizosphere acidification should be accounted for when considering soil formation processes (pedogenesis) and plant nutrition.

Beside the release of protons by roots to counterbalance the excess uptake of cations over anions, the production of organic acids by rhizosphere microorganisms has received considerable interest, especially so in the context of ectomycorrhizal fungi associated with forest trees. For instance, Wallander and Wickman (1999) and Bakker et al. (2004) showed the impact of such a process on the weathering of various silicate minerals. Wallander and Wickman (1999) stressed that in such symbioses, the ectomycorrhizal fungal partner was largely responsible for the observed elevated concentrations of citric and oxalic acids and the consequent release of K and Mg from biotite. Ectomycorrhizal fungi, and especially

![Figure 10.10](attachment:image.png)

**Figure 10.10** Bioavailability of P supplied as phosphate rocks of varying particle size as a function of the root-induced change of pH in the rhizosphere (negative values indicate rhizosphere acidification, positive values indicate rhizosphere alkalinisation) of perennial ryegrass (black symbols), oilseed rape (white symbols) and a nitrogen-fixing legume (grey symbols), the annual medic *Medicago truncatula* (Béhi, Arvieu, Benhassaine and Hinsinger, unpublished data).
some species and strains, thus play a major role in these processes, rather than the host plant itself (Landeweert et al., 2001; Hoffland et al., 2004). Casarin et al. (2003, 2004) clearly showed that only one of the two tested ectomycorrhizal fungal species was responsible for significant rhizosphere acidification as a consequence of secretion of oxalic acid, while the non-mycorrhizal host plant did not cause any significant pH change. Fungal hyphae of the so-called rock-eating fungi have been shown to dissolve silicate minerals, thereby forming tunnel-like biopores inside mineral grains, presumably as a consequence of the production of organic acids (Jongmans et al., 1997; Landeweert et al., 2001; Hoffland et al., 2004; Van Schöll et al., 2008). More recent works have stressed that when considering the role of mycorrhizae in biological weathering, the potential contribution of the bacteria associated with ectomycorrhizal roots and hyphae should also be taken into account as it may sometimes be large (Calvaruso et al., 2007; Uroz et al., 2007). While Frey-Klett et al. (2005) reported on the so-called mycorrhiza helper bacteria belonging to the group of fluorescent pseudomonads associated with ectomycorrhizae, Calvaruso et al. (2007) showed that different ectomycorrhizal fungal species may exert a selective influence on bacterial communities that are potentially implicated in the production of either acidic or complexing metabolites for promoting the weathering of soil minerals. In the case of arbuscular mycorrhizae, Berthelin and Leyval (1982) and Leyval et al. (1990) showed that the host maize plant was responsible for the largest portion of the biologically mediated release of K and Fe from biotite. Based on current knowledge of the physiology of arbuscular mycorrhizal fungi, these are not expected to release significant amounts of organic acids; this contrasts with ectomycorrhizal species. Biopores of hyphal origin in silicate grains have been reported in boreal forest ecosystems where ectomycorrhizal species are abundant and diverse (Landeweert et al., 2001; Hoffland et al., 2004; Schöll et al., 2007). In contrast, a major proportion of soil bacterial species, the phosphate solubilising bacteria, have been reported to dissolve Ca phosphate minerals and phosphate rocks as a consequence of their ability to produce protons or organic acids (Kucey et al., 1989; Richardson, 2001; Rengel and Marschner, 2005; Richardson et al., 2009). Thus, in the case of non-ectomycorrhizal plants such as crops, roots are expected to be major contributors to acid-promoted weathering of minerals, with a potential contribution of rhizosphere bacteria, but with little or no effect of arbuscular mycorrhizal fungi, which are not to be considered as rock-eating fungi.

10.4.4 Cation complexation or chelation as related to plant activities

Hinsinger et al. (2001) used a microcosm experiment to study the role of different crop species in the weathering of a basalt rock which was leached with a nutrient solution in the presence or absence of plants. In the presence of plants, the dissolution rates of most studied elements (Ca, Mg, Si) increased from two to tenfold, compared with unplanted systems where only leaching was responsible for the observed weathering. The impact of the crops on Fe dissolution rates was much greater, up to 500-fold larger than the minute rate measured in the absence of plants. The latter was expectedly small as a consequence of the neutral pH and oxidising conditions. A significant root-induced acidification was observed which possibly explained the kinetics of release of Ca and Mg from the basalt, but obviously not that of Fe (Hinsinger et al., 2001). Hinsinger et al. (2001) could not elucidate the underlying processes. However, it is likely that the much greater plant-mediated promotion of basalt dissolution measured for Fe compared to other elements was due to rhizosphere processes other than root-induced acidification, such as Fe chelation by root or microbial siderophores.
Bertrand et al. (1999) and Bertrand and Hinsinger (2000) drew similar conclusions in interpreting the results obtained during the experimental dissolution of a poorly soluble Fe oxyhydroxide (goethite) used as a sole source of Fe in the rhizosphere of various crop species. Reichard et al. (2005) showed that phytosiderophores secreted by wheat were implicated in the ligand-promoted dissolution of goethite and that microbial siderophores were also involved (Reichard et al., 2007). These authors also reported a synergetic effect of (phyto)siderophores and carboxylates such as oxalate in the dissolution of goethite.

In the numerous works conducted on the role of carboxylic acids produced by roots and/or rhizosphere microorganisms on the weathering of minerals and rocks, few have attempted to distinguish the direct pH effect from that of the complexing carboxylates. Over the past 50 years, numerous investigations have been conducted on the effects of organic acids on the weathering of rocks and minerals in order to elucidate the possible role of these biochemical agents in soil environments and pedogenesis (e.g. Huang and Keller, 1970; Schnitzer and Kodama, 1976; Razzaghe and Robert, 1979; Tan, 1986; Barman et al., 1992; Eick et al., 1996a, b). In most of these studies, organic acids were implicitly assumed to originate from the decomposition of organic matter by soil microorganisms. Some research showed a more direct implication of soil microorganisms in the weathering of minerals and rocks via the excretion of organic acids or siderophores (e.g. Banfield et al., 1999) and possibly other metabolites which influence pH and redox conditions (see review by Robert and Berthelin, 1986). Until recently, there was little recognition of the potential direct effect of root exudates and secretions of higher plants in the ligand-promoted dissolution of silicate minerals and rocks (see Robert and Berthelin, 1986; Hinsinger et al., 2001). By the end of the twentieth century, the direct implication of land plants in the weathering of rocks and minerals was still a question for debate (Drever, 1994; Jackson, 1996). Mostly this debate focused on lichens which are symbiotic associations of algae and fungi. Lichens are characterised by low growth rates and nutrient requirements that enable them to play the role of pioneer vegetation in the colonisation of fresh rocks (Chapin, 1980). Several studies have shown the role of chelating organic acids released by lichens (e.g. lichenic acid) in the early stages of biological weathering of rocks (Robert and Berthelin, 1986).

The research of Jones and co-workers on carboxylates (e.g. Jones et al., 1996; Jones, 1998) and that of Takagi’s group in Japan and Marschner-Römheld’s group in Germany (Takagi, 1976; Takagi et al., 1984; Römheld, 1991; Marschner and Römheld, 1994) on phytosiderophores has produced an abundant literature on complexing and chelating agents by plant roots. However, phytosiderophores are only secreted by the Poaceae, which include all the cereals and numerous pasture species (grasses), suggesting that phytosiderophore-promoted dissolution of Fe-bearing minerals and rocks is restricted to these species (Reichard et al., 2005; Robin et al., 2008). In contrast, exudation of carboxylates is a ubiquitous process that all plant species share in common as it is largely the consequence of a passive diffusion of these metabolites which occur at elevated concentration in root cells (Jones, 1998). As stressed, plant species differ considerably in their carboxylate exudation pattern, both qualitatively and quantitatively, and this process is also influenced by the biotic/abiotic environment and nutrient status of the plant (Jones, 1998). Moreover, several of these carboxylates (e.g. the many monocarboxylic anions) do not induce any metal complexation or chelation. Among the most common root exudates, tri-carboxylates such as citrate are by far the most efficient carboxylates in ligand-promoted dissolution of rock-forming or soil-forming minerals, notably via their role in Fe (Robin et al., 2008) and Al chelation. Oxalate is exuded in minor amounts by plant roots except for a few species such as buckwheat (Zheng et al., 1998; Ma et al., 2001), but is of great significance in forest
Plant-induced changes in soil processes and properties

Plant-induced changes in soil processes and properties as it is produced by many ectomycorrhizal and saprotrophic fungal species (Cromack et al., 1979; Dutton and Evans, 1996; Wallander, 2000; Casarin et al., 2004). For citrate, most plant species exude small amounts but some have been shown to produce large amounts, often combined with other carboxylates such as malate: oilseed rape (e.g. Hoffland et al., 1989) and, among legumes, chickpea (e.g. McClay et al., 1997) and white lupin (e.g. Gardner et al., 1983). The physiology of white lupin has been extensively studied over the past two decades because it releases much larger amounts of citrate than any other crops, thanks to its specialised, so-called cluster roots, which exhibit singular anatomy, architecture and fate (Dinkelaker et al., 1989; Neumann et al., 1999, 2000; Uhde-Stone et al., 2003). This combination of unique morphological and physiological root traits has been described in other cluster-rooted plants, and especially in species of the Proteaceae which are native species occurring predominantly in the highly weathered landscapes of Australia and, to a lesser extent South Africa (Lambers et al., 2006, 2008, 2010). The so-called proteoid roots of these species are capable of exuding citrate and many other carboxylates at rates that are far above those of white lupin and thus orders of magnitude greater than those reported in most crop species (Roelofs et al., 2001). In such environments, Verboom and Pate (2006a) reported on spatial coincidence between the occurrence of ferricretes (pedogenic concretions of Fe oxides in the soil profile) and proteaceous shrub-heathland, while calcretes (pedogenic concretions of Ca carbonates in the soil profile) coincided with myrtaceous woodland. Such differences in soil properties were observed over short (decametric) distances, and thus in conditions of similar climate and parent material in a flat landscape, except for the vegetation cover. This strongly suggests that the vegetation was responsible for such differences. Verboom and Pate (2006b) thereafter developed the ‘phytotarium’ concept to emphasise the major role of the vegetation on the development of the soil profile. Based on the observation of Fe oxide concretions around root channels in lateritic or deep podzolic soil profiles, they hypothesised that such ferricretes were bioengineered by plants, and especially so Proteaceae, due to the ability of their proteoid roots to exude large amounts of Fe-mobilising carboxylates such as citrate. This mobilisation of Fe favoured the vertical redistribution of Fe in the profile and its re-precipitation along root channels at depth (Verboom and Pate, 2006a). In addition to their role in weathering and mineral formation in soils, complexing and chelating compounds produced by roots and associated microorganisms play a key role in plant growth and nutrition, through their (1) positive impact on the bioavailability of metal micronutrients, especially iron (e.g. Marschner, 1995; Robin et al., 2008), but also major nutrients such as P (e.g. Jones, 1998; Hinsinger, 2001); and (2) negative impact on the bioavailability of toxic metals, as especially documented for Al (e.g. Ryan et al., 2001, 2005; Jones et al., 2003; Kochian et al., 2004).

10.5 Conclusion

The contributions of higher plants to the weathering of primary silicate minerals and rocks and to the formation and fate of secondary minerals (e.g. carbonates, oxides or clay minerals) are manifold and largely based on rhizosphere processes. This weathering and mineral formation also results in the release of plant-available nutrients. Some models of weathering of silicate minerals and rocks now account for the additional effect of the uptake of nutrients by plants (e.g. Taylor and Velbel, 1991; Bormann et al., 1998, Moulton et al., 2000). A number of models of Earth geochemical history have stressed that the considerable drop in atmospheric pCO$_2$ which started in the early-mid Devonian (410–380 million years ago)
Soil Conditions and Plant Growth

coincided with the advent of deep-rooted vascular plants and suggest that such plants were responsible for this dramatic global change (Berner, 1992, 1997; Retallack, 1997; Beerling and Berner, 2005). According to Retallack (1997), ‘the early Paleozoic greenhouse may have been curbed by the evolution of rhizospheres with an increased ratio of primary to secondary production and by more effective silicate weathering’. Beerling and Berner (2005) speculated that the decrease in atmospheric pCO₂ was due to enhanced photosynthesis and the subsequent sequestration of C into the soil, burial of C into sediments and enhanced weathering of Ca- and Mg-bearing silicates. However, none of these geochemical models fully accounts for all the other rhizosphere processes that have been reported, in spite of their ability to dramatically influence mineral and rock weathering rates, and thereby the rate of soil formation. Besides the role of roots and rhizosphere microorganisms on elevated soil pCO₂ via respiration which is well acknowledged, the potential impact of protons, carboxylates and siderophores released by roots and rhizosphere microorganisms on the weathering of rocks and pedogenesis is little accounted for. Furthermore, the various plant-mediated processes responsible for (1) the uplift of nutrients such as K or non-essential elements such as Si and subsequent changes of soil mineralogy and (2) the precipitation of calcretes, ferricretes and silcretes need to be better accounted for as they provide explanations for situations where the conventional views of soil geochemistry and pedogenesis accounting only for physically or chemically mediated processes clearly fail (Hinsinger et al., 2001; Verboom and Pate, 2006a; Velde and Barré, 2010).

Plants are thus to be considered as ecosystem engineers (sensu Jones et al., 1994), modifying the bioavailability of soil resources for themselves and other living organisms (Lambers et al., 2009). While plants have been long regarded as passive receivers in the acquisition of water and nutrients in ecosystems, compared with animals, this current knowledge about the ecophysiology and ecology of plants shows, conversely, that they are very active. While animals can freely move when encountering adverse conditions and thereby avoid resource depletion, plants have to cope/thrive with the environment where they develop, as a matter of survival. This major constraint probably explains why higher plants have evolved so much plasticity in their root systems and so many strategies to cope with the adverse physical, chemical and biological conditions of soils (Marschner, 1995; Lambers et al., 2006, 2009, 2010). These many processes occur in the rhizosphere, which for some of these represents a very narrow (micrometric-millimetric) layer of soil around roots, while for some other processes it expands centimetres away from roots (Hinsinger et al., 2005; Gregory, 2006b). These rhizosphere processes which are related to the many physiological functions of roots, such as respiration, water and ion uptake, acidification/alkalisation, exudation and rhizodeposition of organic compounds have a number of positive or negative effects on plant growth and nutrition, via the alteration of the distribution and concentration (availability) of resources (water and nutrients) and the alleviation or promotion of toxicities and diseases. Several of these are the consequence of the considerable changes in the structure and activity of microbial communities occurring in the rhizosphere (Dessaux et al., 2010), with rhizodeposition of C-rich compounds being a major driver of such changes (Jones et al., 2009). These rhizosphere processes which all occur, by definition, as a consequence of root activities not only have an impact on the plant itself but thus also on other living biota.

Plant-induced changes in the availability of water and nutrient resources, organic matter and, ultimately, soil fertility are manifold and complex. Through these changes and others (e.g. mechanical effects), plants also alter the habitat of soil biota, or create new habitats for these essential components of soils and ecosystems, and their diversity (Hinsinger
et al., 2009). Plants are thereby unique ecosystem engineers, and the challenge for the sustainable management of agroecosystems is to make use of the accumulated knowledge on plant functions and functioning, accounting for their manifold and complex interactions with soils.

References


Plant-induced changes in soil processes and properties


11 Sustainable management of soil and plant health by optimising soil biological function

Dominic Standing¹ and Ken Killham²

¹ School of Biological Sciences, University of Aberdeen, Aberdeen, UK
² Honorary Fellow, The James Hutton Institute, Dundee, UK

11.1 Introduction

11.1.1 Concepts of soil and plant health in relation to management and land use

Soil health and plant health are easily understood concepts and, for plant health, relatively easy to quantify. Is the plant growing well? Do its leaves show signs of chlorosis? Is it likely to produce a full harvest? These are simple questions that any agronomist and agriculturalist will be able to answer without difficulty. Plant health, however, relies on soil health, and soil health is far less easy to quantify as it can be interpreted in numerous ways. Broadly speaking, soil health is the measure of the capacity of that soil to underwrite specified functions. Included here are plant and animal production systems, natural landscape and ecosystem maintenance, regulation of water flow and quality, air quality and nutrient cycling, and the attenuation of many of the negative impacts associated with human habitation and industry. Thus, depending on one’s point of view, soil health is primarily an expression of the soil to produce maximum harvests at maximum profit, or cheap and plentiful food, or to support a landscape of diverse plant life and associated wildlife, or provide suitable ‘ground’ for habitation and industrial activities. However, not all of these are in line with the sustainable use of soil or the maintenance/improvement of soil health.

As national and international policy guidelines tend towards sustainability of resource use, understanding determinants of soil health is essential. Land managers in both the agricultural and industrial sectors can monitor the effects of specific practices on soil health and relate these to its sustainability.

To gain maximum crop yield from their land, farmers have often relied on the input of inorganic chemical fertilizers, particularly NPK, or organic additions such as manures and composts. This results in the alteration of the available and potentially available soil nutrient pools. The pH of the soil is key to determining the level of biological activity and nutrient availability, and management practices must ensure maintenance of soil pH within a range suitable for biological functions which underpin a particular land use.
Sustainable management of soil and plant health by optimising soil biological function

Soil health is a summation of dynamic biotic and abiotic processes and feedbacks. While these are simply presented in Figure 11.1, they can interact in highly complex ways. Consider, for example, that a 60% clay-rich soil may have a specific surface area of nearly 500 m² per g of soil. This surface area is packed into an extremely complex system of solids and pore spaces. The pore spaces may be water or gas filled and may be connected or not. The entire surface will have a degree of electro-chemical activity influencing the sorption and desorption of ions and molecules. Superimposed on this will be approximately one billion bacteria, largely attached to pore walls, as well as a large mass of fungal mycelia and a wide range of microfauna (Killham, 1994). Some of the fungal mycelia are mycorrhizal and connected to plant root systems. All of these biotic components interact with the physico-chemical environment. Thus, there is a great level of complexity in terms of the actual and potential interactions between the biota and the soil habitat. However, in terms of land use and management, there is less interest in the microscale of soil (and its biotic interactions) and more in the macroscale (field and landscape), and, specifically, how management practices impact on soil use and health (Table 11.1). The biological (microbial and faunal) components are key to the creation and maintenance of soil organic matter (SOM). SOM plays a vital role in soil structure, water relations, as well as nutrient status and turnover. Soil health is dependent on physical, chemical and biological qualities (Table 11.1). The physical components of the soil often determine water flow and thus the movement of nutrients, contaminants and even micro-organisms. For example, a sandy soil will drain and transport water rapidly in comparison to a silty loam or clay soil (assuming no major cracks). The structure of a soil is a measure of its texture, porosity and bulk density, and while a soil’s texture is not easily altered, changes in porosity and bulk density can dramatically impact crop yield.

Figure 11.1  Simple schematic of the tripartite interaction between soil, plants and microbes and how this relates to soil health. Plants act as carbon pumps into the soil driving microbial heterotrophic functions and nutrient turnover. This impacts on nonbiological aspects of soil function, in turn impacting on plant growth. Each element of the tripartite interaction can involve multiple feedbacks. Changes in any one of the principal axes will have repercussions on the remaining two with a concomitant change in soil health.

Soil health is a summation of dynamic biotic and abiotic processes and feedbacks. While these are simply presented in Figure 11.1, they can interact in highly complex ways. Consider, for example, that a 60% clay-rich soil may have a specific surface area of nearly 500 m² per g of soil. This surface area is packed into an extremely complex system of solids and pore spaces. The pore spaces may be water or gas filled and may be connected or not. The entire surface will have a degree of electro-chemical activity influencing the sorption and desorption of ions and molecules. Superimposed on this will be approximately one billion bacteria, largely attached to pore walls, as well as a large mass of fungal mycelia and a wide range of microfauna (Killham, 1994). Some of the fungal mycelia are mycorrhizal and connected to plant root systems. All of these biotic components interact with the physico-chemical environment. Thus, there is a great level of complexity in terms of the actual and potential interactions between the biota and the soil habitat. However, in terms of land use and management, there is less interest in the microscale of soil (and its biotic interactions) and more in the macroscale (field and landscape), and, specifically, how management practices impact on soil use and health (Table 11.1). The biological (microbial and faunal) components are key to the creation and maintenance of soil organic matter (SOM). SOM plays a vital role in soil structure, water relations, as well as nutrient status and turnover. Soil health is dependent on physical, chemical and biological qualities (Table 11.1). The physical components of the soil often determine water flow and thus the movement of nutrients, contaminants and even micro-organisms. For example, a sandy soil will drain and transport water rapidly in comparison to a silty loam or clay soil (assuming no major cracks). The structure of a soil is a measure of its texture, porosity and bulk density, and while a soil’s texture is not easily altered, changes in porosity and bulk density can dramatically impact crop yield.
Drivers for managing soil and plant health include political/legislative (soil protection/conservation, watershed protection), economic as well as environmental. Managing soil and plant health is driven by political/legislative, economic and environmental factors. The scope of this chapter is to introduce managing soil and plant health for a variety of services involving soil biological function. Management for biofertilisation is considered. Reference is first made to chemical and organic fertilisation; however, despite success of microbial N₂ fixation and some partial success with P-solubilising micro-organisms, the majority of crop fertilisation for high yields is still achieved through inorganic, chemical applications. The chapter also deals with the restoration of soil after degradation. Here, understanding the role of plants, micro-organisms and animals in maintaining soil structural stability is emphasised. Also highlighted are the bioremediation of soil contamination (increasingly being used as a sustainable solution to pollution problems worldwide) and the many factors determining its success. Finally, crop protection is considered and, despite some successful applications to physically protected (e.g. glasshouses and polytunnels) high-value crops, the field application of biocontrol agents is still in its infancy.

With all of the aspects of managing soil and plant health covered in this chapter, the interested reader is advised to follow-up the references supplied, which provide a more comprehensive picture of current management systems as well as the future potential and ongoing research.

### 11.2 Managing soil and plant health

#### 11.2.1 Improvement of nutrient supply/status

**Chemical/organic fertilisation versus biofertilisation**

Globally, management of the soil/plant system to improve nutrient status/supply is largely carried out through application of chemical fertilisers and to some extent by the application
of manures and composts. Production of fertiliser nitrogen is currently estimated to be about 170 Tg per year (Galloway, 2003), with the global use of N fertilisers increasing by about 15 Tg per year in order to meet the ever increasing demand for food. Nitrogen fixation in agricultural systems represents about another 40 Tg of reactive nitrogen being added globally. This contribution from fixation is mainly due to legumes and cyanobacteria/rice paddy systems (Killham, 1994). It is hard to estimate the proportion of this contribution from fixation due to biofertilisation as some legumes (e.g. clover in temperate pasture systems) are not inoculated and inoculation of rice paddies is far from routine. However, it does indicate the considerable potential of biofertilisation, a potential likely to be increasingly fulfilled as fertiliser costs escalate and crop/land management is driven towards sustainability.

Fertiliser efficiency in terms of nutrient supply to the target crop is continuing to be increased through innovative management of the plant/soil system such as precision agriculture (Haneklaus and Schnug, 2006) and better transfer of knowledge to farmers, while integrated soil management is increasingly exploiting the combined use of inorganic and organic fertilisation. However, there is a wealth of information on fertiliser and nutrient management regimes. The principles underpinning these systems are well understood and incorporated into existing management practises. Comprehensive references are widely available (e.g. Havlin et al., 2004). Improving nutrient supply using biofertilisers, however, is a very different story and is a challenge that needs serious consideration for sustainable management of the plant/soil system.

**Biofertilisers for N₂ fixation, P solubilisation, S oxidation and to introduce plant growth promoting rhizobacteria (PGPRs)**

Biofertilisation to improve crop N supply largely revolves around legume inoculation with *Rhizobium* and growth of cyanobacteria with paddy rice, and these are the most important forms of biofertilisation worldwide. Inoculation of various legumes for enhancing N₂ fixation has been highly successful (mainly in tropical latitudes, where gains from fixation tend to be highest). Biofertilisation of rice with cyanobacteria, usually growing in association with the water fern *Azolla*, is also highly successful and supplies fixed N to some 86% of global rice production (Vaishampayan et al., 2001). Other attempts to increase soil fertility through use of free-living nitrogen fixation have been far less successful. The key, fundamental problem is that the energy required to drive the fixation is high and any free living inoculum (other than cyanobacterial systems which are driven by solar energy) must compete for available C with all other heterotrophic micro-organisms in the soil (Killham, 1994). For a temperate cereal crop, there is a rhizosphere carbon flow of approximately 0.5 ton C ha⁻¹ y⁻¹. If all of this soil carbon flow were available to a population of the free-living diazotrophic bacterium *Azotobacter*, it would still only be enough to generate about 10 kg N ha⁻¹ y⁻¹ (Figure 11.2; Killham, 1994). In reality, the *Azotobacter* will only successfully compete for a rather small proportion of this carbon, and hence free living N₂ fixation will not significantly contribute to the needs of a growing crop, and biofertilisation will not be economically viable.

In view of the carbon limitation on free living N₂ fixation illustrated in Figure 11.2, for all cropping systems except the rice paddy, the legume-*Rhizobium* system offers a far more encouraging plant/soil management tool than free living fixation because the rhizobial symbiont carries out N₂ fixation within the plant, and the root nodules formed receive the energy for fixation from translocated, photosynthetically fixed carbon from the plant host leaves (Killham, 1994).
Great improvements have been made in terms of management of rhizobial biofertilisation of legumes over recent decades. The soils in tropical latitudes where heavy inoculation is carried out often represent a hostile environment (particular in terms of water stress and adverse pH) when the rhizobial inoculum is introduced, and without protection, loss of viable rhizobial cells can be high. The technology of inoculum delivery to the legume root has therefore become more efficient through the use of effective carriers and seed coats, and better inoculum production and storage (Deaker et al., 2004).

Inoculation programmes with mycorrhizal fungi, phosphate solubilising and S oxidising micro-organisms have encountered much more limited success than the rhizobial/legume and cyanobacterial/rice paddy systems and have only delivered benefits under certain conditions (Siddiqui et al., 2008; Banerjee and Yesmin, 2009; Kahn et al., 2009). Therefore, widespread and routine biofertilisation of field crops/trees with these forms of microbial inoculation has yet to be realised, although it clearly offers considerable potential.

In terms of mycorrhizal systems, inoculation with ectomycorrhizal basidiomycotina has offered promising biofertilisation management systems since the fungi can be readily cultured and bulk inoculum produced and delivered to tree seedling hosts in nurseries (Rincon et al., 2005). However, inoculation has failed to generate a consistent, positive response in the tree host across the wide range of edaphic conditions required/found in the field.

There are numerous examples of trials with P-solubilising rhizobacteria and fungi where successful crop biofertilisation has been achieved (Rai, 2005), and there is also clear evidence of a diverse range of plant growth promoting rhizobacteria (PGPRs), eliciting impressive, increased crop yields in certain soils (Egamberdiyeva, 2007).

In the work by Alogaidi et al. (2011), rice cultivars IAC25 and Azucena were grown in sterile, P-limited subsoil with a biofertiliser inoculum of 5% w/w live topsoil. Controls contained sterile subsoil and sterile topsoil. The photographs of representative root systems (Figure 11.3) clearly show that the presence of a rhizobacterial community is important for IAC25. Interestingly, this does not appear to be the case for the Azucena cultivar.

Figure 11.2 The limitation to N₂ fixation by free-living heterotrophic bacteria in soil is the supply of available carbon, which is mainly through rhizosphere flow. The numbers in the figure refer to typical fluxes for a temperate cereal crop. For a colour version of this figure, please see Plate 11.1. From Killham (1994).
consequences for P-acquisition and growth were that IAC25-sterile was significantly smaller than IAC25-live and took up significantly less P. This was not translated into a lower P-use efficiency. The PGPR response then is by no means universal and this, with appropriate screening, opens up opportunities to match soil conditions to crop cultivars for biofertilisation.

P-solubilising bacteria (PSB) are a widely studied subset of the soil microbiota due to their obvious importance in liberating P for plant uptake. The most efficient PSBs are found in the *Bacillus* and *Pseudomonas* genera (Tilak et al., 2005). Of particular interest are the pseudomonads – an abundant and versatile group of Gram negative, aerobic, chemoheterotrophs. Goldstein and Rogers (1999) demonstrated the solubilising mechanism, extracellular quinoprotein glucose dehydrogenase, oxidises glucose (primary component of seedling exudate—see Yeomans et al., 1999) to gluconic acid. Fungi such as *Aspergillus* and *Penicillium* can solubilise bound phosphates, often through secretion of organic acids and/or phosphatase enzymes. In the case of P solubilisers and other PGPRs, large-scale, consistent, commercial application has yet to be realised, even though, when used as biofertilisers under favourable conditions, microbial inoculants can substitute up to 25% of plant phosphorus requirements (Rai, 2005). As with mycorrhizal systems, the main reason for this is that the soil ecological conditions required vary from one inoculum/crop host/soil combination to another and the rules of engagement are poorly understood. It is widely thought that plant species exert selective pressures on rhizosphere bacterial community structure, resulting in identifiable communities associated with that plant (Berg and Smalla, 2009). Marschner et al. (2001) have demonstrated that rhizobacterial communities change not only with species but also with plant growth and development. In that study, community changes were associated with nutrient cycling and root growth after winter rains. They suggest that the key drivers behind rhizobacterial community change and activity are quantitative and qualitative changes in root exudates.

