Environmental Management in Practice

Volume 2

Compartments, Stressors and Sectors

Edited by B. Nath, L. Hens, P. Compton and D. Devuyst

ROUTLEDGE

ENVIRONMENTAL MANAGEMENT IN PRACTICE: VOLUME 2

Methods of environmental management, and especially the 'tools' of environmental management, are increasingly being relied upon worldwide to deliver a degree of sustainability in all human activities. A thorough understanding of the nature, capabilities and limitations of these 'tools', as well as the conditions under which they can be best applied, is essential for students, researchers and practitioners within the field of environmental management.

Environmental Management in Practice presents three comprehensive volumes containing the most up-to-date research and practical applications in the field. Spanning the four main aspects of environmental management: instruments, compartments, sectors and ecosystems, this three-volume work contains over sixty contributions from leading specialists in each field and offers the first major source of contemporary international research and application within environmental management in practice.

Volume 1: Instruments for Environmental Management, focuses on the instruments and tools currently available to the environment manager. A theoretical background to the instruments is given together with an overview of those instruments that are in common use today, with particular attention to the physical, economic, legislative and communication instruments.

Volume 2: Compartments, Stressors and Sectors, deals with the problems that occur in the three 'compartments' of the environment—namely, air, water and soil. The contributors also address the socio-economic sectors of industry, traffic, energy, agriculture and tourism.

Volume 3: Managing the Ecosystem, focuses on those ecosystems in which human intervention has been or continues to be predominant, specifically within cities and rural areas.
Packed with accessible and up-to-date information, these three volumes provide a comprehensive overview of environmental management for those studying, researching and practising in the field.

Bhaskar Nath is Director of the European Centre for Pollution Research, London. **Luc Hens** is Professor and Head of the Human Ecology Department at the Free University of Brussels. **Paul Compton** is an environmental and demographic consultant. **Dimitri Devuyst** is Co-ordinator of the programme of Environmental Impact Assessment in the Department of Human Ecology, Free University of Brussels.

ENVIRONMENTAL MANAGEMENT IN PRACTICE: VOLUME 2

Compartments, Stressors and Sectors

EDITED BY
B.NATH, L.HENS,
P.COMPTON AND D.DEVUYST

London and New York

First published 1999
by Routledge
11 New Fetter Lane, London EC4P 4EE

This edition published in the Taylor & Francis e-Library, 2002.

Simultaneously published in the USA and Canada
by Routledge
29 West 35th Street, New York, NY 10001

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

Library of Congress Cataloging in Publication Data
Environmental management in practice/edited by B.Nath…[et al.].
p. cm.
Includes bibliographical references and index.
Contents: v. 1. Instruments for environmental management —v. 2. Compartments, stressors and sectors—v. 3. Managing the ecosystem.
1. Environmental management. I. Nath, Bhaskar.
GE300.E577 1999 97–52071
363.7–dc21 CIP

ISBN 0-415-14907-X (Print Edition)
ISBN 0-203-02320-X Master e-book ISBN
ISBN 0-203-17290-6 (Glassbook Format)

CONTENTS

FIGURES

TABLES

BOXES

CONTRIBUTORS

Giancarlo Barbiroli is Professor of Technology and Economics of Production Cycles at the Faculty of Economics, University of Bologna, Italy. He holds a degree in economics and a Ph.D. in commodity science and technology. Dr Barbiroli's research has concentrated on relationships between technological dynamics and natural resources (energy, materials, agro-food and agro-industrial resources), and between technological dynamics and industrial production. Professor Barbiroli's aim in analysing relevant effects on industries and economic systems has been to draw general conclusions which help to identify and implement sustainable production options. His approach to research has always been inter- and multidisciplinary, frequently using quantitative methods. He has published ten books and around one hundred and twenty papers in scientific journals. From 1984 to 1993 he was Dean of the Faculty of Economics at Bologna. He is a member of the American Association for the Advancement of Science.

Philip M.L.Barnes has been in the international energy business for 38 years, largely with Shell International. He has worked in a variety of countries including Germany, the USA and the UK, and has covered many aspects of the field, from strategic planning to crude oil sales. He is now a senior research associate with the Oxford Institute of Energy Studies, Oxford, UK.

Dimitra Brillaki is a chemical engineer at a postgraduate level at the Computational Fluid Dynamics Unit of NTUA. She has research experience in air quality/air pollution issues.

Alfons Buekens, Ph.D., is a chemical engineer and Professor at the Free University, Brussels. Over the years he has developed processes for the thermal treatment of waste and conducted numerous studies in the field of the environment and safety.

Mickel Christolis is a research collaborator at the Computational Fluid Dynamics Unit of the National Technical University of Athens (NTUA) on environment and risk analysis issues. He has 15 years' experience in air pollution monitoring, pollutants dispersion modelling, industrial accidents impact assessment and in the implementation of the Seveso directive in Greece. He is a member of the New York Academy of Science.

Paul A.Compton, M.Sc., Ph.D., formerly Professor in the School of Geosciences at the Queen's University of Belfast, Northern Ireland, is an environmental and demographic consultant. His academic interests lie at the interface between the natural and social sciences.

James G.Cruickshank, B.Sc. (Hons), was a senior lecturer at the Department of Geography in the Queen's University of Belfast (1960–1987). Since then, he has been the Principal Soil Surveyor at the Department of Agriculture in Northern Ireland. He has published over thirty scientific papers, written a popular book on soil geography and edited two major texts dealing with environmental resources in Northern Ireland.

Dimitri Devuyst received his Ph.D. degree in human ecology from the Free University of Brussels (Belgium), as well as a Master's degree in human ecology and a Bachelor's degree in botany from the same university. His main interests are in the field of environmental impact assessment. Currently he is postdoctoral fellow of the Fund for Scientific Research, Flanders (Belgium).

Sophia Everett is a senior lecturer in the Institute of Transport Studies, Graduate School of Business at the University of Sydney (Australia). She holds a Master's degree in Public Policy from the University of New England and an MA and Ph.D. from the University of Wollongong. She was formerly a post-Doctoral Fellow at the University of Wollongong and Associate Director of the Maritime Programme at the Institute of Transport Studies.

Kerry Garrett, Ph.D., is former Head of the Department of Agricultural and Environmental Science at the Queen's University of Belfast and also Deputy Chief Scientist in the Northern Ireland Department of Agriculture. He is a past-Fellow of the Salzburg Seminar and President of the Fertiliser Society and has engaged in collaborative research with fellow scientists in the USA, Mexico, China, Hungary and Bulgaria.

Hans W.Gottinger has worked in operations research and energy, environmental and industrial economics. He earned a Ph.D. degree in economics at the University of Munich. He has been a visiting professor at the University of California at Berkeley and the University of Oxford, Professor of Economics at the University of Bielefeld and is Professor of Management Science at the University of Maastricht. He is also the Director of the International Institute for Environmental Economics and Management (IIEEM), Bad Waldsee, Germany.

Luc Hens graduated as a biologist and later received his Ph.D. in biology from the Free University of Brussels (Belgium), where he is at present Professor and Head of the Human Ecology Department. He also lectures at Antwerp University (Belgium), the Technical University of Sofia (Bulgaria) and Ankara University (Turkey). His specific area of research concerns the elucidation of interdisciplinary instruments for sustainable development. Professor Hens acts as an expert in environmental policy in several councils in Belgium. He is the European Editor for the *International Journal on Environmental Pollution*.

David A.Hensher is an economist and Professor in the Institute of Transport Studies at the University of Sydney (Australia). He holds a Ph.D. from the University of New South Wales and is a Fellow of the Australian Academy of Social Sciences. His professional career has involved the government sector and consultancy as well as the Universities of Oxford, Melbourne and Macquarie. He has numerous publications in the field of transport economics and the environment.

Paul Hooper is an Associate Professor and Deputy Director of the Institute of Transport Studies in the University of Sydney (Australia). He holds a Ph.D. in management from the University of Sydney and a Masters in transport economics from the University of Tasmania. He was formerly employed in government as a policy adviser and has acted as a consultant on transport and tourist matters to numerous organisations.

Robert E.Jonckheere, Doctor of Applied Sciences, began his career as engineer in the industrial sector (in the United States, Belgium and France). He is retired full professor in honour at the Department for Mechanical Engineering and Acoustics of the Vrije Universiteit Brussel.

Nicos Manalis works in the Air Quality Department of the Ministry of the Environment in Greece. He is responsible for the air pollution laboratory of the Ministry of the Environment and has much experience in air quality monitoring issues.

Nicolas Markatos earned his M.Sc. degree in Athens and his Ph.D. degree in London (Imperial College). He was a research fellow at Imperial College, Professor and Co-Director at the Centre of Process Analysis at the University of Greenwich, London, and is currently Professor of Chemical Engineering at the National Technical University of Athens. He has been Rector of the University since 1991. Dr Markatos is the author of three books and over one hundred scientific publications in computational fluid dynamics, transport phenomena, turbulence, environmental flows and two-phase flows.

Atanaska Nikolova is an environmental engineer and took a European Masters Degree in Environmental Protection and Sustainable Development at Sofia University of Technology (Bulgaria) in 1995. In 1996 she worked in environmental impact assessment in tourism projects at the Human Ecology Department of the Vrije Universiteit Brussel (Belgium).

Maurizio G.Paoletti graduated in natural history at the University of Padova (Italy) and has received degrees in biological control from the University of California at Berkeley and in human ecology from the University of Padova. He currently holds a research position at the University of Padova and has been a visiting professor at Cornell and Ohio State Universities, as well as at the Free University of Brussels and Torun University in Poland. Dr Paoletti's research interests cover soil vertebrates, ecology, agro-ecology and ethnobotany, areas in which he has published 170 scientific papers and 15 books.

Mainwaring B.Pescod, B.Sc., SM, C.Eng., is Professor of Environmental Control Engineering and Head of the Department of Civil Engineering at the University of Newcastle upon Tyne (UK). He has held academic positions in Sierra Leone and Thailand and served as consultant to the FAO and World Health Organization. Professor Pescod has numerous publications in the field of environmental engineering and was made an Officer of the Order of the British Empire (OBE) for his services to technical education.

M.J.Peterson is Associate Professor of Political Science at the University of Massachusetts (USA). She is the author of *Managing the Frozen South: The Creation and Evolution of the Arctic Treaty System* (1988) and of articles on the International Whaling Commission and on international fisheries management.

Ivan Raev received a Ph.D. degree in forest hydrology in 1974 and a D.Sc. degree in forest ecology in 1990. He has been head of the Department of Forest Ecology since 1986 and since 1991 has been Professor of Forest Ecology. In 1991 and 1992 he was President of the Forestry Committee of the Bulgarian Council of Ministries. He has been Director of the Forest Research Institute of the Bulgarian Academy of Sciences since 1995. Professor Raev is the author of more than one hundred and twenty-five scientific publications.

Ross Robinson is an Associate Professor in the Institute of Transport Studies at Sydney University (Australia) and Director of its Maritime Programme. He was awarded a Ph.D. by the University of British Columbia and holds an MA from the University of New South Wales. During his professional career, he has been involved with the Port Development Programme, UNCTAD, Geneva and is a former Director of the Port Development Programme UN/ESCAP, Bangkok, Thailand. Dr Robinson is also former Director of the Centre for Transport Policy Analysis at the Universty of Wollongong.

Marc Van Overmeire is Head of the Department of Mechanical Engineering and Acoustics and Professor at the Free University of Brussels (Belgium). He obtained his Doctorate in Applied Sciences from the same university in 1983. He is a member of the Steering Committee of the Belgium Society of Mechanical and Environmental Engineering and also a member of the Belgium Acoustical Society (ABAV).

Filip J.R.Verbandt is a mechanical engineer and Doctor in Applied Sciences from the Vrije Universiteit Brussel (VUB). He is a specialist in noise and vibration problems and acts as an acoustic consultant, owning his own company, EVA-International. Dr Verbandt also lectures on noise and vibration in VUB, EHSAL (Economische Hogeschool Sint-Aloysius) and KUL (Katholieke Universiteit Leuven) programmes on human ecology and environmental impact assessment.

Alyson C.Warhurst is Professor on Environmental Strategy at the School of Management, University of Bath. She is also Director of the Mining and Environment Research Network (MERN) and of the International Centre for the Environment (ICE) at the University of Bath. Professor Warhurst has an established academic reputation in business research at the public policy-corporate strategy interface, and has published widely in this field. As an international expert in the mining sector she has worked *inter alia* in South America and China.

Paul L.Younger studied geology as an undergraduate and holds a Ph.D. in civil engineering. He was a Harkness Fellow at Oklahoma State University between 1984 and 1986 where he was awarded his MS. He has worked for the UK National Rivers Authority as a hydrogeologist and is currently Lecturer in Water Resources Engineering at the University of Newcastle upon Tyne. He was Honorary Secretary of the British Hydrological Society from 1994 to 1996.

PREFACE AND ACKNOWLEDGEMENTS

Environmental management draws its knowledge base from across the spectrum of disciplines—the natural, social and medical sciences, the humanities and engineering. It aims to maintain a harmonious relationship between the environment and human society, and in its approach to this adopts a holistic, interdisciplinary stance. Since value judgements are an integral part of environmental management, it is as much an art as a science in its methodology and application.

The growth of interest in environmental management is relatively recent. It reflects a widely held perception of accelerating environmental deterioration caused by the pressure of human activities, as evidenced by worsening problems of pollution and the destruction of natural landscapes and habitats. These concerns can be traced back to the 1960s, when the interconnectedness of nature was vividly demonstrated by the way in which seemingly benign activities such as the chemical control of pests could, by diffusing through the food chain, produce adverse environmental effects in regions ostensibly untouched by man's activities.

As our knowledge of the global environment has grown, other worrying effects have come to light. The emission of greenhouse gases is linked to global warming and climate change. Although we do not fully understand the probable effects of this, it may well result in greater temperate aridity and so jeopardise the world grain supply, with potentially disastrous consequences. Moreover, resultant changes in sea level could submerge major coastal sites of population.

There is also the well-established connection between CFC emissions, the depletion of upper atmosphere ozone, and increased ultra-violet radiation at the planet's surface. This has negative implications not only for human health but also for the well-being of other species. Likewise, the destruction of the tropical rain forest is seen as a grave threat to biodiversity and the world's gene pool. The fact that these hazards are the subject of internationally agreed measures of amelioration (albeit implemented with variable commitment by individual countries) testifies to the potential gravity of global warming, ozone depletion and loss of biodiversity.

These global issues also raise concerns at the level of ecosystems. The effects of modern agricultural practices on environmental quality are a case in point. Pesticide and fertiliser residues pollute the groundwater; animal and plant habitats are destroyed as hedgerows are removed and wetlands drained in the interest of intensive cultivation; soil structure is broken down, creating problems of soil erosion. Now, in addition, intensive rearing of plants and animals is even causing concern for the safety and wholesomeness of the food produced. Populations are no longer willing to accept assurances from experts that genetically engineered crops are safe, or that it is right to feed natural herbivores, such as cattle, protein supplements derived from the rendered remains of other animals.

Of course it is not only agriculture that is problematic. Urban living and its associated activities can be just as destructive of the environment; not least, the creation of built environments where residential, commercial and industrial areas and communications infrastructures either obliterate or radically change pre-existing landscapes and ecosystems. Moreover, urban systems depend upon the mobility of people and goods for their effective functioning, and so create the traffic problems associated with further detrimental effects on the environment. The problems caused by excessive use of energy and natural resources in production and consumption, and their implications for future generations, also have to be tackled. Measures to ensure effective waste disposal and the curbing of air and water pollution are vital for the maintenance of environmental quality.

When viewed over a longer time-scale, however, the environmental picture is somewhat less gloomy. For example, popular coverage might lead one to suppose that human activity is the only cause of climatic change: the evidence does, after all, appear compelling, with the atmospheric content of the major greenhouse gas, carbon dioxide, having risen progressively since the start of industrialisation in the seventeenth century. But the record also shows that the world's climate has fluctuated markedly both in the recent geological and even historical past, and scientists still disagree on whether the rise now observed in global temperature should be attributed solely to greenhouse emissions. That this rise may be part of a natural progression readily absorbed by global systems cannot be ruled out at this stage.

It is also worth bearing in mind that ever since the domestication of plants and animals and the discovery of fire, human beings have moulded their natural surroundings to suit their purposes. During the medieval period, for instance, much of the forest that once covered the continent of Europe was destroyed (a process analogous with the present destruction of the tropical rain forest), without any apparent harmful consequences in the long term. Moreover, it was the agricultural and industrial revolutions that created those environmental conditions we consider 'natural' and with which we are comfortable. It is debatable, in other words, whether any truly natural landscapes and ecosystems remain—they have all to a greater or lesser extent come under human influence.

For most of human history our attitude to the environment has been purely exploitative: nature was there to be conquered, and the resource endowment to be used in the furtherance of human development. Little or no attention was paid to the possibility of detrimental environmental impacts—indeed, in most instances these were simply not appreciated because the complex relationships and linkages of environmental systems were not understood. It is only in this century that this attitude—of man as the conqueror of nature—has changed (indeed in Eastern Europe it persisted, with disastrous consequences, right up until the demise of communism). We now think more in terms of stewardship, whereby humans owe a duty of care to the environment, and in terms of sustainability. However, it is still invariably the case that when choices have to be made economic self-interest wins the day.

The broad scope of environmental management creates its own particular problems. The information on which it relies is scattered across disciplines isolated from one another by the traditional boundaries that demarcate major branches of academic endeavour. It follows from this that relevant advances in the natural sciences may not be appreciated by those working from a social science perspective, and so on. It is therefore a major objective of this book to bring together the expertise found within the diverse fields of environmental management, with the aim of providing an accessible overview of its content and methods. The treatment is biased towards environmental management as practised at the regional level—the so-called meso-scale. Global issues such as climate change and loss of biodiversity lie outside the scope of this book and so receive only incidental mention.

The idea for this book came initially from the involvement of the four editors in environmental training programmes in Eastern Europe, and a book was duly published by the Free University of Brussels Press in 1993. This publication is a revised, improved and extended edition of that earlier version and is presented in three volumes. The theoretical principles of environmental management are illustrated with the use of up-to-date examples and case studies, and self-assessment questions are included to aid students who may wish to use it as a textbook. It should also be of interest to policy-makers and researchers seeking information about the management of today's environmental problems.

Volume 1 considers the instruments for environmental management under four main headings —predictive and scientific instruments; economic instruments; legal instruments; and instruments for environmental communication and education. It not only covers relatively long-established instruments of management, such as environmental impact assessment and risk assessment, but also introduces more recently developed approaches such as material flow analysis and life-cycle assessment. Volume 1 aims at up-to-date, comprehensive coverage and includes discussion of such important topics as the concept of sustainable development, environmental legislation in the European Union and the USA, and the management of environmental conflict.

Volume 2 is devoted to the environmental management of compartments, stressors and sectors. It covers the impact of population and the way environmental information is processed and interpreted through the filter of human culture. Soil, air and water are the environmental compartments referred to in this volume. They are subject to stress through over-use and pollution, and the manner in which these stressors should be managed constitutes a major strand of enquiry. The sectors referred to are agriculture, forestry, industry, transport and tourism, and how these should be managed in the interests of preserving environmental quality.

The theoretical and practical considerations involved in the appropriate management of major natural ecosystem types—wetlands, tropical forests, desert areas, the coastal margin, river and inland water environments—are discussed in Volume 3. These are, of course, 'natural' ecosystems only in a relative sense; the question of management arises precisely because of human impacts. Of equal importance are those environments created entirely by human activities, and in recognition of this Volume 3 also deals with rural and urban environments, as well as human ecosystems under threat, and the management of archaeological sites.

The three volumes of this textbook contain over sixty chapters written by more than eighty authors. This large project would have been impossible without the support and active contributions of many colleagues whose names are not mentioned in the individual chapters. We would like to thank especially the secretarial staff of the Human Ecology Department, Free University of Brussels (VUB) for their excellent work; especially Mr Glenn Ronsse, who was responsible for the final formatting and overall secretarial management of the project. Sincere thanks are due to the peer reviewers of the chapters.

For Volume 2 these were: W.Bauwens, Free University Brussels, Belgium; A.Buekens, Free University Brussels, Belgium; P.Calow, University of Sheffield, United Kingdom; L.Candela, Civil Engineering School, Barcelona, Spain; E.Cohen, The Hebrew University of Jerusalem, Israel; A.Cops, Catholic University Leuven, Belgium; J.Cramer, TNO, Apeldoorn, Netherlands; C.Cuijpers, Catholic University Leuven, Belgium; R.Dams, University of Ghent, Belgium; W.De Breuck, University of Ghent, Belgium; J.De Ruyck, Free University Brussels, Belgium; M.Earle, European Parliament, Brussels, Belgium; R.M.El Asswad, Addis Ababa, Ethiopia; M.K.Garrett, Queen's University of Belfast, Northern Ireland, United Kingdom; P.Janssens, Catholic University Leuven, Belgium; M.Jolánkai, Hungarian Academy of Sciences, Budapest, Hungary; J.Korsmit, Free University Brussels, Belgium; J.Kretzschmar, VITO, Mol, Belgium; E.Leppäkoski, Abo Akademi University, Turku, Finland; L.W.Lockhart, University of Winnipeg, Manitoba, Canada; F.Lox, Free University Brussels, Belgium; N.Lust, University of Ghent, Belgium; J.Matilla, Dauphin Island Sea Lab, Alabama, USA; B.McCay, Rutgers University, Newark, USA; E.Mészárös, University of Veszprem, Hungary; A.Montanari, Italian Geographical Society, Rome, Italy; S.Moolenaar, Landbouwuniversiteit Wageningen, Netherlands; J.Nijs, IMEC vzw, Leuven, Belgium; W.Passchier-Vermeer, TNO-Prevention and Health, Delft, Netherlands; P.Ponce, Free University Brussels, Belgium; G.A.Peterson, Colorado State University, Fort Collins, Col., USA; G.M.Pierzynski, Kansas State University, Manhattan, Kan., USA; L.Reynders, Universiteit Amsterdam, Netherlands; J.Schrijvers, University of Ghent, Belgium; R.D.Schuiling, Universiteit Utrecht, Netherlands;

B.Shutes, Middlesex University, London, United Kingdom; O.Teichmann, Universität-Gesamthochschule Siegen, Germany; J.Teigland, Western Norway Research Institute, Lillehammer, Norway; E.Tollens, Catholic University Leuven, Belgium; A.Tukker, TNO-Centre for Technology and Policy Studies, Apeldoorn, Netherlands; J.Unruh, University of Arizona, Tucson, Ariz., USA; R.Van Der Duin, KUB-Vakgroep Vrijetijdsweten-schappen, Tilburg, Netherlands; W.Verstraeten, University of Ghent, Belgium; K.F.Wiersum, Landbouwuniversiteit Wageningen, Netherlands.

We are deeply indebted to Mr K.Van Scharen, Executive Director of the VUB University Press, publishers of the first edition of this textbook in 1993, who was most supportive in the publication of this second edition. We thank Routledge for guidance during the establishment of this project and in its final production.

The publishers have made every effort to contact copyright holders of material reprinted in *Environmental Management in Practice: Volume 2*. However, this has not been possible in every case and we would welcome correspondence from those individuals/companies we have been unable to trace.

LIST OF ABBREVIATIONS

a.q.s.	Ambient Quality Surveillance
a.s.l.	Above Sea Level
ADR	European Agreement Concerning the International Transport of Dangerous Goods by Road
AIT	Action Initiation Time
AQS	Air Quality Standard
AWC	Available Water Capacity
BAT	Best Available Technology
BATNEEC	Best Available Technology not Entailing Excessive Cost
BGH	Bovine Growth Hormone
BIC	Business Innovation Centres
BOD	Biological Oxygen Demand
BOD_5	5-day, 20°C Biochemical Oxygen Demand
BT	Bacillus Thuringiensis
BTX	Benzene, Toluene and Xylene
C	Carbon
C/N	Carbon/Nitrogen
CANMET	Canadian Mining Research Institute
CAP	Common Agricultural Policy
CETEM	Brazilian Mining Research Institute
CFC	Chloro-fluorocarbon
CFP	Common Fisheries Policy
CMR	Convention relative au contrat de transport international de Marchandises par Route (Convention on the transport of goods by road)
CNOOC	China's National Offshore Oil Corporation
CO	Carbon Monoxide
CO_2	Carbon Dioxide
COD	Chemical Oxygen Demand
CPA	China Pacific
CSO	Combined Sewer Overflow
DC	Developing Country
DDT	Dichlorodiphenyltrichloroethane
DG	Directorate General
DICE	Dynamic Integrated Climate Economy
DNA	Deoxyribo Nucleic Acid
DOAS	Differential Absorption Spectrometry
DQO	Data Quality Objective
EC	European Community
ECOMOST	European Community Model of Sustainable Tourism
ECS	Environmental Care System
EEZ	Exclusive Economic Zone
EFTA	European Free Trade Association
EIA	Environmental Impact Assessment
EPA	Environmental Protection Agency
ESA	Environmentally Sensitive Area
EU	European Union
EUS	Environmental Utilisation Space
FAO	Food and Agricultural Organization (of the UN)
FC	Field Capacity
FDA	Food and Drug Administration
FID	Flame Ionisation Detector
FOC	Flag of Convenience
GCM	Global Circulation Model
GDP	Gross Domestic Product
GHG	Greenhouse Gas

GIS	Geographical Information System
GMO	Genetically Modified Organism
GQA	General Quality Assessment
HDPE	High Density Polyethylene
HOV	High-Occupancy Vehicle
ICAO	International Civil Aviation Organization
IFM	Integrated Fertiliser Management
IIASA	International Institute for Applied Systems Analysis, Luxemburg
INC	Industrialised Country
IPM	Integrated Pest Management
IPPC	Integrated Pollution Prevention and Control
IQ	Individual Quota
ISO	International Organization for Standardization
ITQ	Individual Transferable Quota
IUCN	International Union for Conservation of Nature
L	Level
LAC	Limits of Acceptable Change
LD_{50}	Lethal Dose for 50 per cent of a Population
LDC	Less Developed Country
L_{eq}	Equivalent Sound Level
LME	Large Marine Ecosystem
L_n	Sound Level
LNP	Noise Pollution Level
LPG	Liquid Petrol Gas
MAC	Maximum Admissible Concentration
MERN	Mining and Environment Research Network
MHD	Magnetic-Hydro-Dynamic
MSY	Maximum Sustainable Yield
NGO	Non-Governmental Organisation
NIMBY	Not In My Backyard
NIMTO	Not In My Term of Office
NMHC	Non-Methane Hydrocarbons
NMVOC	Non-Methane Volatile Organic Compounds
NO_x	Nitrogen Oxides
NPL	Noise Pollution Level
NR	Noise Rating
NRA	National Rivers Authority
OECD	Organization for Economic Co-operation and Development
OECD	Pacific OECD Countries of the Pacific Including Japan, Australia and New Zealand
OY	Optimum Yield
P	Pressure
PAH	Polycyclic Aromatic Hydrocarbon
PCB	Polychlorinated Biphenyl
PCDD	Polychlorodibenzo-p-dioxin
PCDD/F	Dioxin
PCDF	Polychlorodibenzofuran
PET	Polyethylene Terephthalate
PIC	Product of Incomplete Combustion
PM-10	Paniculate Matter (10% smallest fraction)
PT	Public Transport
PVC	Polyvinylchloride
PVE	Product Volume Efficiency
PWP	Permanent Wilting Point
R&D	Research and Development
RQO	River Quality Objective
SADC	South African Development Community
SIA	Strategic Impact Assessment
SME	Small- and Medium-Sized Enterprises
SPM	Suspended Paniculate Matter
SSO	Storm Sewer Outfall
SSZ	Source Sewer Zone
TAC	Total Allowable Catch
TCM	Tetrachloromercurate
T.C.M.	Transport Control Measure
t_1	Actual time exposed to a given noise level
T_i	Maximum permitted time for a given noise level
TNC	Transnational Company
TSP	Total Suspended Particulates
TSS	Traffic Separation Scheme
TURF	Territorial Use Rights in Fisheries
UAM	Urban Airshed Model
UNCED	United National Conference on Environment and Development (Rio de Janeiro, Brazil 1992)

UNCLOS	UN Convention of the Law of the Sea
UNEP	United Nations Environment Programme
USEPA	United States Environmental Protection Agency
UWC	Underwater Clearance
VAM	Vesicular Arbuscular Mycorrhiza
VOC	Volatile Organic Compound
WCED	World Commission on Environment and Development
WHO	World Health Organization
WRI	World Resources Institute
WTO	World Tourism Organisation
WWF	World Wide Fund for Nature (formerly World Wildlife Fund)

LIST OF UNITS

Prefixes to the names of units

G	giga (10^{9})
M	mega (10^{6})
k	kilo (10^{3})
d	deci (10^{-1})
c	centi (10^{-2})
m	milli (10^{-3})
µ	micro (10^{-6})
n	nano (10^{-9})
p	pico (10^{-12})
f	femto (10^{-15})

Units

a	annum
Å	Ångstrom (0.1 nm)
atm	atmosphere
bbl	billions of barrel
boe	barrels of oil equivalent
Bq	Becquerel
°C	degree Celsius or centigrade
cal	calorie
d	day
dB	decibel
g	gram
Gt	gigatons (=109 tons)
Gtce	gigatons of coal equivalent
GWe	gigawatt electricity
h	hour
ha	hectare
hrs	hours
Hz	hertz
J	joule
K	degree absolute (Kelvin)
l	litre
lpcd	litres per capita per day
m	metre
M	molar (mol/litre)
min	minute
P	phon
Pa	pascal (unit of pressure; 100 kPa=1 bar)
pe	percentage
PM_{10}	fraction of particulates in air of very small size (⩽10 µm)
ppm	parts per million
ppmv	parts per million (volume)
S	sone
s	second
t	tonne
Tcf	tonnes of carbon fuel
TW	terawatt
Twyr	terawatt per year
V	volt
W	watt
yr	year

Other abbreviations

kg_{bw}	kilogram body weight
ln	logarithm (natural, base e)
log	logarithm (common, base 10)
n or N	total number of individuals or variates
ppb	parts per billion
s^2	sample variance

PART I

ENVIRONMENTAL COMPARTMENTS AND STRESSORS

INTRODUCTION

Paul A.Compton

Part 1 of this volume introduces readers to the strategies employed in the management of the three environmental compartments of soil, air and water, and defines the major problems involved. The stressors that operate on these compartments arise from the growth and activities of the human population. Management strategies must therefore seek to reconcile the needs of the human population for environmental resources with the preservation of environmental quality. Many of these strategies are designed to combat the adverse consequences of pollution, and the management of solid waste disposal and noise pollution are also discussed here.

The proper management of soil, air and water is integral to the well-being of the human population. Our food is dependent upon maintaining the soil in good heart. Water is not only needed for personal consumption and hygiene but also to support the basic activities of society, including the proper functioning of settlements, industry and agriculture. None of these can be taken for granted and their management is becoming increasingly complex and costly. As for air, this is ubiquitous and treated as a free good, but here, too, it is obvious that air quality cannot be left to look after itself and that proper measures of pollution control go hand in hand with healthy living. The importance of air, soil and water is not, of course, to be measured solely against the yardstick of human well-being; they are also the environmental compartments within which ecosystems exist. The sensitivity of ecosystems to alterations in their environmental surroundings as a result of human activities is well known and profound changes have already been recorded.

The aim of soil management is the preservation of soil structure and fertility. As Cruickshank shows in his chapter, these are maintained, in the absence of human intervention, through the processes of nature. But agriculture disturbs the natural chain of events and good husbandry is needed to prevent soil deterioration and erosion. What constitutes good husbandry varies according to climate, topography and underlying geology. Agricultural practice is therefore a function of environmental conditions and historically humans have adopted a range of different strategies to maintain the soil in good condition. In tropical areas, for instance, the device of 'slash and burn' was used, whereby the soil is cultivated for a short time and then abandoned to allow fertility and structure to recover naturally. Fallow periods were also part of traditional agricultural practice in Europe, which evolved later into systems of mixed farming, with crop rotation, periods of temporary grass, and the return of animal waste to the land to maintain soil structure and fertility.

Pressure of population and the quest for greater production and profitability have, however, made these 'environmentally friendly' forms of agriculture unattractive to modern farmers. Sometimes the response has been to extend farming to uncultivated areas; but since this has invariably been done with only the haziest prior knowledge of local environmental conditions (which have subsequently turned out to be marginal for agriculture) the extension of cultivation to such areas has usually been accompanied by the adoption of inappropriate farming practices. In the case of the US Dust Bowl in the 1930s, and as we are currently seeing in the Amazon rain forest, these activities have resulted in rapid deterioration of soil fertility, the breakdown of soil structure and rapid erosion of the soil, in which the top layers are literally washed or blown away.

The other response has been the intensification of existing farming, in which soil fertility is maintained through the use of artificial fertilisers. Combined with the application of other chemicals to control weeds and pests, this has led to the adoption of virtual cereal monocultures in many temperate regions. The problem here is not one of maintaining soil fertility (which is actually enhanced through the use of chemicals) but the adverse effect on soil structure of the abandonment of mixed farming systems. The use of heavy machinery on the land compacts the subsoil, and the absence of farm animals from these arable systems means that organic matter is no longer returned to the top soil. Subsoil compaction impedes drainage, leading to increased runoff; this together with the breakdown of the structure of the top soil again leads to enhanced rates of soil loss through the erosive action of wind and water.

The management of water resources and the monitoring of air quality overlap in the sense that both are concerned with pollution control. They also share a common set of principles for controlling pollution, either by setting maximum permitted emission levels for different sources of pollution or by establishing minimum standards of air or water quality. The former approach is associated with direct regulation and the latter with indirect regulation, in which economic mechanisms and incentives may play a leading role.

However, as Pescod and Younger show, the management of water resources involves a much broader remit than that of merely controlling pollution. Basic to this is an understanding of the hydrological cycle which describes the transport of water in its various forms from atmosphere, to earth's surface, to groundwater and its return to the oceans. Water resource management is also concerned with ensuring the availability of adequate water resources for human use. As consumption goes on rising, society is increasingly coming up against the limits set by water availability. Consequently conservation is now a pressing issue in many countries. Although the use of water for industrial purposes may well have declined in recent decades, this has been more than offset by rising domestic use and the use of water in agriculture.