While it may be expected that, over evolutionary time, different plant species can select for specific rhizobacterial communities, it is fascinating to consider this phenomenon in terms of the plant-breeding history. Traditionally, the key driver for plant breeding has been to maximise yield, and this has most often been achieved through high nutrient input. However, if specific crop cultivars can select for rhizobacterial communities with different

---

**Figure 11.3** Effects on root growth of the addition of a live or sterile soil (1:20 mix with sterile subsoil) as a biofertiliser inoculum source (Alogaidi et al., 2011).
functions such as nutrient cycling, then these traits may be exploited for plant growth promotion under the reduced fertiliser input regimes that are becoming increasingly necessary. Germida and Siciliano (2001), in an investigation of modern and ancient wheat cultivars, demonstrated that rhizobacterial community structure was greater on land races than modern cultivars. However, roots of modern wheat cultivars were aggressively colonised by fast growing endophytic bacteria, particularly pseudomonads. It is suggested that either root morphologies or root exudate profiles were responsible for this (Siciliano et al., 1998; Ferrieira et al., 2008), although Briones et al. (2002) found that microscale differences in O₂ availability along rice root cultivars were the key determinant in community profiles of ammonia-oxidising bacteria. In a study that investigated rice root exudate and plant development, Aulakh et al. (2001) confirmed that the chemical composition of exudate is under cultivar-specific control. Intracultivar differences were found as well as a shift from sugars to organic acids as the rice plants developed.

Furthermore, in an experiment that investigated the organic acid profiles of rice cultivars grown with and without access to P fertilisation, Elliot (2011) found the profile response to fertiliser P differed between cultivars (Figure 11.4a and b).

If the chemical signature of root exudate is cultivar specific (as evidenced in Figure 11.5), it can be hypothesised that cultivars will have different effects on rhizosphere chemistry and pH. In turn, this has implications for rhizosphere biological function such as P-mobilisation. The potential for linking cultivar-specific exudates to P-solubilisation and therefore crop nutrition is considerable. A promising way forward is to identify appropriate cultivar genetic markers (Quantitative Trait Loci or QTLs) linked to specific exudate patterns.

Figure 11.5 highlights that for the consistently successful use of microbial inocula, favourable edaphic conditions (and their associated management) are required at the point of inoculum delivery and during microbial inoculum establishment in the soil, and this is challenging to manage at the commercial scale. One of the most important aspects of this relates to the soil physical conditions, in particular, the matric component of soil water potential (Figure 11.6).

Figure 11.6 demonstrates that for most soils, with the exception of periods shortly after rainfall/irrigation or longer periods in soils with impeded drainage, active micro-organisms are confined to water in the soil micropores, and connectivity between these habitats is low. For bacteria introduced into soil, therefore, movement through the soil by chemotactic motility (over short distances) and mass flow (over greater distances) for colonisation of a root system or substrate is not possible without assistance from some other agent. These agents may take the form of motile soil animals (e.g. earthworms, nematodes and protozoa) or the apical tip of a developing plant root.

11.2.2 Restoration of soil after degradation

Restoration of soil structural integrity/soil stability and the role of soil organic matter

Loss of soil structural integrity/soil stability and consequent soil erosion represents one of the greatest challenges to meeting the food and fibre needs of the ever-increasing world population. Salinisation/desertification and soil mismanagement are the main causes of this loss of soil and soil fertility worldwide (Zhao et al., 2009). The loss in crop productivity due to soil degradation (from erosion) is estimated at 18 Tg of food staples per year at the 1990 level of yields for sub-Saharan Africa and 272 Tg for the world at the 1996
Sustainable management of soil and plant health by optimising soil biological function

This may represent a yield loss at the landscape level or even total crop failure at the farm level. Much of the soil’s structural integrity comes from its organic matter, including the living fraction (the soil microbial biomass).

Soil organic matter plays a key role in soil structure and consequently in stability in terms of resistance to erosion (Figure 11.7), with soil structural stability mediated by three main soil organic matter fractions (Cresser et al., 1993). The microbial biomass itself, both living and necromass, contributes through the enmeshing of soil particles. This mainly involves filamentous micro-organisms such as fungi and actinomycetes, which produce networks (mycelia) of tubes called hyphae. Physically protected organic matter, largely polysaccharidic in nature, also plays a vital part in stabilisation as the polysaccharide acts like glue holding fine soil particles together in aggregates. Organic matter moieties which carry a net negative charge are also important in soil structure, contributing to the development of stable soil aggregates through electrostatic interactions.

Figure 11.4 Radar plots of organic acid profile of five rice cultivars grown in sterile sand. Yoshida’s nutrient solution with P amendment (a) or no P amendment (b), n=4–5 (Elliot 2011). Note that in Figure 11.4a cultivar IAC25 shows a remarkably different organic acid profile from the other cultivars grown under P amendment. In Figure 11.4b, with limited P conditions, the cultivar Kasalath shows a different strategy from the other tested cultivars.
Figure 11.5  Factors controlling the success of a microbial inoculum in soil. Note that many of these factors can strongly interact – e.g. Predation of microbial inocula by protozoa and nematodes is regulated by matric potential as continuous water films are required for movement of these predators which are considerably larger than their microbial prey.

Figure 11.6  The relationship between soil matric potential and the neck diameter of the largest water-filled pores. For a colour version of this figure, please see Plate 11.2.
charge will bind to other components of the soil matrix which carry a charge. This is often by cationic bridging since much of the mineral and other organic fraction of the soil is also negatively charged (Killham, 1994).

Without the returning of adequate crop and animal residues, continuous agricultural production on the same land reduces the soil organic matter fraction, mainly through the process of mineralisation. Mineralisation without replenishment of stocks represents the ‘mining’ of soil organic matter. This type of agriculture is common worldwide and complete collapse of some soils can occur when organic matter levels fall below a critical threshold (Russell, 1977), as organic matter serves as both a short-term and more persistent binding agent. However, it should be borne in mind that the soil organic matter threshold value below which structural stability can be lost will vary from one soil type to another mainly because of the complex role of organic matter–clay interactions in the aggregate stabilisation process.

In addition to the problem of declining organic matter, loss of soil structure can also occur through the processes of slaking and dispersion. These processes are often linked to intensive cultivation (particularly ploughing), causing compaction and vital loss of the pore size distribution needed for maintenance of soil fertility (Greenland and Pereira, 1977). Because of these different processes, involving organic matter, soil mineral particles (particularly clays) and water, the mechanisms of soil structural collapse and degradation vary climatically and from one soil type to another, but the impact is generally devastating and the benefits of restoration are therefore considerable. An increase of 1 ton of soil carbon pool to degraded cropland soils may increase crop yield by 20–40 kilograms per hectare (kg/ha) for wheat, 10–20 kg/ha for maize and 0.5–1 kg/ha for cowpeas (Lal, 2004).

Figure 11.7  The mechanisms by which organic matter contributes to soil structural stability.
Soil management to restore soil structural integrity/soil stability involves the application of a range of technologies and management approaches. Some of these technologies and approaches enhance soil structure and include residue incorporation and mulching, use of appropriate soil conditioners, addition of manures and some forms of agroforestry. A number of additional technologies and approaches are designed to conserve soil and include cover cropping, contour ploughing, riparian zones, minimum tillage and efficient and appropriate irrigation. Álvaro-Fuentes et al. (2009) have demonstrated that in the semiarid Mediterranean agroecosystems, for example, the adoption of zero tillage together with the suppression of long-fallowing period can significantly increase the amount of SOC (particularly the hydrophobic components of humic material) stabilised in the soil surface and this can significantly improve soil structure and aggregation.

Management approaches and technologies to promote/restore structural stability through increasing soil organic matter have to be selected on a soil/site-specific basis but offer the practitioner of plant–soil management tools of soil restoration from physical degradation worldwide.

**Plants and algae for soil stabilisation**

Because plants are involved in both building and maintaining soil structure, plant management for soil structural stabilisation can be carried out in a number of ways.

Because plant root systems, particularly those with extensive rhizomes, can have a strong influence in stabilising soils, some plants are introduced to stabilise sands and degrade soils. This is a common practice to minimise coastal erosion in Europe and North America where *Spartina* (marram grass) and *Ammophila* (beach grass) have long been planted on unstable sand dunes, which are prone to erosion (Nordstrom, 2008). The mechanisms of stabilisation are threefold. The roots and rhizomes themselves enmesh the sand/soil particles, the carbon flow from the roots drives adhesion of particles both directly and indirectly through microbially produced polysaccharides, and mycorrhizal fungal hyphae associated with the roots also contribute to soil aggregation (Koske and Gemma, 1997).

Photosynthetically driven stabilisation of soils is also mediated by algae, particularly filamentous cyanobacteria such as species of *Microcoleus*, *Phormidium* and *Scytonema* and *Nostoc*. Threshold wind velocities for erosion of fine desert sands are increased by the presence of these cyanobacteria, particularly those that form thick, cohesive cryptogamic crusts (Hu et al., 2002). With further development of the algal crust, more and more fungi can grow, facilitating the formation of lichens. This further improves soil conditions which can encourage mosses and hence much greater protection against erosion. So, the algae have a key role as primary colonisers and in the early stages of formation of the crusts, which protect desert and other unstable, otherwise bare soils from erosion, as well as provide important inputs of C and N (Orlovsky et al., 2004).

Managing soils to improve structural stability through either encouragement of indigenous algae or the application of algal inocula (algalisation) is a new approach compared to the use of similar algae in rice paddies for biofertilisation. However, if the constraints of adverse water potential and pH can be overcome (most unstable soils in arid and semi-arid zones are too dry and often too acid or alkaline for ready colonisation by algae), then considerable improvements in aggregate stability can be achieved through algalisation (Nisha et al., 2007).
Soil animals – role in soil structure, ecosystem engineering and managing systems lacking soil animals

The soil mesofauna can amount to 2.5 ton of biomass in a single hectare of soil (Killham, 1994), and key members of the soil mesofauna have long been linked with maintenance and improvement of soil structure. This is particularly true of mesofauna such as earthworms and dung beetles. The effects of these can be very considerable, and some soils and their associated organic matter inputs (animal and plant detritus) are heavily processed by these animals. Along with plant roots, as discussed in the previous section, earthworms and other burrowing animals of sufficient size move soil particles to create and then stabilise (with mucilaginous stabilizing agents and faecal materials) cylindrical section pores which are key to good soil structure (Oades, 1993). It is not unusual for some top-soils in farming systems with lower inputs of fertilisers and particularly insecticides to entirely comprise earthworm cast material (Edwards, 2004).

Management of soil structure involving animals is of considerable importance in certain parts of the world. There are two broad approaches. One involves managing soil conditions to control faunal activity. Application of organic residues and manures, for example, encourages earthworm activity in many soils, assuming pesticidal inputs are low. This encourages incorporation of the organic inputs, cycling of the bound nutrients and modification of soil structure through burrowing/soil processing and lining of tunnels. The second management approach involves the actual introduction of fauna and may also comprise managing soil conditions to encourage faunal establishment and activity. In South Australia, for example, earthworms have been reintroduced into wheat belt soils where they had been lost through continuous fertilisation and associated acidification to beyond threshold pH limits for lumbricid worms (Edwards, 2004). Reintroduction involved incorporation of turves of earthworm-rich soil along with liming the soil to enable their successful colonisation/re-establishment. Dung beetles, belonging to the superfamily Scarabaeoidea, have also been used as ecosystem engineers and their introduction in pasture soils to cycle nutrients bound in faecal material from grazing cattle and enhance soil structure is well documented in a number of countries (e.g. Bang et al., 2005).

11.2.3 Remediation of soil contamination

Bioremediation versus conventional remediation (in situ and ex situ)

With strong regulatory and other environmental drivers, management of soil and plant health is becoming an increasingly important part of remediation of chemically contaminated land.

Even though estimates of the extent of chemically contaminated (involving organic contaminants and/or metals) land will vary with definition/standards as well as with methods of chemical analysis, and a global figure is unknown, the extent of the problem is very considerable indeed. Estimates of the extent of chemically contaminated land in England and Wales alone range from 50 000 to 200 000 ha (Royal Commission, 1996; EA, 2000), with up to a further 9 000 ha in Scotland (Scottish Government, 1998). The main areas of heaviest chemical contamination are associated with long histories of industrial development. However, the problem is worldwide, with chemical contamination from environmentally persistent pesticides and from diffuse atmospheric pollutants almost ubiquitous.

Bioremediation increasingly provides an attractive alternative (and complement) to the conventional ‘dig and dump’ remediation of contaminated land. The attraction is often based on lower costs and regulatory drivers particularly tied to reduced dependence on land-filling,
but also on the linkage between soil and plant health end points and bioremediation. This linkage is fundamental since bioremediation draws on the activity of the plant/soil system, either for the degradation or immobilisation of many of the target contaminants.

Two broad management approaches to bioremediation can be adopted, depending on whether contaminated material is excavated (*ex situ* bioremediation) or not (*in situ* bioremediation) for biological treatment. Both *in situ* and *ex situ* bioremediation of contaminated land require integrated management of the plant, soil and water system. The key objective is to remove any constraints to the activity of the micro-organisms which carry out bioremediative processes and then optimise the physiochemical conditions, which determine the rate of these processes (Figure 11.8). In managing bioremediation, issues of nutrient status/supply, water potential and oxygenation are of paramount importance to ensure optimal rates of remediation where the aim is degradation of organic pollutants, particularly most
petroleum hydrocarbons (Piotrowski et al., 2006). Other issues such as toxicity of contaminants and co-contaminants, varying bioavailability of target contaminants and pH-dependent toxicity (particularly for heavy metals) are more poorly understood and often require careful consideration to ensure successful bioremediation.

The identification and removal of underlying toxic constraints to bioremediation is illustrated in Figure 11.9 and represents a powerful approach to exploiting the full and diverse bioremediative potential of soil micro-organisms in the restoration of contaminated soils.

Bioremediation may aim to restore plant/soil health to its former status, but this is often unrealistic and unnecessary for many types of end use. However, indicators of soil health provide a useful tool for monitoring the bioremediation process. Margesin et al. (2000) showed that hydrocarbon biodegradation could be monitored well by using soil biological parameters, with residual hydrocarbon content correlating positively with soil respiration, biomass-C (substrate-induced respiration) and with activities of soil dehydrogenase, urease and catalase. Robust and ecologically representative indicators of soil health clearly need to be agreed so that relevant end point targets for remediation, whether biologically driven or otherwise, can be set. Useful efforts are now being made towards identifying suites of possible indicators for future monitoring of soil to ensure sustainable management (e.g. Ritz et al., 2009), acceptable to policy makers.

**Bioaugmentation versus biostimulation**

It is often necessary to carry out some kind of intervention to obtain satisfactory rates of bioremediation. Two key types of bioremediation intervention which offer considerable promise for effective management of plant and soil health are biostimulation and bioaugmentation. Biostimulation is based on the concept that suitable organisms/genes are present for bioremediation but that activity is restricted by a constraint that can be alleviated. It therefore involves the addition of electron acceptors, nutrients or electron donors to increase the numbers or stimulate the activity of indigenous biodegradative micro-organisms (Widada et al., 2002). Bioaugmentation is based on the concept that bioremediation cannot proceed because of a lack of suitable organisms and therefore involves the addition of (indigenous or nonindigenous) laboratory-grown micro-organisms capable of biodegrading the target contaminant (Vogel, 1996; Widada et al., 2002) or serving as donors of catabolic genes (Top et al., 2002). For most contaminated soils which retain reasonable biological activity, organisms for bioremediation are present, but their activity is constrained. So, biostimulation will be commonly applied as a management tool in bioremediation. However, in contaminated sites where suitable organisms for bioremediation are at low population densities or are absent, real benefits can be derived from bioaugmentation.

The advantages of bioaugmentation and biostimulation have been combined to remEDIATE pesticide (atrazine) contamination successfully (Silva et al., 2004). Bioaugmentation requires the degrading inoculum to reach the contamination in the soil/fill/stratum as well as carry out biodegradation of the target organoxenobiotic. In the field, this aspect of ‘delivering’ the degrader micro-organisms to their site of action may well be the most challenging part of a bioaugmentation programme as many organoxenobiotics become bound in micropores of the soil and these are not pathways for the percolating water which can carry the degraders. This constraint to bioaugmentation will often be an important factor in applying a combined approach of bioaugmentation and biostimulation as indigenous degraders may already have colonised these more inaccessible parts of the soil, and the activity of such degraders may be enhanced through biostimulation.
Soil Conditions and Plant Growth

**Rhizoremediation**

Synergistic enzymatic and chemical activities of plant and microbial metabolism in the rooting zone of contaminated soil can transform and degrade/immobilise contaminants and restore soil health.

**Figure 11.9** (a) The declining bioluminescence response of a *lux* gene marked bacterial biosensor to soil increasingly contaminated with copper (concentrations are given in pore water) demonstrates toxicity and explains why mineralisation of toluene, the target organic contaminant for bioremediation in this case, is increasingly inhibited by high concentrations of copper. (b) The relatively stable bioluminescence and toluene mineralisation across the same range of copper concentrations with the addition of EDTA is because this chelating agent complexes strongly with the copper and makes it unavailable (and hence non toxic) to the toluene mineralising bacteria. Sousa and Killham (2011).
In the case of metals, rhizoremediation usually involves their immobilisation in the soil or plant, while in the case of organic pollutants, it usually involves their degradation. Rhizoremediation of metal contaminants is generally most effective when the metals are taken up by the plant, potentially for removal when the plants are harvested, but is constrained by the fact that many of the strongest metal accumulators (particularly the hyperaccumulators) tend to be small plants and so future strategies may involve introducing their key genes involved in transporting and tolerating considerable metal loads (such as the genes encoding the cysteine-rich, metallothioneins metal-binding proteins) into other plants (Milner and Kochian, 2008). In the case of organic pollutants, the aim of rhizoremediation, as introduced earlier, is usually their degradation, either to nontoxic forms or through complete mineralisation to carbon dioxide.

The terms phytoremediation (Harvey et al., 2002; Morikawa and Erkin, 2003) and rhizoremediation (Kuiper et al., 2004) are commonly applied when plant metabolism contributes to the bioremediation process in some way. In certain cases, such as the elevated rates of hydrocarbon degradation in grassed compared to bare soils with similar physico-chemical properties (Gunther et al., 2010), it is the degradation of plant residues which is the main key to driving biodegradation or immobilisation of the target contaminants, while in others, the plant and/or its symbionts are more actively/specifically involved. This is exemplified by deep-rooted poplars which can access mobile, organic contaminants and degrade/detoxify them either by nonmodified plant metabolism, introduced pollutant transformation genes or by the action of microbial endophytes in the tissues of the trees (Doty, 2008).

The mycorrhizal symbiosis involving plant associations with fungal root symbionts (mycobionts) infers a number of properties to the host plant which are key to bioremediation of soil contamination and offer interesting options for plant/soil management. The ectomycorrhizal association, common to most trees, offers considerable metal immobilisation potential, for example, and enables phytoremediation with trees that can colonise sites to greater depths than hyperaccumulators or other metal tolerant plants. There are at least two mechanisms for this – one involves complexation of the metals with oxalates released from the mycorrhizal roots of conifers into the ‘mycorrhizosphere’ and one involves their uptake via the mycorrhizal fungal hyphae and subsequent immobilisation in the tissue of the tree (Meharg, 2003). Furthermore, because metal contaminated sites select for the mycorrhizal fungal community that can best cope with the presence of potentially toxic metal ions, the associated fungi are able to confer diverse physiological attributes to the host to enable colonisation of a wide range of sites (Meharg, 2003).

The mycorrhizosphere appears to be a key micro-habitat of the plant/soil system – where not only metals can be phytoremediated but organic contaminants also, through a variety of possible mechanisms (Joner and Leyval, 2003) – and offers additional potential for management of the plant/soil system for bioremediation.

There are numerous advantages to rhizoremediation, many of which are associated with the ability of the root system to penetrate otherwise inaccessible contaminants and make them accessible to plant or microbial metabolism. It is also a process ideally suited to the remediation of contaminants such as radioactive (Macek et al., 2009) and explosive (Gerhardt et al., 2009) materials which pose obvious risks to traditional ‘dig and dump’ approaches. One of the main drawbacks, however, is that it is usually a relatively slow and seasonal process and so it is often necessary to allow rhizoremediation to proceed for a number of years before clean-up soil/groundwater targets are reached.
**Indicators for monitoring bioremediation/restoration of soil health**

Management of bioremediation to restore soil health after damage from pollution requires robust indicators which can be monitored to assess any restoration of soil quality/health. A great deal of recent research has focussed on identifying which indicators will be most suitable, and although this work is incomplete (due both to the complexity of soils and the wide range of services that soil must supply and hence could be assayed), it is generally accepted that emphasis should be on biological properties of soils to complement existing and well understood physicochemical properties (Lilburne et al., 2004).

A wide suite of indicators was identified in a UK study by Ritz et al. (2009) and consists of molecular fingerprinting techniques as well as more traditional enzymatic and organism-based approaches, addressing the major biological trophic levels which mediate soil function. Reliable measurement of these indicators will require standard operating procedures, and confirmation will be required of their inherent sensitivity. It will also be necessary to confirm their capacity to discriminate between soil and land-use combinations, as well as their ecological/agroecological relevance. The question arises: whether the diversity of soil health indicators available can be integrated into a meaningful index for general application? To attempt to answer this, the approach of using a wide suite of indicators, as Ritz et al. (2009) have suggested, has been demonstrated by the development of a Soil Quality Index (SQI) (Alrumman et al., 2011) using test agricultural soils across a wide textural range (Equations 11.1–11.6). Two key bioassay types were loosely classified as either metabolic assays (soil microbial biomass, soil enzyme activities and respiration) or toxicity assays (higher organisms including plant assays, soil invertebrate assays and genetically modified bacterial sensors representing functional response of soil microbial communities). Bioassay data were transformed into an SQI modified from Canadian soil quality guidelines (www.ccme.ca) using the following formulation and equations:

A test pass/fail is decided, *a priori*, on evidence (or chemical determination) that the test soil is contaminated when compared to the control soil.

\[
F_1 = \frac{\text{Number of failed test}}{\text{Total number of tests}} \times 100
\]  

(11.1)

Then, the *Excursion* was determined using Equation 11.2 as follows:

\[
\delta_v = \left| \frac{\text{Mean of treated sample}}{\text{Mean of control}} - 1 \right|
\]  

(11.2)

The *ase* (average sum of excursion) was calculated using Equation 11.3:

\[
\text{as} \delta_v = \frac{\sum \delta_v}{\text{Number of failed test}}
\]  

(11.3)

Then, \(F_2\) was calculated using Equation 11.4 was applied to obtain range between 0 and 100

\[
F_2 = \frac{\text{as} \delta_v}{0.01 \times \text{as} \delta_v + 0.01}
\]  

(11.4)

Finally, the soil quality index (SQI) was calculated using Equation 11.5

\[
\text{SQI} = 100 - \frac{\sqrt{F_1^2 + F_2^2}}{1.414}
\]  

(11.5)
The value of the divisor (1.414) was calculated as follows:

$$\frac{\sqrt{100^2 + 100^2}}{100} = 1.414$$ \hspace{1cm} (11.6)

Toxicity assays were found to be more sensitive than metabolic assays. Conversely, metabolic assays were found to be more robust than toxicity assays. The SQI highlighted inter-soil differences according to physical class. Toxicity assays overestimated soil quality values compared to the General SQI. Assay type showed different aspects of soil quality with respect to contamination or treatment with implications for Risk Assessment. Models such as this show promise but should be further validated with a wide range of soil physical classes, history of use and range of contaminants. An enlarged data set will define the specificity of the SQI and highlight key variables contributing to soil health. A recent example of this methodology has been the identification of the recovery of soil following perturbation by hydrocarbon contamination (Alrumman, 2011).

In order for management of the soil/plant system to achieve a more global perspective, indicators for monitoring the effects of soil management should be internationally robust, but it is yet to be seen how readily identification of such a set will be achieved. Soil characterisation has traditionally focused on distinguishing key physical characteristics. This has helped provide data for soil maps that aid policy decisions for soil use. However, the identification of key soil (bio) indicators for contaminant remediation and post-remediation land use is likely to become increasingly useful. It may be possible to identify a core set of these (bio) indicators that are internationally transferable. Other less transferable indicators may also have to be adopted to meet the specific policy/soils requirements of a particular country or region.

### 11.2.4 Protection of plants from pests and diseases

**Conventional protection versus bioprotection**

As with managing nutrient supply, globally, protection of crop plants remains a chemically dominated industry, and this is reflected in the well-documented management systems that prescribe pesticide application. This section will explore management of the plant/soil system involving bioprotection.

With a few exceptions, commercially viable management systems exploiting consistent bioprotection of crops against soil-borne diseases and pests are largely confined to protected crops where soil and other environmental conditions are controlled. There are around 30 commercial biocontrol agents being used worldwide, but with a market value of less than £5 M (Yoshihiro, 2006), they are clearly not that widely used. Nonetheless, technical, economical and environmental drivers are forcing agriculture and forestry to consider new and sustainable methods, such as the use of biocontrol, for protection against soil-borne pests and pathogens.

Bioprotection and biocontrol management systems can either involve direct application of protection/control agents or the management of plant-soil conditions to enhance the activity of indigenous agents. The latter can involve suppression of activity of pests or disease causing agents or enhancement of activity of antagonists of those pests or diseases. Field management of potatoes, for example, can involve careful control of soil pH as the potato pathogen *Streptomyces scabies* is most aggressive under near neutral or alkaline conditions (Killham, 1994), while application of stable, mature composts can be highly
suppressive to a range of crop pathogens (Hoitink et al., 1996). Management of soil redox conditions through flooding of potato fields (this allows the decomposition of potato haulms to drive down the redox potential and generate anaerobic conditions) between crops can effectively inhibit potato cyst nematodes (McSorley, 1998).

In terms of introducing commercial agents, the most promising results to date for field managed, crop pest and disease control have mainly been derived from antagonistic fungi and bacteria. *Bacillus thuringiensis* is the most widely applied species of bacteria used for commercial, biological control, with at least four subspecies used to control lepidopteran (moth and butterfly caterpillars), coleopteran (beetles) and dipteran (flies) insect pests (Entwistle et al., 1993). *Trichoderma viride* has been used successfully against Dutch Elm disease and to treat the spread of fungal and bacterial pathogen growth on tree wounds, while other species in the genus *Trichoderma* are used to manage some soil-borne plant pathogens (Arora, 2003). However, even these more successful systems have not provided consistent, field-based control.

In terms of managing the plant/soil system for more reliable control of pests and diseases, biocontrol agents can be combined with other chemical products, with physical methods (solarisation or steam sterilisation) or agronomical practices, such as enhancement of suppressive soils, and as part of integrated pest management programs (Spadaro and Gullino, 2005). Because biological control is less consistent, less effective and slower acting than control normally achieved with pesticidal chemicals, successful management may well often depend on integration with other such control measures.

Although biocontrol through manipulation of the soil biota, whether exploiting indigenous populations or introduced inocula, suffers from inconsistent performance across the wide range of edaphic conditions involved, biological control forms a critical part of a ‘systems approach’ to integrated, sustainable soil pest management to combat insecticide-resistant pests and to minimise the future use of pesticides (Bale et al., 2008). It is becoming clear that successful and consistent biological control beyond the glasshouse requires further research to understand the biology/ecology of the pest/control agent/soil interaction so that soil management can adequately underpin the interaction. One such area where progress is being made is understanding the underlying mechanisms behind ‘conducive’ and ‘suppressive’ soils. It has been known for some considerable time that some soils become ‘suppressive’ over time. In these soils, protection of the crop from soil-borne pathogens is strongly driven by microbial action. This consists of PGPR bacteria and fungi colonising the root system. Much of the protection conferred by many strains of fluorescent pseudomonad PGPRs is attributed to the production of antibiotic secondary metabolites such as 2,4-diacetylphloroglucinol (DAPG) and/or phenazines. There have been many advances in understanding the biochemical and molecular underpinning of antibiotic (particularly DAPG) production. However, our understanding of antibiotic production *in situ* lags behind this. Duffy and Defago (1999) and Shanahan et al. (1992) showed the selective *in vitro* effects of different C sources and minerals on DAPG production. However, how DAPG production is affected by qualitatively changing complex C (i.e. mixtures of sugars, organic acids and amino acids, characteristic of root C flow) has not received much attention. Effects of this root-derived C on the growth and physiology of the soil microbial community is of great interest and a greater understanding of these presents an opportunity for effective bio-control. Figure 11.10 shows the response of a DAPG producer, *Pseudomonas fluorescens* CHA0 to defined C sources. It should be noted that DAPG production is minimal under 100% reducing sugars and that organic acids stimulate antibiotic production (Standing and Killham, 2007).
While soil microbe–iron relations and the production of siderophores have been implicated in biocontrol, they do not contribute as much as the antibiotics. Having said this, Mazurier et al. (2009) provide experimental validation for the following *Fusarium* wilt disease model: in suppressive soils a high microbial biomass provides a highly competitive environment with respect to C and iron sequestration. This physical lack of resources may suppress the saprophytic growth of the fungus *Fusarium* which leads to reduced pathogenicity and susceptibility of *Fusarium* to fluorescent pseudomonad-produced phenazine. Of the two antibiotics produced by microbial biomass DAPG appeared to be far less effective than phenazine. While the molecular mechanisms underpinning antibiotic production are well characterised in the laboratory the ‘field’ soil–plant–microbe system is far more complex and many putative signalling systems remain unknown (Haas and Defago, 2005).

Janvier et al. (2007) state that ‘Healthy soils are suppressive soils, thus disease suppressiveness can be considered as an indicator of soil health’. Thus, effective biological control and management of soil-borne plant pathogens and pests require a holistic approach – selection, optimisation and regional evaluation of biological control agents and improvement in the understanding of the influence of cultivars, environments and cropping systems on biological control (Janvier et al., 2007). Only full biological control evaluations will provide the critical information on performance of selected control agents under the diverse agroecological conditions of farming and forestry worldwide.

**Figure 11.10** Boxplots of DAPG production (μg mL⁻¹) by *P. fluorescens* CHAO in response to reducing sugars (90%, 70%, 50%) amended with different organic acid:amino acid ratios. Significant differences between treatments (n = 3) are indicated by different letters.
11.3 Conclusion

This chapter has demonstrated that sustainable management of soil and plant health through optimising soil biological function offers an increasingly promising way forward for our land but requires a thorough knowledge of the underpinning factors (many of which were introduced in Figures 11.5 and 11.8) which control this biological function. These factors rarely act singly and generally interact in a highly dynamic way. Despite this complexity, however, we are becoming better able to understand and hence exploit the rules which govern soil–plant–microbe interactions for benefits such as biofertilisation and bioprotection of crops, for restoring and maintaining soil structure, for bioremediation of contaminated land and for minimising climate change.