There is also the matter of climate change to be taken into account. For instance, the climate of Western Europe now appears to have entered a phase of lower rainfall, with drier winters and therefore inadequate recharge of the groundwater table. Hence the need for conservation through measures such as wastewater reclamation, artificial recharge and proper pricing structures is not an exclusive feature of areas subject to drought, but also occurs in areas where rainfall might otherwise appear to be adequate. A strong case can be made for building grid systems that enable water to be moved from water surplus to water deficit areas in similar manner to the transport of electricity. The desalination of sea water may also become a more attractive proposition.

Although Christolis *et al.* focus on the technical aspects of air monitoring and management, they also introduce the reader to the general aspects of air pollution and the impact on human health and the environment. Emphasis is placed on monitoring because the effective management of air quality depends, above all, on the availability of regular and reliable information. Much has, of course, been accomplished over the last few decades in the improvement of air quality. A reduction in the emission of particulates has accompanied the shift away from the burning of coal as an energy source towards the use of cleaner fuels; the classic urban smog is now a thing of the past. Sulphur dioxide emissions, a major cause of acidification, are also coming under more effective control as power stations install technology to remove the sulphur produced during the generation process or switch to the use of more efficient and cleaner gas turbines.

But offsetting these improvements is the increase, resulting from the growth of road traffic, in pollution involving nitrogen oxides, carbon monoxide and hydrocarbons. This is now perhaps the most pressing air quality issue awaiting resolution. Not only does it damage the built environment, but it is also hazardous to human health: it is associated with diseases of the respiratory system, and contains carcinogens. However, the efficient working of modern societies is almost totally dependent upon the motor vehicle, and it seems unlikely that the problem of road traffic pollution can be solved in the foreseeable future. It is bizarre, to say the least, that governments should be attempting to curb car usage through piling on additional costs at the same time as they actively compete to encourage the multinationals to locate new vehicle manufacturing plants within their borders, in the interests of boosting employment opportunities and export earnings. One might be tempted to believe that this is prompted more by the need to raise general revenue than by a desire to improve environmental quality.

When excessive, noise, like air pollution, may fall into the category of environmental stressor. In the workplace it may become an issue because of its adverse effects on the health of the workforce. But of more concern here is the effect of noise in the broader environmental context. Obvious instances are the noise generated around airports and alongside arterial motorway routes, where the environmental impact results in general loss of amenity as expressed in terms of depressed property values. These matters, along with measures of noise control and the legislative framework regulating noise in the European Union, are discussed by Marc Van Overmeire and his co-authors.

Alfons Buekens' chapter is a comprehensive treatment of waste management. The aim of waste management is the safe disposal of waste without causing harm to the environment. This requires the separation of hazardous from non-hazardous waste and a thorough understanding of the consequences and implications of the various methods of disposal. For instance, landfill sites should be selected with regard to geological conditions so as to prevent the leachate infiltrating the groundwater table. Care is also needed when operating such a site so as to minimise the production of methane gas during the process of material breakdown. Similarly, care should be exercised in the location of incinerators and appropriate technology employed so as to avoid the release of toxic substances into the atmosphere. But attention is increasingly being redirected away from the mechanics of disposal and towards reducing the amount of waste produced by the use of economic and other instruments. Within this, recycling is an obvious and attractive strategy, although this needs to be thought through carefully: the overall amount of energy expended in implementing some of the most obvious candidates for recycling, such as glass and paper, may well be greater than in traditional methods of disposal.

1

SOIL MANAGEMENT

James G.Cruickshank

SUMMARY

Soil management is presented as a selective review of the practices and problems of managing certain elements of soil, such as the physical survival of soil, soil conservation, soil structure, soil water, soil organic matter and plant nutrients in soil.

Soil is introduced as a central part of the natural environment, and is itself a complex, sensitive and reactive system. Soil management is seen as the management of a vital environmental resource within the framework of the whole natural environment, and particularly within ecological and economic constraints for sustained agricultural production. The broadening of this concept of soil management into a soil protection policy is considered with reference to the European Soil Charter (1972) and the European Soil Protection Policy (1987). Europe is taken as the geographical context of this chapter, and the present as the time and technological context for this review of soil management.

The chapter starts with soil reclamation as exemplified by Dutch polders and Danish heathlands, but devotes more space to the physical conservation of soil against the erosive forces of wind and running water, giving examples from areas in Europe. The management of soil water examines the availability of water to plants, and demonstrates how drainage or irrigation may improve soil water status. Soil structure and soil organic matter are described and, more importantly, discussed as soil properties critical for the development and maintenance of a fertile soil for agricultural production. Chemical fertility is seen as the need to supply chemical nutrients in amounts close to the requirements of plants, but avoiding levels reaching a pollution state in soil. The contemporary chemical problems in Europe of soil acidification, salinisation and metal contamination in soil are considered briefly in conclusion. The review of soil management is selective in topics and restricted to European examples.

ACADEMIC OBJECTIVES

This chapter shows that soil management entails management of land for optimal and sustained agricultural production. It is possible to have soil management for other land uses, notably for horticulture and forestry. Soil management involves the adoption of certain management practices that are understood to be designed to improve the condition of the soil for agriculture, and to allow the land to sustain long-term production. Subsidiary aims include defining the differences between soil management and other, similar concepts such as soil protection and soil conservation; placing soil management in a context of time and space and showing that ecological and economic constraints limit practices and achievements of soil management for agriculture. To achieve these aims, it is proposed to show how selected management practices contribute to better management and so improve the condition of soil for crops and livestock enterprises.

CONCEPTS

Soil is central to the natural terrestrial environment. No part of the environment is more important than any other, either in economic or ecological terms, but soil has the most linkages to other parts of the total natural environment. The interaction of other elements of the environment at any one place is expressed in the character of the soil. In itself, soil is a very complex environment, highly variable over space because it is produced by the interaction of environmental factors at any one point in space.

Soil is a three-dimensional mantle of organic and inorganic material over most of the earth's land surface. It is a multi-function medium, providing a rooting place for the anchorage and growth of plants, a habitat for soil flora and fauna, an environment for the decay of organic litter, a reservoir and drain for soil water, a store and supply of plant nutrients, a sink and pathways for pollutants, a foundation for buildings and roads, as well as being a vital natural resource for agriculture. Soil functions as an open system of many parts. To modify one part of the soil system, possibly through management, may create changes in other parts of the system. All parts are interconnected and interactive as defined in an open system. Frequently, mistakes of bad management occur simply because of lack of understanding of the functioning of the soil system and, conversely, recovery may also be due to the resilience of its system character.

Management of soil for agriculture has the joint aims of growing crops for profit and of maintaining or even improving soil fertility. Both must be within the long-term constraint of sustaining agriculture for future generations. Good management can make constructive improvements to soil fertility, such as long-term drainage and irrigation works, maintaining levels of organic matter by introducing manure, building terraces and barriers for soil conservation, as well as short-term measures such as the use of artificial fertilisers, chemical herbicides and pesticides. All these management practices are merely manipulations of the natural soil system, and always and everywhere on earth, soil management is constrained by natural, ecological controls. But it is usually a concern for economy that prevents the

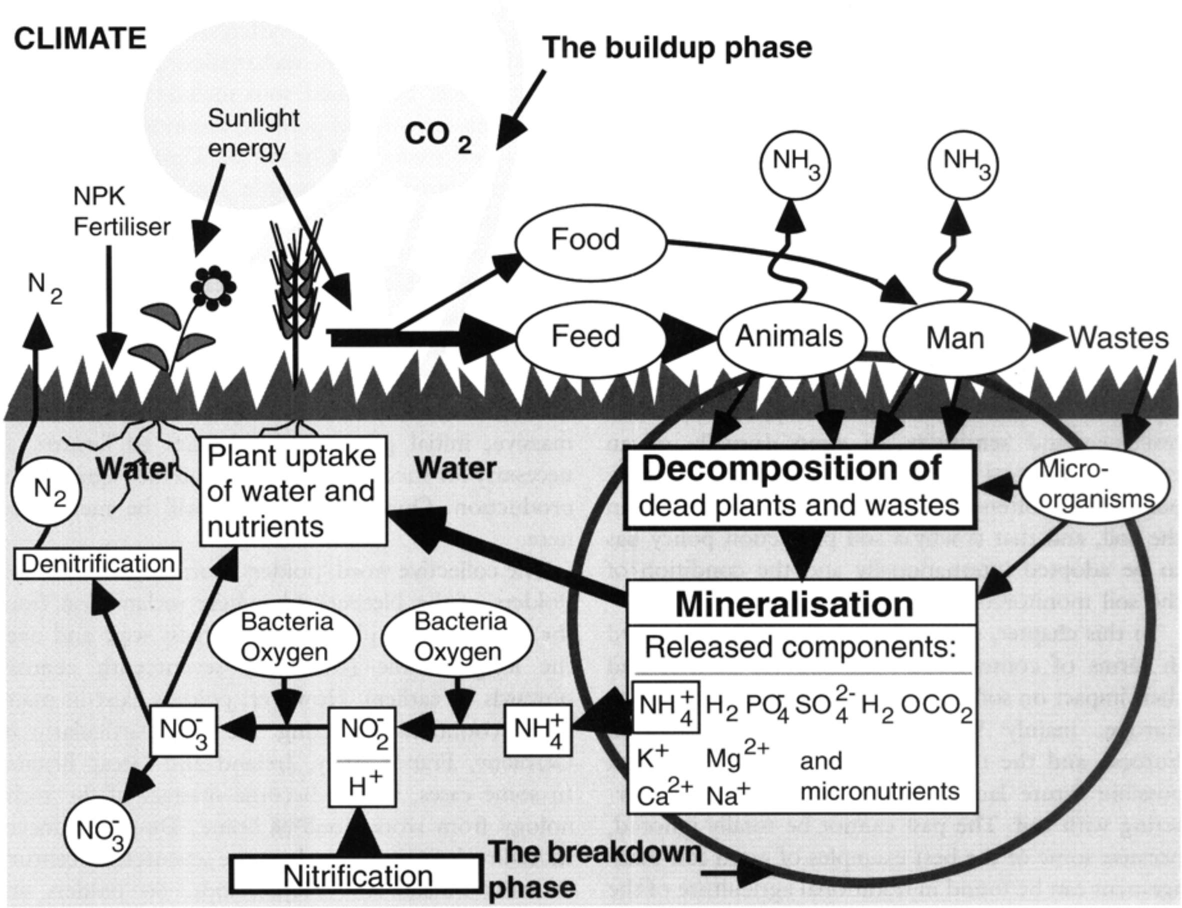


Figure 1.1 The soil system and the environment

creation of an artificial agricultural system (Davies *et al.*, 1993).

During the past century in Russia and the United States, and over the last 50 years elsewhere in the developed world, soils have been surveyed, sampled, analysed and mapped to provide a record of their character and distribution. In Europe, many countries have used soil maps as a guide to form secondary classifications of land quality for agriculture, and as guidance for potential land use. Soil maps have also been used in soil management to help advise on types of drainage, irrigation needs, risk of erosion and in particular soil properties that limit crop production (see Hodgson, 1991).

The concept of soil management has been enlarged in recent years to become a soil protection policy that recognises all the various users and managers of soil and the impacts they have. This is an extension beyond the concept of soil management for agriculture alone and also includes industrial and domestic activities that release pollutants which eventually enter and can damage the soil system. A European Soil Charter was adopted as long ago as 1972 and now includes elements that were developed in a soil protection policy by 1987 (see Barth and L'Hermite, 1987). Attention is still on the soil, but the condition of the soil is now evaluated in terms of quality for current and possible uses, soil degradation by erosion, acidification, salinisation and compaction, the soil's chemical buffering capacity, and the soil's resilience and sensitivity to stress (usually of an extreme, chemical nature). The source of damaging pollutants is often very distant from their impact in the soil, and that is why a soil protection policy has to be adopted internationally and the condition of the soil monitored frequently.

In this chapter, soil management will be discussed in terms of contemporary agricultural policies and their impact on soil. The geographical context will be Europe, mainly Western, Northern and Eastern Europe, and the time focus will be on present and possible future land uses and farm practices interacting with soil. The past cannot be totally ignored, because some of the best examples of good soil management can be found in traditional agriculture of the last two centuries, following scientific developments in the eighteenth century. The beneficial effects of grass were discovered and so a high proportion of grass was included in crop rotations. Nutrients and organic matter came from animal manures and weeds were controlled by cultivation. At the end of the twentieth century, large-unit monocultures prevail. Large machines are used on the land, and chemical solutions have been found for nutrient supply and the control of weeds. Management reflects a massive change in technology.

RECLAMATION OF SOIL

Reclaimed soils have been brought into agricultural production from a previously unused, unproductive state, or even from a non-existent state. A great variety of types of land are involved, with reclamation possibly their only common feature (Harris *et al.*, 1996). Likewise, a variety of soil management techniques are used for reclamation and subsequent maintenance. Reclaimed soils include land formerly below the sea, such as polders, mountain and upland areas (see Volume 3, Chapter 5, 'Upland and mountain environments'), unstable sand both as coastal sand dunes and inland heaths, marshland above sea level which requires extensive drainage, deserts, very steep slopes which have to be terraced and the land of industrial spoil or tips. Most of these are areas of large size and in hostile natural environments. The extreme environmental conditions mean that state intervention in the reclamation operation, through massive, initial grants and subsidies to farmers, is necessary for success in bringing land into agricultural production. Only two examples will be mentioned here.

The collective word 'polders' normally refers to the Polders of the Netherlands where reclamation from the sea has taken place on the largest scale and over the longest time (from the seventeenth century onwards or earlier). However, polders exist in many other countries bordering the sea, particularly in Germany, France, Italy, Ireland and Great Britain. In some cases, having become masters of the technology from work done at home, Dutch engineers started reclamation as early as the seventeenth century in areas outside the Netherlands. So polders are found extensively in the coastlands of Europe, but only where the submarine soil has been worth reclaiming. The soil must contain adequate clay so that soil structure can be developed and nutrient storage can be expected. Soils composed entirely of sand or silt would not be considered for reclamation because they lack the clay necessary for soil development.

The process of reclaiming land from the sea is long, costly and complex, and benefits to the human community through new agricultural production must be measured on a scale of decades and centuries. The process is undertaken in stages starting with enclosure of the proposed polder by building sea walls, pumping out the sea water, drainage into deep channels and continuing to pump out saline and later fresh water. The process continues by leaching out saline salts from the plant rooting zone, usually by applying calcium-rich salts like gypsum ($Ca SO_4 2H_2O$) so that the calcium ion will exchange with and displace sodium and magnesium. The natural levels of the latter are far too high for the tolerance of cultivated plants and must be reduced in a desalination process. Early in the plant reclamation stage the soil will be inoculated with bacterial micro-organisms and planted with salt-tolerant grasses so that soil structure will develop slowly. Thereafter there is a sequence of cultivated plants from lucerne and clover to peas and beans of the leguminous group, eventually to cereals and vegetable crops. The sequence is designed to use salt-tolerant plants and nitrogen-fixing (in their root system) plants early in building fertility in the new soil. Polders are not uniform, and even these general rules will be altered by local conditions, notably by the initial salinity levels of the soils.

Reclamation of the sandy glacial outwash plains of West Jutland in Denmark is the second largest single unit of reclamation in Europe, and one with very different soil management problems. Again, the state was involved—through the Danish Heath Society—when the work started in the late 1860s, following the territorial loss of the southern province of Jutland to the Prussian Empire. The work has involved a great deal of tree planting to stabilise the coastal sand dunes in the west as well as the planting of shelter belts of trees throughout the area. The sandy soil has produced many management problems, and irrigation has to be used on a large scale.

SOIL MANAGEMENT AND SOIL EROSION

The world-wide problem of soil erosion is one of great antiquity and complexity, but probably there has never been a time when this problem has been of such socio-economic and environmental importance as it is now. Although soil erosion on a disastrous scale is a product mainly of desertification in the tropics, it is also a problem in the agricultural lands of several continents, including Europe, where spectacular examples of soil erosion occur. The reason for a particular soil erosion problem depends on the regional or local interaction of physical-environmental and socioeconomic factors; causes of extreme examples of soil erosion can themselves differ extremely as, for example, in the vineyards of the Mediterranean or intensive cereal production in Great Britain. However, soil management factors do play a significant role in all examples of soil erosion on agricultural land, and have been placed in the overall context of the causes of soil erosion and exemplified in a comprehensive volume, *Soil Erosion on Agricultural Land* (Boardman *et al.*, 1990). This is the source for much of the information in this section.

Soil erosion is the physical removal of soil, either by wind or running water, and is therefore of fundamental importance to agriculture. Cultivated crops must have an adequate depth of soil for rooting and nutrition, so that satisfactory yields can be attained. In the context of Northern and Western Europe, soil depth must be at least 20 cm to give adequate yields; for example, 6 tons per hectare for winter wheat is financially profitable for the farmer and will ensure a sustained yield for the future. Ideally, soil thickness for cultivated plants should be more than 20 cm, and in places in England, where the soil commonly reaches a depth of 50 cm in thickness, a loss of 5 cm through soil erosion may reduce cereal yields by 0.5 tons per hectare. Crop yield can therefore be related to soil depth, though reliable quantitative examples are few, and rational soil management aims at conserving soil uniformly over a managed area. Where soil is being removed, or even redistributed by soil erosion, loss at a rate greater than the rate of soil development cannot be tolerated for very long. If agriculture is to have a future, soil conservation measures must be adopted to reduce the rate of soil erosion to an acceptable, normal rate, where removal is balanced by soil development.

Soil erosion is nearly always the result of an inappropriate choice of farming system for local environmental conditions, but can also arise from unsuitable management of an existing system by introducing new soil cultivation techniques. At the extremes of environmental risk for soil erosion, as found at the climatic limits for agriculture, it may become particularly difficult to find a farming system in harmony with the physical environment. Those that have survived over the centuries are usually grazing systems of low intensity, often associated with nomadic societies on desert margins or in mountain environments. Increase of human and animal populations has frequently led to overgrazing, soil surface exposure and soil erosion. In Europe, there are only a few examples of this kind of soil erosion problem, and these are found in semi-arid parts of the Mediterranean. It is in that region also that intensive systems of vineyards, cereals and olive groves have been a traditional way of farming; they have survived the high risk of erosion in the past because of high human inputs in building and maintaining terraces and irrigation schemes. When the latter are neglected or abandoned, the aftermath of such a highly artificial agriculture has been rapid soil erosion, a pattern that has occurred in the Mediterranean lands in several periods from the end of the Roman Empire to the present century.

The mechanics of soil erosion take place in two stages: the first is the separation of individual particles from the soil mass, the second is their transport by erosive agents such as running water or wind until the lack of energy forces deposition of the particles. Exposed soil surfaces are the most vulnerable, and at the start of the erosion process the impact of raindrops detaches individual soil particles. The effect is increased by the length of time the soil surface is exposed to intense rainfall and is also increased where the soil mass is weakened by wetting and drying, freezing and thawing, and by man-management through tillage operation, ploughing, harrowing and rolling the soil surface. The soil is made more vulnerable if these tillage operations are badly timed and take place when the soil is saturated or very wet. Soil in a weak or unstable state is unlikely to be able to resist the detaching impact of intense rainfall.

In a similar way, wind can act as a detaching agent, particularly on dry, exposed soil surfaces where individual particles are loose, small and lightweight. Most vulnerable are organic particles from dried-out peat and mineral particles in the fine sand range (0.1– 0.2 mm diameter). Transportation by wind is by repeated bouncing close to the ground (called saltation) of the larger particles, as well as the blowing of smaller silt and clay sized particles at high levels (which eventually form loess deposits). Transportation by water may either produce fairly uniform sheetlike movement by overland flow or, if more severe, will tend to be concentrated into rills, gullies and rivers. Technical details about soil erosion and conservation are available in many specialist publications (see Morgan, 1986), but the main task here is to assess the role of soil management in the complex of factors that may initiate and accelerate soil erosion processes.

A complex of four groups of factors controls soil erosion: (1) energy or crosivity of the eroding agent (wind or water); (2) the resistance of the soil to both detachment and transportation of particles (which includes aspects of soil texture, aggregate stability, shear strength, infiltration capacity, soil moisture status, organic matter content and certain chemical properties); (3) slope steepness and slope length (increasing each increases soil erosion by water); and lastly (4) the effect of plant cover (both natural/semi-natural vegetation and agricultural crops) in intercepting the eroding agent and protecting the soil surface. Soil management practices can affect the soil's resistance to erosion. In some cases management can also alter slope angle and length, but the most common practice is to provide plant cover protection for differing periods of the year.

Outside the Mediterranean region, soil erosion problems in Europe are associated (mainly) with soils of weak aggregation and unstable structure such as the sandy plains of Iceland, west Denmark and north Germany, as well as in western Hungary (see Box 1) and eastern England. In these areas, wind erosion is the main hazard for sand and certain peat soils on fairly level ground. Slope angle is not a factor, and climatically these wind erosion risk areas range from

BOX 1.1 HUNGARY: A COUNTRY WITH MAJOR SOIL MANAGEMENT PROBLEMS

György Vàrallyay (1993) has written that 'soils represent a considerable resource for Hungary, and consequently their rational utilisation has particular importance for the economy'.

Hungary is a relatively small country of just over 91,000 km^2 land area, and despite the agricultural value of its soils, three major problems of soil degradation (discussed in this chapter) are found within its boundaries. These are soil erosion, soil acidification and salinity-alkalinity, and the most severely affected soils are found in three separate regions of the country. Soil erosion affects 15.6 per cent of all soils to a moderate or severe degree in the west of the country in a region centred on Lake Balaton and lying west of the River Danube. Soil acidification is on the increase in Hungary, and of the affected soils, those soils with values of less than pH 5.0 increased from 19 to 24 per cent from 1980 to 1985. Strongly acidic soils cover 13 per cent of Hungary, and a further 14 per cent are considered highly susceptible because of their low buffering capacity. The areas worst affected by acidification are in the mountains near the north-east border of Hungary. Lastly, salt-affected soils, produced by the process of salinisation and alkalisation, are found mainly in the Great Hungarian Plain, east of the Danube, and cover at least 10 per cent of the country. In this area, annual potential evapo-transpiration exceeds annual precipitation by 200–250 mm, or 70–75 mm in each summer month, and as it is an enclosed basin, groundwater becomes stagnant and salty.

With approximately 50 per cent of soils in Hungary suffering quite severely from the effects of these three types of soil degradation, good soil management practice is extremely important. It has been noted that agricultural yields in Hungary have increased dramatically at considerable financial and environment cost, with the application of fertilisers. It is considered that soil acidification, which can be managed by controlled liming, is caused much more by the high rate of mineralisation of nitrogen-based fertilisers than by acid deposition from the atmosphere. The difference is of the order of five times greater from mineralisation, although occasional high acid deposition may come in acid-stress episodes.

Management measures for the control of soil erosion are discussed elsewhere in this chapter, but the study of salinisation-alkalisation processes and their control have a long tradition in Hungary. The prevention of salt accumulation requires the use of good quality irrigation water on to the soil surface, and groundwater is kept low by horizontal drainage techniques. If the process is well established, sodium Na^+ ions should be replaced by calcium Ca^{++} ions, which in turn are more easily leached out of the soil by acid solutions. In both situations, the control of groundwater by horizontal drainage is essential to prevent salinisation.

cool, humid to hot, semi-arid summers. Wind force may be more important on the north-west coastlands of Europe, and drying out of the surface may make a greater contribution in hot continental interior regions. The main areas of water erosion risk are the upland margins of Central Europe, areas affected by snowmelt in southern Scandinavia, and a variety of lowland sites of both undulating and level ground in Germany, Belgium, northern France and eastern Britain. In these areas, a combination of soil properties (mainly fine sand loams and silt loams with weak structure) and arable cropping produce the erosion problem.

Soil management for conservation starts with the choice of crop, which is often based on economic rather than ecological considerations. Grass, permanent pasture more than rotation grass, is by far the best crop for soil conservation. Grass protects the soil surface better than any other crop, and the organic waste from leaves and root systems builds soil organic matter, so developing a strong and stable soil structure in the cultivated top soil. Among arable crops, spring sown cereals are the next best because the protective stubble from the previous year remains on the surface over winter. The remaining arable crops are of decreasing value in soil protection, particularly in the high risk winter season when autumn sown cereal crops provide no cover. In Western Europe, there has been an increase in winter wheat and winter barley due to EU policy (in Great Britain winter cereals increased threefold from 1970 to 1983), and particularly in the period 1975 to 1985, during which time there was an associated increase in soil erosion by winter rains. Root crops like sugar beet and potatoes likewise give no protection during winter, while vegetables and fruits, which are row crops/row-planted, may encourage water erosion because of the planting arrangement. The combination of unprotective crops and unstable soils further increases the risk of erosion.

In a comprehensive review (Quine and Walling, 1991), a large number of soil erosion experiments on arable land in Britain have been assessed, including some using a caesium-137 (^{137}Cs) technique which can give a retrospective measure of erosion over the last 30 years, as well as many other measurements by conventional techniques. The results were in general agreement and showed that about 40 per cent of arable land in England and Wales had been affected by soil erosion to some degree, although in a majority of cases erosion was restricted to less than 10 per cent of the area of the fields studied. Most importantly, it showed the expected relationship between soil erosion rates and soil texture, ranging upward from 0.6 t/ha/year on clay soils, to 2.0–4.3 t/ha/year on silty soils, and finally to the highest rates of 3–10.5 t/ha/year on sandy soils. Only rates of less than 1 t/ha/year are considered tolerable and within the range of being balanced by soil supply or renewal. Although these measurements refer only to parts of fields and are not particularly serious in the short term, there is a growing concern about the long-term productivity of arable land (particularly if used for winter cereals) in Great Britain and other parts of Western Europe. The exact causes of the erosion may be difficult to isolate and evaluate, but continuous cereal cropping has been shown to result in dramatic falls in soil organic matter, so that after a period of about 50 years, the organic matter content will have fallen to a level too low to sustain stable soil structure. Arable cultivation techniques in recent years have been changing to allow the use of larger machinery in larger fields. This often leads to problems of soil compaction at plough depth and capping or crusting of the soil surface, which increases runoff by overland flow. In turn, this leads to higher water erosion rates, which can also be encouraged by ‘tramlines’ downslope created by the tracks of farm vehicles, which function as gullies for soil erosion.

Soil conservation measures are seldom present at farm management level in Western Europe, and very few countries have a state policy on soil conservation. Care should be taken to maintain soil organic matter between 4 and 8 per cent weight of the Ap or ploughed surface horizon. The best form of input is farmyard manure, as decomposition is relatively slow and the beneficial effect long-lasting. The value of organic matter is enhanced by base minerals such as calcium and magnesium, and also by iron, which bond humus and clay into stable structural aggregates. There follow from this the further advantages of crumb or granular structures which increase soil pore space, and lead to increased water-holding capacity and infiltration rates. The overall effect is to minimise surface water erosion rates and improve the fertility of the soil for cultivated crops. Conventional tillage operations of ploughing, harrowing, rolling and planting will not damage stable and fertile soils, but can cause problems on fine sand or silty soils especially when dry, on dispersed saline soils, and on sticky, swelling clay soils (with moisture beyond the plastic limit). On such problem soils, it is recommended that some form of minimal tillage, or conservation tillage involving surface organic mulch, be practised. One option is to adopt direct drilling of seed into the stubble of the previous crop, and to control weeds by chemicals. Another is to till alternating strips contoured around the topography, leaving the intervening strips unfilled, or possibly direct drilled. Minimal disturbance of the soil is good conservation practice.

Soil erosion has to be of catastrophic proportions before there is state intervention and the formulation of a state conservation policy. The Dust Bowl in the southern USA in the 1930s is the classic case, although American conservationists had warned of this catastrophe for the previous half century. In Iceland, it took several decades of wind erosion of the sandy plains before legislation was first introduced in 1907 culminating in a Soil Conservation Act in 1965. Here the policy includes aerial fertiliser application and aerial seeding, as well as strict controls through incentives to reduce sheep numbers to within the grazing capacity of the land. In Denmark, the woodland clearance and subsequent reclamation of the sandy heathlands of West Jutland have included planting and renewal of shelter belt forestry for more than the last 100 years. Elsewhere in Europe, conservation policies exist for regions within states, but conservation is a national policy only in Germany and Hungary (Hinrichsen and Enyedi, 1990). Hungary is affected both by water erosion in its western and northern hill and mountain lands and by wind erosion on the sand and peaty soils in the eastern plains (see Box 1.1). Legislation for soil conservation at national, regional and farm levels was unified in the Land Codex Act in 1987 (see Boardman *et al.*, 1990).

SOIL WATER

Soil water is the most vital element in agriculture. A supply of water is essential for plant growth, nutrition and support. Plants wilt and die from lack of water, but there are practical problems in providing plant water in an available form at times of peak demand. This demand is greatest during a plant's growing season, which coincides with the warm period of the year when soil water losses through evaporation and transpiration usually exceed the inputs from rainfall. The supply of water to the plant depends on the store of water in the soil and on its availability in a form that can be attracted into the plant root system. The total amount of water is not so important, but it is critical that the water be in an available (capillary water) state as skins or films of moisture held loosely on the surfaces of soil particles and soil aggregates. Lying between the gravitational or drainage water that moves freely through the pore space of the soil and tightly held hygroscopic water, capillary water is held by surface tension at between 1 and 15 times atmospheric pressure. Plants can exert a suction pressure of up to 15 bar (or 15 atmospheres) to attract water into the root cells by what is known as capillary action. Diffusion into the root cells is by the process of osmosis, whereby water moves through a permeable membrane from lower to higher dissolved salt concentrations.

After water has drained freely through the soil, and the gravitational or drainage water has been removed, the water remaining in the soil at field capacity (FC) is mainly capillary water available to plants. The importance of soil drainage as part of soil management for agriculture is obvious from that fact alone. Soil must be drained of excess water to allow the capillary water to be available to plants—in addition to the other advantages of soil drainage discussed below. During the warm or hot season, capillary movement takes this water upward to the drying ground surface, where moisture is lost through evaporation. The available water capacity (AWC) of the soil is reduced by both evaporation and transpiration through plants, and unless the supply of soil water is replenished by rainfall or irrigation, the permanent wilting point (PWP) will be reached, that is the point beyond which soil water (now hygroscopic water) is held too tightly (by more than 15 atmospheres pressure) to be attracted into plant roots.

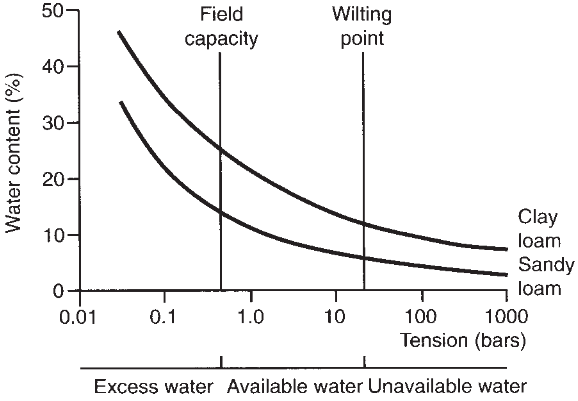


Figure *1.2a* The relationship between soil texture, water content on mass and water tension
Source: Simpson, 1983

The available water capacity (AWC) of a soil is a very important part of its fertility. AWC is linked with the average rainfall and evaporation at the site, while the soil factors affecting AWC are soil texture, structure, depth and drainage. Sands and loamy sands hold the least water at AWC because of their small surface areas per volumetric unit of soil. Medium AWC is found in sandy loams, silt loams and sandy clay loams, while a high AWC is associated with clays and highly organic soils. The reason for the range of AWC in different textures is connected with the internal surface area over which films of capillary water are stretched. Similarly, soil structure affects AWC through the variation of internal spaces or pores in which capillary water can be stored and moved. Platy structures in compacted soil have the least AWC, and open crumb or granular structures the most. Soils with low AWC are most at risk from drought, particularly in dry climates.

Plants require between 200 and 500 units of water for every unit of dry matter they produce. A cereal

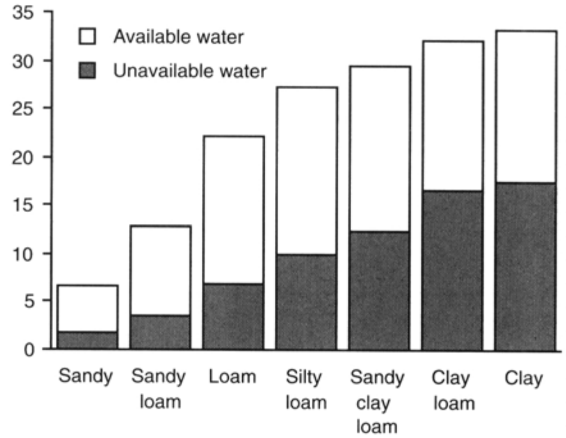


Figure 1.2b Soil texture, available water capacity and unavailable water capacity
Source: Simpson, 1983

crop yielding 10 t/ha of dry matter requires 2000–5000 t/ha of water, or the rainfall equivalent of 200–500 mm after drainage, runoff and evaporation losses. Such amounts of water are required within the growing season for optimum yield, and even where it is provided, soils with low AWC will lower yields of sugar beet and potatoes to about one-third of the potential with high AWC, and of cereals by about one-half. Irrigation can make up the difference.

SOIL MANAGEMENT, DRAINAGE AND IRRIGATION

Saturated or waterlogged soil cannot be tolerated by any kind of agricultural enterprise for prolonged periods. In a permanent state, waterlogged soils are not compatible with productive agriculture, and only a low output extensive grazing system would generally be possible. Any agricultural system designed to be productive and self-sustaining requires favourable physical conditions, and this includes an adequate supply of soil moisture in a form attractive to plant roots. It implies also that there is an adequate amount of pore space for rooting, and that there is a suitable balance between moisture and air (oxygen) in that pore space. It may be necessary either to introduce water by irrigation or to remove water by drainage to maintain that moisture-oxygen balance, since only a small minority of soils can achieve it naturally. These are freely draining soils in a moist, cool climate; otherwise the desired balance has to be brought about and maintained by management of the soil (Keane, 1986).

The benefits of soil drainage to any agricultural system are many and various, but are often interconnected. It is essential to have an adequate water supply for crop growth and transpiration to keep plants in a 'turgid' state, and so able to function at an optimal level. It is also essential to have a supply of moisture in a capillary form that enables the movement of dissolved nutrients and oxygen into the root system through a diffusion process. In wet climates it is very often necessary to remove excess water from wet soils as rapidly as possible by some type of installed drainage. Although cultivated plants vary in their capacity to tolerate waterlogging for short periods, species vary in that tolerance through their life cycle. For almost all crops the most vulnerable time is at or after sowing the seed, at germination and emergence. At later stages, cereals are thought to be the most tolerant of crops and can survive a month of waterlogging with only about 10–15 per cent loss of yield. At the opposite extreme, drought can be even more damaging, and if prolonged will destroy most crops, but again there is a variety of tolerance levels. In all forms of arable agriculture, it is desirable to maintain a supply of capillary moisture into the root systems.

The main problems with saturated or waterlogged soils are that they are low in oxygen (even though there is dissolved oxygen in soil water), and that the soil is colder and takes longer to warm up at the start of the growing season. Germination is retarded and some of the subsequent growing season may be lost. Most soil organisms require aerobic soil conditions for optimal functioning, so processes like the decay of organic matter and nitrification are inhibited. Wet soils encourage a number of plant root diseases, and encourage competing weeds to dominate and smother the cultivated plants. Nor do wet soils respond as well in passing nutrients to plants from applied fertiliser, especially when nitrogen fertiliser is applied, because denitrification is increased. Consequently, soil fertility is impaired, which leads to lower crop yields. So poor soil drainage has a multiple and interacting effect on soil fertility and crop yield. Likewise, poor drainage reduces the physical stability of soil structure with the result that the trampling of livestock or the weight of machinery on wet soil will cause at least temporary damage to soil structure and pore space, leading in turn to a loss of soil fertility (Henderson and Farr, 1992).

It is clear that farmers have been aware of the general benefit of soil drainage since the beginnings of agriculture in Western and Northern Europe. In early plant agriculture in wet climates, adequate drainage was achieved first of all by avoiding clay soils, and then by digging cultivation ridges for plants and allowing the shallow ditches or furrows between them to act as open drains. The spectacular examples of drainage for soil reclamation in the Dutch Polders, the Fens in England and marshlands in Germany, Poland and Italy were undertaken by Dutch engineers much later, in the sixteenth, seventeenth and eighteenth centuries. The first field drainage in Europe was an integral part of the Agricultural Revolution in the late eighteenth century, when drainage improvement was part of field enclosure. Drainage at this time involved open ditches around fields to accept the outfall from under field drainage (that is drains made of wood, sods or stones placed in the subsoil) and then pass the excess water into the streams and rivers of the country's arterial drainage.

In Europe over the last 200 years a variety of types of field drains have evolved, from the early wood and stone drains, through ceramic or clayware tile drains, and on to perforated plastic pipes and mole drains created by moling machines in clay soils. The variety also extends to special drains for particular terrain and soil types so that today almost any mineral or organic soil can be artificially drained. Improvement of drainage in the most sticky clays has been possible only in the last 25 or 30 years, but by now it is probable that over 70 per cent of agricultural land in all the northern and western countries has been improved by field underdrainage, from Ireland and Britain to Scandinavia, Germany, Poland and Hungary. It is impossible to say exactly how much, and published figures should not necessarily be believed because no records were kept of installed drains prior to about 1930. In any case, it is not certain what proportion of early drains are still functioning. In Europe, the situation is that all the land that can be sensibly and profitably drained has been drained. A limit has been reached for agricultural advantage, notwithstanding the 'conservation argument' that drainage should stop in the interests of preserving most of the remaining wetland habitats.

Irrigation, the introduction of water into the soil to raise available water capacity, is a technique complementary to drainage. Irrigation is not necessary in the wettest climates of Western and Northern Europe, but is still necessary for anticipated crop yields during the growing season even in parts of eastern Britain, southern Scandinavia and the North German plain. Intensive agriculture in Central and Eastern Europe depends very greatly on irrigation. But irrigation brings a number of its own problems which require particular care and awareness in soil management. In any of the regions where irrigation is necessary, sandy textured soils are the first type to require it, as they are least able to retain soil moisture in the absence of rainfall. Normally, there are few problems in the application and infiltration of water into sandy soils by sprinkler irrigation. However, great care should be taken in the application of water to a bare surface where soil structure is weak and unstable, or to clay soils if the clay content is more than 35 per cent the weight of fine earth. In clay soils, irrigation brings doubtful advantages combined with a high risk of waterlogging and reduction in crop yield. Machinery damage to soil structure may also result.

The most common cause of crop damage and loss through irrigation is the salinity of the irrigation water. In the coastal areas of Northern Europe, salinity can result from using bore hole water which may be high in natural salts from contamination with sea water or recycled river water. In continental interior situations, such as the Great Hungarian Plain, the salinity of the groundwater is derived from geologically salt-rich strata, the enclosed formation of the region, poor outward drainage and salty groundwater rising to the ground surface during rainy periods. The control of salinisation in this situation is dependent on applying good quality irrigation water to the ground surface to encourage leaching and using calcium ions to replace (or exchange with) sodium ions (see Box 1.1).

MANAGEMENT AND THE IMPORTANCE OF SOIL STRUCTURE

Soil structure is like an internal 'scaffolding': it is the product of individual soil particles sticking together in the process of aggregation, in which the 'glue' may be humus, clay, or compounds of calcium or iron. Small aggregates adhere together to make larger aggregates which become the structural units. The fact that the surface horizon of the soil can contain up to 60 per cent void space or pore space implies that the 40 per cent solid material must have a structure or scaffolding to keep it in place. The most desirable type of soil structure for plants is one that is stable and has maximum internal pore space for plant rooting and supply of moisture, nutrients and oxygen to plant roots. This is most likely to be found in the organic-rich surface horizon where humus helps to build structure and where most plant roots are found, though rooting may extend to over 1 metre. It is part of good management that this natural condition of the soil be maintained in agricultural systems.

There are four main shapes of structural unit, as displayed in Figure 1.3, with local conditions within the soil influencing which of these shapes predominates. A columnar or prismatic shape is associated with the wetting and drying of subsoil clay, a horizontal platy structure is a product of local compaction in the subsoil, blocky or cubic shapes are associated with loamy soil, and crumb or granular structures are found in organic-rich surface horizons. Structural stability is a very important attribute, and becomes increasingly so under the physical pressures of modern farming practices. Stability is usually measured by the soil's aggregate ability to survive immersion in water, as the weakest structures, developed by wetting and drying or freezing and thawing, are often destroyed or dispersed by such treatment. More stable structures survive because the bonding of the particles has come from humus, clay or iron and calcium compounds. Stable structures in the topsoil also protect against erosion.

The building of soil structure is a secondary process of soil development and is restricted to the soil horizons above the parent material. Farmers, through soil management practices, have some control over the building, maintenance and possible destruction of soil structure, but only in the surface horizon— the horizon most affected by tillage practices or the lack of them. Little can be done to alter or improve the structure of the middle horizons, normally called the subsoil, except to shatter compacted horizons with a subsoiler. Clay subsoil, especially when wet, is a permanent problem for structure and for drainage.

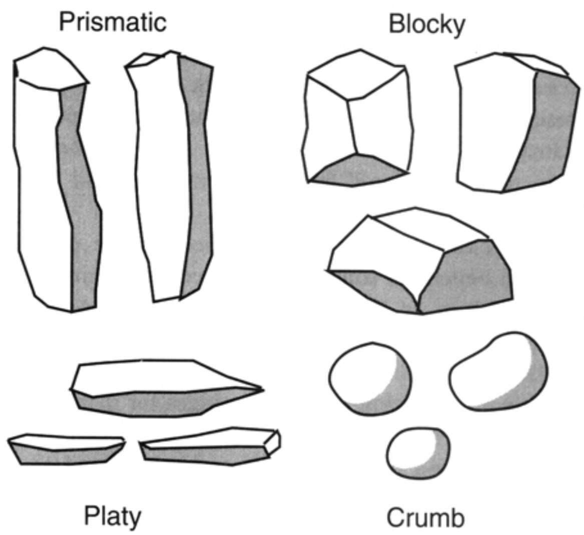


Figure 1.3 Soil structure is divided into four categories, in addition to the structureless condition. Structural units (peds) are three-dimensional. Prismatic or columnar structure (vertical dimension longest) is associated with clay-rich soil. Blocky or cubic structure, of almost equal dimensions, is associated with loamy soil. Platy structure (horizontal dimension longest) is associated with compacted or cemented horizons. Crumb structure (spheroidal shape) is found in organo-mineral soil usually with a loam texture.

The size, shape and therefore the packing of structure units affect internal pore space and capillary channels in the soil. Structural units start at about 2 mm in diameter, just above the size of large sand grains. Loose, coarse sand is not regarded as having structure, unless called single-grain structure. Crumb and blocky shapes are usually between 20 and 50 mm in dimension, and normally have the most internal pore space, at about 35–55 per cent of total volume. Microscopic root channels are also found passing through structure units, as well as within the meso- and macro-pores around them. The largest structural units are the columnar or prismatic units which may be up to 50 cm in vertical dimension. Platy units are much smaller, 5–10 mm, but because of the interlocking of the units in compacted horizons, they contain the minimum of pore space (about 20 per cent of volume) and are not useful to plant roots. Good management should aim to prevent such platy structure developing within the rooting zone of cultivated soil.

A stable, crumb or blocky, soil structure is the basis of the physical fertility of the soil. It reveals the great value of the presence of the bonding agents, particularly organic matter and lime, and shows that soil structure survives best where there is minimum disturbance. Reduction of organic matter levels and long periods of soil exposure to the impact of rain drops, conditions associated with arable farming, act against the maintenance of stable structure. Wet and waterlogged soils are the most vulnerable to structural damage or collapse. Particles can be separated or dispersed from each other (slaked), and with the impact of rain can be washed away to form impervious surface capping layers. The impact of vehicles or machinery on wet and weakened soil structure can cause slaking and compaction of soil particles. Similar damage, called ‘poaching’ or ‘puddling’, is created by excessive trampling of wet soil. Ploughing a soil when wet can cause compacted ‘plough pans’ below the plough depth, and the preparation of a seed bed, by breaking large units into a fine structure, is a skill of good management also dependent on soil moisture status and weather conditions at the time. Minimal disturbance is the ideal treatment, and includes a wide range of options from grassland farming to direct drilling and minimal cultivation in arable systems.

Soil structure is extremely important for soil fertility and maximum crop production. It also affects properties of a fertile soil such as rooting space, water and nutrient supply to plants, and the very conservation of the soil itself. However, if structural damage is inflicted, it is usually only temporary. Management, through increasing organic inputs, liming and encouraging the activity of soil organisms, can rebuild soil structure over time. It is also in the interests of management to break up any undesirable compacted platy structure by subsoiling.

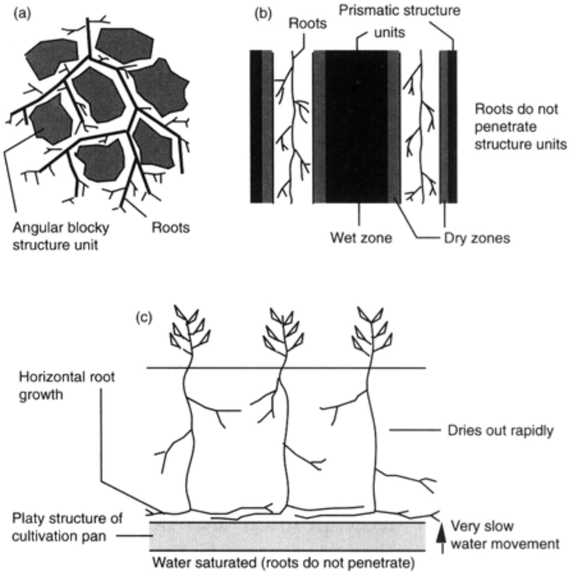


Figure 1.4 Root penetration and water availability

THE ROLE OF SOIL ORGANIC MATTER

Soil organic matter is the key to soil fertility, in that its beneficial influence affects several soil properties which are measures of fertility. In particular, organic matter contributes to structure-building, water-holding capacity, plant nutrition and to the cultivation of the soil. It is in the interests of good soil management to maintain the level of soil organic matter in the ideal range of between 4 and 8 per cent of total sample weight to maintain soil fertility in agricultural systems. In arable farming this is difficult, as deficiencies in the level of organic matter are caused by the removal of crop material, but in grassland systems there is usually no problem in maintaining at least 8 per cent by weight. Because of the very low density of organic matter, 8 per cent by weight represents about 20 per cent by volume. This is clearly visible by its dark brown colour and feels soft and smooth to the touch.

The desirable form of soil organic matter is not raw vegetable litter, whether in the form of fresh straw, wood chips or animal manure, but rather is well decomposed, soft, amorphous and rich in residual humus. This humus is both the continuous and residual product of organic decomposition and involves the repeated digestion and excretion of organic materials by populations of macro-organisms (e.g. beetles, mites and earthworms) in the early stages of decay of raw organic litter and later on by populations of micro-organisms (e.g. fungi and bacteria). The former will attack almost any type of organic material, including tough lignin, oils and waxes, but micro-organisms thrive best on sugars, starches, proteins and cellulose. The speed and effectiveness of the decomposition process depends on having favourable conditions for the activity of the macro- and micro-organisms. Ideally, the soil should be moist but not wet, well aerated with plenty of oxygen available, alkaline or very slightly acid (pH 5.5–7.0), with a temperature above 6°C, similar to plant growth temperature.

The rate of decomposition also depends on the composition of incoming organic litter, particularly the ratio of carbon/nitrogen (C/N) in the material. Lowest ratios of around 15–20 occur in young cereal and grass leaves, slurry and some root materials, while the highest ratios of around 100 are found in straw, sawdust, woody material and leaf needles from conifer trees. To digest material with high C/N ratios, micro-organisms have to draw on nitrogen from elsewhere in the soil system and this represents a loss from the nitrogen cycle. In such circumstances, extra nitrogen fertiliser inputs may be necessary. In any event, high C/N material is much slower to decompose and will take months or over a year. One way of viewing decomposition is to remember that an indefinite equilibrium is reached when the organic litter has been changed to constituent mineral elements and humus with a C/N ratio of about 10. Approximately one month may be required to decompose organic litter for every multiple of 10 that the original C/N is above 10.

Humus has many of the advantages of clay, and few of its disadvantages, in the beneficial effect it has on other soil properties. Its binding, cohesive character is best demonstrated on soils with a sandy texture. It should also be noticed that it is possible to introduce organic matter, but not clay, in field management. Humus is important for its role in promoting structure in surface horizons, where the result is usually a stable structure with a crumb or granular shape enclosing a large proportion of pore space. The humus-mineral mixture also makes the surface soil relatively easy to cultivate, avoids the problems of plasticity in wet clay soils, and gives the surface some protection from erosion. Humus also gives soil increased water-holding capacity without becoming sticky or plastic and it is a store of certain plant nutrients like nitrogen, phosphorus and sulphur. Through careful management it is possible to build up the organic matter of the soil by keeping crop cover at a maximum, including as much grass-clover cropping as possible, importing manures into the soil and, most importantly, maintaining favourable conditions of aeration, moisture and alkalinity to encourage the activity of macro- and micro-organisms in the decomposition processes.

SOIL MANAGEMENT AND PLANT NUTRITION

Soil is the main source of essential nutrients for plants. Carbon, in the form of carbon dioxide, is exceptional in being absorbed through the leaves in the process of photosynthesis. But some 21 major and micro (trace) elements are assimilated from soil, in a dissolved state as soil water enters the plant root system, or by ionic exchange between the plant root and soil materials. All the major elements, with the exception of potassium, are components of the fabric of the plant, and are returned to the soil as organic litter or green manures. Peat, which is partially decomposed organic litter, contains reserves of nitrogen, while unweathered rock may contain large amounts of calcium and potassium, but in both cases the elements are unavailable to plants. The two most important points about the supply of nutrients from soil to plant are, first, that each nutrient element is individually required and compensation by one another is not possible, and, second, that it is not the total amount of each element that is critical, but the amount in a form available to plants. For most but not all nutrient elements, the optimum chemical status for nutrient availability to plants is between pH 6.0 and pH 7.0. Even for those, particularly trace, elements with different optimal pH ranges, the range 6.0–7.0 still allows adequate intake. This is why farmers aim to manage soil pH at that level. However, there are many other environmental and physical conditions affecting the inputs, throughflow and losses of nutrients from the soil, which indirectly affect their availability to plants. For the optimal nutrient benefit to plants, especially the expensive application of artificial fertilisers, soil management must aim to minimise losses and increase availability of the essential elements between soil and plant (Bøckman *et al.,* 1990).

Because of the independence of all plant nutrients, no single element can be said to be more important than any other. However, since some of the major elements (N, P, K and Ca) are required in large amounts, particular attention should be given to the more efficient use of these in fertiliser application. The most expensive and difficult to manufacture into fertilisers is the N, P, K group, and because these elements also provide collectively the greatest response in growing plants and the greatest effect on crop yields, special consideration should be given to their cycling in nutrient management (Addiscott *et al.,* 1991). This is not to relegate the rest to a less important status, as all are equally important, but the inefficient and ineffective use of N, P, K may also cause serious environmental damage to soil, drainage water, rivers and lakes, by the 'fertilising' effect (or eutrophication) on these water systems (see Box 1.2).

Nitrogen (N)

Nitrogen is the dominant nutrient element in terms of the quantity absorbed and the growth and yield response it creates in cultivated plants. It is also the nutrient that is most often deficient in soil, but is difficult to apply in fertiliser in the right quantity and form. There are several ways of losing nitrogen from

BOX 1.2 EXCESSIVE LEVELS OF PHOSPHORUS AND NITROGEN IN AGRICULTURAL SOILS, PARTICULARLY IN IRELAND

Phosphorus and nitrogen are critical nutrients in agricultural productivity, but if applied in excess, may lead to losses from the agricultural system and consequent pollution of rivers and lakes. The emphasis in the European Union (EU) in the 1990s on sustainable agriculture as an essential part of environmental conservation means that more attention must be devoted to monitoring the sinks and fluxes of phosphorus and nitrogen. This has already happened in Scandinavian countries, but there is a strong case for extending control of phosphorus and nitrogen to all countries in Europe.

The case of Ireland (Culleton *et al.,* 1994) has been described as possibly the worst in Europe, but the phosphorus problem may be as serious in Belgium, Germany, France, Italy and the Netherlands, where higher levels of phosphorus fertiliser application were found in the 1980s. The problem in Ireland is made worse by unknown and unmeasured inputs of phosphorus in pig and poultry slurry. What is known is that over the past half-century in Ireland phosphorus in chemical fertiliser has been applied at the rate of 10 kg P/ha/year. Outside mountain areas, almost all agricultural land in Ireland has phosphorus levels above 6 mg/kg and 26 per cent of Irish soils have P levels above 10 mg P/kg soil, which is high enough to allow optimum crop production for a number of years without further fertiliser application. Levels of phosphorus much lower than this can cause eutrophication and pollution in rivers and lakes; lake pollution is now common in areas of intensive agricultural production in Ireland. The average for European countries is 15 kg P/ha/year, which is three times higher than the average for the USA, and shows how much excessive phosphorus in soil is a European problem. Tunney (1990) has shown that in Ireland inputs of phosphorus are at least twice as great as the outputs from the agricultural system.

Nitrogen is the nutrient required in the greatest quantity for healthy plant growth, and the volume of agricultural production is closely related to the use of applied nitrogen fertiliser (see main text). Sherwood and Tunney (1991) have shown that in Ireland total output or recovery of nitrogen is only 16 per cent of input, and even if about 12 per cent may be immobilised in the soil, 72 per cent is lost to water and the atmosphere. This is both a financial loss to agriculture and potentially damaging to the environment. Nitrogen fertilisers are the most expensive of all fertilisers, both in energy and finance production costs, so there is a strong economic case for more efficient use. Leaching losses are greatest from arable and fallow systems, and here conservation alternatives are most difficult to find. In grassland systems biological fixation by clover offers a small improvement in the efficiency of nitrogen use, but much research remains to be done to understand the circulation of nitrogen.

Source: Culleton *et al.,* 1994

circulation (by leaching in solution, by volatilisation as ammonia gas, and through denitrification by bacteria changing nitrate into nitrogen gas), all of which are undesirable for environmental and financial reasons. The main flow lines of the nitrogen cycle are shown in Figures 1.5 and 1.6.

Nitrogen is vital to plants as a constituent of protein. It is an abundant gas in the atmosphere, but is not available to plants in that form. It enters the soil as part of plant litter or animal manure, but has to be converted by bacteria to ammonium (NH_4^+) (termed mineralisation), and to nitrate (NO_3^-) to be available to plants. The second part of this process is termed nitrification. Some plants of the leguminous group (peas, beans, lupins and clover) are able to fix nitrogen directly from the atmosphere with help from bacteria in nodules on their root systems. These crop plants, as well as their relatives in the wild, are used in pioneering or colonising, or in crop rotations, to increase soil nitrogen levels. Nitrogen is introduced as fertiliser in soil management, and much has been written about the increasing and very large amounts of nitrogen applied in this way in Europe during the past 60 years. The crop yield advantage has encouraged this trend, but the more that is applied, the higher the quantity that is lost into the soil and water systems. It has often been estimated that only about half the applied nitrogen reaches the plant. The common forms of nitrogen fertiliser are ammonium nitrate, both the ammonium and nitrate ions being available to plants, and urea, where conversion to ammonium is required for availability. Organic manures—namely farmyard manure, green manures and slurry—are other inputs of nitrogen where bacteria conversion to ammonium and nitrate is again required. Application in the form of manures brings the usual benefits associated with additions of organic matter to soil, but it is a disadvantage that the amount of nitrogen added is not precisely known.

Although a significant part of soil research in agriculture has been devoted to furthering the understanding of the working of the nitrogen cycle, there are still large areas unexplained. In some field trials (Watson *et al.*, 1992) attempts have been made to measure nitrogen losses, and while leaching, denitrification and volatilisation losses vary according to the weather conditions of a given season, they

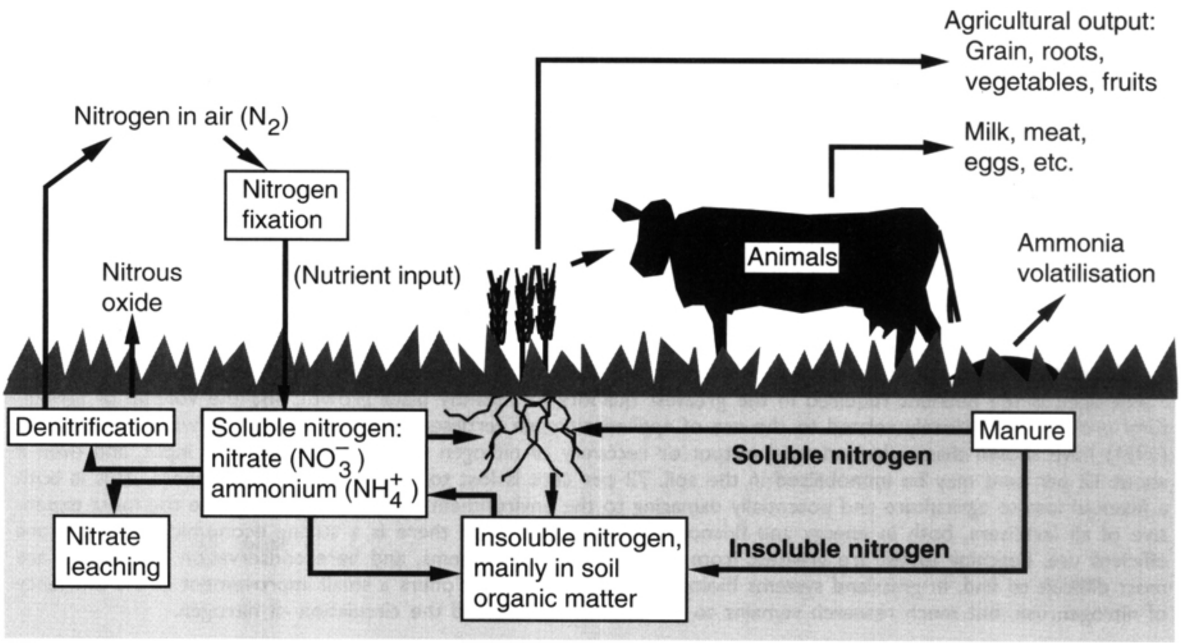


Figure 1.5 The agronomic nitrogen cycle

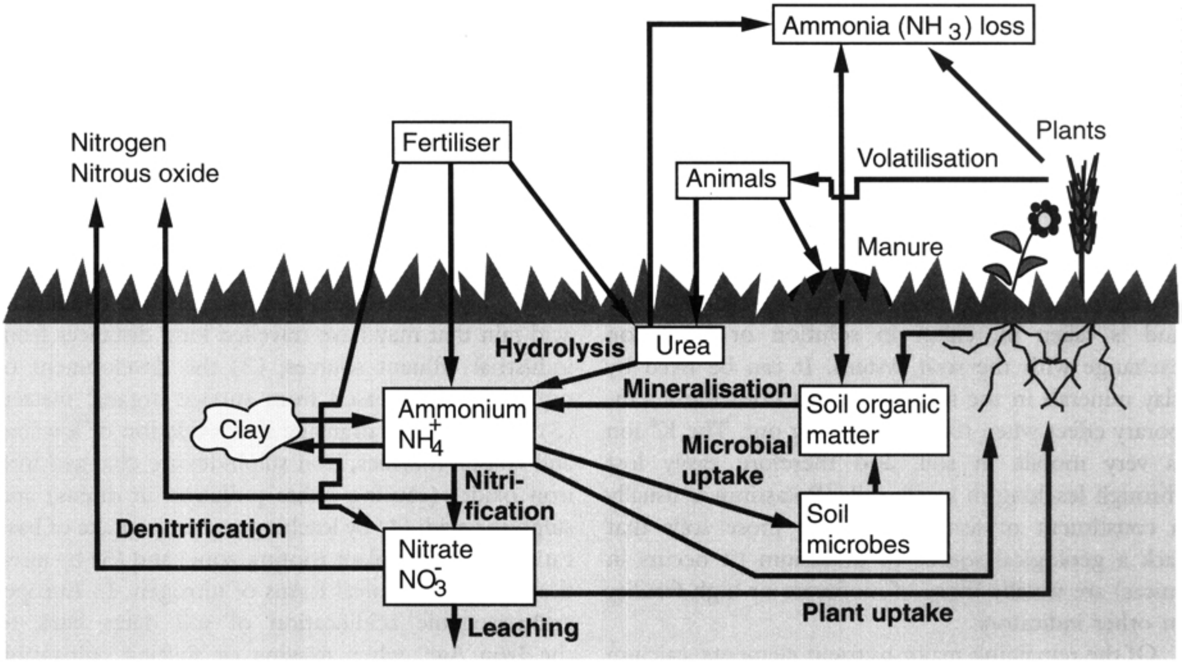


Figure 1.6 Conversion of nitrogen forms in the soil

remain in that order of magnitude as losses at likely application rates of nitrogen fertiliser in the climate of Ireland. Of even greater concern, however, is the fact that these three principal losses accounted for only about half the total loss of nitrogen from the cycle. The unexplained losses lie outside the present scope of soil management.

What is known about the circulation of nitrogen can and does play a part in soil management, where the aim should be increased efficiency through the reduction of nitrogen loss (Kristensen, 1995). Leaching loss is high in wet climates and well-drained soil, and is increased further in sandy porous soils, in soils low in organic matter, in arable systems and if nitrogen application is followed by heavy rain. In vulnerable areas, possibly as much as 25–30 per cent of the land area of Europe, nitrate lost through leaching may enter underground aquifers and eventually pollute drinking water to levels above the EU limit of 50 mg/litre. Denitrification is high in gleyed or waterlogged soils and where bacteria draw oxygen from nitrate, and also in acid and wet soils. Liquid manures, such as slurry, should be applied only to soils and in weather conditions that allow slurry to infiltrate the soil, and not in conditions where it will be washed over the soil surface or where it will dry and cake on the soil surface. Compacted soil, drought, poor drainage and infected root systems will all reduce the levels of nitrogen intake by plants. Increasing nitrogen fertiliser application levels may be counter-productive unless it is from very low N-levels, and may well make surface waters unhealthy for fish life and pollute underground aquifers (see Box 1.2).

Phosphorus (P)

Phosphorus is available to plants as the phosphate salt, or anion (H_2PO_4), and not in ion (cation) exchange between roots and soil. The phosphate ion is not as soluble as nitrate and does not move or leach out of soil in the same way. The problem in this case is that a very large proportion of soil phosphate—more than 90 per cent—is unavailable to plants, or chemically ‘fixed’ at any one time. Because so much soil phosphate is in store, even from applied phosphate fertiliser, what becomes available to plants within a narrow pH range of 6.0 to 7.0 may have been in the soil for a long time. It is a very much the aim of soil management to maintain the pH in this range to avoid fixation either above or below the range. Fixation involves the formation of insoluble calcium phosphates above pH 7.0; below pH 5.0 in acid soils fixation can be on the surfaces of clay minerals or in chemical complexes of iron and aluminium.

Potassium (K)

Potassium is available to plants as the cation of K^+, and is taken in, either in solution or by cation exchange with the root system. It can be fixed by clay minerals in the soil, but this is normally a temporary effect when the soil is drying out. The K^+ ion is very mobile in soil, and therefore easily lost through leaching in sandy soils. Potassium is usually a constituent of fertiliser because those soils that lack a geological source of potassium (it occurs in micas) are usually those of moderate or high fertility in other indicators.

Of the remaining major nutrient elements, calcium and magnesium behave in a similar way to potassium, entering the plant by cation exchange with the root system. There is often an available geological source for these elements, and if not, they are not too expensive to supply as fertiliser. Calcium (through lime) is used mainly to bring pH values within the optimum range for all nutrients, but is also a necessary plant tissue constituent itself. Sulphur behaves in a similar way to nitrogen and phosphorus, being cycled mainly through the soil organic matter.

PROBLEMS WITH SOIL CHEMISTRY

In Europe over the past two decades there has been a marked increase in problems caused by extremes of soil chemistry which have depressed crop yields and have even threatened the existence of plant life. The problems are related to various types of soil acidification due to a variety of causes, but also to salinisation of interior basin areas like the Hungarian Plain (see Box 1.1) and other areas inappropriate for irrigation. Soil contamination by toxic substances from sewage sludge disposal and effluent discharges from waste sites has also been increasing. Altogether, these problems of soil chemistry (along with soil erosion) comprise a large part of what is known as soil degradation, widely regarded as one of the greatest challenges facing mankind at present.

Soil acidification is a state of excess acidity reached when the pH falls below 4.5 and is associated with mobilisation of the hydrated ion of aluminium . The problem may be as much one of aluminium toxicity as soil acidity, and the effect is to reduce soil fertility and seriously damage crop growth. Excess soil acidity is encouraged by underlying geology, and can be caused in at least five ways: (1) from inputs of acid rain that may have travelled long distances from industrial effluent sources; (2) the development of organic acids released from surface organic matter; (3) following the drainage and oxidation of lowland and coastal marshes, iron sulphides are changed into iron oxides (causing ochre pollution in drains) and sulphuric acid; (4) by leaching and crop intake of base cations from the plant rooting zone; and (5) by nitrification of ammonical forms of nitrogen. In Europe, anthropogenic acidification of soil dates back to the Iron Age, when moving or shifting cultivation gradually gave way to an increased use of lime and lime-rich marls as natural improvers of soil acidity. However, the problem has become far more acute since the Industrial Revolution, and measurements in a wilderness area in southern Britain (Johnston *et al.,* 1986) show a fall of 2.9 pH units in that time. Liming is still the remedy in most cases, and during periods of widespread acidification in agricultural land the government has intervened with lime subsidies (mainly in the 1950s and 1960s).

Following legal moves to stop sewage sludge dumping in the North Sea, there has been increasing pressure to dispose of sewage waste on land (Giller and McGrath, 1989). Unfortunately, waste from industry usually contains amounts of heavy metals toxic to plant life. In Britain, about 60 per cent of sewage sludge is used on farmland, which can allow toxic metals to enter the food chain. Although there are now strict EU guidelines to limit the levels of heavy metals in sewage sludge, it is nevertheless potentially a major problem for soil management. Of the metals found, zinc and cadmium are most likely to be absorbed by plants. Higher than permitted levels of zinc will kill plants, but dangerously high concentrations of cadmium can pass through plants without effect and thus on to livestock and humans. Other metals in sewage sludge that are normally required only in trace amounts by plant life are copper, nickel, lead and chromium. If they are allowed to enter the soil, they will remain there indefinitely and accumulate. The risk is that levels will be reached that are toxic for soil micro-organisms, at a time when management is becoming more dependent on them for the cycling of nutrients from organic matter. It may well be that even stricter limits are required for heavy metals in sewage sludge.

CONCLUSIONS

This review of soil management has not been comprehensive. It has been limited to only one of the major uses of land, namely agriculture; even there it has been selective in topic and explored in the context of north-west Europe. Within this defined scope, there have been major omissions in areas of the chemistry of micro- or trace element nutrients, of the types and uses of a wide range of farm machinery, on some of the problems of cultivation, and many of the problems arising from intensification in farming. The soil management problems of other agricultural systems in other environments and continents have not been mentioned. Amongst what has been covered, the most serious problems are those of long-term potential damage, such as the complete loss of soil through erosion, leakages from the soil nitrogen cycle and the accumulation of heavy metals from sewage sludge. Because of the system resilience of soil, good management can repair other problems of a more temporary nature.

REFERENCES

Addiscott, T.M., Whitmore, A.P. and Powlson, D.S. (1991) *Farming Fertilisers and the Nitrate Problem,* Wallingford, UK: CAB International.