The rules mentioned in the previous paragraph, which appear to govern soil–plant–microbe interactions, have invariably been deduced from studies carried out at small scales and usually under controlled conditions in the laboratory. The challenge for soil biologists will be to explore how to scale-up and address the temporal and spatial variability of the field and landscape levels. Land managers can then consistently benefit from the increasing knowledge base of soil biology by systematically adopting more sustainable strategies of integrated management of soil and plant health, rather than the ‘trial and error’ approach often used in the past.

References


Managing the soil–plant system for the delivery of ecosystem services

David W. Hopkins\textsuperscript{1} and Edward G. Gregorich\textsuperscript{2}

\textsuperscript{1} School of Life Sciences, Heriot-Watt University, Edinburgh, UK
\textsuperscript{2} Agriculture and Agri-Food Canada, Central Experimental Farm, Ottawa, Ontario, Canada

12.1 Introduction

Ecosystem services are benefits humans gain from ecosystems either directly or indirectly. The \textit{Millennium Ecosystem Assessment} (MEA, 2005) provides a major synthesis of the ways in which the supporting, provisioning, regulating and cultural services combine to support humanity (Figure 12.1). Soils and plants have explicit, generally fairly obvious and central roles in the supporting services through, for example, soil formation; in the provisioning services through food, wood and fibre production, water supply and in some cases fuel production such as peat for burning; and in the regulating services through climate, flood and disease regulation, and water purification. Soils and plants also contribute to some of the cultural services as outlined in the MEA (2005). Our objective in this chapter is to summarise these roles and indicate how they can be managed with the aim of increasing the sustainability of soil–plant systems. We deliberately extend our coverage beyond the traditional areas of soil and plant science, such as agriculture, ecology and forestry, to include some aspects of cultural ecosystem services (Figure 12.1). Monitoring soil quality is also included because of the ongoing need to be able to assess the ability of the soil–plant system to deliver ecosystem services and the need for early warnings of change and thus potential lack of sustainability (Section 12.3). Sustainability is a widely used term, but for the present purposes we are following the definition in Gregorich et al. (2001): the long-term viability of a particular soil–plant (or land-use) system, which includes steps to ensure natural resource conservation, economic viability (including productivity and security) and social acceptability (including the cultural services). Clearly, for the soil–plant system to be truly sustainable in the sense of being able to continue indefinitely without any resource depletion or degradation is a more exacting and probably unachievable aim because, for example, considering an agricultural system, there is no escape from the thermodynamic principles. To be truly sustainable, all the nutrients and water removed with crops and any soil lost during cultivation and harvesting would need to be replaced by equivalent inputs of regeneration of soil by weathering and recycling. Despite the difficulties of deriving an absolute
definition of sustainability, the term is in general usage and has particular value in the context of the expectations placed on the soil–plant system as the human population expands towards the anticipated plateau of about 9 billion by around 2050. By this time, the amount of land will have declined to 1.6 ha per person, compared with 2.0 ha per person in 2005 and 5.2 ha per person in 1950.

There have been several attempts to place a monetary value on ecosystems goods and services, including those derived from soils. They are united in two different ways: whatever approach is used, the monetary values are astonishingly large (Table 12.1) and the concept that they have a real monetary value is an illusion because ecosystem services such as soil formation cannot be traded because there are no realistic alternative sources. Despite this second criticism, one of the most widely quoted sources (Costanza et al., 1997) places a value of US$17.1 × 10^{12} on soil formation and substantial values to other processes in which the soil–plant system participates, such as nutrient cycling and water regulation and supply (Table 12.1). The likely development of carbon trading over the next few decades as a response to the need to move to a lower carbon economy will lead to economic values being placed on carbon storage and sequestration in soil. The MEA (2005) suggests that the global value for carbon will be between US$10 billion (for the avoidance of doubt, 1 billion = 1000000000 = 10^9) and US$44 billion in 2010 and that a total of 4.5 billion tonne of carbon (or equivalent) will be traded giving a value of between US$2.2 and US$9.8 per tonne. This places the value of topsoil carbon (assuming 20 cm soil depth, 1.2 g cm^{-3} bulk density and 2% carbon concentration) between about US$100 and US$500 per ha. Pimental et al. (1997) also list a range of biological processes that occur either partly or predominantly in soils as amongst the most valuable benefits of maintaining biodiversity. These include waste disposal, soil formation, biological nitrogen fixation and bioremediation, with global values of US$760 billion, 25 billion, 90 billion

![Figure 12.1](link) Linkages between ecosystem services and human well-being, redrawn from the Millennium Ecosystem Assessment (MEA, 2005). For a colour version of this figure, please see Plate 12.1.
and 121 billion per annum, respectively. The apparent discrepancy between the value for soil formation and that of Costanza et al. (1997) is in part because Pimental et al. (1997) have considered only the biological contribution, not the geochemical and geophysical processes.

As an alternative to cash values, the concept *emergy* (as distinct from *energy*) can be used (Odum, 1996; Odum and Odum, 2000), although it is not without its critics. *Emergy* (or energy memory) is the amount of energy in total required both now and in the past to produce a particular service. In the context of plant production for food, for example, the *emergy* comprises the direct energy inputs harnessed in photosynthesis but also the energy required in cultivation, fertiliser production and distribution, crop harvesting, food processing and distribution, and even the past energy input required for cultivar development, soil formation, biogeochemical and hydrological cycling. Voo ra and Thrift (2010) estimate the annual *emergy* input from soil to global processes is about $1.4 \times 10^{19}$ J year$^{-1}$, which is about one-tenth of the amount from oil.

Leaving aside the difficulties of quantifying ecosystem services, the management of the soil–plant system, or any aspect of the ecosystem, for the simultaneous delivery of multiple ecosystem services is extremely challenging because of conflicting effects of different management options and the lack of sufficiently sensitive decision-making tools to account for the multiple interactions. The example in Table 12.2 summarises some of the conflicts for lowland agricultural grassland. Increased grass productivity and therefore milk or meat yield can be at the cost of a range of other ecosystem services, including air and water quality, biodiversity conservation and potential landscape quality (which depends in part on an individual’s quality judgement), and can be accompanied by deleterious processes such as unsustainable erosion and losses to groundwater (Table 12.2).

Clearly, assigning a value to different ecosystem services helps prioritise the importance of different ecosystem services in situations where they interact, but because they cannot be traded interchangeably, compromises are still essential, especially with the increasing demands being placed on the soil–plant system by the large and growing human population.

**Table 12.1** Estimates of the monetary value of ecosystem services.

<table>
<thead>
<tr>
<th>Ecosystem service</th>
<th>Value (trillion US$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Soil formation</td>
<td>17.1</td>
</tr>
<tr>
<td>Recreation</td>
<td>3.0</td>
</tr>
<tr>
<td>Nutrient cycling</td>
<td>2.3</td>
</tr>
<tr>
<td>Water regulation and supply</td>
<td>2.3</td>
</tr>
<tr>
<td>Climate regulation</td>
<td>1.8</td>
</tr>
<tr>
<td>Habitat</td>
<td>1.4</td>
</tr>
<tr>
<td>Flood and storm prevention</td>
<td>1.1</td>
</tr>
<tr>
<td>Food and raw materials</td>
<td>0.8</td>
</tr>
<tr>
<td>Genetic resources</td>
<td>0.8</td>
</tr>
<tr>
<td>Atmospheric gas balance</td>
<td>0.7</td>
</tr>
<tr>
<td>Pollination</td>
<td>0.4</td>
</tr>
<tr>
<td>All other services</td>
<td>1.6</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>33.3</strong></td>
</tr>
</tbody>
</table>

Note that US$1 trillion = $1 \times 10^{12} and that the total global gross ‘national’ product was US$18 trillion at the time of the assessment.

Source: Adapted from Costanza et al. (1997).
Table 12.2  Relationships between ecosystem services in lowland agricultural grassland with drivers in the rows and responders in the columns compiled from a review of peer-reviewed literature.

<table>
<thead>
<tr>
<th></th>
<th>Agricultural production</th>
<th>Climate regulation</th>
<th>Air quality</th>
<th>Water quality</th>
<th>Nutrient cycling</th>
<th>Biodiversity conservation</th>
<th>Landscape quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agricultural production</td>
<td><strong>↑</strong></td>
<td>*<strong>↑</strong></td>
<td><strong>↑</strong></td>
<td><strong>↑</strong></td>
<td><strong>↑</strong></td>
<td><strong>↑</strong></td>
<td><strong>↑</strong></td>
</tr>
<tr>
<td>Climate regulation</td>
<td><strong>↑</strong></td>
<td><strong>↑</strong></td>
<td><strong>↑</strong></td>
<td><strong>↑</strong></td>
<td><strong>↑</strong></td>
<td><strong>↑</strong></td>
<td><strong>↑</strong></td>
</tr>
<tr>
<td>Air quality</td>
<td><strong>↑</strong></td>
<td><strong>↑</strong></td>
<td><strong>↑</strong></td>
<td><strong>↑</strong></td>
<td><strong>↑</strong></td>
<td><strong>↑</strong></td>
<td><strong>↑</strong></td>
</tr>
<tr>
<td>Water quality</td>
<td>↔</td>
<td><strong>↑</strong></td>
<td><strong>↑</strong></td>
<td><strong>↑</strong></td>
<td><strong>↑</strong></td>
<td><strong>↑</strong></td>
<td><strong>↑</strong></td>
</tr>
<tr>
<td>Nutrient cycling</td>
<td><strong>↑</strong></td>
<td><strong>↑</strong></td>
<td><strong>↑</strong></td>
<td><strong>↑</strong></td>
<td><strong>↑</strong></td>
<td><strong>↑</strong></td>
<td><strong>↑</strong></td>
</tr>
<tr>
<td>Biodiversity conservation</td>
<td><strong>↑</strong></td>
<td><strong>↑</strong></td>
<td><strong>↑</strong></td>
<td><strong>↑</strong></td>
<td><strong>↑</strong></td>
<td><strong>↑</strong></td>
<td><strong>↑</strong></td>
</tr>
<tr>
<td>Landscape quality</td>
<td><strong>↑</strong></td>
<td><strong>↑</strong></td>
<td><strong>↑</strong></td>
<td><strong>↑</strong></td>
<td><strong>↑</strong></td>
<td><strong>↑</strong></td>
<td><strong>↑</strong></td>
</tr>
</tbody>
</table>

0, no relationship; ↓, responder declines as driver increases; ↑, responder increases as driver increases; †, no conclusive trends (i.e. both increases and decreases in the responder have been observed as the driver increases); ↔, insufficient published information. The intensity of shading indicates the nature of some of the interactions: light grey, win–win and dark grey, lose–lose. The strength of the interaction is indicated by the number of asterisks: ***, high confidence; ***, medium confidence; *, low confidence.

Source: Redrawn and adapted from Pilgrim et al. (2010).
12.2 Roles of soils

12.2.1 Food and fibre production

The concept of ecosystem services is well established, but perspectives of the relevance of particular services (Figure 12.1) differ depending on which service is most acutely needed. Where food supply is not secure, the provisioning services, particularly food production, tend to be prioritised. In large parts of the world, some of the other ecosystem services may be regarded as luxuries that can only be afforded when the food supply has been secured. The primary ecosystem service derived from the soil–plant system worldwide is the production of food and fibre. In view of the burgeoning world population and the increasing demand for food, one of the most critical aspects of managing the soil–plant system is the management of soil fertility and the capacity of soils to support crop production. With continued pressures to produce more food, there will be increasing pressures to develop strategies that increase fertiliser efficiency and improve the recycling of organic resources. With the growing public awareness of the need for environmentally sustainable practices in food production, there is increasing recognition of high efficacy and ecologically sound fertiliser practices. Fertilisation of soils can be intensified through agro-ecological methods using biological and technical knowledge in order to make more effective use of locally available resources and processes. This includes the management of soil fertility and innovative methods for recycling of organic matter and nutrients from farms, food processing, household waste and compost, and sewage disposal systems that produce biosolids (Figure 12.2a).

Availability of nutrients is considered a universally important attribute of soil quality and, because of its direct link to food production, it is one of the ecosystem services most studied. A part of soil fertility is the potential of the soil to supply nutrient elements in the quantity, proportion and form needed to support plant growth. In natural ecosystems, the release of these nutrients from organic materials present is often mediated by micro-organisms through decomposition, mineralisation and solubilisation reactions, or as a result of parent material being gradually broken down by weathering. In managed systems, this release of nutrients is supplemented by the addition of fertilisers to offset those removed from the soil–plant system during cropping at a rate that cannot be replaced by the natural processes. In intensively managed systems, this supplementation is achieved by the addition of inorganic fertilisers, manures, slurries, composts and other recycled ‘wastes’ and by management practices such as acceleration of the decomposition/mineralisation process through tillage as well as the use of plants with nitrogen-fixing symbioses. The fertility of soil can also be diminished by human activities, through use of management practices that promote soil erosion, nutrient mining (harvesting of nutrients at a rate in excess of their rate of replenishment), alteration of the soil biota or structure, salinisation through improperly managed irrigation, acidification through inappropriate fertilisation and deposition of acidifying pollutants.

The usage of inorganic nitrogen, phosphorus and potassium fertilisers (Figure 12.3; FAO, 2003) indicates increasing demands for both nitrogen and phosphorus, and approximately steady demand for potassium. The options for supplying nitrogen include biological fixation that globally contributes about 60% of the nitrogen used in agriculture and industrial fertilisers produced by industrial fixation (the Haber–Bosch process), which contributes about 40%. There are small additions of fixed nitrogen as a result of fixation during lightning strikes. The industrial fixation of nitrogen from the atmosphere by the Haber–Bosch process is an energy-dependent process, requiring high temperature and pressure, which has been possible in the developed world because of a ready supply of relatively inexpensive energy.
However, its long-term sustainability is not certain, but biological fixation remains a viable alternative to industrial fixation. However, the supply of phosphorus for fertilisers is likely to become limiting before restrictions on energy supply limit industrial nitrogen fixation. The global supply of phosphorus from phosphate rock for fertilisers is dwindling and likely to become exhausted in a few decades. This poses a more serious challenge globally and is
leading to considerable attention being focussed on reducing losses and inefficiencies in phosphorus usage, developing new approaches to exploit the reserves of organic phosphorus in soils and recovering phosphorus from sewage waste.

The fact that the supply of nitrogen and phosphorus limits plant productivity and the relative cheapness of fertilisers in some parts of the world has resulted in localised or regional over-application of fertilisers. This has led to subsequent nutrient leakage from agricultural soils with significant environmental implications. Fertiliser applications to agricultural land have more than doubled the quantities of reactive forms of nitrogen and phosphorus entering the biosphere (Galloway et al., 2008). Whilst these inputs have clearly helped to keep world crop production ahead of human population growth, the environmental costs of nutrient pollution from agriculture have been substantial, including the degradation of downstream water quality, eutrophication of aquatic ecosystems, production of photochemical smog and rising global concentrations of nitrous oxide (Vitousek et al., 1997).

An evaluation of nutrient balances in three maize (corn)-based agricultural systems illustrates that nutrient additions in intensive agriculture systems in different world regions range from inadequate to excessive and that both extremes have substantial human and environmental costs. In some developing countries, such as Kenya (Table 12.3), the application of fertilisers often does not replenish soil nutrients removed in the harvest crop, drawing down the nutrient capital of what may once have been high-fertility soils (Table 12.3). By contrast, policy-driven increases in fertiliser use have contributed to rising crop yields as China has striven for increased food security (Table 12.3). Fertiliser additions of nitrogen and phosphorus to many fields far exceed that taken up by the crop and much of the excess is lost to the environment, degrading both air and water quality. In Midwestern USA, increased nitrogen and phosphorus fertilisation has contributed to large increases observed in yields since the 1940s. Consequent hydrologic losses have caused eutrophication of freshwaters and the coastal Gulf of Mexico, which has increased the biological oxygen demand of gulf waters (the so-called hypoxia problem) as a result of the decay of algal and cyanobacterial blooms depleting the dissolved oxygen.

![Figure 12.3](#) World nutrient usage from 2007–2008 to 2011–2012 (FAO, 2010).
Reducing the risks of fertilisers having detrimental environmental effects will require a combination of better agricultural practices that raise fertiliser efficacy and efficiency and increased efforts to trap agricultural runoff before it leaches into watercourses. Better placement and timing of fertiliser applications including widespread adoption of precision technology (such as GPS-referenced yield mapping and fertiliser placement) can reduce the amount of fertiliser wasted in the field. Other strategies to keep excess fertiliser from contaminating watercourses include establishing vegetation in riparian buffer strips around crop and pasture lands and restoring natural wetlands. Over the long term, controlling water pollution from agriculture will require better coordination between agricultural and environmental objectives.

### 12.2.2 Water storage and purification

The global hydrological cycle describes the transfer and temporary storage of water in different reservoirs, which include the atmosphere, lakes, streams, oceans, glaciers, vegetation and soil, and the transfers between them by evaporation, transpiration, sublimation, condensation, precipitation, infiltration, runoff and groundwater flow (Table 12.4). Humans have substantially modified the global hydrological cycle in terms of both water quality and water quantity with the development of widespread intensive agriculture necessary to support the global population; 60–70% of the water used by humans is used for agriculture (FAO, 2003). The supplies of surface and groundwater (sometimes referred to as ‘blue water’) are currently essential inputs to agriculture through irrigation, which is used in systems supplying approximately 40% of the world’s food (Gleick, 1993). Water retained in soil and vegetation is sometimes referred to as ‘green water’ and it is estimated that 80% of agricultural water use comes from this source (Molden, 2007). In the future, explicit consideration of green water is likely to become much more important in the global water budget. For example, some forecasts suggest that 6 billion people will experience water shortages by 2050 if only blue water is taken into account, but this estimate falls to 4 billion people when green water is also accounted for (Rockström et al., 2009), indicating the importance of efficient management of water in the soil–plant system.

Between 15% and 20% of the water used globally goes to producing products that are exported and consumed out of the region/country of production. This water is exported in

---

**Table 12.3** An evaluation of nitrogen and phosphorus balances in three corn-based agricultural systems (kg ha\(^{-1}\) year\(^{-1}\)).

<table>
<thead>
<tr>
<th>Inputs and outputs</th>
<th>Kenya</th>
<th>China</th>
<th>USA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>P</td>
<td>N</td>
</tr>
<tr>
<td>Fertiliser</td>
<td>7</td>
<td>8</td>
<td>588</td>
</tr>
<tr>
<td>Biological nitrogen fixation</td>
<td>62</td>
<td>62</td>
<td>62</td>
</tr>
<tr>
<td>Total agronomic inputs</td>
<td>7</td>
<td>8</td>
<td>588</td>
</tr>
<tr>
<td>Removal in grain and/or beans</td>
<td>23</td>
<td>4</td>
<td>361</td>
</tr>
<tr>
<td>Removal in other harvested products</td>
<td>36</td>
<td>3</td>
<td>361</td>
</tr>
<tr>
<td>Total agronomic outputs</td>
<td>59</td>
<td>7</td>
<td>361</td>
</tr>
<tr>
<td>Agronomic inputs minus harvest removals</td>
<td>-52</td>
<td>+1</td>
<td>+227</td>
</tr>
</tbody>
</table>

Low input corn in Western Kenya, high input wheat and corn double-cropping systems in Northeast China, and corn–soybean rotations in the upper Midwestern USA.

*Source*: Derived from Vitousek et al. (2009).
Soil Conditions and Plant Growth

An increased awareness of how much water is used in production has led to the concept of virtual water (also known as embedded water or hidden water). This refers to the water used in the production of a good or service. In general, crop products have lower virtual water content than livestock products. For example, the global average virtual water contents of maize, wheat and rice are 900, 1300 and 3000 m$^3$ tonne$^{-1}$, respectively. In contrast, the virtual water contents of roasted coffee, chicken meat, pork and beef are 2100, 3900, 4900 and 15500 m$^3$ tonne$^{-1}$, respectively (Chapagain and Hoekstra, 2004).

Strategies to increase the amount and residence time of water in the soil involve manipulating the soil and/or crop residue, and the soil organic matter content by tillage or residue/mulch management, both of which increase the water-holding capacity of the soil and reduce evaporation losses. These strategies are relatively easy to implement because they are often part of the overall soil-crop management. Incorporation of organic materials to increase or maintain good structure will enhance the water-holding capacity of soils, particularly in light-textured soils. This incorporation of organic matter supports the soil micro- and macro-biological community. Invertebrates are particularly notable because of their role in water infiltration and runoff, as well as water movement within the soil profile (Swift et al., 2004).

Agricultural management practices that focus on altering green water significantly affect the predictions of water shortages. For example, it is estimated that modifying tillage/mulching management can reduce soil evaporation by 35–50% (Rost et al., 2009). Construction of ponds, dykes and subsurface dams to increase on-farm water storage allows farmers to redirect water to crops during droughts and can help recover more than 50% of the water normally lost.

Climate change is likely to induce greater variability in rainfall in some parts of the world which could lead to greater risk of droughts and floods (IPCC, 2007). In addition, higher temperatures will increase water demand (IPCC, 2007). Thus, under a changing climate, more intensive and extensive measures involving water harvesting will be required. Water-harvesting methods involve inducing, collecting and utilising runoff and have been practised since antiquity. These are complemented by modern techniques to increase the total yield of water per unit area and decrease the threshold of rain needed to form runoff. Other indirect methods that can be used to manage water more conservatively and improve the efficiency of water use are to develop crops by breeding or genetics that are more drought tolerant.

### Table 12.4 Inventory of water at the Earth’s and residence times.

<table>
<thead>
<tr>
<th>Reservoir</th>
<th>Volume (km$^3$ × 10$^6$)</th>
<th>% of total</th>
<th>Average residence time (years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oceans</td>
<td>1370</td>
<td>97.25</td>
<td>3 200</td>
</tr>
<tr>
<td>Ice caps and glaciers</td>
<td>29</td>
<td>2.05</td>
<td>20–100 (glaciers)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>20 000 (Antarctic ice cap)</td>
</tr>
<tr>
<td>Groundwater</td>
<td>9.5</td>
<td>0.68</td>
<td>100–200 (shallow)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>10 000 (deep)</td>
</tr>
<tr>
<td>Lakes</td>
<td>0.125</td>
<td>0.01</td>
<td>50–100</td>
</tr>
<tr>
<td>Soil moisture</td>
<td>0.065</td>
<td>0.005</td>
<td>0.1–0.2</td>
</tr>
<tr>
<td>Atmosphere</td>
<td>0.013</td>
<td>0.001</td>
<td>0.02</td>
</tr>
<tr>
<td>Streams and rivers</td>
<td>0.0017</td>
<td>0.0001</td>
<td>0.1–0.5</td>
</tr>
<tr>
<td>Biosphere</td>
<td>0.0006</td>
<td>0.00004</td>
<td></td>
</tr>
</tbody>
</table>

Source: Derived from Pidwirny (2006).
Filtration of water through soil provides potable water and water for agricultural (irrigation) and industrial uses, wildlife and aquatic habitat, and recreation. Pollutants such as metals, pathogens, organic contaminants and excess nutrients are adsorbed, modified and filtered out as water moves through soils in agricultural land, wetland areas, forests and riparian zones. Thus, there are human health, environmental, economic and recreational benefits derived from this water purification by soils. There are increased pressures to dispose municipal wastes on soil, and biosolids derived from sewage treatment plants are often spread on land. Liquid municipal biosolids may be applied as a source of nutrients for crops and organic matter to improve soil structure as well as to reduce sewage disposal in landfills. Soils with light texture and with large macropore networks facilitate rapid transport of contaminants from the surface or vadose zone to deep within the soil profile into tile drains and shallow groundwater. This rapid transport often bypasses the soil matrix where the contaminants can be sequestered, stabilised, degraded or available for uptake by plants. In the future, with increased pressures to use farmland as a repository for urban waste, it will be important to develop land application techniques that reduce the risk of contaminants such as pathogens, nutrients and metals from reaching groundwater, adjacent surface water source and/or subsurface artificial drainage networks.

12.2.3 Waste recycling and disposal

In all terrestrial ecosystems, with the exception of some urban settings, soils are the waste-disposal system for the residues, wastes and biomass after death (necromass). Although the estimate of US$760 billion per annum by Pimental et al. (1997) for the total monetary value of organisms involved in waste disposal includes both aquatic and terrestrial organisms, the value soils contribute to this process is inevitably large. Decomposition and recycling of natural organic wastes through the decomposer food chain relies on micro-organisms, which have major catabolic activities, and the soil fauna which contribute to the comminution, mixing, fragmentation and redistribution of organic materials in soils and the recycling by, for example, grazing on the primary microbial decomposer organisms. The decomposer, or detritus, food chain is therefore a complex set of loops in which material and energy are recycled between different groups of soil organisms, each assimilating some of the resources, releasing other components as metabolic wastes such as carbon dioxide, oxygen and ammonium (the process is widely referred to as mineralisation) and leaving some modified residue behind for another group of decomposers. The decomposition process carried out in soils is regulated by a combination of the biophysical conditions in the soil (e.g. temperature, water availability, aeration and acidity/alkalinity), the community of organisms active in decomposition (i.e. which organisms are present, what capacities they have and how active they can be under the prevailing environmental conditions), and the biochemical and physical properties of the material being decomposed (i.e. the nutrient and energy that can be obtained by the decomposer organisms) (Swift et al., 1979).

Environmental conditions

The combined physical and chemical conditions in which decomposition occurs both regulates the rate of decomposition and, in conjunction with the characteristics of the resources, imposes a selection pressure on the decomposer community thereby influencing its composition. Soil temperature, water content, acidity or alkalinity, oxygen supply and texture are amongst the most significant controlling environmental factors.
Organisms

Only 1–5% of soil organic carbon is contained in the living biomass, which predominantly comprises bacteria and fungi. Although probably not strictly correct, it is often assumed that all the micro-organisms necessary to complete the decomposition of any natural compounds (and many anthropogenic compounds) are present in soil and that collectively they are infallible. This assumption holds for many decomposition processes because of the large functional redundancy in decomposer microbial communities (see Section 12.2.5).

Substrate quality

The physiochemical characteristics of organic material – for example its solubility in water or hydrophobicity, water content, nitrogen and other nutrient content, biochemical recalcitrance and toxicity, and physical protection – strongly influence decomposition. Thus, the substrate quality of SOM can be regarded as a suite of combined properties that influence the supply of energy to heterotrophic soil organisms. Although this is a simple concept, the ability to assess substrate quality is not easy. It has long been recognised that different components of plant litter were decomposed at different rates and that this is a reflection of their resource value to decomposer organisms. The primary source of SOM is plant litter at various stages of degradation. Plant residues usually enter the soil as litter, including leaves, stems, roots and root exudates. Excreta from livestock is another source of organic matter in soils with faeces containing not only the undigested diet but also intestinal-dwelling bacteria and endogenous wastes from animal metabolism.

Nitrogen mineralisation

Nitrogen mineralisation is the result of a set of processes allied to decomposition which lead to the conversion of nitrogen in organic compounds to inorganic nitrogen, usually ammonium, carried out by a wide range of micro-organisms (Harmsen and van Schreven, 1955). This is significant because for many crop plants inorganic nitrogen is the main form in which nitrogen in taken up from the soil. Nitrogen mineralisation operates in effective opposition to nitrogen immobilisation because, depending on the nitrogen supply in the substrate being utilised by decomposer organisms, there may be a surplus of nitrogen in relation to carbon, or a deficit. Jansson and Persson (1982) describe these processes as the mineralisation–immobilisation turnover process in which inorganic nitrogen is released into the soil solution and assimilated from the soil solution. There is more recent evidence that soil micro-organisms assimilate nitrogen from organic sources in the apparently more efficient (from the micro-organisms’ perspective), direct immobilisation route with only the nitrogen surplus to the micro-organisms’ requirements being released into the soil solution (Figure 12.4).

If the substrate offers a nitrogen surplus to the decomposer organisms, inorganic nitrogen will be released and, if the substrate offers a nitrogen deficit to the decomposer organisms, they will uptake nitrogen from the external (soil) environment, i.e. immobilisation. Thus, the carbon–nitrogen ratio of the substrate strongly influences whether inorganic nitrogen is released during decomposition or whether the decomposer organisms make a demand on the soil nitrogen pool, whereby nitrogen is assimilated into microbial tissues. Typically, a carbon–nitrogen ratio of about 20 is the threshold for nitrogen mineralisation above which it is immobilised, because the substrate offers a nitrogen deficit to the decomposers, and
below which nitrogen is mineralised, because there is a surplus of nitrogen relative to in the substrate. However, the substrate quality and the composition of the decomposer community influence the relationship between carbon–nitrogen ratio of the decomposing substrate and mineralisation and immobilisation (Figure 12.5). Wheat straw has a carbon–nitrogen
ratio typically in the range 80–100:1 and will lead to net nitrogen immobilisation, whereas leaf and stem residues from a nitrogen-fixing legume with a carbon–nitrogen ratio in the range 12–15:1 will lead to net mineralisation. This provides an opportunity to manage the synchrony of nitrogen supply to plants, or to remove inorganic nitrogen from the soil pool and thus reduce the chance of nitrogen loss by denitrification and leaching (Figure 12.5).