Barth, H. and L'Hermite, P. (eds) (1987) *Scientific Basis for Soil Protection in the EC,* London and New York: Elsevier Applied Science.

Boardman, J., Foster, I.D.L. and Dearing, J.A. (1990) *Soil Erosion on Agricultural Land,* Chichester, UK: Wiley.

Bøckman, O.C., Kaarstad, O., Lie, O.H. and Richards, I. (1990) *Agriculture and Fertilisers,* Oslo, Norway: Agricultural Group, Norsk Hydro a.s.

Culleton, N., Tunney, H. and Coulter, B. (1994) 'Sustainability in Irish Agriculture', *Irish Geography* 27, 1: 36–47.

Davies, B., Eagle, D. and Finney, B. (1993) *Soil Management* (5th edition), Ipswich, UK: Farming Press Ltd.

Garrett, M.K, Watson, C.J., Jordan, C., Steen, R.W.J. and Smith, R.V. (1992) 'The Nitrogen Economy of Grazed Grassland', *The Fertiliser Society Proceedings* No. 326: 1–32.

Giller, K. and McGrath, S. (1989) 'Muck, Metals and Microbes', *New Scientist* 124, 1689: 31–32.

Harris, J.A., Birch, P. and Palmer, J. (1996) *Land Restoration and Reclamation, Principles, and Practice,* Harlow, UK: Longman.

Henderson, W.C. and Farr, E. (1992) *Drainage Design,* Glasgow and London, UK: Blackie.

Hinrichsen, D. and Enyedi, G. (eds) (1990) *State of the Hungarian Environment,* Budapest, Hungary: Hungarian Academy of Sciences, Ministry for Environment and Water Management, and the Hungarian Central Statistical Office.

Hodgson, J.M. (ed.) (1991) *Soil Survey—A Basis for European Soil Protection,* Luxembourg: Commission of the European Communities.

Johnston, A.E., Goulding, K.W.T. and Poulton, P.R. (1986) 'Soil Acidification', *Soil Use and Management* 2: 3–10.

Keane, T. (ed.) (1986) *Climate, Weather and Irish Agriculture,* Dublin, Ireland: AGMET, Meteorological Service.

Kristensen, L. (ed.) (1995) *Nitrogen Leaching in Ecological Agriculture,* Oxfordshire, UK: Academic Publishers.

Morgan, R.P.C. (1986) *Soil Erosion and Conservation,* Harlow, UK: Longman Scientific and Technical.

Quine, T.A. and Walling, D.E. (1991) 'Rates of Soil Erosion on Arable Fields in Britain', *Soil Use and Management* 7, 4: 169–176.

Sherwood, M. and Tunney, H. (1991) 'The Nitrogen Cycle —A National Perspective', *Irish Journal of Agricultural Research* 30, 1: 75.

Simpson, K. (1983) *Soil,* London and New York: Longman.

Tunney, H.I. (1990) 'A Note on the Balance Sheet Approach to Estimating the Phosphorus Fertiliser Needs of Agriculture', *Irish Journal of Agricultural Research* 29, 2: 149–154.

Vàrallyay, G. (1993) 'Environmental Aspects of Soils and Land Use in Hungary', *Zeszyty Problemow Postepow Nauk Rolniczych—Polska Aka Nauk* v, 400: 53–72.

Watson, C.J., Jordan, C., Taggart, P.J., Laidlaw, A.S., Garrett, M.K. and Steen, R.W.J. (1992) 'The Leaky N-cycle on Grazed Grassland', *Aspects of Applied Biology* 30: 215–222.

SUGGESTED READING

Brady, N.C. (1990) *The Nature and Properties of Soils* (10th edition), New York: Macmillan.

Briggs, D. and Courtney, F. (1989) *Agriculture and Environment,* Harlow, UK: Longman.

Dent, D. (1986) *Acid Sulphate Soils: A Baseline for Research and Development,* Publication 39, Wageningen, The Netherlands: International Institute for Land Reclamation and Improvement.

Foth, H.D. (1990) *Fundamentals of Soil Science* (8th edition), New York: John Wiley and Sons.

Jarvis, M.G. and Mackney, D. (1979) *Soil Survey Applications,* Soil Survey Technical Monograph 13, Harpenden, England.

Lee, J. and Diamond, S. (1972) 'The Potential of Irish Land for Livestock Production', *The National Soil Survey of Ireland Bulletin* No. 26, Dublin.

Marshall, T.J. and Holmes, J.W. (1988) *Soil Physics,* Cambridge, UK: Cambridge University Press.

Soil Use and Management (1985–1996) Volume 1 onward, Oxford, UK: British Society of Soil Science and Blackwell Scientific Publications.

SELF-ASSESSMENT QUESTIONS

1 What is included in soil protection policy in the EU?
2 How is soil protection different from soil management (as defined here)?
3 What are the main features of soil reclamation in (a) the Dutch polders and (b) the Danish heathlands?
4 In what environmental conditions would you expect the risk of soil erosion to be due to the agent of (a) wind or (b) running water?
5 In what form is water used by plants, and how may society act to maintain this water supply?
6 What is the best type of soil structure for agriculture, and why?
7 Why is organic matter important in soil management?

2

SOLID WASTE MANAGEMENT

Alfons Buekens

SUMMARY

Traditionally, waste management was devoted to the hauling, landfill, pulverising, composting, recycling and incineration of municipal solid waste (MSW) and specific other waste streams. Today, however, waste has evolved into a major societal issue, a topic embracing almost all disciplines, from philosophy to business.

Moreover, in waste management, a legislation-driven business, the rules of the game as well as their implementation are time- and location-dependent. Nevertheless this text attempts to be of use both in the western world and in developing countries. Statistical data and reference to local legal contexts are kept to a minimum. Methods for defining, characterising and classifying waste and its related hazards are briefly introduced.

Waste management is treated in a context of sustainable development, by ranking the various options and at the same time explaining the practical difficulties of implementing the approach in question. Hazardous waste prevention and elimination is covered, as are the problems of introducing waste management into industrial enterprise.

Finally, a brief description is given of the major technical methods used in waste management, that is logistics, landfill, composting, incineration and recycling.

ACADEMIC OBJECTIVES

This chapter introduces the reader to the realm of waste, a material difficult to define, without value, and one that aroused little interest until the 1960s. Since then it has developed into a topical subject profoundly affecting government, corporate strategy, service industries, industrial design, the packaging of commodities and numerous other issues of economic, social and political importance. Students should gain an understanding of some of the consequences that actual waste management policy may have, and sometimes should have, on enterprise, industrial systems and the individual consumer.

The strategies followed in waste management and their local implementation in enterprise, as well as their limits, should be understood after reading this text. Some self-assessment questions, which may serve as introductions to open-ended discussion, are given at the end of the chapter.

INTRODUCTION

Waste arises during all human activities: *production*, *distribution*, *consumption* and even *leisure*. Its collection and elimination are vital for public health, hygiene, safety and the environment. Waste management is a major *service activity*. Together with schools and road building it takes a large share of the budget of local communities.

Waste is a *regulation-driven* business. Legislation, administration, inspection, economy, finance, health and safety, national, corporate and business environmental policy are major and interrelated factors.

'Waste' is difficult to define. The term connotes a material of negligible value, together with a desire to discard it from its original surroundings. There is no suitable definition yet devised that satisfactorily encompasses the variable value and eventual fate of all kinds of waste materials and commodities, with their specific legal and administrative status, and with the environmental and safety hazards associated with their generation, storage, transportation and elimination. Because of legal obligations, such as notification and the payment of levies, the term 'waste' is often a source of cumbersome and costly administrative procedures. This is an impediment to the free trade of goods destined for recycling. Conversely, distinctions in the definition of (hazardous) waste may create undesirable and uncontrolled transboundary movements. Nations with high standards in waste management often shift the burden of their domestic waste to their neighbours or even, more dramatically, towards poor and technically ill-equipped countries.

Waste can be classified on the basis of its origin, composition, physical aspects, chemical or hazardous properties and preferred method of disposal. Classifications are rarely comprehensive or entirely comparable, for waste management can be addressed from a variety of viewpoints, each of which requires different types of (forcibly incomplete) information. Moreover, because of its very nature, waste is rarely constant and predictable in form, size or composition and a precise determination of its properties is tedious, expensive and of limited use.

The responsibility for public cleansing and waste management used to lie with local authorities, but the planning, financing and operation of large landfills, incinerators or hazardous waste treatment centres has called for the creation of larger bodies and jurisdictions. Legal codes tend to be established on a nation-wide basis, when it comes to basic principles or on a regional basis for their practical implementation; the European Communities have issued a number of Directives to promote greater uniformity and raise standards in some member states. The protection of the marine environment from hazardous waste streams also requires an international approach.

Waste storage and handling should proceed hygienically and safely, taking into account the nature and volume of the waste and the activities that led to its generation. *Collection* and *transportation* are a problem of logistics; they involve the selection of suitable receptacles, grouping methods, collection vehicles, and transfer stations and modes towards the site of final disposal.

Traditionally, waste *elimination* is based on landfill (tipping), composting, incineration or recycling. Much research and development (R&D) work has been devoted to innovative technologies in mechanical, physico-chemical and thermal waste treatment, but many projects were halted at the pilot stage. Most of those that continued as (generally public supported) demonstration plants did not perform well, either technically or economically.

Most countries have created a specific legal basis and administrative organisation to manage, monitor, control and plan waste-related activities.

In many countries, abandoning waste is an offence. The generation, storage, transport and disposal of waste is subject to rules, which are more stringent when the waste is defined as 'hazardous'. Numerous operations require a *licence;* others have to be notified or are subject to environmental *taxes* or levies. This administrative and legal context requires a clear definition of (hazardous) waste. There is general consensus that waste generation should be prevented or reduced, and that re-use and recycling should be stimulated. If material usage is no longer possible, incineration is advocated, with recovery of heat. When all other options have failed, final disposal should be conducted without environmental damage. Innovative waste management may be applied at several levels: that of a household, a shop or an enterprise. Methods for auditing the existing situation and taking measures for waste reduction, re-use, recycling, conversion or elimination are discussed below.

The responsibility and liability of manufacturers and distributors has been extended considerably, not only with respect to manufacturing waste, but also with respect to the fate of the product at the end of its lifetime. Waste generation should be minimised, not only during the production of goods, but also at the level of their ultimate disposal. In the future, systems for collecting, dismantling and recycling materials and spare parts may be expected to gain momentum, but the associated economic factors and the question of who should pay the costs are still unclear and unsettled issues.

POLICIES AND STRATEGIES IN WASTE MANAGEMENT

Ranking of possible options (WCED, 1987)

In industrialised countries, waste management is generally regarded as a major problem which cannot be solved, as it was in the past, by technical means alone. A fundamental modification in current patterns of production, consumption and waste elimination is required in order to develop a new type of society, based on the principle of sustainability, which stops compromising the environment and the well-being of future generations. This implies less depletion of non-renewable natural resources (fossil fuel, ores) and reduction of pollution at the source, by converting linear production systems into cyclic structures.

Conceptually, the following hierarchy of options has been put forward:

- waste prevention;
- waste re-use;
- recycling of materials from waste;
- using the energy content of waste;
- safe elimination.

The above ranking is, at present, generally accepted, but its practical implementation is still hampered by numerous problems. It is a real challenge to reconcile a long-term ecological vision with the casual lifestyle and materialistic attitude of the average consumer on one hand, and the prevailing industrial-economic-financial system on the other hand. Today's citizens may be genuinely worried by environmental damage, but it is unclear how willing they are to sacrifice buying power, diversity of products or the use of disposables for the sake of the environment. Moreover, a socially and environmentally correct market economy based on free enterprise, free trade in goods and price establishment on a basis of supply and demand has almost universally been hailed as an effective and efficient economic system. However, this system features a low present-day value of long-term benefits, such as those implied in a sustainable society.

Waste prevention (ECE, 1978; Meadows, 1972; Meadows et al., 1992)

The first commandment of waste management is prevention, for its combination of resource conservation and waste avoidance makes it cost-effective. Realistic methods of prevention are:

- for the consumer, avoiding superfluous consumption and adopting healthy and thrifty lifestyles;
- for the manufacturer, designing for long life and easy servicing, repair and dismantling;
- for industry:
  (a) *loss prevention* by improved materials handling and storage, preventive maintenance, adapted procedures and controls, proper training and improved motivation of operators, etc.;
  (b) use of *clean technologies*, with increased yield, less by-products and side streams and a conversion from linear towards cyclic or integrated production structures.

Non-waste technology has been actively promoted for two decades, spurred by the reports of the Club of Rome. Some examples of clean technology are:

- using larger unit packs;
- converting from packed product to bulk handling;
- standardising automotive parts or beer bottles (Germany);
- closing the white water circuit in paper making, saving process water and fibres;
- counterflow rinsing of electroplated objects in the galvanising industry;
- continuous casting of steel.

Substitution also makes it possible to replace *hazardous* with *less dangerous* materials. Monitoring acquisition and consumption may reduce throw-away and misuse, beneficial on both environmental and safety scores.

A major obstacle to waste prevention is that it often requires a profound change in plant lay-out, process technology, operating procedures and human habits. Moreover, valuable materials are often discarded because in the short term their recovery would cost more in terms of capital, labour and innovative initiative than savings or recoveries can generate in monetary terms. When a truckload is discharged incompletely it may be more economical to dump the remains rather than haul them back over a considerable distance, and meanwhile immobilise the truck with this fractional load. Any such hampering of environmental improvement urges a review of the monetary value of these products.

Table 2.1 Some negative consequences associated with waste prevention

Waste management measure	*Undesirable consequence*
Less consumption	Reduces industrial production, trade, individual income and general prosperity; increases unemployment
Long lifetime	Countered by technical progress; old appliances and cars consume more and are more polluting
Ease of repair	Excessive cost of servicing and limited reliability
Standardisation	Less consumer choice; against accepted marketing principles
Less (one-way) packaging	Increases distribution cost, losses and amount of damaged goods
Ban disposables	Reduced convenience
Re-use white and brown goods	Problems of wear, obsolescence, reliability, hygiene, corrosion, past improper use or overvoltage
Materials recycling	Small-scale activity with a major manual component at initial stages; product demand and profitability vary cyclically; bad neighbour because of stockpiling, smell, noise, soil pollution; less constant quality than for virgin material

The term 'waste prevention' is often used incorrectly when what is really meant is the re-use or recycling of materials. Thus cars, white or brown goods, electronic equipment, etc. will have to be designed for re-use or recycling, which might even involve manual dismantling. The technical mechanisms and economic measures needed to induce manufacturers to take products back after the end of their lifetime are still in a phase of definition.

Some potential societal problems involving prevention, re-use and recycling are listed in Table 2.1.

The concept of *mandatory utilisation* (in German: *Verwertungsgebot*) as a legal means to impose re-use or recycling finds its place in environmental thinking, as long as its cost is not *prohibitive*.

Re-use and materials recycling (Buekens, 1992)

All goods, whether long- or short-lived, take their share of a national or a family budget, consume natural resources and sooner or later become waste. Consequently, a general feeling has developed that something must be done about this apparently irreversible growth of the mountain of waste we generate. Increasingly, legislative and administrative interventions have been called for in the form of mandatory re-use or recycling, or of economic incentives. Citizens are invited to collaborate in source segregation and resource recovery schemes on the basis of:

- resource recovery centres, where bulky and reclaimable wastes can be delivered free of charge;
- kerbside collection, directly combined with sorting, as in the Laidlaw developed Canadian *blue box* system. The collection crew directly sorts a number of easily recognisable items into separate compartments or in ancillary containers, attached to the collection vehicle. Undesirable items remain in the box. Kerbside collection attains a higher recovery yield and is more convenient but is also more expensive.

In affluent western society a simple glass bottle has two seemingly opposite features. On the one hand it is a high tech, strict quality product, on the other it is also a low cost, widely available commodity. The question therefore arises of whether to:

- *re-use,* i.e. retrieve, group, clean and refill it;
- *recycle,* i.e. re-use it as cullet in bottle manufacturing; or
- *eliminate,* i.e. discard it together with general waste.

The answer up to now has been made on the basis of objective logistic and economic factors on one side, and of subjective and personal preference factors (expressed in concert by bottlers, distributors and consumers) on the other. Lighter packages strongly support sales of mineral water and soft drinks. Large crates with refillable bottles are the most economical solution in catering and large households. One-way containers of plastic

BOX 2.1 'WORK FROM WASTE' PROJECT

Work from Waste brings together appropriate technologies employed all over the world for recycling paper, iron and steel, tin, non-ferrous metals, plastics, textiles, rubber, minerals, chemicals, oil, and human and household wastes.

All these materials are suitable for labour-intensive processing, often requiring little capital and providing a cash income plus other environmental and community benefits. Jon Vogler describes how to set up and run a small waste recycling business, primarily for use in developing countries, but also of interest to communities and groups practising local self-reliance in industrialised countries (Vogler, 1981).

Work from Waste is the result of the Wastesave *project,* set up by Oxfam in 1974 to raise money for poor world development by recycling rich world wastes, an approach to solving the waste problem and at the same time creating jobs.

How to start a waste business is explained in simple words in *Work from Waste*.

or glass have a higher grade of convenience and, as long as waste collection and elimination are well organised and inexpensive, the public at large has followed the easy way, i.e. the one-way approach.

Energy use (Buekens and Schoeters, 1984)

Numerous waste streams are combustible and can be fired in a dedicated furnace, the incinerator, with either (furnace-) integrated or waste (=subsequent) heat recovery. Although it is done less often, refuse can also be co-fired together with other fuels (coal, lignite, peat, oil, gas) in power plants, industrial boilers, cement kilns or even domestic stoves (the latter operating, however, without adequate operating or air pollution controls). There is a large variety of stokers that can accommodate waste. Moreover, it is possible to convert combustible waste into Refuse-Derived Fuel (Alter and Dunn, 1980), into flammable gas, oil, char or specific chemicals by methods such as pyrolysis, gasification and partial oxidation.

The value of the heat generated is rarely more than a fraction of the cost required to pre-process and fire the waste, to amortise the much-larger-than-usual boiler and to clean the flue gas from particulate, acid gases and dioxins. Old units have been a notable source of dioxins and heavy metal deposition. It should be noted that burning a heterogeneous mixture of wastes products inevitably results in the products of incomplete combustion, for example, polycyclic aromatic hydrocarbons. These emissions are higher when the technology used is poorer.

Elimination of useless waste

Strategies in final disposal

The principal benefits and limitations of various options in waste reduction and elimination are enumerated in Table 2.2.

Table 2.2 Benefits and limitations of waste elimination methods

	Potential benefits	*Potential limitations*
Re-use	Highest value in waste	Individuals prefer new goods
Recycling	Recovery of raw materials and their intrinsic energy content	Market specifications and limited absorption potential
Composting	Recycling method, converts organic waste into a useful soil conditioner	Only for biodegradable, organic materials; smell and litter problems; quality and marketing problems; composting residue to be landfilled or incinerated
Incineration	Important reduction in volume; the residue is sterile and no longer toxic, flammable or putrescible; heat recovery is feasible	Technically complex and expensive process Deep treatment of flue gas, scrubbing liquors and residue is normally required
Sanitary landfill	Simple, relatively inexpensive	Land requirements; site becomes improper for numerous applications; potential soil and ground-water pollution

Final disposal is problematic, especially for hazardous and radioactive waste. Three strategies have traditionally been proposed to tackle the latter, each with an appropriate mnemonic tag:

- *Delay and, decay,* involving temporary storage of highly radioactive waste in a suitable cooled vessel, until a sizeable fraction of radioactivity has been decayed and radiation and heat generation have declined to more tractable levels.
- *Concentrate and contain,* whereby waste is buried in such a way that any non-controlled contact with the environment is believed to be excluded for extremely long (possibly even geological) time periods, so that radioactive or hazardous compounds cannot be taken up in biological cycles. Concentrate and contain, even for ordinary waste, became a mandatory strategy for all types of waste in numerous countries.
- *Dilute and disperse,* whereby low radioactive effluents can be discharged into the atmosphere or into water resources, after appropriate control. Hazardous waste can be mixed with ordinary refuse, the absorptive power of which serves to retain, dilute and gradually decompose or neutralise the waste. According to this view certain amounts of acids, bases, sludges, tars, cyanides or heavy metals may be co-disposed with household refuse, without harming the environment. (The only proponent of such dilute and disperse is the United Kingdom.)

Each of these strategics entails obvious environmental risks. Whatever the strategy adopted, careful planning and mapping of waste depositories is required for reference, to avoid grave errors such as that of Love Canal (New York), where a housing project was inadvertently located on a former chemical landfill. A survey of the methods to prospect for *contaminated sites,* evaluate their risk and, if necessary, start a sanitation programme is given by Buekens and Nieuwejaers (1987), and the liability for clean-up costs is discussed by Bocken (1987).

Disposal at sea

The oceans are an enormous reservoir. Seafaring nations in particular regarded them as a suitable sink for waste streams. Disposal at sea has been used extensively for sewage sludge, for municipal solid waste (e.g. by New York until 1935 and Istanbul until about 1975) and specific industrial wastes (red mud from titanium oxide production, acid and phenolic waste, low-radioactive waste, stored in drums and enclosed in concrete). Since the Oslo Convention (1977) the environmental dangers, mainly of accumulation in the food chain, have been worked into regulations and there has been a ban (black list) on the dumping at sea of numerous chemicals (mercury, cadmium- and organohalogen-bearing waste, and non-degradable plastics). Grey list substances are subject to relatively stringent administrative controls. There is a world-wide movement towards freeing the sea of all waste-related operations, including seagoing incineration vessels.

Lagooning

Liquid waste, slurries and suspensions are often treated and stored in lagoons, which act as sedimentation basins and often as biological reactors as well. Aerobic oxidation may be enhanced by the use of surface aerators.

Lagoons may be either natural or excavated, and either lined with plastic film or situated in an impervious layer, to avoid undesirable soil infiltration. The environmental dangers of this procedure should be evaluated on a case-by-case basis.

Disposal on land

Disposal on land, also termed (sanitary) landfill or (controlled) tipping is the most important final disposal method. Composting and incineration both leave solid residues to be landfilled. Numerous inorganic bulk waste streams, such as fly-ash and cinder from coal-fired power plants, phospho- or sulpho-gypsum, dredging spoils and metallurgical slag can only be recycled (market outlets permitting) or landfilled.

Mine tailings, colliery spoils and metallurgical slag are three examples of waste that in many cases have gained value and been exploited more than a century after being tipped. Different techniques may be applied for crude, shredded, baled or precomposted waste. Shredding of refuse reduces the need for cover material; preliminary composting (in German: *Rottedeponie*) reduces the generation of landfill gas and of high COD-leachates.

Almost all waste materials are amenable to being landfilled, including:

- household and trade waste;
- construction and demolition waste;
- industrial waste.

Rubber tyres are often unwelcome because of their tendency to float. Sometimes they are shredded and used as a drainage layer. Sludge and filter cakes adversely affect the stability of a fill; they should be spadable, a quality determined by moisture content and torque experiments. Agricultural waste is highly fermentable and leads to huge gas generation and fast subsidence.

Salt mines, geological storage, deep well injection

Some operations in land disposal are highly specific, such as:

- the permanent (and retrievable) storage of numerous chemicals, e.g. spent cyanide hardening salts or transformers once filled with PCBs in galleries of salt mine caverns;
- the permanent storage of radioactive waste in geologically suitable strata (clay) or formations (salt domes, granite);
- deep well injection of brines and liquid waste into porous structures, separated from the surface by impervious layers.

A thorough risk analysis to determine the potential for disruption of the protective layers shielding the environment is required in each particular case.

Administrative structures

According to the first waste-related European Union Directive (1975) each member state should appoint specific authorities in charge of the planning, organisation, licensing and control of waste management operations. Planning encompasses the preparation of a scheme which details:

- the nature and amount of present and future wastes to be disposed of;
- the technical standards for doing so;
- suitable sites for these purposes;
- specific measures required for special wastes;
- the related expenditures;
- suitable administrative, legislative, technical, and financial measures to rationalise collection, sorting and disposal.

Measures should be taken to ensure that waste generators either deliver the waste to appropriate public or private enterprise, or ensure adequate disposal by their own means. All facilities for treating, storing or disposing of waste should be licensed, with due specification of:

- the nature and quantity of waste to be treated;
- the general technical requirements;
- the prospective measures taken;
- the necessary data related to origin, treatment and final destination of the waste treated.

All operations of such licensed facilities or enterprises should be subjected to periodic control in order to ascertain adherence to the conditions stipulated. The cost of treatment should be borne by the waste generator (the ‘polluter pays’ principle).

Compulsory notification is an important tool for compiling reliable statistical data on the generation, flows and final destination of the waste stream. Such data are important to support inspection and controls as well as measures to prevent or reduce waste generation, or encourage re-use or recycling.

Opinions differ on whether waste collection and elimination should be handled as a public service, by private initiative or in a mixed mode. In the past, public inspection was sometimes confronted with inadequate or even illegal operation by public waste operators; in other cases public waste companies had difficulties in securing sufficient funds, or capable and politically independent management (see Box 2.2). Such factors help explain the present shift towards private operation. However, a good *modus vivendi* between public and private forces has been fruitful in overcoming problems of profound sectoral change and overcoming the NIMBY problem.

BOX 2.2 APPROPRIATE SOLID WASTE MANAGEMENT FOR DEVELOPING COUNTRIES: THE FIRST SANITARY LANDFILL OF TURKEY: IZMIR HARMANDALI

General Information

Izmir, the third largest city in Turkey, is located on the west coast of the Anatolian Peninsula on the Aegean Sea. It was the first Turkish city to construct its own sanitary landfill. This site, inaugurated in April 1992, was designed and constructed entirely by Turkish engineers and technicians, proving that the construction and operation of a sanitary landfill is not impossible for developing countries. The site covers an area of 900,000 square metres, was designed to serve the city of Izmir for 15–20 years, and cost only US$ 1.6 million, including land expropriation, geological studies, mapping of the site, fencing, road construction, reception facilities, administrative building, repair shop, etc.

Wastes generated

The amount of solid wastes generated in Izmir is approximately 2,000–2,100 ton/day. Furthermore 260 tons of sludge from wastewater treatment plants with an average water content of 85 per cent are generated on a daily basis. Dependable data for the industrial wastes are not available. During a recent survey conducted in 135 factories it was observed that 595 tons of process wastes and 1,570 tons of ashes are produced on a weekly basis. Taking into consideration the number and size of the remaining factories, the daily amount of industrial wastes can be estimated as 500 tons. Furthermore 27 ton/day of infected hazardous wastes were generated from hospitals in Izmir. The composition of the wastes vary considerably according to the season. The table below gives the results of the systematic studies conducted during the winter and the summer seasons.

Solid waste composition in Izmir

Item	*Winter (%)*	*Summer (%)*	*Item*	*Winter (%)*	*Summer (%)*
Food remains	61.41	86.59	Metal	2.95	1.17
Glass	15.15	4.25	Textiles	–	1.00
Paper & cardboard	9.02	3.92	Miscellaneous	1.71	0.80
Plastics	9.76	2.27			

Selection of the site

In the selection of the site common factors like health, safety, accessibility, climatic conditions, hydrogeological conditions, hauling distance, etc. were considered, as well as:

(a) public-neighbourhood acceptance;
(b) low cost of the land;
(c) low construction cost.

Taking these factors into consideration and after an extensive search the site near the village of Harmandali, 25 km from the centre of Izmir, was selected. The reasons for the selection of this site were as follows:

(a) *Public-neighbourhood acceptance:* The site selected for the construction of the landfill and the surrounding areas were mainly used by the inhabitants of Harmandali village as an animal dung drying field (dried animal dung is used as a low calorie fuel) so dumping of solid waste in the nearby area did not disturb the public.
(b) *Cost of land:* The chosen site was mainly owed by the Turkish Treasury, and was transferred to the Municipality at almost no cost. Some pieces of land belonging to private owners were expropriated.

(c) *Low construction cost:* One of the greatest costs involved in the development of sanitary landfills is that of constructing the impermeable layer at the bottom. In the area selected, however, the hydraulic conductivity varied between 10^{-7} to 10^{-10} m/s with an average value of 10^{-9} m/s. For the most part, the thickness of the existing clay layer was more than 20 m. Thus the requirement of impermeability was satisfied by nature and there was no need for any further insulation. The situation was such that the National Water Authority readily gave the approval required for the construction of the landfill in the area.

Use of the site

The selected site was divided into zones for the disposal of (a) domestic-commercial wastes; (b) hazardous industrial wastes; (c) hospital wastes; pre-treatment of sludge; and (d) the administrative zone.

Design and construction

The approach taken in design and construction was to achieve maximum economy without reducing the safety of the site. This led to some clever simplifications, examples of which are given below:

(a) *Control of leachate:* Collection of leachate at the base of the disposal site was achieved by locating leachate collection pipes only on the dry stream beds and their branches, since the topography of the land in this area forms a natural channel. It was further recommended that leachate storage tanks of capacity of 50 m^3 should be located at approximately 300 m intervals, from where the accumulated leachate was to be pumped back on the deposited solid wastes. A leachate flow equal to the maximum rainfall recorded at this area was used as the design flow. During the construction phase of the project, however, this design was changed so that pipes were not used for the construction of the leachate collection system. Instead the natural river beds were cleared of vegetation and other materials, compacted and filled with gravel and broken rocks from the site, and used as leachate-collecting conduits. Furthermore, instead of constructing the recommended leachate accumulation tanks, a small earthen dam was constructed to collect the leachate to be pumped back.

(b) *Control of biogas:* Given the impermeability of the soil surrounding the site, it was decided to install biogas ventilation chimneys at 50 m intervals without any blockage around the site. Hard perforated PVC pipes (ID=15 cm) were to be used. The plan had been to place the pipes at the middle of 1×1 metre pit supported with gravel. The diameter of the gravel layer surrounding the pipe should have been 1 metre. While interim solid waste layers were squeezed, a second PVC pipe was envisaged, to be placed over the first one. Thus the ventilation chimney would have been elongated. During construction, however, the loction of the ventilation pipe was simplified in the following way: holes were opened on the sides of metal barrels from which the top and bottom lids had been removed. By locating these barrels on top of each other a pipe was formed. The perforated hard PVC pipe was located at the centre of this pipe, and the space between the barrel and the PVC pipe filled with gravel. This simplification has proved highly effective.

(c) *Bird control:* Bird control was achieved partly by locating nylon fishing lines above the dumping site.

(d) *Control of light movable wastes:* A fisherman's net was fixed on 6 m high poles around the landfill area and used as a movable curtain to control the drifting of light wastes such as papers, plastics, etc. by wind.

Source: The late K.Curi, President, Turkish National Trust Committee on Solid Waste, Bogaziçi University, Bebek, Istanbul, Turkey

Legislation and administrative trends

Decrees of the French Revolution declared public hygiene and cleansing to be official responsibilities at the local level. These include municipal solid waste (MSW) collection and elimination, street sweeping and related menial yet essential activities. Also in the field of waste, environmental concerns have been a prolific source of new codes and measures. Their implementation is, however, time-consuming, gradual, and faces numerous conceptual and practical problems.

Legal texts are often incomplete and insufficiently detailed with respect to definitions, fields of application and means of practical execution. Typical examples are the definition of *waste* and of *toxic and hazardous waste*, the methods of sampling and analysis or of assessing (eco-)toxicity and other (environmental) hazards. There are at least 30 different leaching tests in regular use and even the spadability of paste-like waste is difficult to test in an incontrovertible manner.

Other environmental codes relate to the operation standards of landfill, composting, incineration or hazardous waste elimination plant and acceptable emission levels. The implementation of even the most modern, pollution-free technology has met with increasingly strong opposition, especially when it comes to finding a final repository for hazardous or radioactive waste (the NIMBY syndrome). The failure to secure treatment opportunities locally, and the high cost of treatment using the most sophisticated means, have both been responsible for 'waste tourism', that is the movement of waste and also recyclables to more waste-hospitable countries. This trend runs against the principle of *proximity*, which calls for waste elimination at a reasonably short distance from its point of generation; this principle figures in the EC Directive that subjects the transboundary shipment of waste (and unfortunately also of *real recyclables*) to administrative controls by the countries of origin, destination and transit.

In the past, the *best available technology not entailing excessive cost* (BATNEEC) was often not available locally, so waste generators had obligations they could not practically meet. Conversely, high-tech waste elimination centres often failed to attract customers in sufficient numbers because of cost factors, announced or published codes failing to be enforced, or because an excessive number of new entrepreneurs entered the market at the same time, without an adequate survey having been made of an unknown waste generation market. The latter is difficult to assess with precision because of insufficient statistical data, the variety of competing disposal options and general uncertainty about the impact of preventive measures, prompted by the impressive escalation of handling and elimination costs. Regulatory action by public authorities may lessen the problem to some extent: these authorities have access to statistical data through compulsory waste notification. They also have both the financial power to build and operate facilities, and the administrative power to persuade waste generator staff to select certain specific routes of elimination, instead of low-tech, low-cost options.

The principle that the polluter pays has remained virtually unchallenged, because its application encourages reduction at the source, rather than a spreading of cost over society at large.

Waste operations should be controllable as well as controlled. Although emission control of waste incinerators in the past was very limited, today most emission parameters of an incinerator plant can technically be monitored continuously, and the analytical data recorded in a tamper-proof way, certifying compliance (or otherwise) with emission standards; at the time the influential German TA-Luft standards were published for the first time (1974), however, such controls were unavailable. Moreover, MSW-incinerators are relatively small in number. But how to test waste elimination practices in many thousands of (say) dry cleaning or printing shops? Of course, compulsory notification of waste generation and administrative proof of its transfer to licensed eliminators (which are much smaller in number, and so easier to monitor) form a major step towards a solution. Enterprise is entering an age of environmental bookkeeping and self-assessment procedures.

Last, but not least, there is the modal consumer. How to devise a system to induce him or her to follow a sustainable lifestyle, without violating privacy or limiting freedom excessively? The only possible answers are motivation, education and group control.

It has always been difficult to evaluate and predict community and individual response to new measures involving waste management and reduction. Billing individual consumers proportionally to the weight of their refuse container or sack may render them more waste-conscious and reduce generation rates, but in other cases it leads to illicit abandoning of waste, and to small-scale waste incineration or burial in gardens. The efforts required to inform and educate the public and ensure its full support have often been underestimated.

Integrated solid waste management

Integrated solid waste management is a systematic approach to managing solid waste. It includes an analysis of all components of a solid waste management system to meet the defined goals (definition by World Resource Foundation, 17 October 1996). Such goals could be circumscribed by responsible management of resources and waste in a way that is feasible from a technical, economic and (not least) a public hygiene viewpoint.

Today many new ventures in waste management are designed in plush offices by imaginative technicians and scientists who are, however, less acquainted with the realities and practicalities of waste. Novelty in this sector should never be evaluated on the basis of conceptual ingenuity, but rather for its potential for simple, foolproof, daily application under adverse and often unexpected circumstances.

PREVENTION AND ELIMINATION OF HAZARDOUS WASTE

Nature and origins of hazardous waste streams

Hazardous waste streams in industry are directly related to the nature of raw materials and processes being used, generally arising as a by-product of a chemical process or as a stream of contaminants to be eliminated from this process. In other cases it is the result of a chance occurrence, such as a fire, a traffic accident, a plant mishap, wrong procedures, or of plain negligence.

Prevention

'Pollution prevention pays' was the famous saying launched by 3M in 1975. Best results are usually obtained by scrutinising current operations, identifying sources of avoidable losses and contaminations and eliminating them if possible by simple means or changes to procedures.

Whenever possible the use of hazardous chemicals should be eliminated or reduced. Great efforts are made to avoid the use of and minimise the losses from chlorinated solvents. Solvent-based paints are gradually being replaced by water-based ones, to eliminate the loss of solvents into the atmosphere. It should be recognised, however, that degreasing with lye rather than with chlorinated solvents or replacing one type of coating by a different product may require radical changes in equipment, operating methods and quality control.

Volume reduction

The cost of waste storage and transportation is largely proportional to its volume. Hence, volume reduction is an important option both on site and in centralised hazardous waste treatment centres. Possible methods are:

- avoidance of unnecessary dilution and mix ing of hazardous and non-hazardous waste streams;
- conversion of sludge to filter cakes using ei ther mobile (trailer mounted) or stationary filter presses, sieve-belt presses or centrifuges;
- incineration of hazardous waste, converting the inorganic part into ash or slag, and the organic into flue gas to be cleaned by removal of dust, aerosols, acid gases and NO_x.

Shredding is often applied to hazardous metal and plastic waste containers which can no longer be recovered, or to bulk up scrap. Baling is used for paper, board, plastic film and bottles (PET has first to be pierced), metal boxes, turnings and shavings.

Detoxification-immobilisation

Hazardous organic waste can be burnt to form innocuous or easily treatable inorganic compounds, according to the reaction:

$$C_xH_yO_zS_aN_bCl_c \ldots + \left(x + \frac{y-c}{4} - \frac{z}{2} + a + \frac{b}{2}\right)O_2$$

$$\rightarrow xCO_2 + \left(\frac{y-c}{2}\right)H_2O + aSO_2 + bNO + cHCl$$

$$NO \rightleftarrows N_2 + O_2$$

Destruction efficiencies of 99.99…per cent and better (six 9s) are attained in specialised incinerator units, treating, for example, chlorinated waste. However, practical application of these theories in practice often shows substantially lower efficiencies. Oily emulsions may also be treated by various other means, including emulsion breaking.

Radioactive and inorganic waste both have an intrinsic dangerous property which is not amenable to destruction or removal. Hence, either immobilisation *(concentrate and contain)* or dilution *(dilute and disperse)* is applied.

Immobilisation of fly-ash or heavy metal containing sludge involves either converting the dangerous compound into an insoluble form (sulphide, hydroxide, phosphate,...) or incorporating it into a structure or matrix that is resistant to leaching. Examples are:

- incorporation into an asphalt or silicate matrix;
- blending with some 6–8 per cent cement, fol lowed by hardening to form concrete blocks;
- melting with appropriate fluxing agents or in special plasma or electric furnaces to yield a glassy aggregate.

Numerous hazardous inorganics become less dangerous after appropriate treatment. The major operations applied to inorganic hazardous wastes are:

- neutralising acid and caustic streams;
- oxidising cyanide solutions with chlorine in caustic media;
- converting cyanide solutions into hydrogen cya nide in a closed vessel, followed by catalytic post- combustion of the highly toxic gas;
- reduction of chromium (VI) to chromium (III) using (bi)-sulphite solutions;
- precipitating heavy metals at an appropriate pH value.

The treatment of heavy metal bearing waste waters by precipitation, electrodeposition, oxidation/reduction, membrane processes, ion exchange, immobilising chelating agents, liquid/liquid extraction and other methods has been discussed by Tels (1986).

Hazardous waste collection and transfer

Hazardous waste is treated either locally or by specialised, licensed enterprises. The latter have to follow strict procedures for safety, technical and administrative reasons. First, representative samples must be taken and analysed for key elements. A description of the waste stream is made, embracing:

- activity and production process of origin;
- various components or compounds;
- physical characteristics;
- required form of packaging;
- quantity per unit load.

Specifications of a form describing a 'waste' stream is given in Table 2.3. After evaluation of the above data, the scope for local pre-treatment (e.g. filter pressing or any other method of phase separation), conditioning or grouping should be investigated and the possible alternative modes of elimination analysed. The latter should be based on existing technologies and capacity available in hazardous waste treatment centres, storage time before acceptance, appropriate administrative and technical procedures, and, of course, cost of treatment.

Acceptance certificates may have to be prepared with regard to the transboundary movement of waste. Importation follow-up and EC documents must be filed, CMR[1] shipping-documents and hazardous material cards prepared, and a 'pro-forma' bill prepared prior to shipping. After the waste has been eliminated, an attestation to this effect must be issued.

Upon delivery, the conformity of waste to its original description and analysis should be tested, either by spot tests or by analytical procedures.

Table 2.3 Required information for acceptance of hazardous waste

- Presence of metallic Na, K, Mg. . . elemental sulphur, phosphorus. . . .
- Presence of cyanides
- Presence of (bonded) halogens F, Cl, Br, I
- Presence of Ag, As, Ba, Be, Cd, Co, Cr, Cu, Hg, Mo, Ni, Pb, Sb, Se, Sn, Te, Th, V, W, Zn
- Presence of radioactive substances, peroxides, perchlorates
- pH value
- Hazards: explosive, oxidant, corrosive, toxic, highly flammable
- Heating value
- Flash and flame point
- Colour, smell, boiling point
- Ash content, moisture content
- Bulk density
- Physical form: powder, solid, paste (spadable), slurry, emulsion liquid, liquefied gas . . .
- Conditioning: bulk, bag, plastic container, metal drum, wooden crate, pressurised containers, big bag. . . . (+ unit volume of container)
- Transport firm and mode: truck with tilting or fixed loadback, container, tank container, tank truck, suction truck . . .
- Protective measures required (gloves, spectacles, suit, filter, dust or gas mask)
- Fire fighting, first aid and spill control measures required

Some hazardous waste may be suitable for recycling, for example, solvents, oils and precious metal-bearing catalysts. Recycling, however, creates numerous hazards because *(inter alia)* of the presence of unpredictable or uncontrolled compounds in waste streams. (Note the required absence of explosive nitro-cellulose, a component of some inks, in a solvent to be recycled.) Other waste may be suitable for use as a fuel, for example, in cement kilns (cheap and contaminated solvents, oil, antifreeze,...). Technical acceptance, however, varies with peripheric equipment (e.g. flue gas scrubbing), as well as with local rules and plant-specific licenses.

Relevant factors are also the consequences of exposure to heat, compression, or mixing (with water, acids, bases, oxidants and reductants). A table of waste streams that should not be mixed is given by Hatayama *et al.* (1980). ADR codes also specify wastes that should not be transported together by road (ADR, 1959).

Table 2.4 Data relevant to solvent recycling

Nature of the solvent + contaminants present
Origin (cleaning, accident, production, unused)
The absence of nitro-cellulose, PCBs, pesticides, bacteria and radioactive substances should be established

WASTE MANAGEMENT IN ENTERPRISES

Management structure and organisation

Waste management requires continuous attention at various levels: the rapidly evolving legislation, reporting to general management, contacts with authorities, the press and the public. Attention must also be paid to internal measures, procedures, problems and external waste transport and elimination. An approppriate internal waste management system should be developed, to be directed by somebody with adequate authority, experience and training. High qualifications are necessary, especially where former practices were unsatisfactory. Problems should be tackled at their roots, for waste elimination becomes ever more expensive and restricted. For uncontrolled mixtures of waste there may be no legal outlet at all.

Changes in waste management systems should be prepared well in advance and preceded by sufficient collection, study and analysis of data as well as discussion with and information from the generating units concerned, in order to gain their advice and support. The implementation of new source separation or administrative follow-up systems are good examples of delicate undertakings relying on the close collaboration of all parties concerned to be successful.

Local waste management is based on the following information and organisational forms:

- dressing inventories;
- characterising waste streams;
- defining appropriate internal procedures for waste collection and elimination;
- defining appropriate procedures for external transport and elimination;
- prevention and reduction strategies.

In large factories a clear line of responsibility should be traced: who is responsible for (a) the waste generation, (b) intermediate storage, (c) centralised local storage, (d) external transport and elimination.

The transfer from the place of generation to central storage is also a transfer of responsibility. Hence, the proper characterisation on ‘ad hoc’ forms and labelling should be checked. Intermediate storage at the place of generation is a potential source of hazard and should be minimised in volume and time.

The environmental manager should study literature and legislation, consult authorities, talk with operation managers, their foremen and workforce, and check on the procedures followed by waste hauliers and disposal facilities.

Data acquisition

Several methods are available for characterising the waste streams generated in a factory or one of its facilities or shops.

Identifying waste streams according to source via the *purchasing department* is especially useful for monitoring and tracking dangerous and polluting products, ensuring their responsible packing and labelling throughout usage and disposal, and establishing a basis for *waste accountancy.*

Material balances can be established for any one production facility and the input of raw materials, solvents, lubricants, catalysts and other auxiliaries can be established on an hourly, daily and yearly basis, or per production batch. A comparison between these various inputs and the product output should allow characterisation of yield and efficiency, as well as the nature and amount of individual directed or fugitive waste streams. The latter are difficult to quantify and can often only be evaluated by determining the closure of the balances or by dedicated monitoring programmes.

Each division or shop should establish an inventory (Table 2.5) of all its emissions, ranging from fugitive ones (evaporation losses, small leaks from couplings, flanges or pumps, breathing losses of tanks, etc.) to both regular and fortuitous solid and liquid waste streams.

Waste management has become so vast and complicated that responsibility should be centralised. Specific but limited responsibility should be placed, however, with the local managers whose division generates the waste. They should ensure that all waste they produce is properly identified, quantified, packed, labelled and stored temporarily in minimal quantities at their premises.

A periodic *physical inspection* of all premises, facilities, backyards, containers and even of derelict land is necessary. Stores and inventories may hide unusable and useless products. Attention should be paid to 'hidden' waste streams, such as PCBs in transformers, or asbestos in thermal or acoustic insulation. Archives should be consulted.

Waste streams are often heterogeneous and regular visual inspection of waste containers is required to ascertain the absence of contraries or of recyclables. Some enterprises, such as the Mercedes car factory at Sindelfingen, have launched systematic analyses of all outgoing containers, whose contents are sorted into numerous streams quantified by weight. The data is used to prepare and implement waste reduction programmes. A merely visual inspection is almost useless, as even the trained eye cannot judge the composition of a waste heap by volume or weight. Plastic film, for instance, tends to float on top of other waste and convey the impression of omnipresence, even when its actual weight contribution is minimal.

Table 2.5 Data required when building a waste stream inventory

Type of waste (classification, nomenclature)
Quantity
Location of origin and of storage
Date and duration of production, local storage, central storage, transportation, acceptance and elimination at a waste facility

A second approach is based on historical and current data with respect to exiting and notified waste streams. The latter are classified according to prevailing nomenclature. Such waste classifications vary from one country to another, and may be based on individual authorised disposal routes (industrial waste, assimilated to domestic refuse; building rubble and demolition waste; industrial waste; toxic and hazardous industrial waste; waste oil; PCBs; solvents, chlorinated or not...). They may also serve the purpose of more detailed data gathering, that is systems to establish a territorial waste database.

Administrative procedures are needed in order to keep track of internal collection and external elimination, including legal documents such as notification forms, proof of transportation, acceptance and destruction in waste elimination centres or the payment of levies.

Data analysis

After sufficient data has been acquired, it can be analysed on the basis of questions like:

- Which raw materials are used?
- Are they hazardous before, during or after usage?
- Can consumption or waste generation be reduced?
- How are they eliminated?
- Is their elimination registered, ensured and lawful?
- Can they be recovered?
- If so, in what amount? At what level of purity?
- How good are contacts with authorities and the press?

Each waste stream should also be examined for the nature and number of its constituents. On the one hand, the collection of all waste in a single container is ideal in terms of the simplicity of the system and its logistics; on the other hand, the elimination of mixed waste is always more cumbersome than that of single components, and indeed is sometimes impossible. Individual components have a better potential for re-use or recycling. A mixture of waste also commands the highest specific disposal fee applied to all components present. Moreover, chemical waste may ignite or even explode, either directly upon mixing or after some unpredictable period of latency. As a rule the following streams should be kept segregated:

- liquid from solids;
- organic from inorganic;
- acids from bases;
- halogen rich from halogen lean;
- hazardous waste from general waste;
- individual hazardous from waste compounds.

Packaging and labelling are subject to specific regulations, based on safety and transport codes (ADR, 1959). Also, the training of truck drivers, technical standards, equipment, identification of trucks, maximum charge and categories of dangerous products (the joint loading of which is prohibited) are stipulated by ADR codes.

When a serious accident occurs it is often found that the enterprise concerned used chemicals whose usage, storage or disposal went unnoticed, or occurred to an extent far exceeding operating licences and indeed actual necessity. Such findings profoundly harm relations with the authorities and the public.

Source separation

Source separation is important in order to identify and retrieve recyclables and to avoid mixing wastes, which makes future identification impossible.

The following components deserve particular attention with respect to segregation of waste at source:

- waste paper and metals;
- packaging;
- canteen waste;
- demolition waste;
- hazardous waste.

Some of these materials are likely candidates for at least partial recovery. For legal, economic or safety reasons, others such as waste oil, chlorinated and non-chlorinated waste solvents and hazardous chemicals should be kept separate.Their recycling potential can be assessed by determining:

- which materials are available (m^3/day, tonnes/day);
- which unit loads and product specifications are acceptable to the trade in secondary raw materials (metals, waste paper or plastics).

It should be noted that 'waste recycling' is a contradiction in terms. Nobody buys waste. Source-segregated materials must satisfy numerous specifications, or be easily amenable to them, in order to have any value at all. There is a market for steel turnings and a bigger one for brass. But steel contaminated with brass is useless!

Education of the workforce and frequent inspection of containers (possibly placed or rented by the trade) are prerequisites for recycling. Fruit peel, cigarette stubs and general dirt should not find its way into containers earmarked for recyclables. This implies that a separate container for each kind of waste is required.

The provision and location of containers for source segregation of different waste streams, the logistics of internal transportation, storage and possible upgrading (e.g. solvent distillation) can only be studied on a case-by-case basis, because quantity, quality, markets, space availability and logistics differ.

Intermediate storage

Measures should be taken to ensure appropriate conditions for storage. For household refuse, canteen waste and other putrescibles, the emphasis should be on limiting storage time and on good housekeeping such as appropriate container selection, litter, odour and pest control and periodic extermination of rodents. Industrial and hazardous waste is often subject to specific codes and regulations related to storage, transport and fire-fighting. Some may be general, whilst others will be specific to site licence and operation. Waste storage areas should have restricted admission, be orderly, segregated storage areas, have a daily register and co-ordination measures for ensuring the periodic elimination of all entries.

Specification of the construction mode and of the storage vessels for materials are directly linked to the nature of the waste (e.g. provisions for heating or cooling, mixing, homogenising, venting of waste gas, overpressure relief valves) and the procedures for their transportation and elimination. The latter may involve stationary or mobile equipment, such as cranes, loaders, bulldozers, conveyor belts, chutes and centrifugal, membrane or moyno pumps. Equipment should be rugged, foolproof and designed for heavy-duty service.

Central storage of waste

Access to waste stores should be restricted, or even locked. A register should contain all the data relating to supply and elimination, and should be updated according to a pre-established schedule.

The storage area should have an impervious concrete floor, with drainage to a receiving pit that should not be connected to the sewer. Liquids should be stored in an enclosed area and the enclosure should be capable of retaining the entire volume of liquids in case of a spill. The height of piled drums should be limited, for example, to a height of two drums. Flammable and special waste should be shielded from direct sunlight.

With respect to leaks and fires it is necessary to separate flammable and non-flammable, toxic and reactive waste. Chlorinated waste and PCBs are barely combustible, but emit dangerous fumes when ignited. Provision should be made to retain the water used for fire-fighting, as became clear after the Sandoz fire which polluted the river Rhine.

Organisation of external elimination

On the basis of the amount and characteristics of the waste, quotations for the cost of disposal are requested, and waste elimination services contracted by the responsible manager. The manager also supervises the grouping, palleting and packaging of the waste, verifies the packaging, labels and transport documents (ADR) as well as the ADR-approved vehicle. Finally he or she co-ordinates the feedback of information and problems during transport or elimination with the waste-producing unit.

Disposal of small quantities of hazardous waste

Small quantities of hazardous waste are a source of specific problems because of the cost and the administrative and logistic difficulties associated with their identification, analysis, collection, handling, packaging and elimination. Typical unit loads such as a single truckload or even a single 200 L drum are rarely attained. There is a basic cost to be considered whatever the size or volume of the parcel or container. Other relevant factors are the loads':

- nature and hazardous character;
- frequency of occurrence;
- re-use and recycling potential;
- ease of identification;
- packaging and storage requirements;
- value on recovery or their cost of disposal.

Branch specific waste includes:

- laboratory waste;
- collected spills;
- residues from air, gas or water cleaning, e.g. collected scum, separated oil or fat;
- combustion residues;
- cleaning waste and residues in tanks.

Non-branch specific hazardous waste includes:

- cleaning products, aerosol cans, paints, coat ings, varnishes, lubricants and pesticides;
- electronic equipment, mercury thermometers, batteries, lamp bulbs, fluorescent tubes and fire extinguishers;
- office waste such as inks, toners, printing tape, glues, magnetic tape, colour markers;
- spent drugs.

Small quantities of hazardous waste are often eliminated together with ordinary waste, or even via drains, because of ignorance, the cost of more adapted elimination channels and the absence of administrative control. It is a serious challenge to establish proper administrative procedures and the necessary and cost-effective logistic and treatment systems to deal with these small waste streams. Possible options include:

- *specific centres,* where small amounts of chemi cals are initially accepted at little or no cost. Gradually, the dump fee is increased as people get used to the possibility and indeed necessity of handing over waste;
- *reverse distribution channels,* i.e. spent chemicals are accepted by the relevant trade (drugs, batteries, tyres,…), which collects the waste

materials and delivers them to waste elimina tion or recycling companies after grouping.

A sectoral approach, for example, in dry-cleaning shops, printing offices or professional paint services, can be used to convince small operators to convert to responsible waste management. Compulsory notification may pave the way for ensuring compliance with legal requirements. Also, trade associations play a vital role in disseminating information on regulations.

TECHNICAL FACTORS IN WASTE MANAGEMENT

Logistic factors (Mantell, 1975; Jünemann, 1991, 1993, 1995)

Waste types

Waste can be classified according to a variety of criteria. In practice, the following classifications are frequently used:

- Household refuse and other urban waste collected at the same time.
- Trade and industrial waste, more or less assimilated to the former.
- Bulky waste.
- Building rubble and demolition waste, i.e. inevitable by-products, residues from effluent end-of-the-pipe treatment, spent catalysts and adsorbents, etc.
- Industrial waste.
- Toxic and hazardous waste.
- (Waste) water treatment sludge cakes.
- Hospital waste.

Definitions, (sub-)divisions and specifications vary from one country to another depending on local codes and their interpretation by inspectorates and waste disposal authorities.

Waste receptacles

A suitable receptacle should be provided at all locations where waste may arise, for example, in a kitchen, workshop, printing office or on a picnic site. The receptacle may have different forms, such as a bag, a bin, a bucket or a drum, and may be open or closable. To discourage littering, it should be conspicuous and strategically located. Its design, shape, volume and material of construction should be adapted to the waste, its origin and to the activities taking place. In a kitchen, for instance, a pedal-operated bucket is useful so that one's hands are free to dispose of the waste without touching the receptacle.

The accumulated waste should be removed periodically for hygiene and safety reasons, according to a schedule based on frequency and rate of waste generation, the volume of the container and the maximum allowable storage time or volume.

The waste should be collected regularly and grouped to form larger unit loads. This involves a variable number of transfers and intermediate storages.

Waste collection

The haulage of trade and industrial waste was traditionally the business of small family enterprises. The activity often developed as an offshoot of transport capabilities, such as a farmer's tractor, a coal merchant's truck, or a trade in bulk building materials. The activity was often synergistic with quarrying, which provided the necessary empty volume for landfill.

Over at least the last ten years there has been a strong tendency towards centralisation. Large firms from the service industries have taken over family enterprises and their operations, and the original entrepreneurs have been scared away by the increasing administrative and legal complexity, or by the required investments and the liability factors of modern waste management. Waste haulage is increasingly undertaken by companies affiliated to suppliers of electric power, district heating, gas or drinking water, and to companies in industrial cleaning and wastewater treatment services.

The service fee required for hauling and disposal of waste includes:

- the distance of transportation between the point of collection and that of disposal;
- the container rent;
- the disposal cost, augmented by levies and taxes, as expressed per unit load, ton or m^3;
- the level of service to be rendered.

The last varies according to frequency of elimination, ease of access to the premises and whether the schedule for exchange is fixed by the waste supplier or the haulier. The waste generator is usually responsible for maintaining the containers in a good condition and ensuring that inappropriate waste is not deposited in the containers.

Waste containers must be inspected regularly for unauthorised components such as oils, paints or chemicals. Once identified, these create their own problems, as either the container will be returned to its place of origin, or the contents of the container will need extra sorting or treatment and a supplemental fee will be charged for subsequent inspection, sorting, transport, treatment and administration.

Landfilling (Patrick, 1985; EC, 1993; Carra and Cossu, 1990)

Scope

Historically, landfill has been the major and often the only waste elimination option; this is still the case in Ireland and Greece, among others (OECD, 1993).

A landfill (or tip) is a place where refuse is tipped. When the operation is well-engineered it becomes *controlled tipping* (UK) or *sanitary landfill* (USA). This technique is still evolving, mainly in terms of its isolation by a combination of barriers, drainage and treatment of internal gas and leachate flows, and final integration into the landscape.

Crude forms of landfill create numerous problems, such as:

- utilisation of space, which becomes unusable for most forms of subsequent land use. Moreover, when the nature of waste and the method of land fill remain undocumented, subsequent use is af fected by numerous uncertainties and hazards;
- windblown litter;
- odour problems;
- spontaneous and other fires (people who scav enge in the fills may make a fire to remove plastic coatings and rubber from reclaimable metals; fires may be fuelled by landfill gas and continue for months or years beneath the tip surface, creating cavities and sudden subsidence and tip instability);
- insects, rodents and birds;
- pollution of aquifers and surface waters by leachates;
- acoustic hindrance;
- traffic problems, road litter and mud tracks on roads.

Those problems can largely be overcome by applying a Code of Good Practice, such as that reported by Patrick (1985).

Site selection

Site conditions differ widely in location, climate and scale of operation and in particular (hydro-) geological conditions.Sites should be selected as a function of factors such as:

- land improvement potential;
- local space planning;
- health protection;
- future land use (an architectural model of the completed site may help reduce opposition);
- groundwater and surface water protection by se lection of a suitable hydrogeological configura tion and implementation of appropriate measures;
- road access for waste hauling traffic, avoiding in habited areas;
- absence of airports (birds are a hazard to aircraft engines);
- distance from the waste-generating centre;
- availability of cover material.

It should be remembered that a landfill site is a source of heavy traffic and noise.

Operating standards (Patrick, 1985)

To form a landfill with a minimum of nuisance and achieve satisfactory standards of land reclamation, the following principles are recommended:

- Concentrate disposal at one area with one work ing front at a time, unless distinct types of waste are to be tipped.
- Deposit waste either at the top or at the lower side of a ramp. It should then be sprayed and formed into a layer using mobile plant with a blade or bucket and the machine should make several passes over it. The angle of the slope should not

Table 2.6 Parameters and other considerations in siting a landfill

- Hydrogeological analysis and risk assessment with respect to water pollution
- Minimum available area: e.g. 50,000 m^2
- Minimum available volume: e.g. 250,000 m^3
- Geometry: filling a cavity or a valley
  creating an artificial hill
- Distance from the waste-generating centre: maximum 30 km or 1 hour driving (for large distances the creation of a transfer station with hauling for barge, train and large trailer vehicles should be contemplated)
- Itinerary: over highways and through industrial centres exclusively. No crossing of villages or residential areas, or use of narrow roads or opening bridges
- Ecological value of the site (rare plants, nesting birds)

exceed 30° to the horizontal. Optimal compacting is obtained by using very heavy (15 to 60 tons) steel wheeled compactors operating on layers not exceeding a thickness of 0.5 m. Thicker layers preclude proper compacting, as does tipping from a height into a precipice.

- Bulky waste should be crushed or broken down to prevent the formation of pockets. For that pur pose it can be pushed upwards over the working face (see Figure 2.1).
- The depth of the initial compacted layer should not exceed 2.5 m.
- The width of the working face should be suffi cient to allow vehicles to discharge without queu ing or interfering with the activities of the bull dozer or compactors.
- Movable screens can be set up to collect wind blown paper and plastic film.
- A layer of at least 15 cm of cover material (earth, foundry sand, cinder) should be applied daily to ensure a tidy appearance, prevent odours, discour age vermin and cover fly eggs.
- Solid waste and cover material should be inclined, to promote rainwater runoff.
- Hospital, animal and fish waste, old food and hospital waste are deposited in front of the work ing face, so that a relatively deep daily cover is ensured.
- Tipping in water as a rule is unacceptable.

When the full filling height is reached, a final cover should be applied. There is a modern tendency to

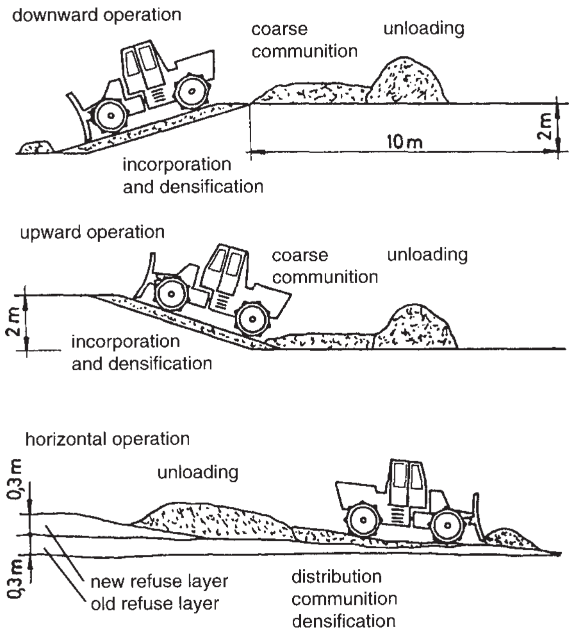


Figure 2.1 Tip operation. Refuse is loaded and the resulting heap is distributed, levelled and grossly communited by, for example, compressing plastic, metal and glass containers. Then the material is densified by repeatedly rolling over it.
Source: Junemann, 1991

cover the tip with impervious liners, sandwiched between drainage layers, themselves separated from soil and waste by geotextiles. On top of liners and superimposed drainage, an intermediate layer should be put down for root development, followed by a layer of top soil. Often, base and side drainage structures, and the provision of intermediate and final cover layers markedly reduce the available fill volume.

Acceptance at landfill sites may depend on prior elution tests. Table 2.7 sets out the parameters proposed in a Draft EC Directive.

Site survey and operational plan

In national planning it is essential that land is allocated for the final disposal or storage of solid waste. Indeed, many types of waste can only be disposed of on land. Moreover, landfill can theoretically make a positive contribution through the reclamation of derelict land and quarries, and by the formation

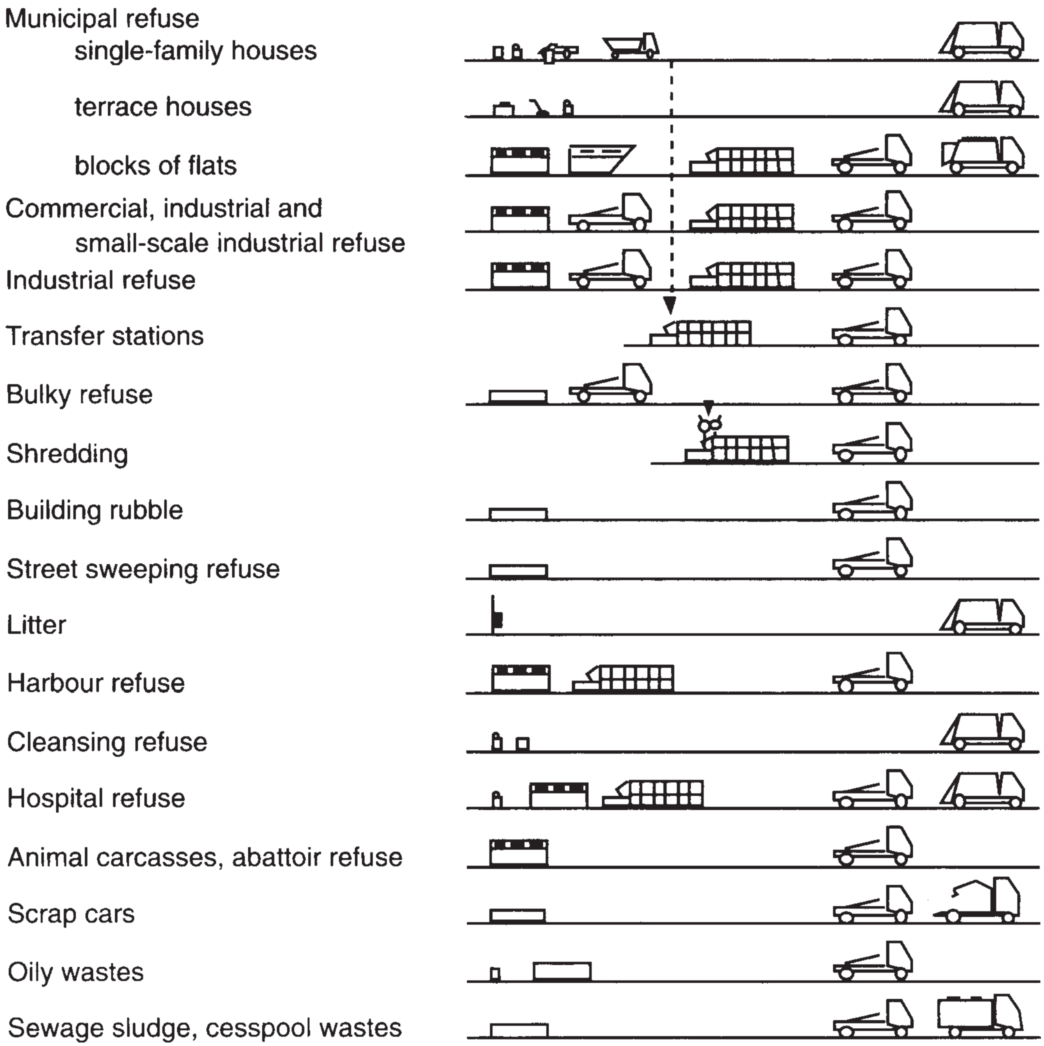


Figure 2.2 Various activities in municipal waste management
Source: Bridgwater and Lidgren, 1981

of new landscape features (useful as golf courses, fun fairs, motor cross tracks or ski ramps). A useful application may alleviate public opposition to the project and the NIMBY (not in my backyard) and NIMTO (not in my term of office) syndromes (priorities tend to change over time). In practice, however, the positive contributions seem to be severely limited. Golf course developers generally favour areas with preexisting natural value over the upgrading of used waste deposit sites.

An official name should be selected and a contour map showing the final levels in 2 m spans should be prepared. The total available volume should be calculated and shown, taking into account initial compacting and final cover material. The nature of the strata underlying and adjacent to the site should be surveyed and their permeability evaluated. Rainfall should be recorded and studied,

Waste should not be deposited in water because this leads to putrefaction and unpleasant smells.

Table 2.7 Parameters for elution tests in the proposed EC Directive on landfill of waste

pH	Fluoride
TOC	Ammonium
Arsenic	Chloride
Lead	Cyanide
Cadmium	Sulphate
Chromium (VI)	Nitrite
Copper	Adsorbed organically-bound halogens
Nickel	Chlorinated solvents
Mercury	Chlorinated pesticides
Zinc	Lipophilic substances
Phenols	

Instead, dams should be built and the water pumped out of the fill so that tipping is completed under essentially dry conditions.

During the planning stage, a land or aerial survey should be made and a scale plan prepared. The topographical and hydrological features should be shown, together with groundwater levels and volumes. The volume of available cover material should be estimated and recorded. Prevailing wind directions should also be recorded. Watercourses and drainage water from surrounding higher land must be diverted or culverted before operations commence.

In some countries systematic land reclamation from the sea is a viable option. After all, tidelands around the Bay of Tokyo have been filled, and artificial waste islands contained by concrete dikes have been constructed to accommodate refuse tipping. Tables 2.5 and 2.6 show some useful data and parameters.