12.2.4 Managing soils for climate change mitigation

Carbon management

Current efforts to mitigate the effects of human-driven climate change have led to a renewed and extended interest in soil carbon as a reservoir of carbon sequestered for the atmosphere. Soils are the third largest store of carbon globally, storing over twice as much as the atmosphere and almost three times the amount that is in vegetation. Carbon sequestration can be achieved by increasing net photosynthesis (e.g. planting new areas to trees) or by slowing the rate of decomposition of organic matter in soil through a change in land management (e.g. reduce tillage intensity) (Table 12.5). However, not all practices that increase soil organic carbon represent a transfer of additional carbon from the atmosphere to soil, because some simply move carbon from one biosphere pool to another, for example from the plant biomass to the soil, with no immediate implications for the atmospheric carbon dioxide content (Powlson et al., 2011). Assessing whether soil organic carbon content has changed as a result of a management change is itself fraught with difficulties arising from the fact that the rates of increase are slow (decade scale) and the actual amount of additional carbon in the soil may be small relative to the initial amount, the spatial variability of soils which challenges sampling protocols and the fact that in order to determine the soil organic carbon stock, the soil depth, density and carbon concentration need to be determined. With the exception of the unusual situation that occurs in long-term experiments, not all of these data are routinely available. Careful assessment of comparisons of management practices is necessary to determine whether true sequestration is occurring. An assessment of the few available studies with time series data has shown that gains in soil organic carbon were often due to a reduction in or cessation of carbon losses, rather than increases in accumulation (Sanderman and Baldock, 2010). They argued that soil organic carbon data from field trials may have limited predictive power when the state of the soil system is unknown and that current IPCC accounting approaches developed from the trial results may not properly credit these management activities. There are limitations to the extent and effectiveness of sequestration in soil or vegetation for mitigating climate change. First, the amount of soil organic carbon that can be sequestered in soil is finite. This is because the increase in soil usually follows an asymptotic relationship; the rate of increase shortly after initiating a new management practice is greatest in the early years but this rate slows, and eventually ceases, in later years as a new equilibrium level of soil is approached (Johnston et al., 2009). Second, the process of gain is reversible and a management practice such as reduced tillage that may lead to an increase in soil organic carbon content will need to be maintained to avoid subsequent carbon release.

Evaluation of management practices for climate change mitigation needs to be conducted with a full appraisal of all greenhouse gases (Table 12.5). Both nitrous oxide and methane have greater global-warming potential than carbon dioxide, so the effectiveness of management practices that lead to increased vegetation and soil organic carbon stocks have to be evaluated with respect to possible increases in nitrous oxide emission or decreases in methane uptake by soils.
Table 12.5  Comparison of GHG mitigation potential for agricultural land management practices in the USA, summarised from scientific literature.

<table>
<thead>
<tr>
<th>Activity</th>
<th>Soil carbon( t \text{ ha}^{-1} \text{ yr}^{-1} )</th>
<th>Land emissions( N_2O ) and ( CH_4 ),( t \text{ ha}^{-1} \text{ yr}^{-1} )</th>
<th>Direct Impact ( t \text{ ha}^{-1} \text{ yr}^{-1} )</th>
<th>Process and upstream emissions ( t \text{ ha}^{-1} \text{ yr}^{-1} )</th>
<th>Average net impact ( t \text{ ha}^{-1} \text{ yr}^{-1} )</th>
<th>maximum area (Mha)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conventional to conservation till</td>
<td>Mean 1.09 Max 1.82 Min 0.00</td>
<td>Mean 0.08 Max 0.38 Min 0.00</td>
<td>Mean 1.18</td>
<td>Mean 0.08 Max 0.10 Min 0.03</td>
<td>Mean 0.12 Max 0.10 Min 0.03</td>
<td>1.25 72</td>
</tr>
<tr>
<td>Conventional to no-till</td>
<td>Mean 1.09 Max 2.60 Min -0.26</td>
<td>Mean -0.18 Max 0.72 Min -0.91</td>
<td>Mean 0.90</td>
<td>Mean 0.14 Max 0.18 Min 0.07</td>
<td>Mean 0.18 Max 0.18 Min 0.07</td>
<td>1.04 72</td>
</tr>
<tr>
<td>Eliminate summer fallow</td>
<td>Mean 0.48 Max 2.35 Min -0.88</td>
<td>Mean -0.03 Max 0.16 Min -0.30</td>
<td>Mean 0.45</td>
<td>Mean -0.13 Max -0.07 Min -0.25</td>
<td>Mean -0.13 Max -0.07 Min -0.25</td>
<td>0.32 20</td>
</tr>
<tr>
<td>Use winter cover crops</td>
<td>Mean 0.83 Max 3.24 Min 0.37</td>
<td>Mean 0.25 Max 1.05 Min 0.00</td>
<td>Mean 1.07</td>
<td>Mean 0.61 Max 0.81 Min 0.41</td>
<td>Mean 1.69 Max 1.74 Min 0.41</td>
<td>1.69 74</td>
</tr>
<tr>
<td>Diversify annual crop rotations</td>
<td>Mean 0.58 Max 3.01 Min -2.50</td>
<td>Mean 0.07 Max 0.33 Min -0.04</td>
<td>Mean 0.65</td>
<td>Mean 0.00 Max 0.00 Min 0.00</td>
<td>Mean 0.00 Max 0.00 Min 0.00</td>
<td>0.65 99</td>
</tr>
<tr>
<td>Include perennial crops in rotations</td>
<td>Mean 0.57 Max 2.20 Min -1.75</td>
<td>Mean 0.03 Max 0.55 Min -0.55</td>
<td>Mean 0.59</td>
<td>Mean 0.17 Max 0.28 Min 0.14</td>
<td>Mean 0.17 Max 0.28 Min 0.14</td>
<td>0.77 56</td>
</tr>
<tr>
<td>Change from annual to perennial crop</td>
<td>Mean 2.26 Max 4.67 Min 0.00</td>
<td>Mean 0.12 Max 0.84 Min -0.55</td>
<td>Mean 2.38</td>
<td>Mean 0.54 Max 0.70 Min 0.39</td>
<td>Mean 0.54 Max 0.70 Min 0.39</td>
<td>2.92 13</td>
</tr>
<tr>
<td>Organic soil amendments (esp. manure)</td>
<td>Mean 1.53 Max 5.10 Min 0.18</td>
<td>Mean 0.05 Max 1.81 Min -1.35</td>
<td>Mean 1.59</td>
<td>Mean 0.94 Max 1.42 Min 0.46</td>
<td>Mean 0.94 Max 1.42 Min 0.46</td>
<td>2.53 8.5</td>
</tr>
<tr>
<td>Manage histosols to reduce GHG emissions</td>
<td>Mean 6.17 Max 15.03 Min 2.75</td>
<td>Mean 2.66 Max 3.09 Min 2.23</td>
<td>Mean 8.83</td>
<td>Mean 0.00 Max 0.00 Min 0.00</td>
<td>Mean 0.00 Max 0.00 Min 0.00</td>
<td>8.83 0.8</td>
</tr>
<tr>
<td>Reduce fertilizer N application rates</td>
<td>Mean 0.00 Max 0.00 Min 0.00</td>
<td>Mean 0.40 Max 1.32 Min 0.14</td>
<td>Mean 0.40</td>
<td>Mean 0.06 Max 0.08 Min 0.04</td>
<td>Mean 0.06 Max 0.08 Min 0.04</td>
<td>0.46 124</td>
</tr>
<tr>
<td>Change fert. N source – slow-release</td>
<td>Mean 0.00 Max 0.00 Min 0.00</td>
<td>Mean 0.46 Max 1.43 Min 0.00</td>
<td>Mean 0.46</td>
<td>Mean 0.06 Max 0.31 Min 0.17</td>
<td>Mean 0.06 Max 0.31 Min 0.17</td>
<td>0.52 93</td>
</tr>
<tr>
<td>Use nitrification inhibitors</td>
<td>Mean 0.00 Max 0.00 Min 0.00</td>
<td>Mean 0.49 Max 1.65 Min 0.00</td>
<td>Mean 0.49</td>
<td>Mean No data Max 0.09 Min 0.03</td>
<td>Mean No data Max 0.09 Min 0.03</td>
<td>0.49 99</td>
</tr>
<tr>
<td>Reduce chemical use (other than N)</td>
<td>Mean 0.00 Max 0.00 Min 0.00</td>
<td>Mean 0.00 Max 0.00 Min 0.00</td>
<td>Mean 0.00</td>
<td>Mean 0.07 Max 0.09 Min 0.03</td>
<td>Mean 0.07 Max 0.09 Min 0.03</td>
<td>0.07 122</td>
</tr>
<tr>
<td>Rice water management for ( CH_4 )</td>
<td>Mean 0.00 Max 0.00 Min 0.00</td>
<td>Mean 1.56 Max 5.22 Min -0.88</td>
<td>Mean 1.56</td>
<td>Mean No data Max 1.56 Min 1.3</td>
<td>Mean No data Max 1.56 Min 1.3</td>
<td>1.56 1.3</td>
</tr>
<tr>
<td>Rice variety development for ( CH_4 )</td>
<td>Mean 0.00 Max 0.00 Min 0.00</td>
<td>Mean 1.17 Max 2.71 Min 0.00</td>
<td>Mean 1.17</td>
<td>Mean 0.00 Max 0.00 Min 0.00</td>
<td>Mean 0.00 Max 0.00 Min 0.00</td>
<td>1.17 1.3</td>
</tr>
<tr>
<td>Improved grazing management, rangeland(^d)</td>
<td>Mean 1.01 Max 4.99 Min -0.10</td>
<td>Mean 0.28 Max 0.31 Min 0.27</td>
<td>Mean 1.30</td>
<td>Mean No data Max 1.30 Min 0.8</td>
<td>Mean No data Max 1.30 Min 0.8</td>
<td>1.30 166</td>
</tr>
<tr>
<td>Species composition on grazing lands</td>
<td>Mean 2.44 Max 4.84 Min 0.18</td>
<td>Mean -0.94 No range(^e)</td>
<td>Mean 1.50</td>
<td>Mean 0.00 Max 0.00 Min 0.00</td>
<td>Mean 0.00 Max 0.00 Min 0.00</td>
<td>1.50 80</td>
</tr>
<tr>
<td>Crop to pasture (LUC)</td>
<td>Mean 2.87 Max 4.70 Min 0.00</td>
<td>Mean 1.00 Max 4.96 Min 0.70</td>
<td>Mean 3.87</td>
<td>Mean 0.46 Max 0.69 Min 0.24</td>
<td>Mean 0.46 Max 0.69 Min 0.24</td>
<td>4.33 Unknown</td>
</tr>
<tr>
<td>Reduced rice acreage (LUC)(^c)</td>
<td>Mean No data Max 4.82 Min 0.00</td>
<td>Mean No range Max 4.82 Min 0.00</td>
<td>Mean No data</td>
<td>Mean 0.00 Max 0.00 Min 0.00</td>
<td>Mean 0.00 Max 0.00 Min 0.00</td>
<td>4.82 1.3</td>
</tr>
<tr>
<td>Cropland conversion to natural (LUC)(^c)</td>
<td>Mean 2.49 Max 4.74 Min -0.15</td>
<td>Mean 1.41 Max 5.06 Min 0.36</td>
<td>Mean 3.90</td>
<td>Mean 1.87 Max 2.63 Min 1.10</td>
<td>Mean 1.87 Max 2.63 Min 1.10</td>
<td>5.77 14</td>
</tr>
<tr>
<td>Convert Histosol cropland to natural (LUC)(^c)</td>
<td>Mean 21.17 Max 73.33 Min 2.20</td>
<td>Mean 1.50 Max 2.49 Min 0.51</td>
<td>Mean 22.67</td>
<td>Mean 1.87 Max 2.63 Min 1.10</td>
<td>Mean 1.87 Max 2.63 Min 1.10</td>
<td>24.54 0.8</td>
</tr>
</tbody>
</table>

(continued)
### Table 12.5 (Cont'd)

<table>
<thead>
<tr>
<th>Activity</th>
<th>Soil carbon ($t \text{ ha}^{-1} \text{ yr}^{-1}$)</th>
<th>Land emissions ($N_2O$ and $CH_4$, $t \text{ ha}^{-1} \text{ yr}^{-1}$)</th>
<th>Direct Impact ($t \text{ ha}^{-1} \text{ yr}^{-1}$)</th>
<th>Process and upstream emissions ($t \text{ ha}^{-1} \text{ yr}^{-1}$)</th>
<th>Average net impact area (Mha)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Max</td>
<td>Min</td>
<td>Mean</td>
<td>Max</td>
</tr>
<tr>
<td>Wetland restoration</td>
<td>3.11</td>
<td>7.70</td>
<td>0.45</td>
<td>-1.35</td>
<td>0.00</td>
</tr>
<tr>
<td>Short-rotation woody crops$^1$</td>
<td>2.71</td>
<td>10.22</td>
<td>0.00</td>
<td>0.76</td>
<td>1.52</td>
</tr>
<tr>
<td>Agroforestry (windbreaks, buffers, etc.)</td>
<td>2.72</td>
<td>4.23</td>
<td>0.84</td>
<td>0.76</td>
<td>1.52</td>
</tr>
<tr>
<td>Switch from dry land to irrigated</td>
<td>1.46</td>
<td>4.77</td>
<td>1.14</td>
<td>-0.42</td>
<td>-0.05</td>
</tr>
<tr>
<td>Irrigation improvements (drip, supplemental, etc.)</td>
<td>0.36</td>
<td>0.58</td>
<td>0.18</td>
<td>0.66</td>
<td>0.94</td>
</tr>
<tr>
<td>Biochar application</td>
<td>3.37</td>
<td>8.92</td>
<td>0.13</td>
<td>1.14</td>
<td>2.93</td>
</tr>
<tr>
<td>Change fert. N source – between common types</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.42</td>
<td>2.80</td>
</tr>
<tr>
<td>Change fert. N timing</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.35</td>
<td>0.52</td>
</tr>
<tr>
<td>Change fert. N placement</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.33</td>
<td>0.47</td>
</tr>
<tr>
<td>Improve manure management ($N_2O$)</td>
<td>No data</td>
<td>0.89</td>
<td>1.22</td>
<td>0.37</td>
<td>0.89</td>
</tr>
<tr>
<td>Improved grazing management, pasture</td>
<td>2.94</td>
<td>5.87</td>
<td>0.55</td>
<td>0.28</td>
<td>0.31</td>
</tr>
<tr>
<td>Rotational grazing, rangeland</td>
<td>Minimal to zero potential</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rotational grazing, pasture</td>
<td>2.20</td>
<td>2.90</td>
<td>1.50</td>
<td>0.05</td>
<td>0.11</td>
</tr>
<tr>
<td>Fertilizing grazing lands</td>
<td>1.05</td>
<td>5.87</td>
<td>0.37</td>
<td>-0.75</td>
<td>-0.60</td>
</tr>
<tr>
<td>Irrigation on grazing lands$^a$</td>
<td>1.04</td>
<td>1.83</td>
<td>0.00</td>
<td>-0.42</td>
<td>-0.05</td>
</tr>
<tr>
<td>Pasture to grassland (LUC) – cease grazing</td>
<td>Minimal to negative potential</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

All GHG units are in equivalents of carbon dioxide ($CO_2$) with 100-year time horizon global-warming potential. LUC, land-use change.

$^a$Cells that are shaded indicate limited scientific data available (i.e., estimate based on expert opinion or on three or fewer field or laboratory comparisons).

$^b$Area for rangeland management does not include any land of federal ownership.

$^c$“No range” is indicated where the value comes from one scientific expert opinion or other estimate, with no indication of variability.

$^d$Impact of reduced rice acreage depends on subsequent land use. These estimates account for elimination of current $CH_4$ emissions.

$^e$For Histosols, the total area farmed is highly variable in the literature.

$^f$For SRWCs and agroforestry, upstream and process emissions do not take into account the end-use for the aboveground biomass, because the GHG impact is highly variable (even though there is a likely GHG benefit). In practice, such lifecycle GHG impacts would need to be assessed.

$^g$n/a = not applicable, national total is not calculated because the net GHG benefit is negative.

$^h$Land emissions assume same $N_2O$ response as for irrigated cropland.

Source: Eagle et al. (2010).
Nitrous oxide

Nitrous oxide is formed as the product/by-product of denitrification and nitrification, both of which are influenced by the aeration status and inorganic nitrogen content of soils, albeit in different ways (nitrification is dependent on ammonium and an aerobic process, whereas nitrate is the main substrate for denitrification, which is carried out by facultatively anaerobic bacteria). Therefore, nitrous oxide emissions from agricultural land are related to the application of nitrogen fertiliser, soil and crop management, and weather conditions. The IPCC (2007) estimated that 70% of the total thermal forcing potential from greenhouse gas emissions from agriculture are attributable to nitrous oxide associated with the use of nitrogen fertiliser; therefore, improvements in nitrogen-use efficiency (i.e. less nitrogen applied for the same crop productivity) have potential to significantly reduce nitrous oxide emissions. Post-harvest residual soil mineral nitrogen concentrations are also positively correlated with leaching of nitrate, which can degrade water quality, as well as nitric oxide emissions and ammonia volatilisation, which also reduce air quality (Mosier et al., 1998). Nitrous oxide fluxes are highly variable temporally and spatially and variations in soil-water content, temperature and concentrations of soil mineral nitrogen are key factors affecting the ‘hot-spots and hot-moments’ of nitrous oxide emissions. These factors account for the very high variability in nitrous oxide emissions. For example, the majority of nitrous oxide emitted over a growing season is accounted for in short periods, very early in the season or during freeze/thaw cycles in winter/spring (Gregorich et al., 2005). Most fertiliser nitrogen management strategies to reduce nitrous oxide emissions are centred on the ‘4R’ framework described by Roberts (2006) – the right rate, right product, right time and right place.

Methane

In cool temperate regions, fertilisation, tillage and compaction can influence soil methane uptake and emissions. Rice soils emit methane because microbial respiration in flooded conditions reduces oxygen potential, creating anaerobic conditions that lead to methane production. Mineral nitrogen status plays a key role in methane uptake by soil and therefore fertiliser and manure amendment practices play a key role in managing for climate change mitigation.

12.2.5 Conservation of biodiversity

It is well established that organisms in soil contribute to a wide range of biological processes which are fundamental to many ecosystem functions delivered by soils (Section 12.1). The processes include the decomposition of organic matter (notably plant and animal wastes and remains, and organic contaminants); the transformation of nutrient elements, releasing them in plant-available, soluble or volatile forms, which may predispose them to loss from soil (most notably in the case of nitrogen); mixing and formation of channels within the soil matrix by soil fauna; stabilisation of soil structure through the production of extracellular peptides and enmeshing filaments and biocontrol of soil-borne plant pathogens and pests. The diversity of organisms is huge and even with the developments in modern molecular biological techniques, the details of the taxonomy, physiology and ecology of only a small fraction of the organisms is understood (Bardgett et al., 2005). The need to conserve soil biodiversity arises from both the roles of the organisms in delivering ecosystems services and because of the reservoir of the biological potential that may be exploitable in the future, provided it is conserved.
There is a view that the number of micro-organisms in soils, the ability of micro-organisms to be transported from place to place and the massive collective metabolic repertoire of micro-organisms mean steps to conserve specific taxonomic or functional groups are unnecessary. These ideas owe much to the proposition that with respect to micro-organisms in natural environments, ‘everything is everywhere, but the environment selects’, attributed to Baas Becking and Beijerinck (de Wit and Bouvier, 2006), coupled with the diverse physiological repertoire of both bacteria and fungi, leading to the proposition that microbial communities are infallible with respect to their biodegradation capacity. Although probably not strictly accurate, these propositions are good enough working assumptions because micro-organisms or microbial consortia capable of degrading virtually any natural compound and many xenobiotic compounds can be enriched from soil communities provided environmental conditions are conducive. For many functions carried out by soil organisms, it is usually sufficient to provide appropriate environmental conditions and resources for the functions to occur. Several of the soil processes that operate to benefit crop production are carried out by a wide range of organisms and there is a high degree of functional redundancy in the community (Figure 12.6). These processes may be described as broad processes because they are distributed across a broad range of different organisms which have different environmental tolerance so that the

**Figure 12.6** Hypothetical models of biodiversity–function relationships. The linear model implies ever increasing functioning with greater biodiversity. The curved model with redundancy (plateau at high biodiversity) implies that eventually no further increase in biodiversity is accompanied with an improvement in function because all the roles necessary for the function are fulfilled and there is spare capacity (redundancy). The corollary of this redundancy is that there is a catastrophic decline in functioning below a certain biodiversity threshold. The curved model with the ‘rivet’ hypothesis implies there is gradual loss of function with declining biodiversity when biodiversity is high, but a much sharper decline in function when biodiversity is low (the ‘rivet’ is derived from the analogy with a machine, such as an aircraft, held together with rivets; loss of some rivets is not catastrophic, but with further losses the machine will fail catastrophically). The optimum model implies that function is related to biodiversity at low levels, but that stress due to competition impairs function at high biodiversity. Finally, the idiosyncratic model represents the situation whereby it is not the biodiversity per se, but the traits of the individual groups of organisms that determine the functioning.
processes can operate over a wide range of environmental conditions. Such processes include decomposition of organic matter, some of the associated transformation of nutrient elements such as nitrogen mineralisation and biological contributions to the stabilisation of soil structure. Even under harsh environmental conditions (e.g. extremes of coldness or dryness), or when soils have been subject to contamination (e.g. metal-loaded sewage sludge) or have been deprived of significant plant inputs for decades (Hirsch et al., 2009), these processes can be maintained because the functions are distributed amongst a sufficiently large range of soil organisms even though the actual biomass of micro-organisms is severely depressed. These are all examples of the tremendous resilience of soil microbial communities arising from their very large functional diversity.

By contrast to the broad processes, narrow processes are those for which the functions necessary are distributed amongst a limited group of soil organisms or which only operate under a very specific set of environmental conditions. Examples include mycorrhizal associations, symbiotic nitrogen fixation, methane oxidation, nitrification, the decomposition of selected xenobiotic compounds and some antagonist interactions with plant pathogens and pests. These processes are much less resilient and more easily decreased or lost through conditions adverse to the limited groups of organisms performing them (Bardgett et al., 2005). In cases where a function is related to a narrow group of organisms, the absence of the correct species can impair functioning, and function can only be reinstated if the organisms are added and the conditions in the soil managed to ensure that the introduced organisms survive. The reinstatement of mycorrhizal associations and the reintroduction of earthworms to land restored after mining or other civil engineering operations are examples of such practices. Their success in restoring functions relies on both the presence of the organisms and the amelioration of soil conditions to allow their survival.

The second reason for conserving soil biological diversity is that soils are a reservoir of potentially exploitable organisms or capacities. This includes the long-established search, or bioprospecting, for novel antibiotic producers amongst soil micro-organisms that can be developed into therapeutic drugs. Probably the most famous example of pharmacological bioprospecting is the discovery and subsequent production of the antibiotic streptomycin from a soil streptomycete (actinobacterium), for which Selman Waksman was awarded the 1952 Nobel Prize for Medicine. Management of this aspect of soil biodiversity falls under the United Nations Convention of Biological Diversity (UNCBD, 1992). The objective of the UNCBD is to ensure conservation of biodiversity, sustainable use of its components and fair and equitable sharing of benefits arising from genetic resources. Implicit in the last of these aims is protection against biopiracy (i.e. the exploitation of indigenous knowledge about nature for commercial gain with no compensation to the indigenous people).

12.2.6 Preservation of soils that deliver cultural ecosystem services

Many systems of soil description regard soils as predominantly natural entities. This may be reasonable in parts of the world which have had low human population densities in the past, but it is not universally valid. Human activities in the past have affected soils by tillage, waste disposal, earth-working (e.g. ditches, dykes, terraces and burial mounds) and the establishment of field systems. Such soils may be regarded as cultural soils, because there is a record of human activity present in the soil structure, composition or morphology. The consequences are that many soils exhibit legacy properties from past land management (Davidson and Smout, 1996).
Soil Conditions and Plant Growth

The terra preta (Portuguese: black earth), or terra preta do índio, soils found in pockets in the Amazon basin are an example of a group of cultural soils that are currently receiving considerable attention. There is an argument that the terra preta soils are sufficiently special to merit conservation status. These soils are rich in charcoal, pottery sherds, and plant and animal remains characteristic of midden sites and date back to at least 2500 years before present. They are cultivated by the indigenous people, are particularly productive when compared with the surrounding soils and have a number of desirable characteristics for plant production, including large concentrations of plant nutrients and a low propensity for nutrient loss by leaching. Whether the incorporation of charcoal into the terra preta soils was a deliberate strategy by the original human inhabitants is uncertain, but its current presence is highly likely to contribute to the favourable soil properties for cultivation. The reason for the upsurge in attention being paid the terra preta soils in recent years is because of the suggestion that they provide an analogue for soils amended with biochar (charred biological materials produced under low oxygen conditions), which is being advocated enthusiastically in some quarters as a carbon sequestration treatment to mitigate climate change because of the inherent stability of biochar. In reality, the terra preta soils are probably an example of an anthropogenic soil, the result of long-term management that did not start as a deliberate strategy, but which was maintained because of the empirically observed benefits. Similarly, the plaggen soils of northern Europe are anthropogenic soils created since the twelfth and thirteenth centuries by artificially deepening through the application of turves and livestock bedding enriched with faeces and urine to surrounding fields.

Cultural soils, especially those containing archaeological evidence, can also deliver ecosystem services. Cultural soils serve as a medium for the preservation of buried archaeological remains, including the occurrence of crop marks that indicate the presence beneath the surface of archaeological features. Although the cultural aspects of soils have been investigated on a site-by-site basis, there are few general principles to draw on when considering their management. Management of cultural and archaeological soils is often subordinate to management of the soil–plant system for the provisioning and regulating services, but some aspects of the management have benefits to the delivery of provisioning, regulatory and cultural services, whilst others are in conflict. Examples of management that benefits the provisioning and/or the regulatory ecosystem services include management to reduce erosion and soil organic loss. Minimising erosion helps preserve the stratigraphic context of cultural soils and protects artefacts in archaeological soils, and conserving soil organic matter can preserve both the biomolecular record of human activity and the evidence of past land management. By contrast, tillage, especially deep tillage, disrupts cultural soils, may lead to direct plough damage to archaeological soils and can promote root penetration deep into the soil leading to physical damage to buried archaeological remains. Similarly, drainage may be desirable for food and fibre production, but it also affects moisture regimes and redox conditions threatening the survival of biodegradable organic artefacts and evidence.

Special conservation status is usually extended to ecosystems, landscapes, cultural features (such as monuments and other archaeological remains) and particular species of plants or animals (rarely micro-organisms). In many countries, soils per se are not the subject of special conservation measures because of the inherent value or rareness. For example, in the UK soils are not considered within existing geological criteria for Sites of Special Scientific Interest (SSSI), but the plant community they support may be. The review of geological conservation criteria for the designation of SSSIs (JNCC, 1996) states ‘the importance of soils, somewhat neglected hitherto in nature conservation, is now gaining increasing
prominence and their study within the concept of sustainability is likely to become an important task in the future’. The importance attached to soils does, however, differ between cultures. The ‘red centre’ of Australia is a clear reference to the soils of a harsh landscape in which Australian aborigine people can find sustenance; there is a strong notion of ‘geospirituality’ amongst many indigenous people including First Nations people in Canada and in the Middle and Far East. Figure 12.2b is a monument to the different soils in China, illustrating the importance of soils in that country and the recognition of their importance for food production. Indeed, the importance of indigenous peoples’ knowledge of soils is an important dimension of agricultural development (Pawluk et al., 1992). The soils of the environmentally extreme Antarctic dry valleys support no vascular plants because of the extreme cold and dry conditions (there has been no rainfall for millennia), but they do support a small active soil biological community (Figure 12.2c). They are considered as having inherent conservation value because of their rareness, extremeness and fragility, and they fall within the designation of either Antarctic Specially Protected Areas or an Antarctic Specially Managed Area. At the other end of a productivity scale, it might be argued that the most productive agricultural soils merit protection from development and degradation, and therefore remain available for food production.

12.3 Sustainable soil use and soil quality

Assessing the capacity of the soil–plant system to deliver ecosystem services is critical to assessing the sustainability. Rapid depletion of soil resources or soil degradation will impair functioning and thereby impair ecosystem service delivery. The concept of soil quality has been developed to address the need for a framework to assess soils. Soil quality is a long-established concept (e.g. Storie, 1932; Jenny, 1961), but the actual term was first introduced in the 1970s (Warkentin and Fletcher, 1977). Since then it has been variously defined and subtly reinterpreted (Table 12.6). However, the ability to perform functions and supply services is common to many of the definitions. Table 12.6 is not intended as a definitive summary of the definitions and interpretations, but it illustrates that the concept of value is common to many definitions, and that the soil properties linked to crop production dominate. This is because of the dominance of agricultural science in the development of soil and plant science during the twentieth and twenty-first centuries. Soil quality is closely allied to the soil health concept, which is an approach to assessing soil quality analogous to assessing human health in which the condition of a soil’s properties are assessed relative to some optimum or the ability of a soil to perform particular functions is assessed against goals placed on it (Gregorich et al., 2001). The important difference between soil quality and soil health is that a soil may be of low inherent quality, but still be in good health if the expectations placed on it are either suitably low or tailored to the environment. A comparison of this interpretation of soil health against the definitions of soil quality in Table 12.6 reveals why the terms are often used synonymously.

Because soil quality is multivariate, frameworks for assessing soil properties from measurable indicators are needed. The development of such systems has been a major activity by many agencies in recent years. Some examples of indicator sets are summarised in Table 12.7. The general principle is that indicators need collectively to cover the biological, chemical and physical attributes as well as often including geophysical attributes (e.g. climatic) and management (e.g. workability) of soils. In some cases, a minimum dataset has been proposed for soil quality assessment (Doran and Parkin, 1996), in others sophisticated
‘calculators’, sometimes web based, have been developed to assist land managers, farmers, policymakers and environmental regulators (e.g. SINDI, 2010). In some other schemes that include a large number of indicators, the importance and desirability of different indicators are ranked (e.g. Aalders et al., 2009). The main challenges in developing indicator sets for soil quality assessment has been to include indicators that are meaningful, readily measurable, cost-effective and which can be compared with data in existing databases to assess change. Scientific innovation has led to a wide range of research-orientated techniques being available, such as molecular biological approaches to assess the soil biological community, but these rarely find a place in routine soil quality assessment because they are often expensive to measure, rely on sophisticated equipment and highly specialised personnel, and because it is rarely possible to interpret the data unambiguously. Such approaches may find their way into future assessment procedures in the same way indicators now regarded as routine were previously cutting-edge research techniques. For this reason, soil archives and sample archives (e.g. −80°C stores of DNA extracted from soil) have an important role in long-term soil

Table 12.6 Examples of definitions and interpretations of soil quality.