Table 2.8 Content of a site specification plan

- All site details
- Landfill plan
- Required plant and equipment
- Types of waste to be excluded
- Measures to control pests and weeds
- Sources of cover material and storage locations
- Leachate and landfill gas collection arrangements
- Water monitoring arrangements
- Final restoration requirements
- Measures to be taken in case of bad weather (frost, prolonged rainfall)
- Storage of topsoil and subsoil

Table 2.9 Preparatory works and facilities required

- Identification and information boards
- Access and site roads, fences and gates
- Screening by earth dams, trees and shrubs
- Water, electricity, telephone services and lighting
- Wheel cleaning (scrapers, rinsers, with clarifiers for dirty rinsing water)
- Movable screens to contain windblown litter
- Containers for public use
- Weighbridge(s)
- Control centre, employee welfare facilities
- Services and stores (tools, equipment), fuel storage, fire fighting appliances

Typical equipment for a landfill site (Patrick, 1985)

Listed below are the typical equipment requirements of a landfill site:

- Tracked bulldozers and loading shovels, which may be fitted with dozing blades or buckets. These are versatile machines suitable for most landfill operations.
- Steel-wheeled compactors, generally with articulated steering and mounted on steel wheels fitted with hardened steel teeth or wedges which tear and crush the waste. These are fitted with dozing blades to spread the waste and are single-purpose machines.
- Wheeled loading shovels mounted on pneumatic tyres are suitable for handling waste and for general work on smaller landfill sites. They operate faster than tracked machines and can readily be moved from one location to another. They have comparatively poor traction when pushing heaps of material on the landfill. Special puncture-resistant tyres are needed for machines for landfill use.
- Scrapers, towed or independently powered, are used for excavating and spreading covering material. They are generally required only for part-time use, except on the largest sites.
- Draglines provide an alternative to scrapers for excavating and covering material, but are only economical for very large-scale operations, e.g. creating a borrow pit.
- Pneumatic-tyred dump trucks are useful for on-site work in moving material and equipment.

Mobile plants used on landfills are subject to special hazards arising from the nature of the material to be handled and should be equipped with special protective devices, such as:

- radiator protection and pusher fan to facilitate removal of dust and litter;
- large-sized air cleaner;
- machine guards to protect engine and transmis sion;
- extended dozer blade (landfill blade) to deal with light and bulky waste;
- reinforced cab, including roll-over safety device;
- air conditioning in taxis (in hot climates);
- front and rear towing hooks.

Leachates

Leachate is a major problem in landfill operation. Because its polluting potential is high and its treatment is difficult, the best policy is to minimise its occurrence, *inter alia,* by providing a natural or artificial impervious bottom and side layers.

Leachates originate in the following ways:

- From the percolation of rainfall through the cover layer. Leachate formation is reduced by provid ing an inclined slope for the last cover, using a relatively impermeable top layer (sandy clay or clayey sand is optimal) and catering for maximal evapotranspiration by plants, shrubs or trees (with a shallow rather than a deep root structure). Plas tic liners, as long as they remain intact, completely halt infiltration and are specified more and more often.
- From a high water table, so that groundwater rises into the waste deposit. This problem can be checked by hydraulic means, such as the creation of impervious bottoms and sides or lowering the groundwater table by selective pumping. The pre ferred measure is prevention through suitable site selection.
- From lateral ingress of groundwater or springs. Lateral impervious liners and groundwater culverting are possible measures.

Household refuse has a fairly large moisture absorption capacity. Due to capillarity, water also tends to rise above surrounding areas in a refuse pile. The presence of moisture and free water accelerate the degradation of refuse, which proceeds anaerobically, with any oxygen included being rapidly depleted. Sometimes degradation is artificially initiated *(Rotte deponie)* or stimulated by recirculation of leachate into the landfill mass.

Leachate control requires various measures, such as:

- construction of a drainage system around the site to intercept and divert surface runoff from adja cent ground;
- construction of tipping bays to isolate clean wa ter from leachate;
- sealing the surface of filled areas with material of low permeability;
- providing suitable gradients to allow runoff;
- soiling and seeding of completed sections of the fill.

Liners of synthetic membrane, protected against damage with a layer of sand, are used at the base and/or the sides of the site. A drainage system is provided by grading the base of the site to allow leachate to flow to collection points or drains.

The drainage system consists of a ramified network of gravel ducts (sometimes shredded tyres) housing perforated PVC or HDPE tubes to drain the various areas of the fill towards a number of central points. The latter are accessible through a set of concrete vertical rings, steel tube or even a telescopic plastic construction that allow for later subsidence. These vertical leachate pumping wells are also subject to strong lateral forces due to uneven subsidence.

Leachate can be recirculated and sprayed over completed sections of the landfill or adjacent grasslands. Another option is biological treatment in an oxidation pond or a sewage treatment plant. Breakdown of BOD is fast, but COD is quite immune to treatment. Physico-chemical treatment methods, such as coagulation/flocculation, concentration by evaporation or reverse osmosis, are increasingly in use.

Water pollution by landfills

The water flow in a horizontal porous layer is described by the Darcy equation:

$$Q = k \cdot i \cdot S$$

with
- Q = water flow (m^3/s)
- S = cross-section of the layer (m^2), perpendicular to the flow direction
- i = hydraulic gradient

$$= \frac{\text{height of the water table}}{\text{width}} \ (m/m)$$

- k = hydraulic conductivity or permeability (m/s)

the value of k varies over many orders of magnitude, from

k = 1 cm/s for gravel
0.01 cm/s for sand, to
10^{-6} to 10^{-8} cm/s for clay.

The flow of vertical water infiltration is usually expressed as a fraction of the annual rainfall. The latter is typically 0.7–0.8 m/year in Western Europe, 0.2–0.3 m/year in arid regions, and several metres in tropical climates.

A large part of the rainfall runs off and evaporates directly or via plant transpiration (evapotranspiration). Only a fraction of the annual rainfall (25–50 per cent) penetrates the soil.

The pollutants follow the general flow of groundwater, that is advection, with some minor spreading because of diffusion. Numerous compounds are, however, retarded or attenuated by mechanisms such as:

- adsorption on clay minerals or organic humus;
- precipitation, immobilisation;
- biodegradation.

Mathematical modelling has become a standard procedure. It is based on the integration of the equations of continuity (mass conservation) and motion. Adjustable parameters are the dimensions and porosity of the water-bearing soil layer. Data on adsorption equilibria and biodegradation rates are becoming increasingly available.

Landfill gas generation (Constant et al., 1989)

Oxygen is rapidly depleted in a landfill and anaerobic fermentation occurs at a rate that is highly variable according to the type of waste and local conditions (moisture, temperature, pH value, presence of toxic materials). Gas is produced, approximately in proportion of 50–65 per cent CH_4 and 35–50 per cent CO_2 Trace contaminants include solvents, volatile acids and chlorinated organics, which render landfill gas corrosive to steel, concrete and engine parts.

Fermentation gas forms explosive mixtures with air, with lower and higher explosion limits of 5 per cent and 15 per cent by volume of methane in air. The gas is produced under over-pressure and follows a path of least resistance out of the fill; since refuse should be spread and compacted into slightly inclined layers, migration over prolonged horizontal distance is possible. Gas may accumulate in closed structures and pose a safety hazard, or suffocate trees, shrubs and grass by displacing soil air. The gas may be contained by means of membranes or other impervious liners, or be exhausted by suction venting. Spontaneous egression may increase dramatically with a sudden decrease in atmospheric pressure; this phenomenon caused an explosion inside a Danish house built on a landfill.

Venting is achieved by providing a gravel-filled trench around the sides of the site, interposed between the waste and the surrounding ground. Alternatively, gravel-filled boreholes with a vertical perforated pipe at the centre may be used to collect any gas.

Landfill gas is moist and tepid, and allowance should be made for the separation of condensation water from transfer lines. The gas can be upgraded to pipeline quality, either by compression and scrubbing out of CO_2 with water, or by use of pressure swing adsorption on solid adsorbents. The Palos Verdes landfill near Los Angeles was an early example of gas purification (1975), but this technique is limited to very large landfills with a connection to a nearby natural gas pipeline.

Typical local uses for landfill gas include fuelling:

- brickmaking furnaces;
- greenhouse furnaces;
- internal combustion engines (corrosion problems have besieged this application, however).

Typically, 250–400 m^3 of landfill gas is generated from 1 ton of household refuse, which declines exponentially over a ten-year time span. This, however, is no more than a rule of thumb. For instance, in a dry climate no degradation at all may occur.

Conclusions and outlook

Landfill is an unavoidable option in waste management, especially in countries with low budgets or waste with too low a calorific value to sustain combustion.

There is a tendency to treat waste as an engineered enclosure, in which waste is stored separated from the biosphere for geological times, and all effluents (gas, leachate) are treated; there is also a tendency to reduce the amounts to be landfilled by source reduction, enhanced re-utilisation of waste, and by prohibiting the tipping of non-mineralised waste.

In developing countries, siting is of the utmost importance to ensure that landfill has a low pollution potential despite having a low budget.

Composting (Brunt *et al.,* 1985)

Scope

Composting is an aerobic biological decomposition process in which a stabilised product containing humic substances is synthesised, while some of the organic material is oxidised to carbon dioxide and water. The heat generated by biological activity raises the temperature to a range of 55–70°C, in which only thermophilic micro-organisms can survive. Pathogens and weed seeds are therefore destroyed.

In recent years anaerobic composting processes have been developed, in which organic waste is digested in a closed tank, generating fermentation gas as well as a compost, for example, the DRANCO (DRy ANaerobic COmposting) system. Earlier attempts (Pompano Beach, Fla.; Valorga, Reims, France) were unsuccessful.

Compostable materials

In principle, all organic waste can be composted, as long as it is reasonably free of biocides or salts. Often, various substrates are mixed prior to composting because of their complementary compositions and structures. Typical substances are:

- the organic fraction of municipal solid waste, sepa rately collected or mechanically separated;
- vegetable, fruit and garden waste;
- straw, bark, sawdust, branches;
- animal dung;
- sewage sludge;
- agro-industrial waste.

The resulting compost is a kind of humus, useful as a soil conditioner or even as animal feed supplement.

Depending on location and climate, some 350–500 kg of compost can be produced per ton of municipal solid waste with between 150 and 250 kg of moisture and organics lost by evaporation and oxidation. The remainder consists of non-compostable materials, such as metals, glass, stone, plastics, textiles, rubber, wood or leather, which must be separated before or after composting.

Microbiological aspects

The composting process is carried out by naturally occurring micro-organisms which grow spontaneously in organic waste if it is kept moist and aerated. In the first stage of composting, mesophilic bacteria as well as actinomycetes, yeasts and fungi break down fats, proteins and carbohydrates. Later, protozoa prey on the bacteria and fungi. As the temperature reaches a range of 40–50°C, most of the organisms that initiated composting are killed and replaced by thermophilic bacteria, which can grow and generate heat up to a temperature of 70°C. At 60–70°C most pathogens, with the exception of a few spores, are killed in a matter of few hours. When the thermophilic bacteria have exhausted the available food, they stop producing heat and the compost cools off. Turning the heap may supply the required food and oxygen for successive new temperature rises. Eventually, new types of organisms dominated by fungi and actinomycetes grow on the residual food, including dead bacteria, and give the compost its final

Table 2.10 Microflora in compost

	Typical number per gram(*)
Bacteria	10^8–10^9
Actinomycetes	10^5–10^8
Fungi/ moulds	10^4–10^6
Algae	10^4
Protozoa	10^4–10^5

(*) the microflora varies as the compost evolves towards maturity ('cured' compost) properties and composition, that is an indigestible, non-odorous, humus-like residue.

Recently, closed anaerobic systems have been developed for similar applications; in addition to a humic type of substance, these also generate a combustible fermentation gas composed of CH_4 and CO_2.

Composting systems

There is an enormous variety of composting systems in use or proposed for use. The following basic operations can be combined in numerous ways:

- Prior removal of non-compostables such as glass, metals, plastics, etc. using a manual sorting belt.
- Size reduction, using a flail or hammer mill, or by attrition upon tumbling in a rotary composting drum or grinding in a large rasp.
- Accelerated composting, using a combination of suitable operating conditions (aeration, moisture, temperature, C/N control) in a single or multifloor tower, a rotary drum or windrow.
- Curing, i.e. maturing of grey or brown unripe compost in a windrow storage area.
- Sizing the compost in various screening fractions on rotary or vibrating screens. • Separation of non-composted materials, such as stones, glass fragments, plastics, by means of stoners, secateurs, air tables or ballistic separators.
- Conditioning (in bulk or bags) prior to sales, possibly with the addition of fertiliser or peat.

Operating parameters

The main parameters of the composting process include average particle size, moisture content, pH value and the C/N ratio of the substrate. Other essential elements, such a potassium, phosphorus and trace elements are sufficiently available and generally not rate-limiting. Aeration, mixing or turning, size reduction, temperature and pH control are also important in producing compost.

Uses for compost

Direct application of waste on to the soil is undesirable, as it creates a nitrogen deficiency. Immature compost is sometimes used for its heat-generating capacity to force early growth in greenhouses. Compost also binds sandy soil and aerates heavy clay. Worked into the earth, it stimulates plant growth, absorbs moisture and controls both erosion and dusting. Much compost is used in vineyards. In Western Europe, compost is mainly applied during two seasons (February to April; October to December) when no crops are grown. Typically 50 tons/ha (5 kg/m^2) are applied.

When used in flower gardens, it is important that compost is free of pathogens, glass splinters and plastic fragments. Compost for growing vegetables or mushrooms should be low in heavy metals, which practically excludes the use of compost produced from household refuse for such applications.

Compost can also be utilised for animal bedding and even as a fodder supplement, and large amounts of it can be used in land reclamation projects, in public gardens and parklands.

Conclusions and outlook

Composting is a valuable method for recycling organic waste and producing a soil amendment. It is of vital importance in fighting soil erosion and improving soil structure and aeration. MSW composting, however, has seen many more failures than successes, as a consequence either of poor technical operation or of insufficient markets. Moreover, today it is considered that MSW-derived compost is of unacceptable quality.

Composting was promised a new future in the 1970s, with the advent of mechanical separation of refuse into dry recyclables and a moist vegetal fraction. Today, the latter is often separately collected, which paves the way for composting at a large and widespread scale, so that developing outlets and marketing again become key factors.

Incineration (Buekens and Patrick, 1985)

Scope

Incineration is defined as 'the controlled burning of solid, liquid or gaseous combustible wastes so as to produce gases and residues containing little or no combustible material' (Patrick, 1980). The process is well controllable in space and time and leaves at most an unrecognisable and sterilised residue. Incineration is widely used in densely populated regions (Japan, Switzerland, the Ruhr area).

Incineration has a number of other important benefits:

- A large reduction in volume, often by a factor of 10 or more.
- Objectionable and hazardous properties, such as putrescibility, flammability, explosivity, toxicity, persistence or pathogen count, can be eliminated completely.
- The heat generated may be put to good use.

On the other hand, incineration is technically a complex process, involving large investment and operating cost as well as good technical skill in maintenance and plant operation in order to conform to modern standards. Flue gas and residue are non-objectionable only when a complete burn-out is ensured. Heat recovery is often inefficient. It generates a huge number and amount of pollutants which are difficult to control. As emission codes become more stringent, operating costs rise and the volume of secondary waste streams increases requiring further disposal with gas scrubbers.

Waste combustion

The combustion of waste is relatively simple when it comes to off-gas, viciated air, or pumpable liquid streams. These can be burnt in simple, stationary combustion chambers operated under proper conditions of temperature, time and turbulence (the 3 ts) to ensure a complete burn-out.

Municipal solid waste is generally burnt on a mechanical grate situated in a combustion chamber that allows the free development of flames. The chamber is refractory-lined and thermally insulated to limit heat losses. When firing high calorific waste, the chamber may even be externally cooled, either using air-cooled side walls or by incorporating the entire combustion chamber into the boiler structure. Sometimes solid waste is burnt in a fluidised bed, a cyclonic or vortex furnace, or as a fuel supplement in a utility boiler or cement kiln. Since industrial waste shows a broad variation in size and consistency, a rotary kiln furnace, followed by a post-combustion chamber, is the usual choice.

Incinerator furnace design is a complex art; it is influenced by various duties and constraints to ensure the following:

- An acceptable operating temperature, generally in the range of 850–1050 °C.
- A suitable residence time, often 2–5 seconds for the flue gas; the residence time required for sol ids depends on their size and on the system se lected. It is of the order of 0.2–1 hour, unless the solids are finely divided or pulverised prior to combustion.
- A proper distribution of primary air (fed through or with the fuel), secondary air (to promote mix ing of the highly viscous combustion gas) and tertiary air (to reduce the temperature of the flue gas).

Furnace design should cope with thermal as well as draft requirements and suitable flow and mixing patterns should also be selected. The design should relate to the heat duty of the furnace and the fuel properties of the waste, as summarised in Table 2.11. Also, the amount and melting behaviour of the ash is a major factor in furnace design and operation.

Environmental aspects (Buekens and Schoeters, 1986)

Refuse incineration gives rise to dirt and grit, to products of incomplete combustion (PICs) and to acid gases. Still, modern incinerators eventually provide extremely clean flue gases by a combination of preventive and operational factors and gas treatment.

Some types of waste are banned from incinerator plants, unless they are specifically equipped to cope with such waste. Two practical examples are:

- volatile metal (i.e. mercury, thallium and cad mium) bearing waste;
- PCB-containing waste, which requires special incinerators with an unusually high destruction efficiency.

Table 2.11 Fuel properties of waste

1 Lower and higher heating value, MJ/kg
2 Short analysis:
  - moisture fraction;
  - ash fraction;
  - combustible fraction.
3 The combustible fraction is further subdivided into:
  - volatile matter;
  - fixed carbon.
4 The elemental analysis of the combustible fraction
  - C, H, O, N, S, Cl, Br, F, ...
5 The composition of the ash fraction, with as main components
  - SiO2 having an acidic reaction
  - CaO, MgO having a basic reaction

  whereas borates, phosphates, vanadates and sodium and potassium salts lower the melting point and increase the tackiness of the ash.
6 The behaviour upon combustion:
  - drying;
  - melting;
  - evaporation;
  - thermal decomposition, yielding volatiles and fixed carbon;
  - gasification of fixed carbon;
  - partial oxidation of volatiles;
  - combustion and post-combustion of refractory substances.
7 Other relevant parameters are:
  - ease of ignition;
  - quality of burn-out at the side of flue gas (halogens inhibit, water vapour promotes) of the residue;
  - melting behaviour of the ash and its interactions with refractory and metallic construction materials.

The main operational factors are:

- the mixing of the waste stream to obtain relatively homogeneous and constant characteristics;
- controlling the operating parameters (temperature, draft,...) relatively close to their set points.

Finally, the flue gas is cooled and treated to remove:

- dust and grit by means of multicyclones, electrostatic precipitators or baghouse filters;
- acid gases by means of wet (water), semi-wet (lime slurry) or dry (lime or soda bicarbonate powder) scrubbing.

Mercury can be arrested by means of acid scrubbing or adsorption on to porous particles of activated carbon often impregnated to collect mercury.

Combustion always remains somewhat incomplete, and as a consequence carbon monoxide, organic products of incomplete combustion (PICs), including carcinogenic polycyclic aromatic hydrocarbons (PAHs) and *dioxins* are formed. However, tremendous progress has been made in reducing the carbon monoxide emissions and PAHs by a better furnace design using higher operating temperatures as well as longer residence times. Dioxins, more precisely polychlorodibenzo-p-dioxins (PCDDs) and polychlorodibenzofurans (PCDFs) are mainly formed during the cooling stage by catalytic reactions, involving either phenolic precursors or residual carbon. They adsorb on to both fly-ash particles and external

Table 2.12 Emission codes: a comparison of current codes

Compound	*EC*	*Germany*	*Holland* (mg/m^3 at s.t.p.)	*Flanders*	*EPA*
Dust	30	10	5	10	30
HCl	50	10	10	10	29
HF	2	1	1	1	
CO	100	50	50	50	
SO_2	300	50	40	50	61
NO_x	300	200	70	100	175–293
Pb+Cr+Cu+Mn	5	2.5	1	1	
Ni+As	1	0.5	incl. (Pb+ . . .)		
Cd	0.2 (Cd+Hg)	0.1 (Cd+Hg)	0.05	0.1	
Hg	incl. (Cd)	incl. (Cd)	0.05	0.1	
Dioxins T.E.Q. (ng/Nm^3)	–	0.1	0.1	–	5–30 (total)

Sources: EC: 89/369/EEC; Germany: 17Blm Sch V; Holland: Richtlijn Verbranden N.L. VROM 1989; Flanders: Ontwerpnorm OVAM 1991; EPA: Clean Air Act 1989.

adsorbents injected into the flue gas, or supplied as a final carbon adsorption bed.

Soil contamination by heavy metals has been shown in the neighbourhood of several incinerators sometimes to the point of making the soil useless for grazing cattle and growing vegetables.

CONCLUSIONS

Solid waste management is a service activity, designed to deal with the collection, treatment and ultimate disposal of waste in a healthy and economical manner that is not at odds with the interests of the environment. Since the 1950s, waste generation rates have been inflated by various factors, such as the introduction of disposables and of pre-packed goods. Moreover, increasing and widespread prosperity has accelerated domestic consumption and industrial production and, consequently, the rate of depletion of natural resources as well as their conversion into waste and other forms of pollution.

Ultimately, all wastes from society end up in one of three environmental depositories: air, water or land. Historically, air and water pollution were first regulated on the basis of emission limits and discharge controls on effluents, whereas the necessity to regulate and control solid waste streams was recognised much later. Prevention, reduction, re-use and recycling of waste are increasingly put forward as logical and cost-effective ways of reducing waste streams. The implementation of these concepts is, however, a slow process, as it calls for profound reorganisation of both industrial production concepts and individual consumption patterns.

Past waste disposal practice has led to the contamination of numerous sites whose cleaning will cost astronomic sums. The level of acceptable residual risk and the required cleaning standards are difficult to define. The problem of liability for such sites has also raised complex legal issues which have been tackled differently in various countries.

REFERENCES

ADR (1959) 'European Agreement concerning the international transport of Dangerous goods by Road', cited in United Nations Committee of Experts on the Transport of Dangerous Goods (1977) *Transport of Dangerous Goods,* Chicago, Illinois: United Nations.

Alter, H. and Dunn, J.J.Jr. (1980) *Solid Waste Conversion to Energy,* New York and Basel: Dekker Inc.

Bocken, H. (1987) 'Black Points, Liability for Clean-up Costs', *Resources and Conservation* 14: 105–118.

Bridgwater, A. and Lidgren, K. (eds) (1981) *Household Waste Management in Europe,* New York: Van Nostrand Reinhold.

Brunt, L.P., Dean, R.D. and Patrick, P.K. (1985) 'Composting', in M.J.Suess (ed.) *Solid Waste Management,* Copenhagen, Denmark: WHO Regional Office for Europe.

Buekens, A. (1992) *Recycling,* Brussels, Belgium: Vrije Universiteit Brussel.

Buekens, A. and Nieuwejaers, J.B. (1987) 'Black Points, Detection, Evaluation and Control', *Resources and Conservation* 14: 39–69.

Buekens, A. and Patrick, P. (1985) 'Incineration', in M.J.Suess (ed.) *Solid Waste Management,* Copenhagen, Denmark: WHO Regional Office for Europe.

Buekens, A. and Schoeters, J. (1984) *Thermal Methods in Waste Disposal, Part I: Pyrolysis and Gasification, Part II: Incineration,* Brussels, Belgium: Vrije Universiteit Brussel.

Buekens, A. and Schoeters, J. (1986) *Environmental Aspects of Refuse Incineration,* Brussels, Belgium: Vrije Universiteit Brussel.

Carra, J.S. and Cossu, R (1990) 'Introduction', in J.S.Carra and R.Cossu (eds) *International Perspectives on Municipal Solid Wastes and Sanitary Landfilling,* London, UK: Academic Press, pp. 1–14.

Constant, M., Naveau, H., Ferrero, G.E. and Nyns, G.J. (1989) *Biogas End-use in the European Community,* Barking, Essex, UK: Elsevier Science Publications.

EC (1993) 'Draft Landfill Directive'.

ECE, Economic Commission for Europe, United Nations (1978) *Non-waste Technology and Production,* Oxford, UK: Pergamon Press.

Hatayama, H.K., Chen, J.J., de Vera, E.R., Stephens, R.D. and Storm, D.L. (1980) *A Method for Determining the Compatibility of Hazardous Wastes,* NTIS, EPA-600/ 2–80 –076.

ICEP 3 (1996) *Environmental Pollution,* London, UK: European Centre for Pollution Research.

Jünemann, R. (ed.) (1991) *Entsorgungs Logistik I: Grundlagen, Stand und Technik,* Berlin, Germany: Erich Schmidt Verlag GmbH & Co. (in German).

Jünemann, R. (ed.) (1993) *Entsorgungs Logistik II: Entwicklung und Bewertung neuer Konzepte und Technologien,* Berlin, Germany: Erich Schmidt Verlag GmbH & Co. (in German).

Jünemann, R. (ed.) (1995) *Entsorgungs Logistik III: Kreislaufwirtschaft,* Berlin, Germany: Erich Schmidt Verlag GmbH & Co. (in German).

Mantell, C.C. (ed.) (1975) *Solid Wastes. Origin, Collection, Processing and Disposal*, New York: Wiley.

Meadows, D. (1972) *The Limits to Growth—A Report for the Club of Rome's Project on the Predicament of Mankind*, New York: Earth Island.

Meadows, D.H., Meadows, D.L. and Randers, J. (1992) *Beyond The Limits*, London: Earthscan Publications.

OECD (1993) *Environmental Data Compendium*, Paris, France: Organization for Economic Co-operation and Development.

Patrick, P.K. (1980) *Glossary of Solid Waste*, Copenhagen, Denmark: WHO.

Patrick, P. (1985) 'Disposal on Land (Model Code of Practice)', in M.J.Suess (ed.) *Solid Waste Management*, Copenhagen, Denmark: WHO Regional Office for Europe.

Tels, M. (1986) 'Advances in Treating Heavy Metals Containing Wastes', *Resources and Conservation* 14: 71–92.

Vogler, J. (1981) *Work from Waste*, London, UK: Intermediate Technology Publications.

WCED, World Commission on Environment and Development (1987) *Our Common Future*, Oxford, UK: Oxford University Press.

SUGGESTED READING

Alter, H. (1997) 'Industrial Recycling and the Basel Convention', *Resources, Conservation and Recycling* 19: 20–53.

Atkinson, W. and New, R. (1993) *Kerbside Collection of Recyclables from Household Waste in the UK—A Position Study*, Warren Spring Laboratory Report LR 946, Gunnels Wood Road, Stevenage, Herts, UK: Warren Spring Laboratory.

Bilitewski, B., Härdtle, G., Marek, K., Weissbach, A. and Boeddicker, H. (1996) *Waste Management*, Heidelberg, Berlin, Germany: Springer-Verlag.

CNCPC (1995) *Enterprise Cleaner Production Audit Manual*, Beijing, P.R. of China: China National Cleaner Production Center.

ERRA (1992a) *Nomenclature: Secondary Materials*, Reference report of the ERRA Codification Programme, Brussels, Belgium: available from European Recovery and Recycling Association.

ERRA (1992b) *Programme Ratios*, Reference report of the ERRA Codification Programme, Brussels, Belgium: available from European Recovery and Recycling Association.

Freeman, H.M. (ed.) (1989) *Standard Handbook of Hazardous Waste Treatment and Disposal*, New York: McGraw-Hill Book Company.

Hester, R.E. and Harrison, R.M. (1994) *Waste Incineration and the Environment*, Cambridge, UK: Royal Society of Chemistry.

Hopfenbeck, W. and Wilg, M. (1995) *Umweltorientiertes Personal-management*, Landsberg, Germany: Verlag Moderne Industrie (in German).

Milner, C. (1996) *Risk to Health Represented by Atmospheric Emissions from Municipal Waste Incineration*, ETSU B/EW/00497/08/REP.

Reijnders, L. (1996) *Environmentally Improved Production Processes and Products*, Proceedings European Round Tables on Cleaner Technology, Dordrect, The Netherlands: Kluwer Publishers.

te Riele, H. and Zweers, A. (eds) (1994) 'Ecodesign', *Waste to Profits: Towards Financial and Economic Dividends from Waste Minimisation in Small Scale Industries in India*, UNIDO, Wien, The Hague, The Netherlands: SDU Publishers.

White, P.R, Franke, M. and Hindle, P. (1995) *Integrated Solid Waste Management: A Lifecycle Inventory*, London, UK: Blackie Academic & Professional.

SELF-ASSESSMENT QUESTIONS

1 Why is waste management a regulation-driven business? Indicate the wrong answer.

(a) Regulations determine numerous legal, administrative and economic aspects of waste management and are still in steady evolution.

(b) Licensing, inspection, notification, tax incentives and levies are important examples of intervention by authorities, the procedures for which have to be defined by regulations.

(c) Industrial organisations, plants and procedures alone are incapable of proper waste management without a legal framework for determining proper standards.

(d) Regulations are required because most waste management activities were criminal in their origins.

2 Supervision of waste operations (1) to (4) can be performed at the various levels (a) to (d): (1) the transboundary shipment of waste; (2) street sweeping; (3) a large, integrated hazardous waste elimination centre; (4) disposal of waste at sea:

(a) a municipality or an equivalent local authority;

(b) the receiving, expediting and transit countries;

(c) the international community as well as the contiguous country or countries;
(d) the public authority responsible for inspection of licensed treatment centres and, to a certain 1extent, the customers of the centre.

Combine (1) to (4) with (a) to (d).

3 Waste prevention is most hampered by:
(a) the squandering attitude of the public;
(b) the lack of standardisation in packaging;
(c) the difficulty of doing without dangerous chemicals;
(d) present lifestyles, which call for the use of disposables, fast food, independent mobility, etc.

4 Waste re-use is difficult to generalise because:
(a) few people buy second-hand goods;
(b) second-hand goods are less dependable and of lower quality;
(c) it is impossible to find good repair services at a reasonable cost;
(d) people do not care for the environment.

5 The aim of the following strategies can be described as follows (one statement is false):
(a) ‘concentrate and contain’ excludes waste from the biosphere for prolonged or even geological time periods;
(b) ‘dilute and disperse’ should be condemned in all circumstances as an unacceptable threat to the environment;
(c) ‘delay and decay’ allows radioactive waste to become more tractable, e.g. for reprocessing of spent uranium bars;
(d) the long-term burial of hazardous waste is unsafe because it may not appear in land planning registers.

3

SUSTAINABLE WATER RESOURCES

Mainwaring B.Pescod and Paul L.Younger

SUMMARY

This chapter aims to give a concise introduction to the sustainability of water resources. To set the scene, the hydrological context and traditional approaches to water resources management are briefly reviewed. The remainder of the chapter is largely devoted to considering the pressures imposed on water resources by changes in water quality. These changes can lead to a degradation in the aesthetic or health-related properties of the water. Such degradation is likely to necessitate enhanced treatment (at increased cost) to ensure continued suitability for existing uses, or, in extreme cases, to necessitate the abandonment of a particular water source. Examples of both eventualities are given where appropriate. Finally, some existing and possible future responses to large-scale, long-term threats to water resources are examined. The chapter concentrates on the engineering aspects of water resources management. The non-engineering aspects are dealt with in Chapter 3 on 'River and Inland Water Environments' and Chapter 4, 'Wetlands' in Volume 3 of this book.

ACADEMIC OBJECTIVES

The primary objective of this chapter is to provide an overview of the distribution, availability and sustainability of water resources, with particular emphasis on those water quality aspects that are central to current decision-making in environmental management. On completion of this chapter the reader should be able to:

- describe those aspects of the hydrological and hydrogeochemical cycles that govern the availability of water resources;
- appreciate the traditional tasks of water resources management and recognise the ways in which these are evolving to take account of the demands of sustainability;
- appreciate the nature and importance of major sources of water pollution;
- distinguish between acute, point sources of pollution and chronic, non-point sources;
- understand the impact of pollution on sustainable use of water resources;
- identify appropriate responses for long-term threats to water resources.

WATER RESOURCES AND THE HYDROLOGICAL CYCLE

The hydrological cycle

The 'hydrological cycle' is a well-known concept. (Should readers be unfamiliar with this concept, or with the basic hydrological terminology used below, they should refer to standard hydrological textbooks, such as Newson (1994), before proceeding.) The hydrological cycle describes the dynamic circulation of water around, on and in the Earth. The most important phases of the hydrological cycle are:

- uptake of moisture from the Earth's surface (es pecially the oceans) to the atmosphere by evapo

ration and allied processes;
- transport of moisture around the planet in the form of atmospheric vapour;
- return of moisture to the Earth's surface in the form of precipitation (rain, snow, dew, etc.);
- transport and storage of water over the Earth's surface in streams and lakes etc. ('surface run off');
- transport and storage of water beneath the Earth's surface in aquifers ('groundwater flow');
- return of water via the latter two routes to the oceans.

The two engines of the hydrological cycle are solar energy (governing evaporation and atmospheric circulation) and the gravitational attraction of the Earth (providing the impetus for precipitation, runoff and groundwater flow).

In terms of water resources, it is instructive to examine briefly the average residence time of water in each of these phases (Table 3.1). As a general rule, the higher the natural residence time of water in a given phase of the hydrological cycle, the greater will be the amount of energy required to exploit that water as a resource. Thus, for those phases in which water is stored on geological time-scales (oceans, ice caps) large inputs of energy will be required to exploit their waters (desalination, and towing and melting icebergs, respectively). Conversely, for the low-residence time, gravity-driven terrestrial fresh waters (streams, rivers, lakes and groundwater), water is stored over smaller time periods, and can thus be gathered using less energetic means.

Surface water resources

In the case of streams and rivers, the means of exploitation can involve simple manipulation of flows with no external inputs of energy. Thus, the simplest water resources development strategy is to divert a proportion of the stream waters to the point of demand (via canals or pipes) using gravity alone. The management strategy for such abstractions of water is also conceptually simple: the cumulative frequency distribution of flows for the stream in question is analysed, and abstraction rates are set so that they do not exceed a stated proportion (often 5 per cent) of the 95-percentile low flow of the stream (Q_{95}, the flow that is equalled or exceeded 95 per cent of the time). The same management strategy applies where waters are actively pumped from a stream, though the details of the hydraulic controls are different. Slightly more elaborate schemes might involve impoundment of surface waters in reservoirs behind various types of dam, to ensure sustained supplies during dry periods. In the latter case, the Q_{95} will be maintained or exceeded by prescribing a minimum release rate (the 'compensation flow') from the reservoir into the downstream channel. In view of the generally low energy inputs required for such activities, it is not surprising that large-scale water resources management was dominated by these simple approaches for the first few millennia of its existence. Indeed, to this day the design of hydraulic structures needed for such tasks (Novak *et al.*, 1996), and the development of control rules for reservoir operation (McMahon and Mein, 1978) remain critically important routine tasks in water resources engineering.

Table 3.1 Approximate average residence time (years) for water in selected phases of the hydrological cycle

	Average residence (years)
Ocean	>4.000
Atmosphere	0.03
Glacial ice caps	>10.000
Streams/rivers	0.01
Lakes	0.04
Groundwater	100

Source: Various, especially Newson, 1994

Groundwater resources

If we look in a little more detail at terrestrial fresh waters, we see that the significant difference in residence time between surface waters (streams, rivers, lakes) and groundwaters is also reflected in a difference in stored volume, such that a full 99 per cent of usable terrestrial fresh waters occur as groundwater. Hence as the more obvious and least energy-demanding surface options for water resources are gradually used up (either in absolute terms, or because of increased opposition to 'drowning valleys' to create reservoirs), the vastly greater volume of water in groundwater storage becomes increasingly attractive. Widespread exploitation of groundwater resources commenced in the industrialised nations in the mid-nineteenth century, as the cost of using steam-driven pumps (developed initially in the eighteenth century for mine de-watering purposes) gradually fell to a level where they could be used economically for water supply schemes.

Because of their longer residence times and their intimate contact with soil and rock materials, groundwaters are generally more highly mineralised than surface waters. On the other hand, because they are protected from pathogen ingress by natural filtration through subsurface strata, groundwaters are generally more bacteriologically pure than surface waters, and often require only precautionary chlorination (to provide a residual disinfection capacity during their movement through distribution networks) before they are passed into public supply. These differences in taste-giving dissolved constituents and the level of treatment required usually weigh against the alternating use of surface water and groundwater sources for the same area of demand. Routine blending of surface waters and groundwaters is certainly possible, and has been practised widely in large conurbations, but the general tendency over the last 150 years has been for separate development of supply networks based on surface water and groundwater.

Since the 1940s it has been increasingly recognised that groundwaters are crucial to most aspects of water resources management, even where they are not directly exploited. This is because of their vast stored volume and their interconnections with surface water bodies. In most humid regions, the majority of streams and rivers are naturally perennial because of the slow release of water from groundwater storage into the surface channels (via springs and as direct inflow through the bed) during periods between rains (Theis, 1940). More recently, it has been recognised that abrupt changes in the pore pressure of shallow subsurface waters can lead to groundwater contributions to peak stormflow in certain upland streams of 80 per cent or more (Sklash and Farvolden, 1979). Hence the fully integrated nature of the land phases of the hydrological cycle is now abundantly clear. This has consequences for the evolution of management strategies for groundwater exploitation.

In former times, issues of sustainable development of groundwater resources were defined as though the aquifer was a simple man-made storage reservoir, and it became axiomatic to equate the 'safe yield' (=sustainable abstraction rate) of a given aquifer with the long-term average annual recharge rate due to infiltration of excess precipitation (precipitation minus actual evapotranspiration). Although erroneous, this idea of 'safe yield' is remarkably persistent (Bredehoeft *et al.,* 1982), and even appears as a formal requirement of some state and national laws (for example, National Rivers Authority, 1992). However, as the twenty-first century begins, it should no longer be acceptable under most circumstances to promote groundwater development without consideration of the likely impact of the proposed development on contiguous surface waters. For all but the most tiny groundwater developments (e.g. scattered shallow wells with hand-pumps) there is no such thing as a simple 'safe aquifer yield' (Bredehoeft *et al.,* 1982). Rather the questions that should be asked when assessing a possible groundwater development are:

- Are the likely reductions in surface water flows which the abstraction will ultimately cause acceptable?
- Are the likely reductions in groundwater storage due to the abstraction likely to cause problems of water quality degradation or land subsidence?

Of these two questions, the first is the most critical, since it applies in virtually all climate zones and in all geological settings. The second question is only applicable where:

1. the aquifer system contains saline water (either at depth, or laterally in coastal areas), or
2. the aquifer strata contain minerals (usually sulphides) which would be unstable if they were left exposed above the water table, and/or
3. the aquifer consists of subsidence-prone lithologies (see Holzer 1984 for a review). Subsidence due to groundwater withdrawal has been reported from London, Venice, Bangkok, Houston, San Joaquin (California) and Mexico City, to name but a few localities.

Because of the lengthy time-scales over which groundwater systems respond to imposed stresses it is often impractical to conduct exhaustive field tests to determine the consequences of a proposed groundwater abstraction programme. Long-term monitoring may ultimately afford hindsight, but forward planning generally demands the application of mathematical modelling techniques (Anderson and Woessner, 1992).

WATER QUALITY AND POLLUTION

The hydrogeochemical cycle

Coupled to the hydrological cycle is a second dynamic system of mass transfer which might be termed the 'hydrogeochemical cycle'. This is responsible for transport of chemical substances through the phases of the hydrological cycle, and for transformations of these substances from one active state to another as they move within and between the various hydrological phases. At its simplest, the hydrogeochemical cycle occurs as depicted in Figure 3.1. In the words of Strömberg and Banwart (1994):

> The atmosphere is characterised as an oxidising, acidic environment relative to the geosphere and its enormous reservoir of mineral bases and reductants.

The bulk of the biosphere thrives at the interface between the oxidised, acidic atmosphere and the reduced, basic geosphere. Most life-forms draw their energy from participation in the redox transformations of carbon and oxygen at the interface, making abundant use of the universal solvent, water. Processes that tip the dynamic equilibrium of the hydrogeochemical cycle too much towards the atmospheric or geospheric extremes can induce stresses at the interface niche occupied by the biosphere. These stresses can prove fatal for specialised organisms, thus reducing biodiversity, damaging the food chain, and ultimately threatening the food sources of higher members of the predation hierarchy (especially homoiotherms like man) which in themselves are fairly tolerant of environmental changes.

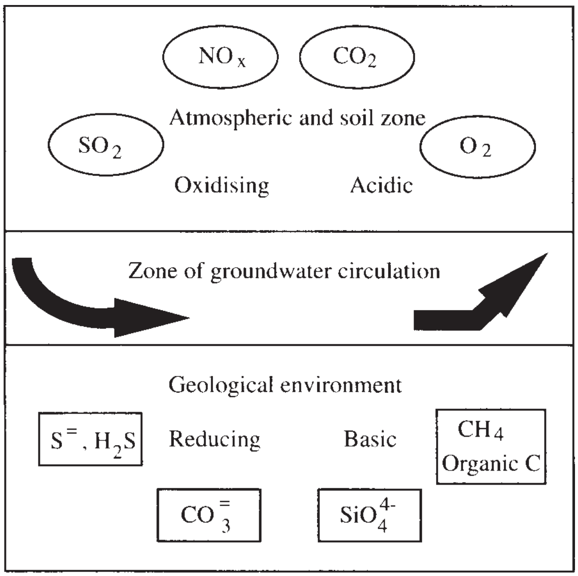


Figure 3.1 The key features of the hydrogeochemical cycle
Source: after Albu *et al.*, 1996

Some of the most potent threats to the long-term integrity of water resources come from man-made stresses acting on the dynamic equilibrium at the atmosphere-geosphere interface; in the following sections, some examples of such threats will be explored. These will span:

- those stresses that are acute (that is, with imme diate deleterious effect) in their impact, and which usually enter receiving waters at discrete points (in which case they are termed 'point-source pol lution');
- those stresses that have imposed their impacts gradually over decades or centuries (hence termed 'chronic'), and which may derive from diffuse processes, such as widespread 'non-point' pollu tion from agro-chemicals or atmospheric pollut ants, and/or global climate change.

Environmental pollution

Pollutants may be regarded as substances or other energy forms (e.g. heat) that are introduced to a local compartment of the environment at concentrations far in excess of those that occur naturally in that compartment. Because the creatures that inhabit that compartment are unused to such high concentrations, they may be unable to cope with the pollutant and may react negatively (flight, disease, death). From this definition it is clear that many 'pollutants' are completely innocuous when they are restricted to their native environmental compartments. This is true, for instance, of many so-called toxic metals: arsenic is an acute systemic poison to most surface-dwelling animals, and yet there are specialised strains of bacteria that thrive by participating in redox transformations of arsenic in certain subsurface environments.

Humans are often the agency by which otherwise harmless substances are introduced 'to the wrong place at the wrong time', becoming pollutants that can cause harm to living creatures, disruption to ecosystems, damage to environmental resources and anthropogenic structures, and even hazards to human health. Perhaps more alarming is the human habit of creating new substances (especially organic molecules) which do not occur naturally in *any* environmental compartment, and which may turn out to have toxic, carcinogenic, teratogenic or mutagenic effects on plants, animals and humans. The environmental significance of many of these molecules is not properly understood before their use becomes widespread.

Pollutants in the hydrogeochemical cycle

Rainwater and acidification in source areas

Water resources might be polluted directly, due to discharges entering receiving waters, or indirectly, through atmospheric pollutants being deposited on land and surface waters. Rainfall carried landwards from the sea transports a significant component of sea water, with elevated concentrations of sodium and chloride ions, some additional calcium and sulphate and traces of other mineral salts. However, when air mass trajectories cross areas where industrial activities are significant, contamination of condensing moisture by atmospheric pollutants can be severe, with increased concentrations of oxides of sulphur and nitrogen, trace organics and particulates.

Rainwater is naturally acidic (pH typically around 5.8), as befits its place in the atmosphere (cf. Figure 3.1). In the source areas of most rivers, where the soils contain base-rich minerals (especially carbonates), then the pH of incoming rainwater will be readily neutralised to around 7. However, where soils are poor in bases (as in most upland peat catchments), then the evaporative concentration of solutes in the rainwater will lower the pH to 4.5 or less without mitigation. Further acidification can then occur as organic acids become soluble. To illustrate the differences attributable to natural soil type, Table 3.2 gives two synchronous analyses of shallow groundwaters from the same small upland source area in northern Britain. One of the piezometers penetrates a peat soil, while the other is in regolith material derived from weathering of sedimentary bedrock (greywacke). Even though the piezometers are scarcely 100 m apart the difference in pH and alkalinity is striking.

Table 3.2 Analyses of shallow groundwaters from Canker Cleugh Catchment, Northumberland, UK, illustrating the contrast in chemistry between groundwaters from peat and sedimentary regolith soils

Determinant	*Peat groundwater*	*Regolith groundwater*
pH	3.56	6.45
Na mg/l	4.74	7.71
Mg mg/l	1.41	17.99
K mg/l	1.32	2.68
Ca mg/l	3.32	35
NO_3 mg/l	0.13	0.17
SO_4 mg/l	2.01	0.28
Alkalinity (mg/l $CaCO_3$)	0.0	84
Colour (T)	450	250
Conductivity (μs/cm)	75	72

Source: Analyses made by J.M.Stunell and I.A.Coleman in July 1995, at the University of Newcastle upon Tyne

However, even where soils have some natural buffering capacity, human activities can result in unnaturally high degrees of acidification. The classic case is the burning of fossil fuels, which can release sulphur dioxide and acid-producing aerosols into the atmosphere. Dry deposition of dust particles from such sources on to the canopies of forests can lead to accumulation of large amounts of acid-generating material, which are rapidly flushed into solution by incoming rainwater (Newson, 1994). Where the forest is itself man-made, purpose-built drainage ditches can result in minimal interaction between base minerals in the soil and the acidified runoff. The consequences of increased acidification of waters in upland catchments include high colour (primarily due to increased solubility of organic fulvic and humic acids) and low pH, which can render water bitter to the taste. The low pH can also enhance the mobility of less common ions such as iron, zinc, lead, selenium, fluoride and boron, with consequences for water treatment requirements.

Pollution and the rural water environment

Water flowing down rivers, or stored in lakes and reservoirs, sustains a complex ecosystem of biota feeding, growing, reproducing, excreting, dying and decomposing. In nature, the whole assemblage of food production and consumption relationships is termed a *food web* in which many *food chains* cross and interlink. In water, through photosynthesis, carbon is fixed in the cell material of phytoplankton, called primary producers, which may be consumed by herbivores (primary consumers such as zooplankton and surface fish), which in turn may be eaten by carnivores (secondary consumers such as middle water fish and birds), which again may be consumed by secondary carnivores (tertiary consumers, such as bottom-feeding fish).

Such activity in natural balance has only marginal effects on the composition of the water but some organisms, for example, the blue-green algae and actinomycetales, can impart unpleasant tastes and/or odours to water, or release toxins. Agricultural activities can sometimes inadvertently encourage the proliferation of these less desirable organisms. Nitrogen and phosphorus are essential elements for most plants, and when added to agricultural soils in their oxidised forms (nitrate, NO_3, and phosphate, PO_4) they can promote rapid and vigorous growth of crops. However, if excess quantities of these nutrients remain in the soil after cropping, they can be flushed into underlying groundwaters, eventually emerging into contiguous streams. In inland lakes, blue-green algae blooms are promoted by high phosphate concentrations. Toxins released by these algae can be injurious to mammals, and the underlying water may become aphotic and anoxic as the covering of algae spreads over the lake surface. In coastal waters (where nitrate is usually the limiting nutrient) bathing beaches can be disfigured by the development of frothy off-white colonies of algae floating on the sea.

It is widely perceived that a high level of nitrates in drinking water can cause stomach cancer and so-called 'blue baby' syndrome (methaemoglobinaemia). However, in the United Kingdom the last blue baby case was in 1972 and there have been no known cases associated with public water supplies (Addiscott *et al.,* 1991). It seems, therefore, that methaemoglobinaemia is nurtured by microbial activity in high-nitrate waters, rather than by processes occurring in sterile (chlorinated) waters with elevated nitrate. The emotive association of waterborne nitrate with cancer remains an unproven theory, lacking in supporting case histories. Nitrates can be removed from waters intended for drinking use by established (though expensive) ion-exchange techniques.

Intensive agricultural practices can also result in the release of hundreds of synthetic chemicals (especially organo-chlorine pesticides) into rivers and groundwater. Many of these chemicals were designed to resist biodegradation, and may thus accumulate in the environment in excessive quantities. Furthermore, they may become bioconcentrated as they pass unmetabolised up food chains. These persistent organics give cause for concern since long-term (chronic) effects of low level exposure to many of them are not well understood.

Less subtle forms of pollution also arise from agricultural activities in the form of sudden, often accidental, releases of vast quantities of reduced, organic materials (such as manure slurries or silage) into streams. This can occur when the bunds of silos or slurry ponds fail due to unforeseen geotechnical problems. The acute impacts on receiving waters can be devastating, resulting in widespread fish kills for many kilometres downstream. Codes of good practice and inspection schemes for storage/drainage features on farms are gradually reducing the frequency and severity of these incidents in several European countries and American states.

Urban pollution

As waters enter urban environments, they are exposed to potential pollution from a wide variety of human activities, ranging from the mundane disposal of sewage to the release of wastes from the most exotic modern industries.

Municipal wastewater can contain a wide spectrum of pollutants, depending on the amount of industrial effluents discharged to the sewerage system. Domestic sewage is primarily water but contains solids of inorganic and organic nature, colloidal and dissolved organic materials (proteins, oils, fats, carbohydrates, hydrocarbons, etc.), dissolved inorganic compounds and biological organisms (including pathogenic viruses, bacteria, protozoa and, possibly, helminths). Recently, concerns have been mounting over the low concentrations of hormones now appearing in sewage effluents; like organo-chlorine pesticides, these can be persistent in the food chain and could possibly lead to reductions in mammalian fertility.

Table 3.3 shows typical levels of the main parameters of a normal municipal sewage and of a very strong sewage from Amman, Jordan, which typifies effluents in water-short areas. The concentration of biodegradable organic matter is typically represented by the five-day, 20°C biochemical oxygen demand, BOD_5.

In most older cities with long-established sewerage systems, sewers usually carry urban storm water in addition to wastewater; these are termed 'combined sewers'. It is rare for a combined sewer to have a capacity in excess of ten times the dry-weather flow. For storms that cause flows in excess of this capacity, therefore, the sewers will surcharge (possibly causing flooding of city streets) unless there are provisions for pressure-relief through overflows to nearby watercourses. Combined sewer overflows (CSOs) may be constructed in various ways, and most are designed to flow only when the discharge rate in the adjacent sewer exceeds a predetermined threshold. Even so, it has been recognised for more than a century that the 'first flush' from an over-spilling sewer can be of very poor quality:

> In urban districts, and especially where there is much traffic, the first storm-washings contain large quantities of putrescible organic matter…they are very foul and often contain as much as the sewage itself.
>
> (Wardle, 1893)

Most types of CSO will require regular maintenance if they are to retain their efficiency. Under-designed or under-maintained CSOs will often begin to flow well before the sewer has reached threshold capacity; indeed some supposed CSOs flow perennially! In such circumstances, raw sewage may be leaked into receiving watercourses for much of the time, causing serious pollution.

In more modern cities, storm waters are often routed to receiving watercourses via separate storm water sewers. Although the separation from foul sewage does afford a reduction in pollutant loadings, many separate storm sewer outfalls (SSOs) still yield water of poor quality. Particular problems are encountered with Pb, Zn (from building materials and automobiles), oil, Na, Cl (from road de-icing salts), BOD_5 and suspended solids.

Waste disposal in landfills can lead to the generation of noxious leachates which can cause severe (if localised) groundwater pollution.

Table 3.3 Major constituents of typical domestic wastewaters, and those from an arid region (Amman, Jordan)

	Concentration, mg/l			
Constituent	*Weak*	*Medium*	*Strong*	*Amman, Jordan*
Total solids	350	700	1200	
Dissolved solids (TDS)	250*	500*	850*	1170
Suspended solids	100	200	350	900
Nitrogen (as N)	20	40	85	150
Phosphorus (as P)	6	10	20	25
Chloride	30*	50*	100*	
Alkalinity (as $CaCO_3$)	50	100	200	850
Grease	50	100	150	
BOD_5	100	200	300	700

Note: * The amounts of TDS and chloride should be increased by the concentrations of these constituents in the carriage water.
Source: UN Department of Technical Co-operation for Development, 1985; Al-Salem 1987

Discharge of leachate and associated groundwaters to surface streams can also cause problems during dry weather. Advances in waste management technology (including construction of low-permeability caps and liners, and installation of leachate collection systems) have significantly reduced the risk of such pollution in countries with strict environmental regulations.

Industrial pollution

Industrial wastewaters are as diverse as industry itself and can contain any of the chemicals used as raw materials or which occur in the products (see Hounslow (1995) for a review). Consequently, industrial wastewaters are more variable in composition than domestic sewage. In general, industrial wastewaters can be classified into those that are predominantly inorganic and those that contain primarily organic pollutants. Some organic industrial wastewaters, such as from food processing operations, can be much stronger (in terms of BOD_5) than normal municipal sewage, and animal wastes tend to be much more polluting than human wastes.

Even after industrial activities cease, contaminated land may remain, giving rise to polluted runoff for many decades or even centuries (National Rivers Authority, 1994a). Problematic pollutants often include oils and tars, organic compounds, 'heavy' metals and soluble salts. Even where contaminated sites are effectively remediated, popular perceptions can prevent future use of waters in formerly contaminated areas for many decades. For instance, a reservoir formerly used for public water supply in north-west England remains out of commission (despite recent drought conditions, and despite the present high quality of the water) because of public memories of a major pollution incident 20 years ago, when the reservoir was contaminated by degreasing agents from a nearby heavy goods depot.

Mining activities can give rise to serious water pollution of both point and non-point source varieties (see Box 3.1). Poorly regulated surface mining will often result in heavy silting of surface waters and increased levels of heavy metals. The de-watering of deep mines will likewise produce contaminating effluents which are costly to treat because of the large volumes involved. When mines are abandoned, water quality may deteriorate further, as flooding of old coal or metalliferous workings containing oxidised pyrite can lead to widespread and persistent acidification of

BOX 3.1 MINEWATER POLLUTION IN THE HIGH ANDES

Bolivia sits astride the Andean Cordillera of South America, and has historically been one of the world's most prolific sources of silver, tin and (latterly) gold. Intensive working of underground mines has been continuous since the Spanish conquest in the sixteenth century. However, exhaustion of the more accessible deposits, combined with unfavourable prices for tin in the international markets since the early 1980s, has resulted in the abandonment of many mines. As the Bolivian population is the poorest in South America, the social impacts of mine closures have been particularly serious. These have only been exacerbated by the emergence of heavily polluted mine waters after the cessation of pumping led to the gradual flooding of underground workings. When the mined voids fill, the oxidised remnants of trace sulphide minerals present in the mine (particularly iron pyrites, FeS_2) are rapidly dissolved. This leads to a dramatic lowering of pH, and to dissolved loadings of ecotoxic metals (Fe, Mn, Al, Zn, Pb, Cd, Ni, As, etc.) which exceed ordinary background concentrations by several orders of magnitude. When such waters emerge at the surface, they result in vivid red, orange and/or yellow staining of streambeds, and the destruction of most aquatic life. For instance the San José Mine (near the city of Oruro) discharges water with a pH as low as 1.8, and with extreme concentrations of iron (2000 mg/l), zinc (88 mg/l), arsenic (44 mg/l) and lead (22 mg/l) (Mendizábal de Finot, 1994). The red stream that receives the San José discharge flows through the city centre, before entering Lago (lake) Uru-Uru, where it badly damages natural fish stocks, which have historically formed the basis for an important rural industry. Similarly, near the capital city of La Paz, some $0.7 m^3/s$ of acidic, metalliferous mine waters emerge from old workings beneath the majestic Andean peaks of Chacaltaya and Huayna Potosí. These waters have turned three (formerly turquoise) glacial lakes blood-red. However, given the scarcity of water resources in the semi-arid Bolivian highlands, such large volumes of water cannot be left to waste: with considerable ingenuity, the mine waters are led over the piedmont in an aqueduct (precipitating iron hydroxide all along the way) to the outskirts of La Paz, where they are blended with similar quantities of unsullied glacial meltwaters to serve for public water supplies.

groundwaters and surface waters, with accompanying transport of large quantities of iron into the surface watercourses, where ferric hydroxide deposition causes clogging of streambeds and a cessation of benthic photosynthesis (Younger, 1993, 1995). Iron, manganese, zinc and sulphate in such waters can necessitate expensive treatment if they impact upon public supply sources.

Sea waters

Pollutants discharged into surface receiving waters from any source will find their way into estuaries, coastal waters and eventually to the open seas. Table 3.4 gives some indication of the degree of pollution in the North Sea as a result of river inputs, coastal inputs, high seas spillage and dumping, and atmospheric deposition. Recent declines in the numbers and health of fish caught by European Union (EU) fleets may in part reflect the pressures imposed on shallow shelf seas by such pollution sources.

Concern over the long-term productivity of the seas, as well as the general objectives of environmental improvement and harmonising trade, have driven the EU to introduce several directives aimed at abatement of sea pollution. These include directives on bathing water, shellfish, discharges of dangerous substances to the aquatic environment and urban wastewater. For instance, the EU Urban Wastewater Treatment Directive obliges improvements in pre-treatment of sewage before disposal by long sea outfalls, and also prohibits (with effect from the end of 1998) the long-standing practice of some EU member states to dispose of sewage sludge by dumping at sea (Wright, 1992). It is hoped that this policy will result in some amelioration of the quality of the North Sea, although it raises further issues with regard to sludge disposal on land and/or to the atmosphere by incineration.

Atmospheric pollution and climate change

Atmospheric pollution by the so-called 'greenhouse gases' (especially carbon dioxide derived from burning of fossil fuels) is thought to be causing global warming, with potentially important consequences for the spatial and temporal distribution of global water resources. While methods for predicting climate change impacts on water resources are still under development, preliminary results (obtained from hydrological calculations based upon the output of Global Circulation Models (GCMs)) suggest that the current distribution of dry and wet climate zones on the Earth's surface may be altered over the next few decades, and that hydrological extremes (floods, droughts) may become more pronounced and more frequent. If these predictions are true, then the future utility of existing water resources infrastructure may be considerably shorter than their initial design lives may suggest. By the middle of the twenty-first century, major changes in management strategies may be necessary to cope with such effects.

Table 3.4 North Sea pollution

		Areas of pollution		
Pollutants	*Levels*	*Estuaries*	*Coastal areas*	*High sea*
Biodegradable substances	mg/l	++	+	(–)
Nutrients – N, P	mg/l	++	++	(–)
Solids	mg/l	+	+	–
Bacterial pollutants (Including E. Coli)	No/100 ml	(+)	(+)	–
Oil & oil products	mg/l	+	+	+
Heavy metal compounds Cd, Hg, Cr, Pb, Zn	µg/l	+	+	+
Persistent organic substances PCBs, DDT	ηg/l/µg/l	+	+	+
Radioactive substances	Bq/kg	?	?	–
Cooling waters, heat input	MJ/d	+	(+)	–

Notes: ++ Considerable; + Slight; (+) Slightly seasonally; - Negligible; (-) Possibly
Source: Collated from various sources

Table 3.5 GQA chemical grading for rivers and canals in England and Wales

	Grade	Dissolved oxygen (% sat) 10 percentile	BOD(ATU) (mg/l) 90 percentile	Total ammonia (mg/l) 90 percentile
Good	A	80	2.5	0.25
	B	70	4	0.6
Fair	C	60	6	1.3
	D	50	8	2.5
Poor	E	20	15	9.0
Bad	F*			

Note: *Quality not meeting the requirements of grade E for one or more determinants
Source: National Rivers Authority, 1994b

Assessing and managing the impacts of polluted discharges

In many countries rivers are classified according to their chemical pollution characteristics. For example, the chemical grading for rivers and canals in England and Wales, under a General Quality Assessment (GQA) Scheme introduced in 1995, is shown in Table 3.5.

GQA grades for the biological (covering invertebrate life) nutrient (relating to nitrogen and phosphorus) and aesthetic (concerned with visual appearance and odour) aspects are currently under development. Ecotoxicological risk assessment is an evolving field of study concerned with assessing the likely incidence and consequence of adverse effects in the aquatic environment arising from actual or predicted exposure to given chemicals (Foundation for Water Research, 1996).

In addition to the EU Urban Wastewater Treatment Directive, which will also necessitate upgrading of most sewage works discharging to inland waters, further EU directives on drinking water, the quality of surface water abstraction, freshwater fish and nitrate are all providing stimuli for improvements in the quality of point and non-point sources of pollution. New directives on the ecological quality of water and on integrated pollution prevention and control (IPPC) are also being developed, and recent announcements have indicated the European Commission's intention to produce a Framework Directive setting out 'a global water policy' for the EU (King, 1996). Such a Framework Directive is expected to address issues of water quantity and quality, together with a comprehensive ground and surface water management policy addressing transboundary water management problems (such as beset the rivers Rhine and Danube). Nevertheless, transboundary water resources management is extremely difficult to achieve in practice, as is obvious from the prominence of this issue in the Arab-Israeli peace negotiations.

SUSTAINABILITY OF WATER RESOURCES

What is sustainability of water resources?

In former times, if sustainability of water resources was considered at all, it was assessed in the spirit of 'sustainable growth' (in the same sense as this term is used by politicians when discussing economics). The notion was that modest but steady growth in demand could be sustained by a rolling programme of water resources developments. This would involve constructing new reservoirs or exploring for new aquifers. However, as the number of catchments available for flooding and the number of pristine aquifers gradually diminishes, and as the consequences of climate change are gradually revealed, a shift in paradigm is becoming inevitable.

'Sustainable growth' in water resources is a contradiction in terms: we are now forced to acknowledge that sustainability in water resources is about wise stewardship of a finite resource; it is about intervening in the hydrological and hydrogeochemical cycles in such a manner that we do not fundamentally damage their natural patterns. Avoidance of damage to these natural cycles is not a matter of mere druidic sentimentality; the emergence of mankind depended and, ultimately, our future survival depends on the healthy continuation of moisture and solute fluxes at and below the Earth's surface. Given that the quantity of water is essentially finite (and that its spatial and temporal distribution might be fundamentally altering), the real scope for positive intervention in water resources lies in the realms of conservation and water quality management. The emerging paradigm in sustainable management of water resources is the maintenance and restoration of the coupling between the hydrological and hydrogeochemical cycles. This is primarily achieved by careful use of water and the control and remediation of pollution. It should always be remembered, however, that activities of using and protecting water resources involve the exploitation of other resources (minerals, chemicals, energy, etc.); consequently, the sustainability of water resources cannot be examined in isolation from considerations on the sustainability of the use of other resources.

Sustainability and water pollution

If the pollution of water is not controlled, water resources will become unfit for legitimate uses, not least as potable water supplies. Given the complexity of modern society, it is virtually impossible to guarantee that some of the pollutants released to the environment will not enter public water supplies. Consequently, indirect re-use of wastewaters commonly occurs where an abstraction from a natural water actually captures some proportion of wastewaters discharged into the same watercourse upstream. It is imperative, therefore, that each country should have in place mechanisms and legislation to ensure the maintenance of water quality in rivers and aquifers which both receive wastewater discharges and also support abstractions for water supply. The essential priority in such situations is planning for the public good (and for ecological benefits), even where this demands some restriction on the activities of individuals or corporations.

In many countries, such regulation has received little attention and the quality of water resources has been deteriorating at an alarming rate. In Bangkok, for example, the Chao Phraya River became anaerobic over an extensive reach within the city as the population increased from approximately two million in the 1960s to the present day estimate of ten million or more. Developing countries often assign low priority to water pollution control, believing that industrial development will be restricted by the additional environmental costs. However, this is a short-term view which many now-developed countries have historically found to be disastrous, not least Japan with the Minamata disease episode resulting from the release of methyl mercury to Minamata Bay in the 1960s.

Where wastewaters are treated prior to discharge, highly efficient removal of major pollutants is usually achieved. However, traditional wastewater treatment methods are less effective in removing some of the trace levels of new synthetic chemicals which now appear in municipal sewage. Constant vigilance is required to prevent the adverse effects of perhaps minute quantities of released chemicals, such as the oestrogenic substances currently causing concern in relation to human fertility (Fawell and Wilkinson, 1994).

The general objective of water pollution control is to limit the concentration of pollutants in a water resource to levels that would not preclude its use for specific purposes. Water quality parameters relevant to specific uses are summarised in Table 3.6. To satisfy such water quality constraints, specific targets might be established, such as River Quality Objectives (RQOs). Once RQOs have been established for each river or reach of river, a water pollution control management system can be introduced to achieve those objectives over reasonable time at affordable cost. This might involve amendment or abandonment of existing polluting practices in the catchment, and/or careful assessment of new development proposals to ensure their compatibility with RQOs.

In the case of groundwaters, the high residence time (Table 3.1) means that retrospective clean-up of pollution is often very expensive, and will frequently fail to produce satisfactory results over normal human time-scales. Consequently, prevention is very definitely better than cure in groundwater systems, and energies are best spent on protection of

Table 3.6 Water uses and quality considerations

Water uses	*Important quality parameters*
Power	Dissolved oxygen, pH
Flood protection	–
Irrigation	Dissolved solids, electrical conductivity, sodium
Potable water supply	Colour, turbidity, hardness, pathogenic organisms, organics
Industrial water supply	Hardness, pH, dissolved oxygen
Navigation (transportation)	Suspended solids, pH
Fishing (commercial and subsistence)	Dissolved oxygen, pesticides, CO_2, pH, heavy metals
Recreation	Pathogenic organisms, pH
Nature conservation (wildlife and aesthetics)	Colour, toxic compounds
Waste disposal	–
Saline water intrusion prevention	–

Source: Updated after Pescod, 1977

groundwater from pollution. This generally involves hydrogeological mapping of aquifer outcrop areas to determine which areas would be most vulnerable to pollution. Local hydrogeological conditions, such as the depth to the water table, the rate of recharge, and the nature of soil and underlying aquifer materials, are the key controls on vulnerability. Legal groundwater abstractions are generally accorded special protection by the definition of concentric 'source protection zones' (SPZs) around each spring or well. Although the details of SPZs differ from one country to the next, they are generally similar in outline. The SPZs used by the UK Environment Agency are as follows (National Rivers Authority, 1992):

- *Zone 1:* Area with 50-day travel time to source, or 50 m radius around source, whichever is the larger. (Land-use restrictions in this zone are fairly comprehensive.)
- *Zone 2:* Area with 400-day travel time to source, or 25 per cent of total catchment area of the source, whichever is the larger. (Land-use restrictions in this zone include: landfilling of anything other than inert waste, storage of certain chemicals, discharges of >5 m^3/d or of trade effluent to soakaways, etc.)
- *Zone 3:* Complete source catchment. (Land-use restrictions include any landfill that is not engineered for containment, certain agricultural practices (such as use of pesticides, fertilisers, etc.) in vulnerable areas.)

Strategies and technologies for sustainability

Water conservation

Water conservation has been part of the culture of people living in water-short, arid regions for centuries. Only in countries (in temperate and tropical zones) that are blessed with high precipitation has profligate use of water resources long been commonplace. Nevertheless, the international political dominance of water-rich countries over the last two centuries has led to an increasing fashion for extravagant water use even in the arid countries. However, with increasing populations and expanding water use by industry, drought conditions have even arisen in the water-rich countries themselves (such as the United Kingdom) and some restrictions on water use have been introduced.

In the poorer developing countries, some rural communities exist on 15 litres per capita per day (lpcd) of water whereas modern society in advanced industrial countries might consume in excess of 400 lpcd. Households in Europe and North America are equipped with modern plumbing and appliances which have not been designed with water conservation in mind. Nevertheless, there are devices that can be used to limit water use in the household, even if economic disincentives have not been successful in reducing water consumption. In addition, greater investment in leakage control measures will improve on the typical 20–50 per cent loss of water in distribution systems. Water demand management is now necessary in most countries if the expectation of the public for a level of service with no water restrictions is to be met within the constraints of available and exploitable water resources.