<table>
<thead>
<tr>
<th>Definitions and interpretations</th>
<th>Country or region of origin</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>… key soil qualities important for crop production have been derived, namely nutrient availability and retention capacity, oxygen availability for roots, excess salts, toxicities and workability See Table 13.6 for specific details of the specific ‘qualities’</td>
<td>Global</td>
<td>FAO (2010)</td>
</tr>
<tr>
<td>… an account of the soil’s ability to provide ecosystem services and social services through its capacities to perform its functions under changing conditions See Table 13.6 for specific details of the specific ‘qualities’</td>
<td>European Union</td>
<td>European Commission Joint Research Centre Tóth et al. (2008)</td>
</tr>
<tr>
<td>… the capacity of a specific kind of soil to function, within natural or managed boundaries, to sustain plant and animal productivity, maintain or enhance water and air quality and support human health</td>
<td>USA</td>
<td>USDA Natural Resources Conservation Service NRCS (2010)</td>
</tr>
<tr>
<td>… value placed on a soil with respect to its fitness for a specific use and … categorisation of the soil for certain use-based ecological aspects</td>
<td>Canada</td>
<td>Gregorich et al. (2001)</td>
</tr>
<tr>
<td>… a soil that is productive and easy to manage under the intended land use; it has biological, chemical and physical properties that promote the health of plants, animals (including humans) … whilst also maintaining environmental quality</td>
<td>Australia</td>
<td>Soil Quality and Health Information for Australia (2010)</td>
</tr>
<tr>
<td>… soil quality is not a single concept, but encompasses aspects of soil physical structure, chemical fertility, nutrient storage, organic matter resources and the biological life in the soil See Table 13.6 for specific details of the specific ‘qualities’</td>
<td>New Zealand</td>
<td>SINDI (2010)</td>
</tr>
</tbody>
</table>
**Table 12.7** Examples of indicators used in soil quality assessment.

<table>
<thead>
<tr>
<th>System</th>
<th>Indicator</th>
<th>Soil quality information</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CHEMICAL INDICATORS</strong></td>
<td></td>
<td></td>
<td>SINDI (2010)</td>
</tr>
<tr>
<td>Olsen P</td>
<td>Plant-available phosphorus indicative of fertility</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pH</td>
<td>Acidity or alkalinity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total carbon</td>
<td>Organic matter reserves with implications for soil structural stability and water retention</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total N</td>
<td>Organic nitrogen reserves</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>BIOLOGICAL/ BIOCHEMICAL INDICATORS</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anaerobically</td>
<td>Availability of nitrogen to plants from the reserve</td>
<td></td>
<td></td>
</tr>
<tr>
<td>mineralisable N</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>PHYSICAL INDICATORS</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bulk density</td>
<td>Soil compaction, physical environment for roots and soil organisms</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Macroporosity</td>
<td>Availability of water and air, water retention and drainage properties</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>System</th>
<th>Soil qualities</th>
<th>Soil characteristics</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>FAO</strong></td>
<td>Soil quality for crop production</td>
<td>Soil texture, Soil organic carbon, Soil pH, Total exchangeable bases</td>
<td>FAO (2010)</td>
</tr>
<tr>
<td><strong>NUTRIENT AVAILABILITY</strong></td>
<td></td>
<td>Soil organic carbon, Soil texture, Base saturation, Cation exchange capacity of soil overall, Cation exchange capacity of the clay fraction</td>
<td></td>
</tr>
<tr>
<td><strong>NUTRIENT RETENTION CAPACITY</strong></td>
<td></td>
<td>Soil organic carbon, Soil texture, Base saturation</td>
<td></td>
</tr>
<tr>
<td><strong>ROOTING CONDITIONS</strong></td>
<td></td>
<td>Soil texture, Bulk density, Coarse fragments, Vertic properties and soil phases affecting root penetration, Soil depth, Soil volume</td>
<td></td>
</tr>
<tr>
<td><strong>OXYGEN AVAILABILITY TO ROOTS</strong></td>
<td></td>
<td>Soil drainage and phases influencing soil drainage</td>
<td></td>
</tr>
<tr>
<td><strong>EXCESS SALTS TOXICITY</strong></td>
<td></td>
<td>Soil salinity, sodicity and factors affecting salt conditions, Calcium carbonate and gypsum (in excess)</td>
<td></td>
</tr>
<tr>
<td><strong>WORKABILITY</strong> (constraining field management)</td>
<td></td>
<td>Soil texture, Effective soil depth/volume, Phases constraining soil management (such as shallowness, rock outcrops, excessive stoniness, gravel, concretions and hardpans)</td>
<td></td>
</tr>
</tbody>
</table>

(continued)
Soil Conditions and Plant Growth

### Table 12.7 (cont’d).

<table>
<thead>
<tr>
<th>Threat</th>
<th>Indicator (referred to as ‘criterion’ in the source document)</th>
<th>Source</th>
</tr>
</thead>
</table>
| Loss of soil organic matter | Soil type  
Soil texture  
Soil organic concentration  
Soil organic content  
Climate (mean annual precipitation and temperature)  
Slope, aspect and position in relief (top or bottom of slope)  
Land use or land cover | European Commission  
Joint Research Centre Tóth et al. (2008) |
| Erosion        | Soil type  
Soil texture  
Soil (bulk) density and hydraulic properties (water retention)  
Topography (slope)  
Land use or land cover  
Climate (annual precipitation and temperature, number of rain days, storm events and potential evapotranspiration)  
Hydrological conditions (catchment information and digital elevation model to give slopes)  
Agro-ecological zone (based on soil, climate and landscape) |                               |
| Compaction     | Soil mapping unit/soil typological unit  
Top and subsoil texture  
Climate (rainfall and evapotranspiration)  
Land use and land cover  
Slope |                               |
| Salinisation   | Soil mapping unit/soil typological unit  
Soil texture  
Soil hydraulic properties (conductivity, water retention and drainage)  
Irrigation area and intensity, salt content, sodicity and alkalinity of irrigation water  
Groundwater information (depth, salt contents, sodicity and alkalinity)  
Climate (annual rainfall and annual potential evapotranspiration) |                               |
| Landslides     | Record of previous landslides  
Bedrock  
Soil texture, structure and permeability  
Slope class (<10°, 10–30° >30°)  
Land use and land cover (including infrastructure, cultivation and mining)  
Climate (storm events)  
Seismic risks |                               |

The New Zealand SINDI and the FAO soil quality for crop production are included to compare the relatively narrow range of indicators needed for a small country with a relatively narrow range of land management and land-use options, compared with the system developed for global coverage. The European Commission Joint Research Centre system is included because it is focussed on identifying the risk of major threats involving soil that would affect its ability to deliver ecosystem functions. The European example also illustrates the widespread applicability of some key soil indicators, such as texture and climate, which provide valuable information for several contexts, as well as the subtle differences in common indicators (e.g. the climate indicator includes annual precipitation and/or number of rain days and/or number of storm events, and potential evapotranspiration depending on the threat).
quality assessment as they provide material that can be accessed when new techniques become more routine and when the purpose of soil monitoring switches between applications, such as changing emphasis from agricultural production to carbon sequestration for climate amelioration. When assessing a soil indicator against a target, or when interpreting any change in an indicator value either with time or in response to a treatment, it is important to realise that different indicators will respond differently; an increase in the indicator does not necessarily imply improved functioning or ability to deliver an ecosystem service. Therefore, the interpretation relies on either expert judgement at the time or on the expert judgement and evidence used to develop soil quality calculators.

Monitoring the condition of soil and how it changes over time is the major reason for collecting indicator data. This requires repeat measurements to judge both changes in soil quality and the efficacy of soil protection strategies (Arrouays et al., 2008). Whilst there are some examples of repeat measurements in the context of long-term experiments (e.g. Johnston et al., 2009), widespread re-sampling and measurement of indicators at a national or regional scale is relatively rare. For example, across Europe, soil at the majority of designated environmental monitoring sites has only been sampled once and therefore most soil monitoring is still at the pre-monitoring, inventory or baseline stage (Morvan et al., 2008). Re-sampling is not as straightforward as it might at first appear. To be useful in monitoring, a soil indicator should not be over-reactive and fluctuate or oscillate too rapidly; the most useful indicators are ones which capture the slow rates of change over years and decades but sensitive enough to initial subtle changes that may provide an early warning of a trend. This means that sampling has to be done on a scale measured over decades. As techniques, personnel and skills change, because records may be lost, and because the requirements of monitoring programmes change over these timescales, particular care and foresight are needed. Without wanting to diminish the efforts of major national scale monitoring and re-sampling exercises, the fact that one of the demands on soil monitoring changed over time from determination of soil organic carbon concentration (i.e. percent carbon by weight) to the need for estimates of total carbon stocks (i.e. quantity of carbon per unit area) has highlighted the difficulties of missing indicators from earlier surveys (e.g. bulk density and depth profiling), changing methods of sample collection and analysis and relocation of precise sampling sites from the map-based era to the satellite-positioning era, especially after land-use change. However, with increasing demands for soil monitoring information by regulators, policymakers and land managers, there will be a continued need for the collection of soil indicator data and the evolution of soil monitoring frameworks.

### 12.4 Concluding remarks

Managing the soil–plant system for the delivery of the range of ecosystem services humans expect is complex because of the competing demands and the opposing effects of different services on the properties of the soil–plant system. The soil–plant system contributes to all the major groups of ecosystems services (supporting, provisioning, regulating and cultural services), and rarely are any straightforward choices available when all the systems have to be managed to deliver all four different types of services. As the human population continues to increase, the demands on the soil–plant system will increase in proportion. Inevitably, this leads to choices and different services being prioritised by different land owners or societies. It has been traditional for the management of the soil–plant system to be addressed from the point of view of delivering provisioning services, and sometimes regulating
services. But the soil–plant system also contributes to other services and prioritisation is more complex than simply choosing, for example, food and fibre production over water quality. The choices and prioritisation will become more intense with increasing human population; yet unless it is possible to reconcile these demands, the soil–plant system will be unable to continue to deliver the full range of ecosystem services. As we explained at the outset, assigning either a monetary or an energetic value to ecosystems services delivered by soils provides an approach to prioritising different attributes of soils, but such attribution is only an analogy and the value cannot be realised. Managing the soil–plant system for the delivery of ecosystem goods and services will always be complex because of the many competing expectations placed on soils. Good knowledge of the physics, chemistry, biology and environmental science dimensions of soils will be essential, but so will a sound appreciation of soil and plant (crop) management, planning, economics and even the social sciences.

References


**13 The new challenge – sustainable production in a changing environment**

Peter J. Gregory¹,² and Stephen Nortcliff³

¹Centre for Food Security, School of Agriculture, Policy and Development, University of Reading, Reading, UK
²East Malling Research, Reading, UK
³Department of Geography and Environmental Science, Soil Research Centre, University of Reading, Reading, UK

### 13.1 Introduction

The last 60 years have seen a large increase in the world’s population from about 2.2 billion in 1950 to about 7 billion in 2011. Accompanying this increase has been a substantial increase in food production and use of resources such as energy, water and land not only to produce food but also to satisfy changed, generally more affluent and resource-demanding, lifestyles. In the last 50 years, production of cereal grains such as maize, rice, wheat and barley has almost tripled, while production of root crops such as potato and cassava has increased by around 40% and that of the coarse grains millet and sorghum by about 20%. Simultaneously, diets have changed to consume more meat so that numbers of chickens have increased 4.5-fold, pigs have increased 2-fold and cattle and sheep have increased by about 30% (Godfray et al., 2010).

Much of this increased production has come about by intensifying production systems and increasing yields (see Section 13.2.1 for more details), but it has often been achieved at some cost to other aspects of the environment. For example, excessive nutrient inputs, especially of nitrogen and phosphorus, have resulted in coastal eutrophication and reduced the quality of water in reservoirs used for drinking water (Vitousek et al., 2009). Similarly, cultivating soils for crop production has often increased the frequency of substantial soil erosion by either water, tillage or wind so that the current mobilisation of soil globally is 35 (±10) Pg year⁻¹ or about 5 t year⁻¹ for every person on the planet (Quinton et al., 2010).

This increasingly common degradation of soils, land and water as a result of agricultural intensification has led to the exploration of notions of land being needed to provide multiple functions simultaneously. These are commonly referred to as ecological or environmental goods and services and include elements of provisioning (e.g. food and fibre production), regulating (e.g. climate regulation and water purification) and cultural and non-material benefits (e.g. recreation and aesthetic and cultural benefits), which are all underpinned by the supporting services of soil formation, nutrient cycling and primary production (Millennium Ecosystem Assessment (MEA, 2005); see Chapters 1 and 12 for details). Such thinking has led to demands for changed agricultural practices that integrate biological and
ecological processes into food production so that agricultural systems become more sustainable and deliver the other essential services (Pretty, 2008; Powlson et al., 2011).

The purpose of this chapter is to examine the multiple environmental changes that humans are inducing and to reflect on the potential contributions that better understanding of soil and plant interactions might have on sustaining the requisite intensive crop production needed in future.

### 13.2 Human-induced environmental change

#### 13.2.1 Land use change – extensification and intensification

Increased crop production has occurred as a result of both extensification (altering natural ecosystems to produce products) and intensification (producing more of the desired products per unit area of land already used for agriculture or forestry). Only about 3 billion hectares of the world’s 13.4 billion hectare land surface is suitable for crop production and about one-half of this is already cultivated (1.4 billion ha in 2008; Greenland et al., 1998). The remaining potentially cultivatable land is currently beneath tropical forests and some grassland areas, but it would be undesirable to convert this to arable land because of the effects on conserving biodiversity, greenhouse gas emissions, regional climate and hydrological changes, and because of the high costs of providing the requisite infrastructure. As in the recent past, future increases in crop production are generally projected to come from increased yield (68%) and higher cropping intensity (12%) with only a small proportion coming from the cultivation of new land (about 20%; Bruinsma, 2003; Smith et al., 2010b).

Table 13.1 shows that while intensification will dominate, extensification will contribute significantly to crop production in sub-Saharan Africa (27%) and Latin America and the Caribbean (33%). There is almost no land available for expansion of agriculture in South and East Asia and the Near East/North Africa (there may, indeed, be loss of agricultural land to urban development) so that intensification will be the main means of increasing production (Gregory et al., 2002; Bruinsma, 2003).

Intensifying crop production through increased yields per unit area will continue to be a major means of delivering the increasing human demand for food and fibre. As cereal production (wheat, maize and rice) has increased from 877 million tons in 1961 to 2342 million tonnes in 2007, the world average cereal yield has increased from 1.35 to 3.35 t ha$^{-1}$ over the same period and is projected to be about 4.8 t ha$^{-1}$ in 2040. Simultaneously, per capita arable

<table>
<thead>
<tr>
<th>Table 13.1</th>
<th>Projected contributions (%) to increased crop production between 1997/1999 and 2030.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Land area expansion</td>
</tr>
<tr>
<td>All developing countries</td>
<td>21</td>
</tr>
<tr>
<td>Sub-Saharan Africa</td>
<td>27</td>
</tr>
<tr>
<td>Near East/North Africa</td>
<td>13</td>
</tr>
<tr>
<td>Latin America and Caribbean</td>
<td>33</td>
</tr>
<tr>
<td>South Asia</td>
<td>6</td>
</tr>
<tr>
<td>East Asia</td>
<td>5</td>
</tr>
</tbody>
</table>

Source: Derived from Bruinsma (2003).
land area has decreased from 0.415 ha in 1961 to 0.214 ha in 2007 (Smith et al., 2010b). Put another way, had the increases in yield of the last 60–70 years not been achieved, almost three times more land would have been required to produce crops to sustain the present population, land that, as indicated, does not exist except by using some that is unsuitable for cropping. Without the increased yields resulting from intensified practices, competition for land for different purposes would have been greater (Smith et al., 2010b). As demands for land to satisfy the required multiple ecosystem services increase, so competition for land will intensify, although there are major uncertainties in such projections (Smith et al., 2010b).

While the immediate loss of biodiversity as a consequence of clearing land for crop production has been well documented; the higher level consequences of increased crop yields on sparing land for biodiversity conservation is a matter of considerable debate. This is because (1) the on-farm losses of biodiversity due to practices giving high yields may outweigh the benefits of sparing biodiverse habitats; (2) high-yielding crops may have negative effects on offsite biodiversity and (3) land-sparing does not occur or is imperfect. The complexity of the factors involved is evident in the findings of Ewers et al. (2009), who analysed the changes in yields of 23 staple crops for 124 countries between 1979 and 1999. While per capita area of the 23 staple crops decreased in developing countries where large yield increases occurred, this was counteracted by a tendency for an increased area of non-staple crops leading to only a weak tendency for land-sparing overall. In developed countries, there was no evidence that higher yields reduced per capita cropped area possibly because of the role of agricultural subsidies in promoting production thereby overriding any land-sparing effects. Ewers et al. (2009) concluded that land-sparing is a weak process, but that improved agricultural technology may have contributed to the maintenance of natural vegetation cover in the past and that future conservation benefits, while debatable, are potentially available if land-use policies are also modified.

Another factor influencing the future use of land will be the changing climate. For example, warming of the climate in Scotland as documented by Barnett et al. (2006) and Gregory and Marshall (2012) could have substantial beneficial effects on the land-use potential for agriculture in Scotland resulting in potential expansion of areas of ‘prime’ agricultural land especially in the east and south of the country (Brown et al., 2008). However, while warming is well established, a crucial and presently largely unknown factor influencing future land use is precipitation. In Scotland, wetter winters are already being experienced leading to waterlogging of some autumn-sown crops such as wheat and rape and to delays in land preparation for spring-sown crops. However, should summer rainfall decline (there is no significant decrease currently identified using long-term data), then this could act in future to reduce land capability (Brown et al., 2011). In parts of the world in which water already limits crop production, the combined effects of increased temperature and decreased rainfall could have severe detrimental effects on land capability and yields (Fischer et al., 2005).

### 13.2.2 Greenhouse gases and biogeochemical cycles

Changes in the terrestrial carbon cycle as a consequence of the burning of fossil fuels are well documented (see Chapter 4), and the contribution of changes in land use to soil organic carbon and emissions of carbon dioxide (CO₂) to the atmosphere have been detailed in Chapters 4 and 12. This section focuses on the role of soils and plants in the production of the greenhouse gases (GHGs) methane (CH₄) and nitrous oxide (N₂O).

GHGs comprise CO₂, CH₄, N₂O, hydrofluorocarbons, perfluorocarbons and sulphur hexafluoride, and all have different degrees of effectiveness in retaining heat in the earth’s
Soil Conditions and Plant Growth

atmosphere. Several surveys have indicated that the direct emissions of GHGs from agriculture contribute about 7% of a country’s total GHG emissions (calculated as mass of CO₂ equivalent; AEA, 2007; EEA, 2007). While this is small in comparison with emissions from the energy sector (85.6% of the total), it is comparable with those from waste and industrial processes combined (7.7%), and attention will continue to be focused on attempting to reduce agricultural emissions. The methodology used to calculate these figures does not take account of the emissions resulting from the use of energy on farms for heating and fuel, nor the energy required to transport and process farm products for markets. When the whole food chain is considered, the contribution to GHGs increases to about 18% (EEA, 2007).

Net GHG emissions from agriculture consist entirely of CH₄ and N₂O because CO₂ uptake in photosynthesis is normally greater than any losses of CO₂ from, for example, soil respiration, unless major clearance of trees occurs (AEA, 2007). Emissions of these gases are very dependent on the form of agriculture practised. For example, in the UK, agriculture is the major contributor to emissions of N₂O in all parts of the UK but to emissions of CH₄ in only Scotland, Wales and Northern Ireland, where animals constitute a relatively larger component of agricultural activity (Table 13.2). Enteric fermentation is the major process contributing to CH₄ emissions (typically 85–90% of the release from agriculture), so stocking levels play a major role in determining the amounts of CH₄ released. This picture in the UK is broadly similar across Europe as a whole where agricultural emissions of GHGs comprise direct soil emissions (26%), cattle (CH₄, 26%), indirect emissions (N₂O, 18%), animal production (N₂O, 7%), cattle (CH₄, 5%), pigs (CH₄, 5%), sheep (CH₄, 5%), solid stores and drylots (N₂O, 5%) and others (4%; EEA, 2007).

### Table 13.2

<table>
<thead>
<tr>
<th>Country</th>
<th>% Total emissions</th>
<th>Enteric fermentation (% agric. emissions)</th>
<th>% total emissions</th>
<th>Fertilisers and manures (% agric. emissions)</th>
</tr>
</thead>
<tbody>
<tr>
<td>England</td>
<td>28</td>
<td>84</td>
<td>61</td>
<td>64</td>
</tr>
<tr>
<td>Scotland</td>
<td></td>
<td></td>
<td>77</td>
<td>63</td>
</tr>
<tr>
<td>Wales</td>
<td>60</td>
<td>90</td>
<td>82</td>
<td>71</td>
</tr>
<tr>
<td>N. Ireland</td>
<td>78</td>
<td>86</td>
<td>79</td>
<td>69</td>
</tr>
</tbody>
</table>

Methane

The concentration of CH₄ in the atmosphere has increased rapidly from its pre-industrial concentration of about 750 ppbv to over 1750 ppbv today. This concentration is greater than that at any time in the last 650,000 years and is estimated to result in an additional radiative forcing of about 0.5 Wm⁻² compared with that in 1750 (IPCC, 2007). There are many natural sources of CH₄ including termites, wild animals and oceans but wetlands (174 Tg year⁻¹) dominate the estimated 238 Tg year⁻¹ emitted (Reay et al., 2010). Anthropogenic sources of CH₄ (336 Tg year⁻¹) are made up of many components but with land-based activities dominant including emissions from ruminants (84 Tg year⁻¹), rice agriculture, landfills and waste (54 Tg year⁻¹ each), and biomass burning (47 Tg year⁻¹; Reay et al., 2010).
More than one-third of CH$_4$ emissions come from soils as a result of the microbial breakdown of organic matter in anaerobic conditions. This process of methanogenesis is commonplace in flooded soils used to grow rice, in landfill sites and in the gut of some species of soil-dwelling termites. Rice is produced on about 160 Mha worldwide on a variety of soils in a range of production systems (Greenland, 1997). Flooded, paddy soils used to produce rice are complex systems comprising both aerobic (the soil–floodwater interface and zones around plant roots) and anaerobic (the bulk soil) regions (Figure 13.1). Respiration by plant roots and micro-organisms depletes the oxygen in the soil–water system so that other chemical species (nitrate, manganese (IV), iron (III), sulphate and CO$_2$) then act in turn as electron acceptors in the respiratory process and are reduced (Ponnamperuma, 1972). Following their reduction, methanogenesis occurs with both hydrogen and acetate acting as major drivers of CH$_4$ production. Both hydrogen and acetate are produced from the degradation of organic materials, but hydrogen is rapidly consumed (in minutes) as an electron donor in many redox reactions, while acetate turnover is typically slower and occurs in hours or days (Conrad, 1989; Yao et al., 1999). Methanogenic bacteria become active as the

**Figure 13.1** The production, oxidation and emission of methane in paddy rice fields. Reproduced from Yagi et al. (1997). With permission from Springer Science and Business Media.
partial pressure of oxygen decreases, the redox potential declines and the partial pressure of hydrogen increases.

Emissions of CH$_4$ occur through several pathways (Figure 13.1; Yagi et al., 1997). In temperate regions, more than 90% of the CH$_4$ can be lost directly from plants following transport of the gas from soil via the aerenchyma, whereas in tropical paddy fields, significant quantities of CH$_4$ are evolved through bubbles (ebullision; Schütz et al., 1989; Denier van der Gon and Neue, 1995). Some of the CH$_4$ produced is consumed by methanotrophs inhabiting oxic regions (such as the soil–floodwater interface), which oxidise the CH$_4$ to CO$_2$.

The production of CH$_4$ from rice fields is substantially affected by the duration of flooding (i.e. the intensity of the anaerobicity) and the quantity and quality of organic matter added; lesser factors include the chemical fertility of the soil and the rice cultivar (Conen et al., 2010). Water and residue management appear as key factors in the potential mitigation of this gas with potential to reduce emissions by about 30% globally (Conen et al., 2010). Draining at least once during the season can reduce emissions by about 15%, while incorporation of straw in the off season before flooding, or composting it off-site, can simultaneously retain most plant nutrients (including most of the nitrogen), reduce the energy content and thereby cut emissions also by another 15%. It may also be possible in future to select rice cultivars to reduce emissions. Dissolved organic compounds released by roots are an energy source for methanogens, but the quantities released are cultivar dependent so that scope may exist to reduce this source of energy (Sigren et al., 1997; Lou et al., 2008). Furthermore, increasing the rate of oxygen diffusion from roots into soil could inhibit CH$_4$ production and stimulate CH$_4$ oxidation at the soil–root interface. Satpathy et al. (1998) found a fourfold difference between ten rice cultivars in seasonal CH$_4$ emissions with similar yield in nine of the cultivars.

**Nitrous oxide**

For most of the last 4000 years, N$_2$O concentrations in dry air were close to 270 ppbv, but since 1850, atmospheric concentrations have risen and now exceed 320 ppbv. This is significant because N$_2$O in the atmosphere has two distinct environmental consequences (Smith, 2010). First, its capacity to absorb infrared radiation is about 300 times greater than that of CO$_2$, so it is a significant contributor to the greenhouse effect (about 6% of the current greenhouse effect; IPCC, 2007) even though its concentration is 1000 times less than that of CO$_2$. Second, when it reaches the stratosphere, it contributes to the loss of ozone and thereby to the penetration of ultraviolet radiation to the Earth’s surface with consequent direct effects on human health.

N$_2$O is a natural product of microbial processes occurring in soils principally via the pathways of nitrification (ammonia oxidation) and nitrate dissimilation (denitrification and ammonification of nitrate), and its atmospheric concentration was kept in check by natural destruction mostly in the stratosphere. Although nitrification requires aerobic conditions and denitrification occurs in anaerobic conditions, there are many circumstances in which both processes can occur simultaneously because of the heterogeneity of soil aeration status (see, for example, Figure 13.1 for the locally variable conditions around rice roots). Emissions of N$_2$O from soils have increased since 1860 mainly because of the increased use of nitrogen to promote crop productivity. Most nitrogen is in the form of unreactive N$_2$ gas, but human activity has generated increasingly large amounts of reactive nitrogen through combustion of fossil fuels, the promotion of biological nitrogen fixation in crop rotations and the manufacture of fertilisers using the Haber–Bosch process (Galloway et al., 2004; Figure 13.2). In 1860, natural nitrogen fixation amounted to between 100 and 200 Tg N year$^{-1}$
and synthetic nitrogen fertilisers about 15 Tg N year\(^{-1}\); by 2005, these values had increased to about 350 and 187 Tg N year\(^{-1}\) respectively (Galloway et al., 2008). Without this introduced reactive nitrogen, it is estimated that over 40% of the world’s present population would not be fed (Galloway et al., 2004, 2008), but Crutzen et al. (2008) estimated that 3–5% of this reactive N input is converted to N\(_2\)O and contributes to the greenhouse effect. Smith et al. (2010a) conducted a top-down analysis and concluded that 4 ± 1% of all anthropogenically released reactive N is transformed into N\(_2\)O; this analysis demonstrates that agriculture is the dominant pathway for reactive N entering the biosphere.

Nitrogen from animal manures is also a source of atmospheric N\(_2\)O. Bouwman et al. (2010) estimated that in 2000 about 65% of the 112 Tg N year\(^{-1}\) in manure was generated by mixed and landless systems and 26% from pastoral systems. Manure collected in barns and storage systems was mainly spread on cropland (33.3 Tg N year\(^{-1}\)) with a smaller amount (5.3 Tg N year\(^{-1}\)) on grassland. Again for 2000, Bouwman et al. (2010) estimated that the combined N\(_2\)O emission from the 79.1 Tg N year\(^{-1}\) applied as fertiliser and 33.3 Tg N year\(^{-1}\) applied as manure was 4 Tg N\(_2\)O–N year\(^{-1}\), a value that is similar to the top-down estimates of Smith et al. (2010a).

Reducing these N\(_2\)O emissions is a difficult problem. Except where over-application of N fertilisers is an issue (e.g. some parts of China), simply reducing N fertiliser application is not an option because this would reduce yields. Substituting N fertiliser with biologically fixed N will also not reduce emissions (Mosier et al., 1998) and may lead to an increase as, for example, when legumes were introduced into a pasture resulting in N\(_2\)O emissions that were two to three times greater (Duxbury et al., 1982). Improved timing and application rates of fertilisers and use of nitrification inhibitors with fertilisers can reduce emissions from arable crops (Smith et al., 1997), but it is also important to recognise that a reduction in emissions for one part of the system may result in an increase in another part of the nitrogen cascade (Figure 13.2).
13.2.3 Urbanisation

In 2008, the world’s urban population exceeded its rural population for the first time. This process of urbanisation has been going on for some time but accelerated significantly during the twentieth century. In 1900, there were, globally, an average of 6.7 rural dwellers for each urban dweller, but by 2025, there will be three urban dwellers for every rural dweller (Satterthwaite et al., 2010). Along with this change has been a decrease in the ratio of food producers to food consumers brought about by rapid growth in the world’s economy and the shift from primary industries (agriculture, forestry, mining and fishing) to manufacturing and services.

Expansion of towns and cities often takes place at the expense of agricultural land not least because most cities have been there for some time and were located where they are because their hinterlands could supply food from fertile soils. However, Satterthwaite et al. (2010) suggest that the loss of agricultural land is often overstated, with only Western Europe having >1% of its land area as urban among all of the world’s regions. Moreover, while land is clearly used for buildings, roads, etc., the remaining land may be used for more intensive production (for example, vegetable and fruit crops) or for a variety of urban and peri-urban agricultural schemes (Smit et al., 1996). Many people living in urban areas rely on urban or peri-urban agriculture for part of their food. For example, in urban areas of East Africa in the 1990s, 17–36% of the population grew crops (often maize) or kept livestock (Lee-Smith, 2010).

In many urban centres, waste materials are efficiently recycled through livestock and composting. For example, in Nairobi, municipal solid waste is used for composting by a range of community-based organisations, and animal manure is imported into the city from the arid and semiarid livestock-producing areas from up to 300km away (Njenga et al., 2007). These represent business opportunities for some urban dwellers. However, the quality of the solid wastes used to produce food is crucial as heavy metals may accumulate causing toxicities in leafy vegetable crops. In Vietnam, Khai et al. (2007) measured the flows of N, P, K, Cu and Zn for intensive small-scale aquatic and terrestrial vegetable systems in two peri-urban areas of Hanoi City. They found that high inputs of chemical fertilisers, chicken manure and irrigation water were used resulting in surpluses of all nutrients (85–882 kg N ha⁻¹ year⁻¹, 109–196 kg P ha⁻¹ year⁻¹, 20–306 kg K ha⁻¹ year⁻¹, 0.2–2.7 kg Cu ha⁻¹ year⁻¹ and 0.6 – 7.7 kg Zn ha⁻¹ year⁻¹). They concluded that irrigation with wastewater contributed to the high inputs and that inputs at these levels constituted a major threat to the soil and water environment.

One major consequence of urbanisation that is little reported is the major shift of nutrients from rural to urban areas. Nutrient budgets for different production systems have been conducted at a range of scales in many countries to demonstrate either gains or losses of nutrients from soils (e.g. Smaling et al., 1997; Vitousek et al., 2009). With the increasing movement of the nutrients present in crops from rural to urban areas, nutrients in the resulting wastes are likely to become concentrated in towns and cities unless effective methods of recycling can be found.

13.2.4 Climate change

There is now little doubt that human activity since industrialisation has led to warming of the globe and that future change is inevitable (IPCC, 2007). In the last 50 years or so, global mean air temperature has increased by 0.13 K decade⁻¹ (IPCC, 2007), with warming
The new challenge – sustainable production in a changing environment

particularly noticeable in northern latitudes. Gregory and Ingram (2008) reviewed literature on climate change and yield and found that many studies noted an increase in mean temperature over the last 30–40 years in northern temperate latitudes of about 1.0–1.4 K (Chmielewski et al., 2004 in Germany; Jaggard et al., 2007 in the UK). They also noted that in some regions, the change in minimum temperature was greater than that in maximum temperature (e.g. Cutforth et al., 2007, in the Great Northern Plains of the USA) but that none of the studies detected any trend in precipitation. Warmer temperatures have resulted in plants developing and flowering earlier in Europe (Menzel et al., 2006), earlier harvesting of cereal crops in the UK (Sparks et al., 2005) and some reports of earlier outbreaks of pests and pathogens (Gregory et al., 2009). However, while past warming can be linked to simple phenological changes in plants, attributing effects to agricultural production is much more difficult because of the many changes in production practices (including new varieties, better weed control, increased fertiliser applications, etc.) that have occurred over the same period. In a German study, the phenology of 78 developmental stages (including sowing, emergence, flowering and maturity) for a wide range of agricultural and horticultural crops was advanced by an average of 1.1–1.3 d decade⁻¹ (or about 4 d K⁻¹) between 1951 and 2004 (Estrella et al., 2007). However, while technological advances or altered farming practices such as new cultivars can influence phenology, de-trending of the data suggested that such effects were only 13% of the response observed (Estrella et al., 2007).

Global circulation models based on well-established physical principles have been developed that reproduce observed changes in climate and project future changes (IPCC, 2007; Jarvis et al., 2010). Current knowledge suggests increases in mean annual air temperature of an additional 1–3 °C by 2050 with some complex, and spatially variable, changes in rainfall (Jarvis et al., 2010). There are considerable uncertainties in these projections arising from uncertainties in future GHG emissions, modelled responses to emission scenarios and misrepresentation of physical processes by the models. Nevertheless, some future warming appears inevitable. Much less certain, and in many agricultural contexts more important, are projections of future rainfall, which influence the likelihood of increased soil erosion by water, the oxidation of soil organic matter (see Chapters 4 and 12) and the yields of crops. Projecting change in rainfall is more difficult than that for temperature because more atmospheric processes are involved and the outcome is therefore more dependent on the representation of interacting processes within the model (Jarvis et al., 2010). This means that the consequences of climate change for rainfed agriculture are presently uncertain. Generally, the effects of climate change on production range from being very negative in areas that are already water limited to more favourable in regions that are temperature limited (Fuhrer, 2006). Typically, warming is projected to have an overall negative effect on yields of many crops because the duration of grain-filling is adversely affected by high temperatures (e.g. Ewert et al., 2005). For example, Lobell and Field (2007) showed a negative response of global yields of wheat, maize and barley to increased temperatures since 1980 equivalent to about 2–3% of production in 2002. The production foregone (about 19 Mt year⁻¹ of wheat, 12 Mt year⁻¹ of maize and 8 Mt year⁻¹ of barley) would have been worth about US$ 5 billion per year and the loss of wheat and maize together was equivalent to that of Argentina (Lobell and Field, 2007).

In some specific regions and for particular crops, though, in which temperature (but not rainfall) currently restricts growth, warming may be beneficial. In the UK, Jaggard et al. (2007) estimated that about 66% of the increased yield of sugar beet over the period 1976–2004 was attributable to changes in weather (mainly warming). About 10% was due to the ability to sow earlier in the year, but 56% was due to within-growing-season changes
in weather. Similarly, in eastern Scotland, Gregory and Marshall (2012) demonstrated that warming over the last 50 years could have contributed up to 25% of the increased yields of potato through earlier emergence and canopy closure allowing a longer period of photosynthesis at maximum rate.

Mitigation of climate change essentially involves the adoption of practices that either reduce emissions of GHGs to the atmosphere and/or remove GHGs from the atmosphere into long-term stores (see Section 4.9.2). Farmers can contribute to the mitigation of climate change by (1) decreasing the use of fossil fuel for farm operations or in inputs having large emissions from fossil fuels during their manufacture (e.g. N fertiliser; see Section 13.2.2); (2) growing bioenergy crops or utilising agricultural wastes to partially replace fossil fuels for electricity production, heating or transport; (3) adopting management practices that decrease net emissions of N\textsubscript{2}O and CH\textsubscript{4} (see Section 13.2.2) and (4) adopting management practices that permit sequestration of carbon (from CO\textsubscript{2}) in soil or in long-lived plants.

There is already some evidence that farmers in many parts of the world are adapting their management to cope with the warmer temperatures and variable rainfall. For example, the date by which 50% of the sugar beet crop in eastern England is sown has advanced by about 15 d since the 1970s because of warmer spring temperatures (Jaggard et al., 2007, 2010). In sub-Saharan Africa, farmers have experienced variable rainfall for many years and are used to varying sowing dates and other practices accordingly; however, these strategies may not be as effective if climatic variability and the incidence of extreme weather events become more pronounced (Haile, 2005).

13.3 Food security and food production

13.3.1 Food security

There are several definitions of food security but the one most commonly used is that it represents the state when ‘all people, at all times, have physical and economic access to sufficient, safe, and nutritious food to meet their dietary needs and food preferences for an active and healthy life’ (FAO, 1996). In the UK, this definition has been modified to give a greater emphasis on consumer preferences and the required safety of food: ‘ensuring the availability of, and access to, affordable, safe and nutritious food sufficient for an active lifestyle, for all, at all times’ (Defra, 2009). Food security is, then, a social construct so that, while production is clearly a critical element contributing to its achievement, access to food and the utilisation of food are also crucial and, in many cases, of overriding importance (Gregory et al., 2005; Ericksen, 2008). Food security is underpinned by effective food systems, which are a set of dynamic interactions between and within biogeophysical and human environments. They include a number of activities (producing food; processing, packaging and distributing food; and retailing and consuming food) that lead to a number of associated outcomes some of which contribute to food security (i.e. food availability, access to food and food utilisation) and others that relate to environmental and other social welfare concerns (Ericksen, 2008). Because food security is diminished when food systems are disrupted or stressed, food security policy must address the whole food system.

Achieving food security for all is no easy matter to achieve, not least because there are distinct, and contradictory, views on the understandings that different communities have about food and the relations between food, society and the environment. Eakin et al. (2010) explore three different perspectives on food, which highlight these different modes of discourse and result in different policies and courses of action. One view perceives food as
a global commodity in which a diverse range of products are bought and sold competitively
and markets operate to reward efficiencies in production, distribution and processing, and
favour product differentiation. McMichael (1994) refers to this as the ‘exchange-value’ of
food in which there is demand without differentiation across a market. Maize is maize, food
is a commodity to be traded and ideas of ‘terroir’ or soils contributing to product quality
have little resonance. Classical soil science, with its emphasis on soil fertility and increasing
crop yields, has arguably, made substantial contributions to this view of food. This perspec-
tive, in which the economic value of food is paramount, has contributed to policies relating
to failures of entitlement as major drivers of food insecurity and to regulations to make
globalised food commodity markets work for all.

As human-induced changes to the environment have accelerated in the last 60 years, a
second discourse has emerged in which food is viewed as an environmental service. This
view was formally stated in 2005 in the MEA (2005) and perceives food as one of several
provisioning, regulating and supporting services provided by natural ecosystems. This view
has rapidly gained credence in scientific discourse (see, for example, the UK’s ecosystem
assessment; UK National Ecosystem Assessment, 2011) and has been referred to several
times in this book (Chapters 3, 4 and 12). However, this discourse is often in conflict with
the commoditisation and globalisation perspective, which seeks to secure production by
controlling the environmental conditions in which food is produced through technological
advances. Resolving these conflicts by the development of systems of food production that
can utilise ecological perspectives and achieve the yields required from food provisioning is
a substantial challenge for the future.

Finally, the globalisation of food as a commodity and the increasing pressures brought
about by global environmental change has given renewed life to a third view of food as a
basic human right. Maslow (1943) proposed a hierarchy of human needs in which food,
clothing and shelter were the underpinning needs for life. Communities to whom this right
is denied are likely to rise up and protest as happened in many countries during the ‘food
crisis’ of 2008. Associated with this basic right are also ideas of food sovereignty and the
right to choose food of particular types, produced in specific ways. Again, this view of food
is at odds with food as a commodity. Eakin et al. (2010) conclude that future food security
requires governance systems that can reconcile these contrasting views and achieve, in addi-
tion to freedoms of availability, access and utilisation, the freedom to make food choices that
support individual, communal (and environmental) preferences.

13.3.2 Food production and yields

While food production alone does not guarantee food security, as stated in Section 13.1, the
growth in the world’s human population from about 2.2 billion in 1950 to about 7 billion in
2011 has been sustained by increased production (Godfray et al., 2010). Simultaneously,
there has been increased demand for fuel and fibre, and this increased production will need
to be maintained if the projected population of 9 billion by about 2040 is to be sustained;
demand for forest products is, similarly, projected to increase to 2030 by 1.4% per annum
for sawnwood and about 3% per annum for paper and wood-based panels (FAO, 2009).
Increases in crop yields have been the major factor supporting increased food production
(Section 13.2.1), and there is much interest in how past success can be sustained.

Yield is a consequence of the interaction of three factors: genotype (G) × environment (E)
× management (M; Hay and Porter, 2006). Evans (1993, 1998) highlights how the synergis-
tic effects of these interactions linked to innovative technologies have contributed to past
Soil Conditions and Plant Growth

increases in yield. Among the important contributors to these have been (1) improved germ plasm able to grow vigorously (e.g. hybrids), resist pathogens and respond to fertilisers without lodging (in particular, the use of dwarfing and semi-dwarfing genes in rice and wheat); (2) the application of fertilisers and particularly the availability of affordable nitrogen fertiliser; (3) the development of chemicals to control weeds, pests and diseases and (4) improved irrigation systems especially in rice-producing countries and for some previously rainfed crops. Together, these technological innovations have modified G, E and M to greatly increase yields. The consequence of these innovations has been that since about 1960, yields of many crops have increased year on year (Figure 13.3). Global yield increases for a number of crops have typically been linear (Figure 13.3a) with values of 53 kg ha\(^{-1}\) year\(^{-1}\) for rice, 41 kg ha\(^{-1}\) year\(^{-1}\) for wheat and 63 kg ha\(^{-1}\) year\(^{-1}\) for maize over the period 1961–2004. Increases in yield have also been linear with time in many individual countries (e.g. Scotland in Figure 13.3b and see Ewert et al. (2005) for wheat yields in several European countries) although in a few instances technological innovations have produced more rapid, stepwise increases in yield (e.g. Australia, Angus (2001); Figure 13.3c). An arithmetic consequence of the linear yield increases with time is that the proportionate increase in yield obtained through breeding programmes has decreased. Of greater concern, though, is that in some countries and regions, absolute increases in yields are no longer occurring at the same rate despite yields being less than the attainable value. For example, rice yields in SE Asia and wheat yields in the UK over almost the last decade have appeared to be approaching an asymptote (Figure 13.3d). The reasons for this are complex and probably include almost constant inputs of nitrogen fertiliser to wheat crops in the case of the UK from 2000 onwards (average N fertiliser application was 188 kg ha\(^{-1}\) in 2000 and 190 kg ha\(^{-1}\) in 2007; Defra 2010) and concerns about the long-term availability of mineralised soil N in paddy crops because of the formation phenolic lignin residues that stabilise soil organic nitrogen (Olk et al., 2006).

Although crop yields have increased globally and throughout North and South America, Europe, Australia and much of Asia, a notable exception has been that of Africa where, for example, per capita food production decreased by about 5–10% between 1980 and 1995 (Sanchez et al., 1997). FAO data analysed by Greenland et al. (1998) demonstrated significant decreases in crop yields in many countries in sub-Saharan Africa, and there are still many countries where yields have improved little. The reasons for the poor performance in Africa relative to other countries are many and include social unrest and war, poor institutions and governance, climatic variability making reliable irrigation difficult and weathered soils that are deficient in nutrients. Greenland et al. (1998) concluded that a major factor behind many of the observed decreases in yield in African countries was the decline of soil fertility accompanied by the lack of fertiliser application. Much has been written about the need to ‘recapitalise’ the soils of Africa, especially with regard to P status, but progress has been limited (Sanchez et al., 1997).

Because the area of cropped land is likely to increase proportionately less than the future demand for food, reducing the gap between current yields and potential yields is a major goal for the future (Jaggard et al., 2010). Potential yield is a theoretical upper limit to yield imposed by solar radiation (affecting growth), temperature (affecting development and growth) and water supply (affecting mainly growth but also development; see Chapter 2 for explanations). Lobell et al. (2009) reviewed data from crops of maize, rice and wheat grown in a range of countries and compared potential yields with measured experimental, maximum farmer and average farmer yields. The gap between potential and the actual yields ranged from about 20% to 80%. In many irrigated cereal systems, yield appeared to plateau
Figure 13.3  Changes in yields with time: (a) global average annual yields (data from USDA) for maize (squares), rice (circles) and wheat (triangles); the lines are linear regressions: (b) national average annual yields in Scotland for barley (circles), rape (squares) and potato (triangles); the lines are linear regressions (data from annual farm surveys conducted by Scottish agencies responsible for agriculture): (c) national average annual yield of wheat in Australia (from Angus 2001 and personal communication); the dashed line joins annual values, the thin solid line is a 5-year running average and the solid black line summarises major trends in yield with respect to changes in crop husbandry; and (d) regional average annual rice yields in Asia (circles – data from USDA) and wheat in the UK (triangles – data from Defra). Figure reproduced from Gregory and George (2011). With kind permission from Wiley-Blackwell.
Soil Conditions and Plant Growth

at or about 80% of potential yield while in rainfed systems, average yields were commonly
50% or less of potential (Lobell et al., 2009). While part of the yield gap is inevitable
because of crop losses during harvest, storage and transport and the way that land areas are
reported (Jaggard et al., 2010), there are still large differences in performance between
adjoining farms. Lobell et al. (2009) concluded that a fundamental constraint in irrigated
systems was uncertainty in growing season weather; this is also a factor in rainfed systems
where interactions between water and nutrient availability are complex (Cooper et al.,
1987). Raising yields above 80% of yield potential is possible, but only if technologies can
be developed and adopted that reduce the uncertainties faced by farmers in assessing soil
and climatic conditions or respond dynamically to these conditions or both (for example,
installation of nutrient and water sensors; Lobell et al., 2009). Such technologies may have
the added benefits of increasing the efficiency of use of inputs and reducing losses off-site
as well as increasing yields.

In a global analysis of wheat, maize and rice production, Neumann et al. (2010) found
that yield gaps were significantly correlated with irrigation, market accessibility and influ-
ence, availability of agricultural labour and slope; the contribution of these factors varied
substantially between regions and generalisations as to the best means of reducing yield
gaps were not possible. In China, the potential to reduce yield gaps of maize was demon-
strated on 66 on-farm experimental plots, raising yields to 13 t ha\(^{-1}\) on average (nearly twice
the typical farmer yield) without any increase in N fertiliser application (Chen et al., 2011).
This was achieved using simulation models to identify appropriate combinations of planting
date, crop density and cultivar at each site based on long-term weather data and then chang-
ing variety, sowing date and spacing as appropriate. This was combined with an in-season
root-zone N management strategy, which matched supply to demand; in particular, a greater
proportion of the N currently applied was applied later in the growing season. This inte-
grated agronomic approach combined with cultivars of appropriate duration increased yields
and reduced off-site nutrient movement (Chen et al., 2011). In an analysis of yield gaps for
rice in South-east Asia, Laborte et al. (2012) examined farmers’ yields in four intensively
cropped regions in Indonesia, the Philippines, Thailand and Vietnam and compared them
with yields simulated in a rice model. Three yield gaps were identified based on climatic
potential (Van Ittersum and Rabbinge, 1997), economic and best farmers’ yields
(Figure 13.4). Average yields varied between 43% and 75% of the climatic yield potential
and 61% and 83% of the best farmers’ yields. The farmers with the best yields were typi-
cally more educated and used fertilisers and labour more efficiently than others. This analy-
sis showed that if the yield gap between the average and the best farmers were closed, then
this would cover the increased yields needed for 2050 (assuming no change in diet) in all
countries except the Philippines, where other institutional changes will also be required
(Laborte et al., 2012).

13.3.3 Biomass and biofuels

Growing crops for biomass and biofuels has been highlighted as a potential competitor for
land with food crops and, thereby, a potential contributor to food insecurity (see, for exam-
ple, Karp and Richter, 2011). It is noteworthy that the area occupied by biofuels and their
by-products in 2004 was only 14 Mha compared with 1500 Mha of crops (i.e. about 1% of
the total cropped area) and 4,500 Mha of pastures worldwide (IEA, 2006). However, there
has been a substantial increase in the area of land used to grown biofuels since 2004, espe-
cially in the USA, with one-third of the maize grown in 2007 used for bioethanol. While the
The new challenge – sustainable production in a changing environment

reasons for growing crops for biofuels are complex (including increased energy security, increased supplies of transportation fuels and a desire to decrease net emissions of GHGs), the use of land for them is likely to increase (FAO, 2008 gives some projections), and the emergence of biofuels as a new source of demand for agricultural commodities may assist in the revitalisation of agriculture in developing countries with positive benefits for personal livelihoods and economic growth (Pingali et al., 2008).

Biomass for energy can be obtained from various sources including specific energy crops (e.g. willow and perennial grasses; Karp and Shield, 2008) and agricultural residues. The potential energy production from these residues is large and globally estimated at over $12 \times 10^{18}$ J (Hall et al., 1993). However, many agricultural residues are reused as feed for livestock so that a waste stream unsuitable for human consumption is upgraded into valuable food products. Nonhebel (2007) demonstrated that, from the perspective of land use, it was better to produce energy from dedicated energy crops and to use the agricultural residues for livestock food rather than use the residues for energy. This was because when dedicated energy crops were grown, less land was required to produce bioenergy than the additional land area required to compensate for protein losses in the human food.

Field et al. (2007) suggest that the interaction of four major factors will determine the success or otherwise of biomass energy in future: (1) the intrinsic productive capacity of land and ocean ecosystems; (2) the alternative uses for the land and water resources used for biomass energy production; (3) the effects of biomass energy technologies on off-site factors such as pollution and invasive species and (4) the conversion processes for increasing the energy yield per unit of biomass and from each unit of land or water. All four factors have implications for the production of both food and GHGs but especially factor two. The obligations imposed by the Renewable Transport Fuel Obligation (RTFO) pose considerable challenges for land use for individual countries if the fuels are to be produced locally. For example, in the UK, the 1.2 billion litres of bioethanol and 1.35 billion litres of bio-diesel required to meet the UK RTFO targets would require about 500,000 ha of wheat (about 20–30% of the current area) and 340,000 ha of oilseed rape (compared with the 670,000 ha

Figure 13.4 A conceptual expression of the relative yield determined by climatic yield potential, economic yield goal and farmers’ yield. Three different yield gaps are identified based on the differences between average farmers’ yield and the climatic yield potential (YieldGap_p), the economic yield goal (YieldGap_e) and the best farmers’ yield (YieldGap_f). Reproduced from Laborte et al. (2012). With kind permission from Elsevier.
grown in 2007), respectively (Cottrill et al., 2007). So about 15–20% of the UK arable area would be required to meet the RTFO targets (Karp and Richter, 2011).

Several studies have attempted to calculate the net effect of growing biofuel crops on GHG emissions after allowing for energy inputs such as fertiliser applications, cultivation and other inputs in the production process (e.g. Hill et al., 2006). In the USA, it has been estimated that, relative to the energy used in their production, ethanol from maize yields 25% more and bio-diesel from soya beans 93% more (Hill et al., 2006). Furthermore, GHGs are reduced by 12% by the production and burning of bioethanol and 41% by diesel relative to the fossil fuels that they replace. However, studies in both the UK and the USA have demonstrated that home-grown crops for biofuels will meet only a small proportion of current fuel needs. Powlson et al. (2005) calculated that making a variety of reasonable assumptions about the land that might be used to produce biomass or biofuels, biomass crops might provide up to 12% of UK electricity demand and sugar beet or oilseed rape crops about 1.2–1.8% of UK diesel or petrol requirements. Similarly, in the USA, using all maize and soya bean production for biofuels would contribute only 12% of petrol and 6% of diesel demand (Hill et al., 2006).

As in the past, land will be used to grow both food and fuel in the future (it is worth remembering that until relatively recently, much land was used to grow forage for horses), so the bigger question is not fuel vs food but how land might best be used to contribute to both food and energy security (Karp and Richter, 2011). Several studies have indicated that the worst option is to clear new land for the production of biomass or biofuel crops (Field et al., 2007; Fargione et al., 2008; Searchinger et al., 2008). Clearing can result in GHG emissions that are much greater than any benefits provided by biofuels because soils and plant biomass are biologically active stores of carbon that are decomposed by microbes when disturbed releasing large quantities of CO₂. For example, replacing a lowland tropical rainforest in Indonesia or Malaysia with oil palm to be used for bio-diesel would produce about 610 Mg CO₂ ha⁻¹, which would take about 86 years to recoup from the use of bio-diesel instead of fossil fuel, while soya bean diesel replacing Amazonian rainforest producing >280 Mg CO₂ ha⁻¹ would take about 320 years (Figure 13.5; Righelato and Spracklen, 2007). Such considerations have led to the conclusion that production of bioenergy crops should be focused either on previously cleared and abandoned agricultural lands or on lower grade, marginal land (Field et al., 2007; Campbell et al., 2008; Karp and Richter, 2011). However, much remains to be done to identify crops and production systems that can yield an economic return from such land.

13.3.4 Use of energy in crop production

Concerns about the amount of fossil fuel energy expended in crop production and food processing are not new (Pimental et al., 1973; Spedding and Walsingham, 1976), but have come to the fore again as energy costs have increased and concerns about CO₂ emissions and the need to develop low carbon cropping practices have emerged (Woods et al., 2010). Pimental and Pimental (2008) provide a variety of examples to illustrate the poor energy returns of many crop production practices ranging from maize production in Mexico using human power and an axe and hoe returning 10.7 times as much energy as consumed in production to a return of 2.2 times for rice production in the USA (Table 13.3). In the Mexican example, the only fossil fuel used was in the production of the axe and hoe; unfortunately, the yield (1.94 t ha⁻¹) is well below that required to sustain the future global population. These figures omit the energy required to convert the grains
into human food – negligible in the Mexican example but substantial enough in the case of the USA to render the overall energy return close to unity.

Although agriculture accounts for only about 2% of total energy usage, Table 13.3 demonstrates that agriculture has become very dependent on fossil fuels to produce the current levels of yield. Fossil fuel is used directly for tillage and crop management and indirectly through the application of energy-intensive inputs such as fertilisers and pesticides. Moreover, the use of energy embedded in buildings, tractors and other infrastructure continues to increase. For the UK, Woods et al. (2010) showed that primary energy inputs ranged from about 1 GJ t⁻¹ for potato to 6 GJ t⁻¹ for oilseed rape; in part, this range reflects the

<table>
<thead>
<tr>
<th>Crop</th>
<th>Country</th>
<th>Tillage</th>
<th>Yield (t ha⁻¹)</th>
<th>Inputs (kJ x 10³ ha⁻¹)</th>
<th>Output (kJ x 10³ ha⁻¹)</th>
<th>Energy ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Groundnut</td>
<td>Thailand</td>
<td>Buffalo</td>
<td>1.28</td>
<td>8,048</td>
<td>20,892</td>
<td>2.60</td>
</tr>
<tr>
<td>Groundnut</td>
<td>USA</td>
<td>Mechanised</td>
<td>3.72</td>
<td>45,817</td>
<td>64,051</td>
<td>1.40</td>
</tr>
<tr>
<td>Maize</td>
<td>Mexico</td>
<td>Human</td>
<td>1.94</td>
<td>2,687</td>
<td>28,881</td>
<td>10.7</td>
</tr>
<tr>
<td>Maize</td>
<td>Mexico</td>
<td>Oxen</td>
<td>0.94</td>
<td>3,222</td>
<td>13,982</td>
<td>4.34</td>
</tr>
<tr>
<td>Maize</td>
<td>USA</td>
<td>Mechanised</td>
<td>8.66</td>
<td>33,961</td>
<td>130,396</td>
<td>3.84</td>
</tr>
<tr>
<td>Rice</td>
<td>Borneo</td>
<td>Human</td>
<td>2.02</td>
<td>4,327</td>
<td>30,626</td>
<td>7.08</td>
</tr>
<tr>
<td>Rice</td>
<td>Philippines</td>
<td>Animal</td>
<td>1.65</td>
<td>7,638</td>
<td>25,126</td>
<td>3.29</td>
</tr>
<tr>
<td>Rice</td>
<td>Japan</td>
<td>Mechanised</td>
<td>6.33</td>
<td>34,405</td>
<td>96,163</td>
<td>2.80</td>
</tr>
<tr>
<td>Rice</td>
<td>USA</td>
<td>Mechanised</td>
<td>7.37</td>
<td>49,542</td>
<td>110,995</td>
<td>2.24</td>
</tr>
<tr>
<td>Soybean</td>
<td>USA</td>
<td>Mechanised</td>
<td>2.67</td>
<td>12,609</td>
<td>40,197</td>
<td>3.19</td>
</tr>
<tr>
<td>Wheat</td>
<td>USA</td>
<td>Mechanised</td>
<td>2.67</td>
<td>17,740</td>
<td>35,354</td>
<td>2.13</td>
</tr>
</tbody>
</table>

Source: Derived from Pimental and Pimental (2008).
Soil Conditions and Plant Growth

different energy contents of the harvested organs in addition to the level of inputs required. Their analysis showed that while organic production systems were often apparently more energy efficient than conventional systems, human energy inputs were frequently higher for organic systems. The principal energy inputs into crops are for diesel, fertiliser and pesti-
cides, but the proportions of the energy used for different elements of the production process differ substantially between crops (Figure 13.6). For wheat, fertiliser (mostly nitrogen) is over 50% on the energy input into the conventionally grown crop, while for field beans that receive little N fertiliser and have lower yields than wheat, more diesel is used per tonne of yield produced. Woods et al. (2010) show that maincrop potatoes in the UK are very energy intensive because they are stored in cool rooms for a long time to provide a year-round supply. This contrasts with early potatoes that are generally not stored on farms.

The energy required to produce fertilisers is still substantial but has decreased significantly as a result of innovation by manufacturers. The main components of nitrogen fertilisers differ in the quantities of energy required to manufacture them (typically 32 MJ kg\(^{-1}\) for ammonium, 22 MJ kg\(^{-1}\) for urea and 13 MJ kg\(^{-1}\) for ammonium nitrate) and are considerably higher than the 5 MJ kg\(^{-1}\) required for P and K fertilisers. These energy inputs are much greater than those required to apply the products via spreaders in the field (Woods et al., 2010).

13.4 Routes to sustainable intensification

13.4.1 Elements of sustainability

The increasing global food demand coupled with the acknowledged environmental change resulting from changed production practices has led many to challenge the technologies that have resulted in today’s intensified agriculture and to call for the development of sustainable production practices (Conway, 1997; Gregory et al., 2002; Pretty, 2008; Foresight, 2011).
While there is general agreement that agricultural sustainability includes elements of profitable production, environmental stewardship and social responsibility, there is much less agreement as to how sustainability is to be achieved in practice beyond the need to integrate biological and ecological insights into the production process. Pretty (2008) reviewed the literature to establish four key principles of sustainability that ‘i) integrate biological and ecological processes such as nutrient cycling, nitrogen fixation, soil regeneration, allelopathy, competition, predation and parasitism into food production processes; ii) minimise the use of those non-renewable inputs that cause harm to the environment or to the health of farmers and consumers; iii) make productive use of the knowledge and skills of farmers, thus improving their self-reliance and substituting human capital for costly external inputs; and iv) make productive use of people’s collective capacities to work together to solve common agricultural and natural resource problems, such as for pest, watershed, irrigation, forest and credit management’. Such principles go well beyond the need for technical innovations including improvements to genotypes and ecological and agronomic management systems and embrace the need to develop important capital assets for agricultural systems including natural, social, human, physical and financial capital. A corollary of this analysis is that many disciplines and ways of thinking will be required to develop sustainable systems and that there is unlikely to be a single solution appropriate to all soils and production systems.

As outlined in Section 13.3.2, the general expectation is that intensive production systems will be the main route by which future food supplies are secured. Cassman et al. (2003) concluded that in order to avoid severe degradation of natural resources and to reduce GHG emissions, such systems would need to produce higher yields on existing cropland, limit expansion of the cultivated area, achieve a substantial increase in N fertiliser efficiency and improve soil quality through increasing soil organic matter. There is little doubt that for most soils, sustainable production is inextricably linked to the maintenance of soil organic matter contents through appropriate additions to offset the losses caused by cultivation and nutrient depletion (Greenland et al., 1998; Lal, 2004).

In a wide-ranging assessment of future food and farming, the UK Foresight (2011) concluded that sustainable intensification was required and that all technological means of achieving this should be assessed and appropriately utilised. How might this sustainable intensification be achieved? Tilman et al. (2011) approached this problem by first defining the global demand for crop production in 2050 and then examining the environmental impacts of different ways of achieving this. Their analysis, based on measured relations between per capita energy and protein demands with per capita gross domestic product and allowing for different economic groups, indicates a 100–110% increase in global crop production; this is larger than the 70% estimated by Alexandratos (2009) using expert opinion only. Using past N fertiliser rates as a surrogate for soil fertility enhancement (though recognising that soil fertility can be enhanced in other ways), alternative pathways to increasing yields to the required levels were investigated. The options examined were (1) current technology in which each economic group retained its 2005 N-dependent yield function; (2) technological improvement in which technological advances continue along existing temporal trends to 2050; (3) technology transfer in which low-yielding countries adopt and adapt the existing high-yielding technologies of high-yielding countries and (4) technology improvement and transfer in which all countries achieve soil- and climate-adjusted yields. Figure 13.7 shows the consequences for yields, land clearance and GHG emissions from adopting these different routes together with trajectories based on minimising land clearance or increasing the efficiency of N use. In summary, with present trends of intensification
in rich nations and extensification in poor nations, by 2050, an additional 1 billion ha would be cleared with GHG emissions of 3 Gt year$^{-1}$ and N use of 250 Mt year$^{-1}$. However, if intensification were concentrated on existing cropland and transfer and adoption of high-yielding technologies were successful, then only 0.2 billion ha would be cleared, GHG emissions would be reduced to one-third (1 Gt year$^{-1}$) and global N use would be 225 Mt year$^{-1}$. Although this analysis omits any effects of future climate change, it indicates what might be possible with investment in innovative technologies and infrastructure.

Figure 13.7 Projections of 2050 values for (a) global yields, (b) global land clearing and (c) global agricultural greenhouse gas emissions and (d–f) the yields and environmental impacts of four hypothetical trajectories along which agriculture might develop by 2050. The emissions of greenhouse gases in figures (c) and (f) are expressed in terms of CO$_2$ equivalents. See Tilman et al. (2011) for details of the methodology used. The four curves shown in figures A–C are based on regression and reflect the different agricultural trajectories assuming magenta, current technology; blue, improvement only; orange, technology adaptation and transfer; and green, improvement and technology transfer. Reproduced from Tilman et al. (2011). With permission from Proceedings of the National Academy of Science. For a colour version of this figure, please see Plate 13.2.
13.4.2 Improved efficiency of resource use

Several chapters (Chapters 1, 3, 6, 9 and 12) have already mentioned the need to improve the efficiency with which resources such as water and nutrients are used in systems of agricultural production. The efficient application of water to crops through sprinklers and drip systems has been dealt with in Section 9.5, so this section focuses on nutrients.

A primary requirement for the future is to produce higher yields with inputs that do not lead to environmental problems either on- or off-site. Nutrient additions that are inadequate relative to crop offtake degrade land through nutrient mining while additions that are excessive degrade land, water and air through leaching, eutrophication and gaseous emissions (Vitousek et al., 2009). Ideally, nutrient additions (whether as mineral fertilisers or manures) and soil biota should be managed to deliver nutrients to crops synchronously with demand (Myers et al., 1994), but this has proved difficult to achieve in practice because applications must normally be made before the demand exists and large canopies do not permit application of solid sources to soils.

Precision agriculture has evolved rapidly in the last decade so that it is now possible to identify agronomic issues, and apply inputs, at the sub-field scale (McCormick et al., 2009; Lawes and Robertson, 2011). Harvesters with weighing facilities and global positioning systems have allowed the production of yield maps on an almost routine basis and these, together with grid and transect sampling of soils to produce maps of nutrients, are allowing farmers to contemplate site-specific nutrient management (Robertson et al., 2008). Such management is knowledge intensive and requires multiple forms of knowledge to be integrated in a way that can be practically managed. For example, should fertiliser be applied according to past yield performance or soil fertility analysis? Khosla et al. (2008) compared the yields and monetary returns from irrigated maize crops grown in northeastern Colorado managed according to management zones based either on site-specific soil properties such as bulk density, texture and organic matter or on yields. They found that, overall, management of N fertiliser based on soil properties produced higher grain yields with net monetary returns ranging from $US 188 to $US 679 ha⁻¹ and concluded that site-specific management zones were an inexpensive and pragmatic approach to N management.

The best approach to adopt, though, is itself likely to be site specific. For example, from an analysis of rainfed yields of 20 fields in Western Australia, Lawes and Robertson (2011) concluded that while applying variable rates of inputs added some value to many fields, the gains were small so that the technology needed to be cheap for it to be implemented by farmers. In their study (a rainfed system with much lower yields and inputs than that of Khosla et al., 2008), the monetary returns were greater for fields with soils of higher fertility and variable yields than where soil fertility was low and uniform across the field.

It has been appreciated for a long time that there is an interaction between the efficiencies which water and nutrients are used, not least because the availability and mobility of nutrients in soils depends on water (see Chapter 6). In rainfed production systems, this interaction is a major constraint to the efficient use of inputs because rainfall is not known in advance. Where soil nutrients are deficient for maximum growth of crops, application of fertilisers and manures may result not only in increased growth but also in increased water use efficiency. This effect of modest applications of fertiliser has been well documented in several studies (Cooper et al., 1987; Latiri-Souki et al., 1998; Bolger and Turner, 1999; Ogola et al., 2002) and is illustrated in Table 13.4. Fertiliser use may increase slightly the total amount of water used (e.g. Cooper et al., 1987 for barley in Syria; Ogola et al. 2002 for maize in the UK), but the principal effect is to increase early canopy growth so that it shades
the surface and thereby reduces evaporation from the soil surface as a proportion of the total water that is evaporated. However, the beneficial effect of fertiliser in increasing growth and reducing $E_s$ is not universal and is dependent on the wetness of the soil surface and the evaporative demand (Gregory et al., 1997).

In semi-arid production systems, the efficiency of N and P fertilisers depends on the amount of water available to the crop so that the response to N, in particular, is variable and limited in dry years. Typically, crop response to N increases with increasing rainfall while response to P decreases on P-deficient soils (Jones and Wahbi, 1992). Studies in the Sahel have also concluded that soil fertility is often a more important factor in rangeland and crop productivity than rainfall, so that effective management of water cannot be achieved without also managing soil nutrient constraints (Payne, 2000). The limiting factors to crop growth at different times during any particular season could be either water or nutrient availability or both. A practical problem to be resolved in many semi-arid regions is how to afford and apply the optimum amount of fertiliser to produce an economically viable yield in a given season. For example, Sadras (2002) analysed financial returns over a 40-year period from fertiliser applications in an area with low and erratic rainfall in south-eastern Australia and found that a very low input of N fertiliser ($5 \text{ kg N ha}^{-1}$) ensured the greatest economic stability at all sites examined. Most scientific analysis is based on hindsight, but farmers must operate without this benefit so that conservative practices tend to dominate.

### 13.4.3 Improved tillage and conservation agriculture

In addition to managing chemical and biological elements of soil fertility, there is increasing emphasis on soil physical fertility and health as a means of improving both root growth and function, and soil ecosystem services. The practical means by which these ends can be achieved often relies on tillage. In much conventional agriculture, soil is often inverted (with, for example, a mouldboard plough or discs) with crop residues and weeds buried. There are many reasons why such tillage practices arose including (1) the preparation of a seedbed allowing uniform germination; (2) removal of weed competition early in the crop’s growing cycle; (3) release of nutrients through mineralisation after oxidation of soil organic matter; (4) burial of crop residues reducing carry-over of diseases and facilitating easier

### Table 13.4

<table>
<thead>
<tr>
<th>Crop</th>
<th>Season</th>
<th>Rainfall (mm)</th>
<th>Fertiliser</th>
<th>Dry matter (t ha$^{-1}$)</th>
<th>Water use (mm)</th>
<th>WUE (kg ha$^{-1}$ mm$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Barley</td>
<td>1981/82</td>
<td>324</td>
<td>+</td>
<td>6100</td>
<td>231</td>
<td>26.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4540</td>
<td>231</td>
<td>19.7</td>
</tr>
<tr>
<td></td>
<td>1983/84</td>
<td>204</td>
<td>+</td>
<td>2880</td>
<td>176</td>
<td>16.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1340</td>
<td>171</td>
<td>7.8</td>
</tr>
<tr>
<td>Millet</td>
<td>1984</td>
<td>260</td>
<td>+</td>
<td>4750</td>
<td>165</td>
<td>28.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2417</td>
<td>163</td>
<td>14.8</td>
</tr>
<tr>
<td></td>
<td>1985</td>
<td>380</td>
<td>+</td>
<td>5000</td>
<td>247</td>
<td>20.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3100</td>
<td>270</td>
<td>11.5</td>
</tr>
<tr>
<td></td>
<td>1986</td>
<td>440</td>
<td>+</td>
<td>3850</td>
<td>268</td>
<td>14.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1140</td>
<td>211</td>
<td>5.4</td>
</tr>
</tbody>
</table>

Source: From Gregory et al. (1997).
sowing; (5) easier incorporation of nutrients; (6) relief from soil compaction and (7) control of some soil-borne diseases and insects (Hobbs et al., 2008). However, by creating a bare soil surface, such tillage also has the undesirable consequences of substantially increasing rates of water and wind erosion and of the oxidation of soil carbon.

Since the 1930s, following the dust storms in the midwest of the USA, there has been a growing movement to reduced tillage systems (Unger et al., 1988); this was given added impetus by the development of herbicides to control weeds. Conservation tillage is a wide-ranging term used to describe the many practices that have evolved, initially to retain at least 30% surface coverage by crop residues, but to varying extents to also conserve time, fuel, earthworms, soil water, soil structure and nutrients (Baker et al., 2006). Several studies have shown that minimum tillage coupled with residue retention increased infiltration and profile moisture storage (e.g. Gicheru, 1994, in Kenya), reduced soil erosion by water by 10–100 fold because of greater aggregate stability and reduced run-off (Rhoton et al., 2002; Montgomery, 2007) and increased soil organic matter content (e.g. Govaerts et al., 2009).

Conservation tillage has been widely adopted in the USA, Brazil, Argentina, Canada and Australia (about 89 Mha in 2004; Derpsch, 2005, cited by Hobbs et al., 2008) and has now been included within a system of production referred to as ‘conservation agriculture’. This system removes the emphasis from tillage alone and combines three principles: (1) reduced tillage – ideally zero tillage but may involve controlled tillage during seeding to disturb no more than 20–25% of the soil surface; (2) retention of crop residues and surface cover – variable rates of retention to reduce erosion and run-off and enhance soil properties associated with long-term production and (3) use of crop rotations – moderation of weed, disease and pest problems coupled with utilisation of beneficial effects of some crops on soil conditions (Verhulst et al., 2010). These principles are applicable to a wide range of production systems from low-yielding, rainfed conditions to high-yielding, irrigated conditions. In practice, though, the techniques used will vary with the specific situation depending on farmer circumstances and biophysical and management conditions. Table 13.5 compares the effects of conventional tillage, conservation tillage and conservation agriculture on a range of soil and crop factors and demonstrates the advantages to sustainable production and other ecosystem services from adopting such an approach.

13.4.4 Managing roots and the rhizosphere

The rhizosphere is a complex zone of soil both influenced by and influencing roots. Some elements of the interactions taking place in this zone have already been referred to in Chapters 3, 6, 8, 9, 10, 11 and 12, but there is increasing interest in the possibilities of managing this interface for the improved acquisition of water and nutrients and prevention of soil-borne diseases in cropping systems (Ryan et al., 2009; Richardson et al., 2011). The architecture of root systems differs markedly between species and within genotypes and has been progress in exploiting this variation to benefit yields (Gregory, 2006; Lynch, 2007). For example, genotypes of common bean with shallow root architecture have been shown to grow and yield better in soils of low P status than genotypes with deep architecture (Rubio et al., 2001; Ho et al., 2004; Henry et al., 2010). Such genotypes have the added benefit that their increased growth results in greater ground cover and less P loss by erosion of topsoil (Henry et al., 2010). In soya bean, too, the most P-efficient genotypes had longer and larger root systems with a greater proportion of the root system in the topsoil (Ao et al., 2010).
Table 13.5 A comparison of tillage, conservation tillage (CT) and conservation agriculture (CA) for various issues.

<table>
<thead>
<tr>
<th>Issues</th>
<th>Traditional tillage (TT)</th>
<th>Conservation tillage (CT)</th>
<th>Conservation agriculture (CA)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Practice</td>
<td>Disturbs the soil and leaves a bare surface</td>
<td>Reduces the soil disturbance in TT and keeps the soil covered</td>
<td>Minimal soil disturbance and soil surface permanently covered</td>
</tr>
<tr>
<td>Erosion</td>
<td>Wind and soil erosion: maximum</td>
<td>Wind and soil erosion: reduced significantly</td>
<td>Wind and soil erosion: the least of the three</td>
</tr>
<tr>
<td>Soil physical health</td>
<td>The lowest of the three</td>
<td>Significantly improved</td>
<td>The best practice of the three</td>
</tr>
<tr>
<td>Compaction</td>
<td>Used to reduce compaction and can also induce it by destroying biological pores</td>
<td>Reduced tillage is used to reduce compaction</td>
<td>Compaction can be a problem but use of mulch and promotion of biological tillage helps reduce this problem</td>
</tr>
<tr>
<td>Soil biological health</td>
<td>The lowest of the three owing to frequent disturbance</td>
<td>Moderately better soil biological health</td>
<td>More diverse and healthy biological properties and populations</td>
</tr>
<tr>
<td>Water infiltration</td>
<td>Lowest after soil pores clogged</td>
<td>Good water infiltration</td>
<td>Best water infiltration</td>
</tr>
<tr>
<td>Soil organic matter</td>
<td>Oxidises soil organic matter and causes its loss</td>
<td>Soil organic build-up possible in the surface layers</td>
<td>Soil organic build-up in the surface layers even better than CT</td>
</tr>
<tr>
<td>Weeds</td>
<td>Controls weeds and also causes more weed seeds to germinate</td>
<td>Reduced tillage controls weeds and also exposes other weed seeds for germination</td>
<td>Weeds are a problem especially in the early stages of adoption, but problems are reduced with time and residues can help suppress weed growth</td>
</tr>
<tr>
<td>Soil temperature</td>
<td>Surface soil temperature: more variable</td>
<td>Surface soil temperature: intermediate in variability</td>
<td>Surface soil temperature: moderated the most</td>
</tr>
<tr>
<td>Diesel use and costs</td>
<td>Diesel use: high</td>
<td>Diesel use: intermediate</td>
<td>Diesel use: much reduced</td>
</tr>
<tr>
<td>Production costs</td>
<td>Highest costs</td>
<td>Intermediate costs</td>
<td>Lowest costs</td>
</tr>
<tr>
<td>Timeliness</td>
<td>Operations can be delayed</td>
<td>Intermediate timeliness of operations</td>
<td>Timeliness of operations more optimal</td>
</tr>
<tr>
<td>Yield</td>
<td>Can be lower where planting delayed</td>
<td>Yields same as TT</td>
<td>Yields same as TT but can be higher if planting done more timely</td>
</tr>
</tbody>
</table>

Source: From Hobbs et al. (2008)
In addition, roots release a wide variety of compounds that modify their immediate environment to increase the availability of nutrients and to keep pathogens at bay (Gregory, 2006; Hiltpold et al., 2010, 2011). Figure 13.8 shows some current and future targets for rhizosphere engineering to increase the efficiency of nutrient recovery and to provide defences against pests and pathogens. Ryan et al. (2009) detail many possible routes to effect these changes, but these are still predominantly in the research stage with development only recently starting. One example of chemical modification with the potential for large-scale benefits is the release of nitrification inhibitors that biologically inhibit nitrification and reduce emissions of N₂O (see Section 13.2.2; Subbarao et al., 2009). Engineering plants to exude organic anions such as malate and citrate have been shown in the laboratory to confer some tolerance to aluminium toxicity (Delhaize et al., 2004; Magalhaes et al., 2007), but the realisation of this potential in field soils has yet to be achieved. In the field, the increased exudation of organic acids may also influence the availability of nutrients and change the populations of soil microbes to metabolise these carbon sources. The net effect of these changes may be to negate the tolerance effect, but more work in different soils is needed to develop reliable systems.

Similarly, the modification of phosphatase activity has been investigated as a means of improving the availability of soil organic phosphorus to plants (Asmar et al., 1995; George et al., 2007; Richardson et al., 2011). Many plants exude phosphatase enzymes from their roots naturally, and this can be associated with a depletion of soil organic phosphorus (e.g. George et al., 2002). Plants can be genetically modified to express an extracellular microbial phytase enabling them to use P from inositol phosphate on simple growing substrates (Richardson et al., 2001; George et al., 2004). However, when grown in soil, adsorption of the enzyme by clay minerals substantially reduces its activity, and the enzyme is a poor

---

**Figure 13.8** Current and future targets for engineering the rhizosphere to improve crop performance. Reproduced from Ryan et al. (2009). With kind permission from Springer Science and Business Media. For a colour version of this figure, please see Plate 13.3.
Soil Conditions and Plant Growth

competitor with existing microbial communities (George et al., 2007; Giaveno et al., 2010). Achieving greater hydrolysis of organic P by plants would be beneficial on many soils, but, again, this is a topic for further research.

The biological and physical interactions occurring in the rhizosphere especially as a consequence of root growth, rhizodeposition, microbial activity, and the repeated wetting and drying of soil at the root–soil interface result in a soil matrix with different physical properties to the bulk soil (Hinsinger et al., 2009). The formation of rhizosheaths around some roots (especially grasses) is a widely recognised structural feature induced by roots and relies on immature xylem vessels, root hairs and the release of water and mucilage from roots (McCully, 1999). Mycorrhizal hyphae are also implicated in the adhesion of soil particles to roots together with the exopolysaccharides produced by rhizosphere microorganisms (Amellal et al., 1998). The release of root mucilage may also change the water relations of the rhizosphere through the release of surfactants that reduce surface tension (Read et al., 2003). Whalley et al. (2005) also found that rhizosphere soil of maize and barley tended to be drier at a given matric potential than bulk soil but suggested that differences in wetting angle and pore connectivity were the likely explanation for these differences. Managing the physical properties of the rhizosphere to stabilise soils, improve soil structure and enable plants to access deep soil water are worthy targets for the future.

References


Index

Note: Page numbers in *italics* refer to Figures; those in **bold** to Tables

acid soils
- anthropogenic drivers, 222–3
- ASS see Acid sulphate soils (ASS)
- designation, 218–19
- distribution
  - arable production, 221
  - global extent, **221**
  - nutrient supply capacity, 222
  - topsoil pH, global distribution, 219, **220**
- plant growth
  - acid tolerant plants, 223–4
  - AI toxicity, 224
  - freshwater flushing, 226
  - highly acidic ASS, 226
  - liming, 224, **225**
  - pH amendment, 224
  - rice and common bean, fertilisation, **226**
- acid sulphate soils (ASS)
  - description, 223
  - highly acidic, 226
  - oxidation process, 223
- ADS see agricultural drainage systems (ADS)
- adverse soil chemical environments
- acid and alkaline soils
  - anthropogenic drivers, 222–3
  - ASS see Acid sulphate soils (ASS)
  - designation, 218–19
  - distribution, acid soils, 219–22
  - plant growth, 223–7
- PTMs see potentially toxic metals (PTMs)
- salt-affected soils
- amelioration and management, 215–18
- origins and properties, 209–11
- plant growth, 211–15
- saline and sodic soils, definition, 206–9
- agar gel techniques, 338
- agricultural drainage systems (ADS)
  - description, 310
  - design and implementation, 312
  - direct and indirect effects, 311
- agricultural science, 6–8, 409
- agro-ecological methods, 394
- agronomic management, crops
  - ‘crop circle’, 44, **44**
  - season, 45
  - site, 44
  - species, 45
  - supplies, 45
  - system, 45
- alkaline soils
  - designation, 218–19
  - plant growth, 227
- ASS see acid sulphate soils (ASS)
- assessment, soil fertility
  - fertiliser requirements
    - bivariate regression, wheat grain yield, 63, **64**
    - management steps, 63
    - ‘precision farming’, 64–5
  - fertilisers, 62
  - industrialised countries, 62–3
  - legume agroforestry systems, 62
  - N fertilisers and pesticides, 63
  - organic ‘wastes’, 63
  - plant analysis
assessment, soil fertility (cont’d)
‘smart plants’, 65
tissue testing, 65
socio-economic factors, 62
soil analysis
base cation saturation ratios, 71
description, 65
indicator measurements, 65–6
K availability, 66–7
mobile nutrients testing, 71
soil solution concentrations, 66
spatial maps, mid-infrared sensors, 66, 67
sufficiency indices/non-limiting thresholds, 68–71
yield and nutrient supply, 67
automated irrigation scheduling
description, 301
edaphic factors, 302
flow chart, irrigation scheme, 302, 303
indicator response, water stress, 302
MAD, 303
time-based automation, 301–2
WUE, 304
bioaugmentation vs. biostimulation, 379
biofertilisers
effect, root growth, 370, 371
microbial inoculum, soil, 372, 374
mycorrhizal systems, 370
N₂ fixation, 369, 370
organic acid profile, rice cultivars, 372, 373
PSB, 371
soil matric potential and water-filled pores, 372, 374
biogenic calcrete formation, plants
calified roots, 334, 334
pedogenesis, 333
tropical soil, 335
biogeochemical cycles and GHGs
methane
concentration, atmosphere, 420
emissions, 421, 422
production, oxidation and emission, 421, 421
nitrous oxide
animal manures, 423
environmental consequences, 422
nitrogen cascade, 422, 423
UK, agriculture, 420, 420
biomass and biofuels
cumulative avoided emissions, carbon, 432, 433
energy, 431
growing crops, 430–431
bioremediation vs. conventional remediation
chemical, 377
‘dig and dump’, 377
factors, 378, 378
identification and removal, toxic constraints, 379, 380
‘blue water’, 397
CA see conservation agriculture (CA)
Calvin-Benson cycle, 29
carbon cycling, plants
allocation
above and below-grounded parts, 327, 328
labelling experiments, 329
root interaction, 328, 329
root length density, soil depth, 328, 330
biogenic calcrete formation, 333–5
grasstreatment soil, 324
rhizodeposition ecological significance, compounds
defined, 330
PGPR activities, 331
plant health, 332
rhizosphere microbial activities, 331
root and rhizosphere respiration,
332–3
and soils, 325
temporal evolution, soil properties, 325, 326
carbon management
GHG mitigation, 402, 403–4
sequestration, soil/vegetation, 402
cation exchange capacity (CEC)
higher value, 175
kaolinite and smectites, 173
sand and silt particles, 173
soil organic matter, 175
cationic elements, plant growth
beneficial and phytotoxic elements
aluminium, 168
Cd and Pb, 168–9
cobalt, 168
sodium, 167–8
genetic strategies
acquisition and utilisation, 184–6
nutritional requirements, plant species, 184
plant traits, 183–4
‘macronutrient’ see ‘macronutrient’
cationic elements
‘micronutrient’ see ‘micronutrient’
cationic elements
movement, root surface
diffusion, 162
function, mineral elements, 160, 161
mass flow, 160, 162
pH effect, phytoavailability, 160, 162
quantities, cations, 162
soil solution
and exchange sites, 175–9
high ionic strength and uncharged complexes, 179–80
properties and formation, 173–5
uptake, rhizosphere
copper, 171
high-affinity transport proteins, 172
iron, 171, 172
magnesium, 171
potassium and calcium, 171
transport proteins, plasma membrane, 170
zinc, 171
CEC see cation exchange capacity (CEC)
CGIAR see Consultative Group for International Agricultural Research (CGIAR)
chemical/organic fertilisation vs. biofertilisation, 368–9
climate change
global circulation models, 425
mitigation, 426
warming, 424–5
cluster roots, 353
conservation agriculture (CA)
effects, 439, 440
practices arose, 438–9
principles, 439
conservation tillage (CT), 439
Consultative Group for International Agricultural Research (CGIAR), 14
contamination, soil
bioaugmentation vs. biostimulation, 379
bioremediation vs. conventional remediation, 377–9
indicators, bioremediation/restoration, 382–3
rhizoremediation, 380–381
crop production
fossil fuel, 432–3
grain and legume, 432, 433
principal energy, 434, 434
CT see conservation tillage (CT)
DAPG see 2,4-Diacetylphloroglucinol (DAPG)
2,4-Diacetylphloroglucinol (DAPG), 384, 385
diffusion, definition, 162
‘Digital Soil Mapping’, 16–17
ecosystem services see soil-plant system, ecosystem services
embedded/hidden water, 397–8
enteric fermentation, 420
FE method see fumigation-extraction (FE) method
fertility
assessment see assessment, soil fertility description, 49
ecosystems dynamics
abiotic factors, 55
agricultural intensification, 54
annual average biomass, 54
ecosystem-driven model, 54
interactions, functioning, 54, 56
long-term Highfield experiment, 54, 55
soil-borne pathogens, 56–7
land capability and suitability assessment
constraints, growth, 50–51, 51
texture and stoniness, 50
management
fertility (cont’d)
integrated soil fertility management, 75–7
targeted inputs, 72–5
twenty-first century, 78–9
nutrient reserve
artificial manures, 51
manuring/fertilising, 51, 52
plant–soil interactions
association and symbiosis, 58–60
plants on soil, 60–62
soil on plant roots, 57–8
soil quality
definition, 52
food and fibre production, conditions, 53
FI see fumigation-incubation (FI)
food and fibre production, soil
biosolids, 394, 395
evaluation, nutrient balances, 396, 397
fertiliser applications, 396, 397
world nutrient usage, 394, 396
food production and yield
crops, 428, 429
factors, 427
irrigated cereal systems, 427, 430
yield gap, 430, 431
food security
defined, 426
‘exchange-value’, 427
globalisation, 427
fractionation methods, SOM
physical approaches
clay and finer silt particles, 111, 113
‘free OM (FR-SOM)’, 114, 114
‘intra-aggregate OM (IA-SOM)’, 115
‘light fraction’, 114
micro-and macro-aggregates, 113–14
organic carbon enrichment factors, 111, 112
solubility, 111
fragmentation, soil
brittle, plastic and liquid behaviour, 253
clods, 254, 254
drivers, 252
LL and PL, 253
thin sections, topsoil, 252, 253
water retention characteristics, 254
friability, soil
classification, 255
defined, 254–5
evaluation, failure stress distribution, 255
management, 255–6
fumigation-extraction (FE) method, 116
fumigation-incubation (FI), 116
Fusarium wilt disease model, 385
GCTE see Global Change and Terrestrial Ecosystems (GCTE)
Global Change and Terrestrial Ecosystems (GCTE), 94
greenhouse gases (GHGs)and
biogeochemical cycles see
biogeochemical cycles and GHGs
‘green water’, 397
groundwater table (GWT)
depth, 273, 310
description, 272
permissible levels, 310
GWT see groundwater table (GWT)
Haber–Bosch process, 394, 422
Hadley Centre General Circulation Model, 124
human-induced environmental changes
climate change, 424–6
GHGs and biogeochemical cycles, 419–423
land use change, extensification and intensification
biodiversity loss, 419
projected contributions, increased crop productions, 418, 418
warming, climate, 419
urbanisation, 423–4
hyphae, 373
integrated water resources management (IWRM), 270
intrusion and extrusion of liquids, soil
structure measurements, 248
irrigation see also automated irrigation scheduling
AquaCrop, 301
categories, 298
‘controlled’ DI, 299–300
CropSyst and WOFOST, 301

deficit irrigation (DI), 299

localised, 299

RDI and PRD, 300

and salinity, 304–5

sprinklers, 298

subsurface drip irrigation, 299

supplemental irrigation (SI), 301

IWRM see integrated water resources management (IWRM)

land application techniques, 399

liquid limit (LL), 253

LL see liquid limit (LL)

‘macronutrient’ cationic elements

calcium

concentrations, 163, 182
deficiency, 163
fertilisers, 182
toxicity, 163

magnesium

cellular, 164
concentrations, 164, 182
fertilisers, 182

potassium

concentrations, 180
deficiency, 163
description, 163
fertilisers, optimisation strategies, 181, 181

MEA see Millennium Ecosystem Assessment (MEA)

methanogenesis, 421–2

microbial biomass, soil

FI and FE, 116

old grassland soil, 116
rhizobia and mycorrhizae, 115

‘whole population’ approach, 115

microbiology, soil

animal and vegetable matter, 10
bacteriology, 11
gaseous nitrogen, 12
leguminous and non-leguminous plants, 12

nitrate formation and soil productiveness, 11

nitrogen supply and plant growth, 12

nitrogen weight, air, 13

‘micronutrient’ cationic elements

copper, 166, 183
deficiency, 183
iron

celates, 182
deficiency, 164–5
description, 164
toxicity, 165
manganese, 165, 183
nickel, 167, 183

phytoavailability, 182
zinc

concentrations, 165–6
deficiency and toxicity, 166
fertilisers, 182–3

Millennium Ecosystem Assessment (MEA)
description, 391

human-induced changes, environment, 427

mineralisation

nitrogen

C:N ratio, heterotrophic bacteria and fungi, 138
description, 137
dissolved organic N (DON), 137–8

phosphorus

accumulation and forms, 139, 140
ectomyorrhizal fungi (EM), 141
Mediterranean Antipodean study, 139
New Zealand soils, 141, 142

phosphatise activity, African soils, 141, 143

molecular fingerprinting techniques, 382
mycorrhiza helper bacteria, 351

neutron moderation technique, 289

nitrogen and phosphorus cycles

agricultural systems, 149–53
characteristics, continents, 136, 137

emissions, air and water
denitrification, 147
leaching, 148
NH₃, 147–8
phosphorus, 148–9

fixation, 139
grazed pastoral systems, 137
land-use change, 155
nitrogen and phosphorus cycles (cont’d)
managed agricultural systems, 133, 133
mineralisation see mineralisation
nitrification and denitrification, 138–9
nutrient losses, water and air
acidified waters, 147
ammonia (NH₃), 147
wider environment, 145, 146
phosphorus fertilisers, 153–5
soil animals, 141–3
soil biodiversity, 143–5
sources, soil-plant systems
biological fixation, 134
pea crops, 134
nitrogen mineralisation
carbon-nitrogen ratio and threshold values, 401, 401–2
micro-organisms, 400, 401
NMR see Nuclear magnetic resonance (NMR)
non-invasive imaging, soil structure measurements, 247
nuclear magnetic resonance (NMR), 247
nutrients
cycling
community changes, 371
crop cultivars, 371–2
dung beetles, 377
supply/status
biofertilisers, 369–72
chemical/organic fertilisation vs. biofertilisation, 368–9
optodes, 338
PGPF see plant growth promoting fungi (PGPF)
PGPR see plant growth-promoting rhizobacteria (PGPR)
phosphorus cycles see nitrogen and phosphorus cycles
photosynthesis
description, 28–9
energy conversion, 29
leaf level light interception
‘light response curve’, 30, 30
saturating irradiance, 31
light interception, crop level
crop age and abiotic stresses, 32
‘harvest index’, 32
‘radiation use efficiency’ (RUE), 31
total dry matter and accumulated intercepted radiation, 31, 31
mechanisms, 29
phytoremediation
description, 201
phytoextraction
advantages and disadvantages, 203
assisted, 204
field trials, 203, 205
hyperaccumulators, 201, 203, 204
phytovolatilisation, 206
PL see plastic limit (PL)
plant and crop science
agronomic management, 43–5
cereals, 23
development
definition, 23
duration, 24
flowering, 24–6
thermal time, 24
growth
definition, 26
solar radiation and water, 26–7
principles, resource capture efficiency, factory, 22
framework, 23
vegetation ‘factories’, 22
solar and water engines, 42–3
solar radiation
‘canopies’, 27
capture, 27–8
conversion, 28–32
sugar, 23
water
capture systems, 37–8
movement systems, 38–9
plant water, 36–7
soil water, 33–6
SPAC, 32–3
plant growth promoting fungi (PGPF), 59–60
plant growth-promoting rhizobacteria (PGPR), 59, 331
Index

plant growth, soil physical conditions and diagnostic measurements effects, 257, 257
local root response
cell wall stiffness, 259
hypoxia, 260
root elongation rate, 258, 259
water content, 260
root shoot signalling and feedback responses, 258, 260
plant-induced changes, soil processes and properties
carbon cycling see carbon cycling, plants
ecosystem engineers, 354, 355
plant-mediated processes, 354
roots, 324
soil acidification, 335–42
vegetation, 323
weathering, minerals see weathering, soil minerals
plants
physiology
agricultural chemistry, 5
ash, 5
carbon and nutrients, 6
‘carbonisation’, 5
respiration, 5
soil acidification
calcareous soil, 335
organic acids, rhizosphere pH changes, 341–2
proton budget and cycle, 336, 337
root-induced changes, rhizosphere, 336–41
plant–soil interactions
association and symbiosis, roots and organisms
acidification and ca-chelation, 60
action modes, PGPR species, 59
AM association, 58–9
nitrogen-fixing, 59
PGPF, 59–60
rhizobia–legume interactions, 59
plants on soil
elements and molecules, 61–2
plant pathogens, 61
‘rhizosphere’, 60, 61
species-specific flavonoids, 61
soil on plant roots
micro-organisms, 58
physiological changes, 57, 58
root branching, 57
water and nutrient uptake, 57
plants protection, pests and diseases
and biocontrol management systems, 383
‘conducive’ and ‘suppressive’ soils, 384
DAPG production, 384, 385
Fusarium wilt disease model, 385
plastic limit (PL), 253
potentially toxic metals (PTMs)
anthropogenic sources, 196
mobility/availability control
adsorption, 197
chelation, metal ion, 197, 198
highly sorptive oxyhydroxides, 198
pH, 196–7
point of zero charge (PZC), 197
speciation, 197, 199
specific and non-specific, 197, 198
natural sources
airborne dust and aerosols, 196
ranges, ‘typical’ uncontaminated soils, 195, 196
phytoremediation see phytoremediation
plant growth
adaptations, hyperaccumulate metals, 201, 202
excessive concentration tolerating mechanism, 201
phytoremediation, 201
toxic effects, 199, 200
precision farming, 17–18, 64, 65
precision technology, 397
proteoid roots, 353
proton budget and cycle, plants, 336, 337
PSB see P-solubilising bacteria (PSB)
P-solubilising bacteria (PSB), 371
redistribution, 272
rhizosphere
pH changes, organic acids, 341–2
root-induced changes
copper bioavailability and rhizosphere
pH, 339, 340
pH-sensitive dyes, 338
protons, 336
rhizosphere (cont’d)
vs. rhizobox experiments, 341
temporal development, pH, 338, 339
and roots management
biological and physical interactions, 442
defined, 439
ingenierie, 441, 441
modification, phosphatase activity, 441
rock-eating fungi, 351
root and rhizosphere respiration, 332–3

salt-affected soils
amelioration and management
saline/sodic conditions, 217
salinity/sodicity prevention, 217, 218
salt tolerant species, 215
sodic soil remediation, 216
tolerant crops, 216–17
origins and properties
saline soils, 209–10
sodic soils, 210–211
plant growth
diversity, salt tolerance, 211, 211
energy input, 214
high pH, 212, 214
ion toxicity and imbalance, 214–15
physiological drought, 212, 214
properties, 212, 213
sensitivity, plants, 212
saline and sodic soils
classification, 208
CROSS, 208
definition, 206
global distribution, 206, 207
’saline sodic’, 209
saturation water content, 206
sodium adsorption ratio (SAR), 208
Sites of Special Scientific Interest (SSSIs), 408–9
SOC see soil organic carbon (SOC)
soil and plant health management
and algae, soil stabilisation, 376
animals, 377
contamination, 377–83
and land use
defined, 366
drivers, 368
indicators, functions, 367, 368
tripartite interaction, 367, 367
nutrient supply/status, 368–72
protection, pests and diseases, 383–5
restoration, soil structural integrity/stability and organic matter
management, 376
mechanisms, 373, 375
slaking and dispersion, 375
soil management, climate change
mitigation
carbon see carbon management
nitrous oxide and methane, 405
soil organic carbon (SOC)
analysis
carbonate C, 87
loss on ignition, 88
Walkley–Black method, 88
content
Broadbalk Wheat experiments, 90
expression, 87
Morrow Plots, Illinois, 91, 91
Nova Vida Ranch chronosequence, Brazil, 91–2, 92
organic materials addition, 123
sandy soil, 93, 94
silty clay loam soil, 92–3, 93
temperature impact, decomposition, 125–6
soil organic matter (SOM)
biological components, 367
chemical structure
black carbon, 106–7
nitrogen, organic forms, 105–6
nuclear magnetic resonance (NMR) spectroscopy, 100–105
polyketides, 107
supramolecular theory, 107
and climate change
carbon losses, 120–121, 121
factors, 120
impact, C stocks, 124–6
sequestration, 121–4
fractionation methods
physical approaches, 111–15
solubility, 111
inputs and outputs, organic carbon
agricultural soils, 88–9
decomposition, 89
management practices
continuous maize treatment, 91
farmyard manure (FYM), 89–90
long-term Broadbalk experiment, 89, 90
Nova Vida Ranch chronosequence, 91, 92
silty clay loam soil, 92–3
soil organic nitrogen (SON), 90
‘space for time’ comparison, 91–2
texture, 93, 94
modelling
application, 98–9
coupling, 98
decomposition, 95–8
types, 94–5
validity, 99
organo-mineral interactions
associations, 108, 109
cationic bridges, 108
dissolved organic matter (DOM)
  decomposition, 108
hydrophobic groups, 109
ligand exchange, 108
sorbed organic matter, 109, 110
physical conditions
arable crops, 118, 118
  Broadbalk Wheat experiment, 119–20
draught, 117
labile fraction, 117
water holding capacity, 116
yields, spring barley, 119, 119
SOC analysis, 87–8
soil microbial biomass see microbial
  biomass, soil
terminology
C content, 87
density map, 87, 88
Soil Organic Matter Network (SOMNET), 94
soil physical environment management,
  plants defined, 238
  physical conditions, plant growth,
  257–61
properties
density, 240
porosity, density and void ratio, 239
water, 240
structure see soil structure
tilth and seedbeds, 249–57
soil-plant-atmosphere-continuum (SPAC)
  short-term and long-term deficits, 33
terrestrial plants, 32
soil–plant interactions
ADAS, 14
CGIAR, 14
chemical and biological constraints, 19
ecosystem services, 18–19
fertilisers, 18
NAAS, 14
nutrients, search
  nineteenth century, second half, 8–10
1800–1860 period, 5–8
phlogistic period, 1750–1800, 4
precision farming, 17–18
‘principle’ of vegetation, 1630–1750
drill and horse-hoe, 3
manures, 3
spearmint, 2–3, 3
Roman times, 1
soil microbiology
  animal and vegetable matter, 10
  bacteriology, 11
gaseous nitrogen, 12
leguminous and non-leguminous
  plants, 12
  nitrate formation and soil
  productiveness, 11
nitrogen supply and plant growth, 12
nitrogen weight, air, 13
soil science, 14–15
surveys, 15–17
water use efficiency, 19
soil-plant system, ecosystem services
  conflicts, lowland agricultural grassland,
  392, 393
  conservation, biodiversity
    biodiversity-function relationships,
    406, 406
    pharmacological bioprospecting,
    407
  Emergy, 392
  estimates, monetary value, 391, 392
  food and fibre production, 394–7
  and human well-being, 390, 391
soil-plant system, ecosystem services (*cont’d*)
management, climate change mitigation, 402–5
preservation
cultural and archaeological soils, 408
geological conservation criteria, SSSIs states, 408–9
human activities, 407
terra preta soils, 408
sustainable use and quality, 409–13
water recycling and disposal, 399–402
water storage and purification, 397–9
soil-plant water relations
drying, soil, 282
leaf water potential, 281–2
soil salinity, 282
stomatal conductance, 281
stomatal regulation, 281
WUE, 281
soil quality and sustainable use
definitions and interpretations, 409, 410
indicators sets, assessment, 409, 411–12
repeat measurements, 413
soil structure
aggregation
‘aggregates within clods’, 244
plant roots, 244
stability, 245–6
densities and porosities, 241, 242
formation
build-up and dissipation, air pressure, 243, 244
hydrological processes, 241, 243
tillage, 243
measures
intrusion and extrusion, liquids, 248
non-invasive imaging, 247
soil thin sections, 247
visual indicators, 248
physical, 241, 242
pore sizes and relevance, properties, 241, 243
soil thin sections, soil structure measurements, 247
soil water balance (SWB)
capillary rise, 273
content determination
electromagnetic (EM) methods, 289–90
gravimetric method, 288–9
neutron moderation/neutron thermalization, 289
drying, 274
estimation
Big-Leaf model, 296
BREB method, 296
Eddy covariance (EC), 295, 295
evapotranspiration, 292
lysimeter set-up, loadcells, 293, 293
micrometeorological methods, 294
modelling, 296
soil evaporation, load cell-based lysimeter, 294, 294
stomatal conductance, 297
surface energy balance (SEB), 294–5
weighing lysimetry, 292–3
evaporation, 273–4
evapo-transpiration, 273
GWT, 272
hydraulic conductivity estimation, 297
hydrological cycle, 271
infiltration rate, 271–2
interception loss, 271
perched water table, 273
potential evaporation, 273
pressure sensors see soil water pressure sensors
redistribution, 272
shortage/excess, water, 270–271
soil-based irrigation-scheduling approaches, 288
steady-state infiltrability, 272
store, 271
surface storage capacity, 272
soil water management
available soil water
clay soil, 277
FC and PWP, 278
relative daily transpiration rate, 277–8, 278
sandy soils, 277
water stress coefficient, 277
well-watered and water-stressed dwarf sunflowers, 279
balance determination
estimation, 292–7
hydraulic conductivity estimation, 297
soil-based irrigation-scheduling approaches, 288
soil water content, 288–90
soil water pressure sensors, 290–292
excess water, 269–70
excess water alleviation
ADS see Agricultural drainage systems (ADS)
causes and remediation solutions, 308–10
design and implementation, drainage systems, 312
moisture contents and retention curve
field capacity (FC), 276
hysteresis, 276
permanent wilting point (PWP), 275
pore size distribution, 276
saturated soil, 275
water release curve, 276
plant indicators, water stress, 284–8
soil-plant water relations see soil-plant water relations
soil water shortage alleviation
agronomic WUE, 306–8
automated irrigation scheduling, 301–4
irrigation and salinity, 304–5
irrigation types and scheduling, 298–301
SWB see soil water balance (SWB)
transpiration, 269
water flow
gravitational potential, 279
gravity and suction, 279
hydraulic conductivity, 280
hydraulic potential, 279
parameter values, 280
pore size distribution, 280
potential diagram, 279
‘preferential flow’, 280
Van Genuchten equations, 280–281
water retention see water, retention
WUE see water use efficiency (WUE)
soil water pressure sensors
Bourdon-type vacuum gauge, 290
electrical resistance method, 292
gypsum, 292
heat dissipation method, 291–2
piezometers, 290
tensiometer, 290–291
thermocouple psychrometers (TCPs), 291
solar radiation
capture, crops
Beer’s law, 28
erectness/flatness, leaves, 28
light transmission, 27
Monsi-Saeki equation, 27
size and longevity, leaf canopy, 27
photosynthesis
description, 28–9
energy conversion, 29
and light interception, 30–32
mechanisms, 29
respiration, 29–30
SOM see soil organic matter (SOM)
SOM models
cohort, 95
food-web, 95, 97
process-based, 94–5, 96
SOMNET see Soil Organic Matter Network (SOMNET)
SPAC see soil-plant-atmosphere-continuum (SPAC)
SSSIs see Sites of Special Scientific Interest (SSSIs)
sustainable intensification
CT and CA, 438–9
elements
global food demand, 434
past N fertiliser, 435
principles, 435
yields, land clearance and GHG emissions, 435, 436
improved efficiency, resource use
applications, fertiliser, 437, 438
nutrients, 437
semi-arid production systems, 438
roots and rhizosphere management, 439–42
sustainable production, environmental changes
description, 417–18
food security and production, 426–34
sustainable production, environmental changes (cont’d)
human-induced see Human-induced environmental changes
routes, intensification, 434–42
SWB see soil water balance (SWB)
terra preta soils, 408
tillage
advantages and disadvantages, 250, 251
conservation, 249
field operations, forms and cultivation, 249, 250
tilth and seedbeds
factors regulation, fragmentation, 252–4
friability, 254–6
physical properties
aggregates, 251
loosening, soil, 250
water uptake, lentil seeds, 251, 252
post-tillage dynamics
changes, tilth, 256, 256
physics, 257
weathering processes, 256
research, 249
tillage see tillage
transpiration, 40
urbanisation, 423–4
vegetation
cation uptake
agroecosystems, 344
potassium budgets and estimates,
long-term fertiliser trials, 346, 347
release, nonexchangeable K, 345
root-induced transformation,
phlogopite mica, 344, 344
weathering rates, micas, 344, 345
fossil records, 325
visual indicators, soil structure
measurements, 249
water
atmospheric saturation deficit, 42
capture systems
root system size, 37–8
speed, root system, 38
evaporation, 39–40
movement systems
flux estimation, 38
root water uptake, 38–9
stem transport, 39
plant
content, 37
plant water content and energy, 36
potential, 36–7
recycling and disposal, soil
decomposition process, 399
environmental conditions, 399
nitrogen mineralisation, 400–402
organisms and substrate quality, 400
terrestrial ecosystems, 399
resource conversion
CO₂ absorption, 40
dry matter/transpired water ratio, 41,
41
dry weight and transpiration, 41
retention
capillary binding, 274–5
capillary fringe, 275
cohesion, 274
description, 274
soil water suction, 275
surface tension, 274
roots, total productivity, 41
soil
available water content (AWC), 35–6
field capacity (FC), 35
hygroscopic coefficient, 35
permanent wilting point (PWP), 35
saturation ratio, 34–5
soil water content (SWC) and
energy, 33–4
SPAC see soil-plant-atmosphere-
continuum (SPAC)
storage and purification, soil
filtration, 399
hydrological cycle, 397, 398
residence time, 398
transpiration see transpiration
water-harvesting methods, 398
water use efficiency (WUE)
description, 41, 283
determinants, 283
‘mulching’, 284
soil texture, 284
use, 282
weathering, soil minerals
cation complexation/chelation
carboxylates, 352
cluster and proteoid roots, 353
microcosm experiment, 351
phytosiderophores, 352
plant growth and nutrition, 353
cation uplift, profile, 347–8
cation uptake, vegetation, 343–7
physical effects, plants, 342
plant-driven pH shift
arbuscular mycorrhizae, 351
bioavailability, P, 348, 350
ectomycorrhizal fungi, 350–351
flooded soils, 349
imprints, root systems, 348, 349
and rocks, processes, 343, 343
WUE see water use efficiency (WUE)
Plate 3.1 Data from the long-term Highfield experiment at Rothamsted, England, shows the impacts of long-term fallowing with intensive cultivation on (a) visible soil physical properties and (b) on some chemical and biological properties. The almost zero inputs of organic carbon via plants and tillage in the fallow plots has very significantly reduced soil organic carbon levels and soil microbial biomass but had no impact on pH in this well-buffered clay soil. Soil microbial diversity indicated by the number of bands in the DNA/DGGE is not reduced by fallowing; in fact, underlying soil microbial diversity is apparently revealed when the dominance of plant-associated bacteria is reduced. From Hirsch et al. (2009). With kind permission from Elsevier.
Low P availability in soil

Responses seen in root cells
- High affinity P transporters are more abundant (Mudge et al. 2002; Smith et al. 2003)
- Organic acids and phosphatases are secreted (Raghothama, 1999; del Pozo et al. 1999; Li et al. 2002)
- Vacuolar Pi is remobilised, fewer P-containing metabolites (Plaxton and Tran 2011)
- Phospholipids are replaced in part by sulpholipids and galactolipids (Drmann and Benning 2002; Kelly et al. 2003)
- Reduced level of RNA (Hewitt et al. 2005)

Responses seen in root systems
- Inhibition in root growth in low soil P zones as a result of reduced cell elongation and meristem activity (Reymond et al. 2006; Svistoonoff et al. 2007)
- Alteration of root architecture including many more lateral roots and long root hairs (Williamson et al. 2001; Lopez-Buccio et al. 2000; Bates and Lynch 1996)

Responses seen in whole plants
- Increased root/shoot biomass ratio
- Some plants may accelerate processes of senescence or flowering (Lynch 1995)

Plate 3.2  Complex and interlinked changes in plant physiology at root cell, root system and whole plant scales occur in response to low P availability in soil as the result of up-regulation of many genes. Photographs show Hakea prostrata R. Br. grown under sufficient and limiting phosphate supply. Cluster roots develop in some species as an adaptation for P acquisition in low P environments (details in lower image). Thanks to Stuart Pearse and Michael Shane from the University of Western Australia for access to greenhouse trials which provided the photographs.
Plate 3.3 Spatial maps of (a) clay and silt content, (b) cation exchange capacity and (c) potentially mineralisable N (where the colours show quartiles ranges as an illustration) over a 10 ha area (200 x 500 m; 180 sampling points on a 25 x 25 m grid) where the properties were determined using mid-infrared sensors. The spatial patterns were very strongly correlated with the same properties determined by traditional laboratory-based methods ($r^2 = 0.91$, 0.91 and 0.82, respectively). Unpublished data from the central wheatbelt, Western Australia provided by D. V. Murphy.
Plate 3.4  Example of the approaches used to develop a soil quality indicator package, highlighting the role of expert groups. These groups drew together a wide range of stakeholders with relevant information and experience including literature review and practical experience. Through facilitated discussion, the groups derived the relationships between yield response and a range of soil properties to inform the interpretation of the indicators used within the soil quality monitoring framework (www.soilquality.org.au) which focuses on soil quality for crop production. Data presentation using the derived categories, descriptors and traffic lights shown for three soil indicators as examples.
Plate 3.5 Developing integrated strategies for effective capture and utilisation of rainfall together with management of soil water is critical to achieve crop growth in dryland cropping systems (Tow et al., 2012). In South Australia, a range of soil management approaches are integrated within farming systems so that the potential yields (defined by available water in any season) can be attained. These include (1) introduction of tree belts (Robinson et al., 2006) and deep rooting perennial pastures to slow a rising (often saline) water table. (2) While wax degrading micro-organisms have been isolated (Roper, 2004), their success as inocula under field conditions has been limited. Instead clay amendment is an economically viable option (Hall et al., 2010) to overcome water repellence where it can be sourced locally (e.g. from subsoil on the same farm). (3) No-tillage and stubble retention practices are now common place to manage erosion and ensure maintenance of soil carbon levels and associated soil biological function (Hoyle and Murphy, 2006). (4) Organic matter stabilises soil aggregates and improves water infiltration into soil by contributing to the development of a more porous soil structure (Hoyle et al., 2012). (5) Plant breeding for improved water use efficiency (more grain per mm plant available water) and tolerance to toxicity (e.g. Al; Tang et al., 2003a). Early vigour lines (Ludwig and Asseng, 2010), which enable roots to track early season water and nutrients in leaching soils, also have the benefit of growing the root tip through surface soil constraints (e.g. diseases) quicker. (6) Development of disease-suppressive soils has occurred through crop management and rotation (Wiseman et al., 1996). (7) Amendments to soil can have either positive or negative effects on the soil biota (Bunemann et al., 2006) which can impact on the level of micro-organisms involved in carbon and nutrient cycling. (8) Deep ripping to remove compaction and at same time including gypsum or lime (Tang et al., 2003b) to depth to ameliorate subsoil constraints. Use of controlled traffic (Kingwell and Fuchsbiicher, 2011) to stop vehicle movement over the majority of the soil surface and tractors with tank tracks instead of wheels to minimise compaction.
Plate 4.1  Map of soil organic carbon stocks (carbon density) in the Brazilian Amazon. Reprinted from Cerri et al. (2007). With kind permission from Elsevier.
Plate 7.1  Global distribution of saline/sodic soils (these soils would fall within the category of salinity/alkalinity according to this map key). http://soils.usda.gov/use/worldsoils/mapindex/stresses.html. With permission from USDA-NRCS, Soil Survey Division, World Soil Resources.
Plate 7.2  Diversity in the salt tolerance of different plant species, shown as increases in shoot dry matter after growth in solution or sand culture containing NaCl for at least 3 weeks, relative to plant growth in the absence of NaCl.
Plate 8.1  Soil physical structure across a wide range of scales. Redrawn from Tisdall and Oades (1982). With kind permission from John Wiley & Sons.

Plate 8.2  Build-up and dissipation of air pressure of soil is a major process leading to slaking of soil aggregates.
Plate 8.3  Thin sections of topsoil from plots after 11 years of zero tillage (left) and conventional tillage by ploughing and discing (right). From VandenBygaart et al. (1999). With permission from Elsevier.

Plate 9.2  Lysimeter set-up based on loadcells in an olive plantation near Seville; changes in weight (with a resolution of 0.01 mm) were measured using three 50 kg capacity load cells per lysimeter from which evaporation loss was derived. (Top left) Perspex lysimeter box placed on loadcells bolted on to aluminium baseplate, (top right) holes to place inner (near tree trunk) and outer lysimeters, (bottom left), placement of lysimeters, now reinforced around edges and equipped with tubes to feed through instrument cables, (bottom right) a photo of the finished installation. Also visible are some meteorological sensors. Photos courtesy of Dr Antonio Díaz-Espejo, CSIC, Seville, Spain.
Plate 9.3  Detail of eddy covariance (EC) equipment, comprising a three-axis sonic anemometer and the adjacent Infra Red Gas Analyser (IRGA). Picture taken at the experimental station of the University of Cartagena (Spain).
Plate 10.1 Mean number of root intersects \( n = 2 \) in each grid cell of \( 25 \text{ cm}^2 \) delimited in vertical trench walls in the irrigated (a) and rainfed (b) sugarcane crops, at 332 days after planting. Reprinted from Battie-Laclau and Laclau, (2009) with kind permission from Elsevier.
Plate 10.2  Calcified roots appearing as calcite crystals shaping the root cortex cells while the stele remains free of precipitates (Jaillard, 1987a). Photograph taken in between limestone plates in the topsoil of a natural ecosystem with typical Mediterranean ‘garrigues’ vegetation cover at St Gély-du-Fesc, close to Montpellier (South of France). The photographed zone is about 2 cm × 2 cm. Photo credit: Philippe Hinsinger.
Plate 10.3  Temporal development of root-induced pH changes in the rhizosphere of growing roots of *Juncus effusus*, half an hour before and after the start of daylight, as measured noninvasively with a planar pH optode. The digital photograph (bottom) shows the two investigated roots. Reprinted from Blossfeld and Gansert (2007). With kind permission from Blackwell Publishing Ltd.
Plate 10.4 Imprints of root systems observed in situ for plant roots growing in between cracks in dense limestone plates (Hinsinger et al., 2003). Reproduced with permission from Springer-Verlag. Photograph taken in a sample collected in the topsoil of a natural ecosystem with typical Mediterranean ‘garrigues’ vegetation cover at St Gély-du-Fesc, close to Montpellier (South of France). The photographed zone is about 5 cm × 7.5 cm. Photo credit: Philippe Hinsinger. From Jaillard and Hinsinger (1993). Reproduced with permission from John Wiley & Sons.
The limitation to N$_2$ fixation by free-living heterotrophic bacteria in soil is the supply of available carbon, which is mainly through rhizosphere flow. The numbers in the figure refer to typical fluxes for a temperate cereal crop. From Killham (1994).
Plate 11.2  The relationship between soil matric potential and the neck diameter of the largest water-filled pores.

<table>
<thead>
<tr>
<th>Pore neck diameter (mm)</th>
<th>Soil matric potential (kPa)</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;0.03</td>
<td>-1000</td>
</tr>
<tr>
<td>&gt;3</td>
<td>-100</td>
</tr>
<tr>
<td>&gt;30</td>
<td>-10</td>
</tr>
</tbody>
</table>
Plate 12.1 Linkages between ecosystem services and human well-being, redrawn from the *Millennium Ecosystem Assessment* (MEA, 2005).
Plate 12.2  (a) Compost comprising composted human (‘nightsoil’) and domestic organic waste applied in small-scale vegetable production. (b) Monument to soils in China (soil of five different colours laid in a geometric pattern). (c) Linnaeus terrace in the upper Wright Valley, Antarctica. Redrawn and adapted from Pilgrim et al. (2010). With kind permission from Wiley-Blackwell.
Plate 13.1  A breakdown of the energy used in the production of wheat, potatoes and field beans in the UK in conventional and organic production systems. Green, fertiliser manufacture; red, pesticide manufacture; blue, post-harvest; purple, machinery manufacture; black, field diesel. Reproduced from Woods et al. (2010) with permission from the Royal Society, London.
Plate 13.2  Projections of 2050 values for (a) global yields, (b) global land clearing and (c) global agricultural greenhouse gas emissions and (d–f) the yields and environmental impacts of four hypothetical trajectories along which agriculture might develop by 2050. The emissions of greenhouse gases in figures (c) and (f) are expressed in terms of CO₂ equivalents. See Tilman et al. (2011) for details of the methodology used. The four curves shown in figures A–C are based on regression and reflect the different agricultural trajectories assuming magenta, current technology; blue, improvement only; orange, technology adaptation and transfer; and green, improvement and technology transfer. Reproduced from Tilman et al. (2011). With permission from Proceedings of the National Academy of Science.
Plate 13.3  Current and future targets for engineering the rhizosphere to improve crop performance. Reproduced from Ryan et al. (2009). With kind permission from Springer Science and Business Media.
More research in Soil Science from Wiley-Blackwell

Sustaining Soil Productivity in Response to Global Climate Change: Science, Policy, and Ethics
Edited by Thomas J. Sauer, John Norman and Mannava V. K. Sivakumar
A multi-disciplinary volume exploring the ethical, political and social issues surrounding the stewardship of our vital soil resources.

Ecology and Management of Forest Soils, 4th Edition
Edited by Dan Binkley and Richard Fisher
ISBN: 978-0-470-97946-4
A completely revised and updated Fourth Edition of this classic textbook in the field. Clearly structured throughout with new material on approaches to the management of forest soils and geostatistics, GIS and taxonomies.

Root Genomics and Soil Interactions
Edited by Martin Crespi
ISBN: 978-0-470-96043-1
Examines the use of genome-based technologies to understand root development and adaptability to biotic and abiotic stresses and changes in the soil environment.

Get your copy today!
Order online at www.wiley.com/go/soil
Find more at www.wiley.com/go/soil