Industrial water usage can also be high in developed countries and measures to conserve water and minimise wastewater release will only be considered by industry if governments apply pressure, particularly in the form of high water charges and strict effluent discharge consent conditions. A major water user in many countries is agriculture and in arid countries with little or no rain-fed cultivation agriculture is often the largest user of water by far. Although water-conservative irrigation techniques have been developed (especially drip irrigation), these have not been adopted in all arid countries, and the selection of crops that use minimal water is not practised as widely as might be hoped. Furthermore, historical water rights often work against water conservation, with farmers traditionally taking advantage of the most convenient water resource without restriction, control or charge. As an example, in the Gaza Strip the only water resource is the shallow coastal aquifer; this is severely stressed by over-abstraction, leading to salinisation and pollution by sewage effluents and agro-chemicals. Nevertheless, farmers in Gaza have always been free to draw water from uncontrolled wells without limit or charge. Hence, newly constituted Palestinian political authorities risk social hostility if they now attempt to introduce rational regulation of abstraction.

If water conservation is to be achieved in any country, there must be strict legislation, effective regulation and diligent monitoring and enforcement. Economic charging for water will usually be essential to promote water conservation and licensing of all abstractions from and discharges to natural waters is generally advisable.

Wastewater treatment

In the case of domestic wastewaters in urban areas, water usage will have a significant impact on the quality of the wastewater collected by the sewerage system. However, once water conservation measures are introduced there is little that can be done to reduce the level of pollutants released, and wastewater treatment in sewage treatment plants must then be applied, normally for environmental reasons. There is a wide range of technology that can be adopted to remove most pollutants but, in general, the higher the efficiency of removal required the greater the cost. Primary treatment to remove suspended solids costs less than secondary biological treatment to remove biodegradable organic material; tertiary treatment, normally introduced to remove nitrogen and phosphorus, has an even greater unit cost. Currently, additional tertiary treatment to remove trace organics (including pesticides) is being applied at considerable expense in areas where there is intensive agriculture. Clearly, the degree of treatment must be chosen so as to balance the cost with the risk of adverse effects resulting from the treated effluent. Secondary treatment is a sensible objective to target in countries where there are many communities still without sewage treatment. In some advanced countries, where secondary treatment of all municipal wastewaters has already been achieved, tertiary treatment should be considered, to overcome the adverse local impacts of any residual pollutants in the secondary treated effluent. Wherever the effluent is to be considered for re-use, the specific quality conditions necessary for the form of re-use will have to be satisfied.

Industrial wastewaters are sometimes discharged to municipal sewerage systems and this can be acceptable if the pollutants contained in the trade effluents are consistent with those which the sewage treatment plant was designed to remove. All undesirable pollutants should be prohibited from discharge to sewers and this is controlled in many countries by licensing all trade discharges and strict monitoring. Even acceptable pollutants are charged for in well-regulated systems, on the basis of the ‘polluter pays’ principle. If charging schemes are punitive, industry will choose to pretreat its wastewaters before release to sewers, and will be forced to remove any pollutants prohibited by the discharge consent conditions. Again, a wide range of treatment technologies is available, to match the wide spectrum of industrial wastewater quality. However, to pursue the conservation theme, industries are increasingly seeking to eliminate or minimise wastewater streams through careful selection, design and operation of processes, and through water conservation and improved housekeeping, rather than resorting to wastewater treatment.

Wastewater reclamation and re-use

In arid, water-short areas conventional water resources are frequently fully utilised, or even overutilised in the case of groundwaters. Wastewaters constitute an additional water resource which can be utilised economically and effectively in industry and agriculture. The potential for irrigation to increase agricultural productivity and the living standards of the rural poor has long been recognised. In the Near East, for example, only 30 per cent of the cultivated area is irrigated, yet this portion generates roughly 75 per cent of the total agricultural production. With increasing water supply and sewerage coverage for urban populations, municipal wastewater becomes a marginal resource worthy of consideration for re-use. In several countries in the Near East (Jordan and Saudi Arabia included) re-use of treated sewage is an integral component of water resources policy.

Properly planned use of municipal wastewater alleviates surface water pollution problems, and not only conserves valuable water resources but also takes advantage of the nutrients contained in sewage to grow crops. The availability of this additional water near population centres increases the variety of crops that farmers can grow. The nitrogen and phosphorus content of sewage might reduce or even eliminate the need for commercial fertilisers. It is advantageous to consider effluent re-use when planning wastewater collection, treatment and disposal, so that sewerage system design can be optimised in terms of effluent transport and treatment methods. It is usually prohibitively expensive to transmit effluent from inappropriately sited sewage treatment plants to distant agricultural land. Additionally, sewage treatment techniques for effluent discharge to surface waters may not always be appropriate for agricultural use of the effluent.

Many countries have included wastewater re-use as an important element of water resources planning (see Box 3.2). In the more arid areas of Australia and the USA wastewater is used in agriculture, releasing high quality water supplies for potable use. Some countries, for example, the Hashemite Kingdom of Jordan and the Kingdom of Saudi Arabia, have a national policy to re-use all treated wastewater effluents, and have already made considerable progress towards this end. In China, wastewater re- use in agriculture has developed rapidly since 1958, so that over 1.33 million hectares are now irrigated with sewage effluent. It is

BOX 3.2 WASTEWATER TREATMENT AND AGRICULTURAL RE-USE IN TUNISIA

Wastewater re-use in agriculture has been practised in Tunisia for several decades and is now an integral part of the national water resources strategy. Many projects have been implemented over the past few years, the most important developments taking place around the Tunis metropolitan area, where 60 per cent of the country's wastewater is produced and 68 per cent of the country's effluent-irrigated area of 6,700 ha is located. The La Chergia activated sludge plant, for example, receives sewage from Tunis and discharges its effluent to the La Soukra irrigation area 8 km distant The development of such systems has taken place in a number of phases since 1981.

In the period 1981–87, the Ministries of Agriculture and Public Health, with assistance from the United Nations Development programme (UNDP), conducted studies to assess the effects of using treated wastewater and dried, digested sewage sludge (from the La Chergia and Nabeul works) on crop productivity and on the hygienic quality of crops and soil (compared with using groundwater, as a control). Among the crops studied were sorghum, peppers, clementines and oranges. Crop yields exceeded those produced by irrigation with groundwater, and bacterial contamination was generally absent Sewage sludge was found to have fertilising potential due to the presence of minerals and organic matter, but was of variable consistency. Application of 30 tonnes dry weight of dried sludge per ha provided an excess of nitrogen (N) and phosphorus (P), but a deficit of potassium (K_2O) in respect of crop requirements. The treated effluent provided an excess of N and K_2O, but a deficit of P_2O_5. Applied together, the effluent and sludge would provide excess nitrogen, which might be of concern in terms of crop growth and groundwater pollution. The application of treated effluents and sludges at the experimental stations over an extended period has not adversely affected the physical or bacterial quality of the clay-rich, sandy soils. However, the chemical quality of the soils has been modified, with an increase in electrical conductivity and a transformation of the geochemical characteristics of the soil solution, from bicarbonate-calcium to chloride-sulphate-sodium. Although trace elements, particularly zinc (Zn), lead (Pb) and copper (Cu) have concentrated in the surface layer of the soil, they have not yet increased to phytotoxic levels. No adverse effects on the chemical or bacteriological quality of shallow groundwater were found.

generally accepted that wastewater use in agriculture is justified on agronomic and economic grounds but care must be taken to minimise adverse health and environmental impacts. The Food and Agriculture Organization of the UN has produced guidelines for wastewater use in agriculture which will allow the practice to be adopted with complete health and environmental security (Pescod, 1992).

Industries can also take advantage of their own or municipal wastewaters for some of the less quality-demanding water uses. Clearly, treated wastewaters might potentially be used for cooling purposes, but the high quality requirements for process water and steam-raising will not normally be met by treated effluent re-use.

Desalination

Where no alternative exists, desalination of sea water or brackish water can provide an additional freshwater resource. However, this is the mostly costly water to produce, even compared with treated wastewater when the costs of sewerage and secondary treatment are considered as environmental protection investments. Most of the world's sea water desalination plants serve the Arabian Gulf area, which is not well endowed with freshwater resources, but is blessed with oil resources. Riyadh in Saudi Arabia, for example, lies some 380 kilometres inland yet is partly supplied by a desalination plant situated on the Gulf coast at Jubail. Desalinated water is also supplied in Abu Dhabi (United Arab Emirates) and the municipal wastewater is treated and used to irrigate landscaped areas in the city. Examples of affordable sea water conversion elsewhere in the world are rare, and even where serious water shortages occur (as in Gaza) desalination is viewed only as a remote solution.

Although the technology has existed for many years, costs have remained high. The largest desalination plants utilise multi-stage flash distillation to produce, essentially, a distilled water which is usually blended before public supply. In recent years reverse osmosis technology, originally applied only to brackish waters, has improved to the point of being competitive with distillation and is being considered for sea water desalination. Nevertheless, for large-scale applications any form of desalination process is expensive as a means of providing public water supplies, and will be a last resort in water resources planning.

Conjunctive use and artificial recharge

The fundamental difference between residence times for ground and surface waters means that peak and minimum storage in the two systems are often out of phase with each other. This offers the possibility of managing the two resources conjunctively, to reduce the overall risk of failure of water supply. Conjunctive use may be practised in several ways, but one of the most popular in Northern Europe is through groundwater-based river augmentation schemes. In such schemes, groundwater is pumped into rivers during dry periods, to sustain surface flows so that downstream demands continue to be satisfied. (These demands can be for public supply, navigation, or for protection of sensitive flora and fauna.) In the winter, when surface flows are typically high, groundwater pumping is discontinued, allowing recharge to raise water table levels once more (Owen *et al.*, 1991). Such schemes have a number of advantages over direct supply, single source schemes. For instance, combining resources with different hydrological characteristics during times of drought reduces 'hydrological risk'; many of the environmental disadvantages and much of the cost of surface reservoir construction are avoided; development and financial commitments can be staged, allowing postponement of full development if projected demands fail to materialise. Finally, it is possible to blend waters with contrasting properties to obtain an optimum water quality for supply, or to sustain environmental conditions suitable for aquatic life.

Artificial recharge is another strategy for optimising the use of water resources. It may be defined as the deliberate introduction of water into the subsurface. This may be done by allowing water to infiltrate from spreading basins, by allowing water to flow down boreholes, or by pumping boreholes adjacent to rivers so that the hydraulic gradient favours ingress of surface waters into the subsurface. Traditionally, artificial recharge has been performed in order to build up strategic reserves of water without constructing a surface reservoir. The classic example of this is the scheme operated by Thames Water plc in north

London (Owen *et al.*, 1991). Artificial recharge could also be employed to replenish aquifers used in conjunctive use schemes. Air conditioning systems for buildings in warm regions can be based upon cyclic abstraction and artificial recharge of shallow groundwaters. In California, artificial recharge has also been used to combat saline intrusion, reduce or stop land subsidence, and (experimentally) to control seismic activity by managing the timing and magnitude of earthquakes.

From the perspective of water resources sustainability, artificial recharge can be used to dispose of wastewaters. This relies on the fact that ingress of polluted waters into aquifers can expose them to natural subsurface mixing/retardation/pollutant degradation processes. In Arizona, carefully managed artificial recharge and subsequent abstraction is used to prepare municipal wastewater for re-use in irrigation (Pescod, 1992). Conversely, in areas with highly polluted groundwater, recharge of water of appropriate quality may provoke chemical reactions *in situ* which will ameliorate the pollution in the subsurface (Nawrot *et al.*, 1994; Morrison *et al.*, 1996).

Addressing catchment-scale problems

Some of the more insidious threats to water resources are those associated with diffuse pollution (from atmospheric sources, agricultural pollution or landuse change) at the catchment scale. Similarly, changes in land use can lead to accelerated erosion of soils, leading to the reduction in storage capacity of reservoirs as they become loci for sedimentation. Hydrological assessment of such extensive problems has only become feasible in the last two decades, since computer technology has developed to enable:

- storage and manipulation of large amounts of data describing the topography, geomorphology, hydrology and water quality of large basins; the relevant software are Geographical Information Systems (GIS);
- simulation of coupled flow and solute/sediment transport in all of the land phases of the hydrological cycle; this demands the use of powerful catchment models such as SHETRAN (Bathurst *et al.*, 1995).

Using such methods, it is possible to develop an overview of the critical features controlling the response of a catchment to land-use change, and to simulate alternative management strategies (Adams *et al.*, 1995).

Appropriate strategies for restoring catchments that have been polluted through atmospheric deposition or land-use change are currently being pioneered by inventive hydrologists and ecologists. The main objective of such techniques is to use natural processes to assist the recovery of healthy, self-sustaining hydroecological systems (Large *et al.*, 1993). This might include:

- reinstating channel patterns (braiding, meanders, etc.) which have been removed by earlier engineering efforts;
- encouraging the development of 'buffer strips' (functional floodplain ecosystems) for habitat enhancement or water quality control (Large *et al.*, 1993);
- deliberate construction of wetlands to passively treat polluted waters before they enter river channels (Hammer, 1992; Younger, 1995).

SUMMARY AND CONCLUSIONS

In an earlier age, water resources management was concerned primarily with the spatial and temporal distribution of quantities of water. In the last few decades, the quality of water has come to the fore as a prime control in the sustainable development of water resources. Pollution from domestic, agricultural and industrial (or post-industrial) sources is limiting the utility of large quantities of water world-wide. Therefore at the start of the new millennium, water resources management is concerned as much with manipulating contrasting water sources, to ensure the attainment of adequate quality for cost-effective treatment, as it is with traditional themes such as reservoir control and conjunctive use. This change in emphasis is well illustrated by practical examples of evolving water resources technologies. The time-scales over which certain diffuse sources of pollution operate to impact surface waters and groundwaters present serious obstacles for long-term water resources planning.

For instance, it is usually many decades before the impacts of certain sources of pollution can be positively identified; this is the case, for example, with pollution of surface waters from atmospheric deposition and from land-use changes, groundwater pollution due to nitrates, pesticides, and the legacy of past industrial activities. Possible climate change impacts further complicate the picture. A forward-looking response to these threats demands a change in thinking as much as a change in technology.

REFERENCES

Adams, R., Dunn, S.M., Lunn, R., Mackay, R. and O'Callaghan, J.R. (1995) 'Assessing the Performance of the NELUP Hydrological Models for River Basin Planning', *Journal of Environmental Planning and. Management* 38: 53–76.

Addiscott, T.M., Whitmore, A.P. and Powlson, D.S. (1991) *Farming, Fertilisers and the Nitrate Problem,* Oxford, UK: CAB International.

Albu, M., Banks, D. and Nash, H. (1996) *Mineral and Thermal Groundwater Resources,* London, UK: Chapman and Hall.

Al-Salem, S.S. (1987) 'Evaluation of the Al Samra Waste Stabilisation Pond System and its Suitability for Unrestricted Irrigation', Paper presented to the Land and Water Development Division, Food and Agriculture Organization of the UN, Rome.

Anderson, M.P. and Woessner, W. (1992) *Applied Groundwater Modelling. Simulation of Flow and Advective Transport,* San Diego, Calif., USA: Academic Press.

Bathurst, J.C., Wicks, J.M. and O'Connell, P.E. (1995) 'The SHE/SHESED Basin Scale Water Flow and Sediment Transport Modelling System', in V.J.Singh (ed.) *Computer Models of Watershed Hydrology,* Col., USA: Water Resources Publications, pp. 563–594.

Bredehoeft, J.D., Papadopolous, S.S. and Cooper, H.H. (1982) 'Groundwater: The Water-Budget Myth', in *Studies in Geophysics; Scientific Basis of Water Resource Management,* Washington, DC: National Academy Press, Chapter 4, pp. 51–57.

Fawell, J.K. and Wilkinson, M.J. (1994) 'Oestrogenic Substances in Water: A Review', *Aqua* 43, 5: 219–221.

Foundation for Water Research (1996) *Ecotoxicological Risk Assessment Manual for Chemicals in the Aquatic Environment,* Report No. SR 3846/1, Marlow, UK: FWR.

Hammer, D.A. (1992) *Creating Freshwater Wetlands,* Boca Raton, Fla., USA: Lewis Publishers.

Holzer, T.L. (ed.) (1984) *Man-Induced Land Subsidence,* Reviews in Engineering Geology Volume VI, Boulder, Col., USA: Geological Society of America.

Hounslow, A.W. (1995) *Water Quality Data Analysis and Interpretation,* Boca Raton, Fla., USA: Lewis Publishers.

King, A. (1996) 'Framework Puts Direction Back into the Directives', *Water Services* 100, 1 202: 8–9.

Large, A.R.G., Petts, G.E., Wilby, R.L. and Greenwood, M.T. (1993) 'Restoration of Floodplains: A UK Perspective', *European Water Pollution Control* 3, 5: 44–53.

McMahon, T.A. and Mein, R.G. (1978) *Reservoir Capacity and Yield,* Amsterdam, The Netherlands: Elsevier.

Mendizábal de Finot, M. (1994) *Oruro: del desastre a la esperanza ambiental,* La Paz, Bolivia: Instituto Latino-americano de Investigaciones Sociales (in Spanish).

Morrison, S.J., Spangler, R.R. and Morris, S.A. (1996) 'Subsurface Injection of Dissolved Ferric Chloride to Form a Chemical Barrier: Laboratory Investigation', *Ground Water* 34: 75–83.

National Rivers Authority (1992) *Policy and Practice for the Protection of Groundwater,* Bristol, UK: NRA.

National Rivers Authority (1994a) *Contaminated Land and the Water Environment,* Water Quality Series No. 15, London, UK: HMSO.

National Rivers Authority (1994b) *The Quality of Rivers and Canals in England and Wales (1990–92) as Assessed by a New General Quality Assessment Scheme,* Water Quality Series No. 19, London, UK: HMSO.

Nawrot, J.R., Conley, P.S. and Sandusky, J.E. (1994) 'Concentrated Alkaline Recharge Pools for Acid Seep Abatement: Principles, Design, Construction and Performance', *Proceedings of the International Land Reclamation and Mine Drainage Conference and the Third International Conference on the Abatement of Acid Mine Drainage. Volume 1: Mine Drainage,* Pittsburgh, Pa. USA: US Bureau of Mines Special Publication SP 06A-94, 24–29 April 1994, pp. 382–391.

Newson, M. (1994) *Hydrology and the River Environment,* Oxford, UK: Oxford University Press.

Novak, P., Moffat, A.I.B., Nalluri, C. and Narayanan, R. (1996) *Hydraulic Structures,* Second revised edition, London, UK: Chapman and Hall.

Owen, M., Headworth, H.G. and Morgan-Jones, M. (1991) 'Groundwater in Basin Management', in R.A.Downing and W.B.Wilkinson (eds) *Applied Groundwater Hydrology—A British Perspective,* Oxford, UK: Clarendon, pp. 16–34.

Pescod, M.B. (1977) 'Surface Water Quality Criteria for Tropical Developing Countries', *Water, Wastes and Health in Hot Climates,* London, UK: John Wiley.

Pescod, M.B. (1992) *Wastewater Treatment and Use in Agriculture,* FAO Irrigation and Drainage Paper No. 47, Rome, Italy: Food and Agriculture Organization of the UN.

Sklash, M.G. and Farvolden, R.N. (1979) 'The Role of Groundwater in Storm Runoff', *Journal of Hydrology,* 43: 45–65.
Strömberg, B. and Banwart, S. (1994) 'Kinetic Modelling of Geochemical Reactions at the Aitik Mining Waste Rock Site in Northern Sweden', *Applied Geochemistry* 9: 583–595.
Theis, C.V. (1940) 'The Source of Water Derived from Wells: Essential Factors Controlling the Response of an Aquifer to Development', *Civil Engineering* 10: 277–280.
UN Department of Technical Co-operation for Development (1985) *The Use of Non-Conventional Water Resources in Developing Countries,* Natural Water Resources Series No. 14, New York, USA: United Nations DTCD.
Wardle, T. (1893) *On Sewage Treatment and Disposal for Cities, Towns, Villages, Private Dwellings and Public Institutions,* London, UK: John Heywood.
Wright, P. (1992) 'The Impact of the EC Urban Waste Water Treatment Directive', *Journal of the Institution of Water and Environmental Management* 6: 675–681.
Younger, P.L. (1993) 'Possible Environmental Impact of the Closure of Two Collieries in County Durham', *Journal of the Institution of Water and Environmental Management* 7: 521–531.
Younger, P.L. (1995) 'Hydrogeochemistry of minewaters flowing from abandoned coal workings in the Durham coalfield', *Quarterly Journal of Engineering Geology* 28, 4: S101–S113.

SUGGESTED READING

Appelo, C.A.J. and Postma, D. (1993) *Geochemistry, Groundwater and Pollution,* Rotterdam, The Netherlands: A.A.Balkema Publishers.
Fetter, C.W. (1994) *Applied Hydrogeology,* 3rd Edition, London, UK: Macmillan.
Janssens, J.G. (1996) *Water Resources Management, A Key to Sustainable Development,* Proceedings of the Conference on Institutional Options in the Water Supply and Sanitation Sector: Private Sector Participation, Johannesburg, South Africa, 8–11 July 1996.
Moshiri, G.A. (1993) *Constructed Wetlands for Water Quality Improvement,* Boca Raton, Fla., USA: Lewis Publishers.
Pescod, M.B. and Arar, A. (eds) (1988) *Treatment and Use of Sewage Effluent for Irrigation,* London, UK: Butterworth.
Twort, A.C., Law, F.M., Crowley, F.W. and Ratnayaka, D.D. (1994) *Water Supply,* Fourth edition, London, UK: Edward Arnold.
US EPA (1992) *Manual. Guidelines for Water Reuse,* EPA/625/R-92/004, Washington, DC: US Environmental Protection Agency and US Agency for International Development.

SELF-ASSESSMENT QUESTIONS

1 Which of the following measures have the potential to increase the natural yield of a surface water resource?
(a) Flow frequency analysis.
(b) Floodplain restoration.
(c) Construction of an impounding reservoir.
(d) Wastewater treatment and re-use.
(e) Conjunctive use with groundwater.

2 In assessing a possible groundwater development, which of the following is the most important consideration?
(a) The pre-development recharge rate.
(b) Likely reductions in surface water flows which the abstraction may ultimately cause.
(c) Likely reductions in groundwater storage due to the abstraction.

3 Which of the following pollution problems may be caused by industrial and urban wastewater discharges?
(a) Saline intrusion in aquifers.
(b) Acid deposition.
(c) Depleted dissolved oxygen levels in rivers.
(d) Algae blooms.

4 Which of the following will wastewater treatment achieve?
(a) Avoidance of sewage sludge disposal.
(b) Reduction of organic pollution loadings on rivers.

(c) Prevention of the presence of oestrogens in drinking water.

5 Water conservation plays a role in water resources sustainability through:
   (a) Expanding desalination provisions,
   (b) Ensuring that water resources developments outpace rising demand.
   (c) Managing water demand to sustain legitimate uses.

4

AIR, QUALITY MONITORING AND MANAGEMENT

Mickel Christolis, Nicos Manalis, Dimitra Brillaki, Nicolas Markatos

SUMMARY

Major air pollutants are monitored continuously in urban areas in order to meet programmes of air pollution management and control. Monitoring networks may have different objectives, and it is important to define a network's aims precisely. Since air pollution is variable both in time and space, a large number of measurement stations are needed, but because of their high cost, a consensus has to be sought. Different systems are used in network design, such as land-use method, but no system is perfect. In order to obtain an accurate representation the site of a given station must be chosen very carefully. The treatment of data and the presentation of the report is of the utmost importance. Furthermore, in order to ensure a high level of performance by the network, a proper selection of instruments must be made and their operational specifications checked. Quality assurance programmes are of great importance to the efficiency of the measurement and calibration procedures that must be frequently undertaken.

ACADEMIC OBJECTIVES

An effective air quality management programme requires that reliable information on air quality be collected, analysed and evaluated regularly and efficiently. The purpose of this chapter is to provide the student and the practitioner with a general overview of air quality monitoring, with the focus on the following topics:

- Air quality monitoring networks.
- Air quality network design systems.
- Air quality monitoring methods and instruments.
- Selection of methods and instruments.
- Calibration.

Although the material presented here is by no means exhaustive, it should provide a solid foundation for further reading and more detailed work.

INTRODUCTION

Major air pollutants (GEMS/AIRI, 1994; AEA Technology, 1996)

Air pollutants may be classified according to their physical state, that is gaseous, liquid or solid. The latter two states imply that the material is present in the atmosphere in particulate form, so that the natural classification by physical state is reduced to either gaseous or particulate forms. Air pollutants may also be classified into two general groups: (a) those emitted directly from sources, called primary pollutants; and (b) those produced in the air either by interaction among two or more primary pollutants or by interaction with normal atmospheric constituents, called secondary pollutants.

The major air pollutants are described below.

Sulphur dioxide (SO_2). This pollutant is a colourless acidic gas with a choking taste. High concentrations of SO_2 can irritate the respiratory system, induce coughing or result in (usually reversible) changes in lung function. SO_2 is associated with asthma and chronic bronchitis and is especially harmful to human health in combination with particulate matter. This gas also corrodes stonework and other materials, can damage plants and, when combined with water vapour in the atmosphere, contributes to the formation of acid rain. It is oxidised more readily by atmospheric oxygen when it is in aqueous aerosols. Natural emissions account for about half of all atmospheric SO_2 (UNEP, 1991). Man-made SO_2 is produced by the combustion of sulphur components which are a natural constituent of coal and oil. Major sources world-wide include domestic fuel combustion, industrial processes and power stations. Small quantities are also produced by diesel-powered motor vehicles. The major sinks of SO_2 in the atmosphere are the precipitation and dry deposition of it as sulphate in rain or as solid particulate matter. Ambient concentrations of SO_2 are usually highest in central city areas, in kerbside environments and around industrial areas.

Nitrogen oxides (NO_x). Oxides of nitrogen include nitric oxide (NO), nitrogen dioxide (NO_2) and other oxides that are non-significant for air pollution. Nitric oxide (NO) is a colourless, odourless gas. Nitrogen dioxide (NO_2) is a reddish-brown, toxic gas which causes throat and eye irritation and may also exacerbate asthma and increase susceptibility to infections. Nitrogen oxides, in the presence of sunlight, can react with hydrocarbons to produce photochemical pollution. They therefore play an important role in the formation of ozone. Their emissions can also be oxidised in air to acid gases which contribute to the production of acid rain. Nitrogen oxides are emitted from natural and man-made sources in nearly equal quantities, but whereas natural emissions tend to be equally distributed world-wide, man-made sources are concentrated in centres of population (UNEP, 1991). NO_x is formed during high temperature combustion processes from the oxidation of nitrogen in the air or fuel. The major outdoor sources of NO_2 in urban areas are fuel combustion in motor vehicles, power generation, heating plants and industrial processes. Relatively limited quantities of NO_2 are emitted directly from these sources; most NO_2 in the atmosphere is subsequently formed by oxidation of NO emissions. Concentrations in urban areas are usually dominated by traffic emissions, being highest in central areas and close to major roads.

Ozone (O_3). Ozone is a bluish gas about 1.6 times as heavy as air. It can irritate the eyes and air passages, causing breathing difficulties, and may increase susceptibility to infection. It is a highly reactive chemical, capable of attacking surfaces, fabrics and rubber materials, and toxic to some crops, vegetation and trees. Ozone is a secondary pollutant produced by reaction between nitrogen dioxide (NO_2), hydrocarbons and sunlight. Whereas NO_2 acts as a source of ozone, NO destroys ozone acting as a local sink. For this reason, ozone levels are not as high in urban areas as in rural areas. As the NO_x and hydrocarbons are transported out of urban areas, the NO is oxidised to NO_2, the breakdown of which produces ozone. In practice, exposure of the population to ozone will tend to be higher in suburban areas than in central parts of cities or in areas downwind of major populated or industrialised regions. Levels also tend to be high in cities with basin-type topographic situations such as Mexico City, Athens and Los Angeles.

Carbon monoxide (CO). This pollutant is a colourless, odourless gas slightly lighter than air. It is a dangerous asphyxiant because it combines strongly with the haemoglobin of the blood and reduces the blood's ability to carry oxygen to cell tissues. It can cause stress, headaches, fatigue, respiratory problems and, in very severe cases, death. Carbon monoxide is produced by the incomplete combustion of fossil fuels or organic materials and its major outdoor source is motor traffic. The spatial distribution of CO concentrations in urban areas is strongly traffic-dependent and therefore tends to follow that for NO_2. Levels are highest at the kerbside but decrease rapidly with increasing distance from the road.

Suspended Particulate Matter (SPM)-PM_{10}. The term 'suspended particulate matter' covers the range of fine solids or liquids dispersed in the atmosphere, as opposed to large size fractions which rapidly settle out due to gravity. PM_{10} represents the fraction of particulates in air of very small size (=10 µm). They are associated with a wide range of respiratory problems including an increased risk of heart and lung disease. In addition, they may carry surface absorbed carcinogenic compounds into the lungs. For this reason, concern about the potential health impacts of PM_{10} has increased very rapidly over recent years. PM_{10} is produced from incomplete combustion processes (e.g. coal and diesel smoke), from industry, construction and natural sources. Emissions from road transport are thought to be the major source of PM_{10} in urban areas.

Hydrocarbons. There are two main groups of hydrocarbons of concern: volatile organic compounds (VOCs) and polycyclic aromatic hydrocarbons (PAHs). VOCs are released in vehicle exhaust gases as burned fuels and emitted by the evaporation of solvents and motor fuels. PAHs are also produced by incomplete combustion of fuels. Of the VOCs, benzene and 1,3-butadiene are of most concern as they are known carcinogens which may cause leukaemia. Other VOCs are important because of the role they play in the formation of the ozone. PAHs comprise a complex range of chemicals, some of which are also carcinogenic.

Lead (Pb). Lead is a cumulative poison and, in sufficient body loading, can cause damage to the central nervous system resulting in behavioural changes and intellectual impairment. Lead is the most common of the heavy metal pollutants, and its largest single source is in emissions from motor vehicles using leaded petrol. There may also be significant emissions over a small local scale from some industrial activities. Since traffic is the major urban source in countries not utilising unleaded petrol, spatial distributions of Pb tend to follow closely those of CO and NO_2.

Air quality standards (Newill, 1976; Stern, 1986; WHO, 1987a and 1987b; EEC, 1996a and 1996b)

Air quality standards (AQS) are the legal limits placed on the levels of air pollutants in the ambient air during a given period of time. AQS are expressions of public policy and thereby requirements for action (WHO, 1987a). Thus, they are based not only on the relationship between pollutant concentrations in the air, with associated adverse effects, but also on a broad range of other considerations. By way of example, factors taken into account in setting AQS are the degree of exposure of population and in particular of sensitive subgroups, climatic conditions, sensitivity of flora and fauna, sites of historic heritage exposed to pollutants, economic and technical feasibility, etc. (EEC, 1996b).

The main steps in the development of air quality standards can be divided into two stages:

1 Scientific stage (ISO, 1981 a)

- Knowledge of air pollutant (characterisation of air pollutant).
- Evaluation of the risk (establishment of the air quality criteria that determine the relationship be tween pollutant concentrations and adverse ef fects).
- Assessment of hazard (determination of routes of exposure and estimation of the number of people exposed).

At the conclusion of this stage it should be possible to determine the air quality goals that are the concentrations of pollutants we believe we can live with, without having adverse effects on health and general welfare.

2 Political and administrative stage

- Determination of acceptable risk: this not a sci entific matter, but a matter of opinion.
- Determination of the groups to be protected, considering not only healthy individuals but also population groups whose state of health needs to be taken into account.
- Choice of control technology, requiring both formulation of strategy and selection of appro priate control techniques.
- Legislation: identification of necessary legal strat egies.
- Economics: striking a balance between costs and benefits.

Standard procedures and measurement methods of ambient air are vital factors in setting and controlling air quality standards. Establishing standards, including air quality standards, is the subject of Chapter 6 in Volume 1 of this textbook series (Hens and Vojtisek, 1997).

AIR QUALITY MONITORING NETWORK DESIGN

Monitoring objectives

The term 'air quality monitoring' is used to signify the process of making repetitive quantitative observations, for defined purposes, of the concentrations of one or more pollutants in the air or of other indicators of the state of the environment, according to pre-arranged schedules in the given space and time.

The first step in designing and implementing any air quality monitoring system is to define its overall objectives. Typical monitoring objectives are as follows:

Establishing a sound scientific basis for policy development. Often little is known about the location and nature of air pollution problems. Monitoring to obtain information on existing air quality helps to determine whether an air pollution problem exists, its extent and its seriousness. This information is obtained from networks of monitoring stations whose locations should be widely accepted as being representative of the conditions under consideration. Policy development can then be appropriately targeted (AEA Technology, 1996; Munn, 1981).

Determining compliance with air quality standards. There may be a legal requirement to check whether air quality standards are being met. Reference to statutory or national air quality standards and guidelines provides an indication of whether air quality levels are acceptable in terms of potential impacts on human health or the environment (GEMS/AIR I, 1994; AEA Technology, 1996).

Assessment of population/ecosystem exposure. Monitoring is required to determine exposure to air pollution. Various kinds of air quality data are needed, so that pollutants can be related to effects on human health (epidemiological studies), ecosystems (vegetation damage) and materials (historical heritage damage, metal corrosion), or to economic impacts, etc. There is a need for studies of both episodic extremes and long-term means.

Assessment of air quality trends. Air pollution levels are being changed over time and it is therefore necessary to monitor long-term trends to ensure pollution problems are not developing. Identifying temporal and seasonal variations in pollution concentrations also helps establish links between air quality and other parameters, such as traffic levels, land use and weather conditions (AEA Technology, 1996). This objective is particularly important in gathering data to estimate the impact of future actions likely to affect pollutant emissions, and in evaluating the efficacy of air pollution control strategies (EEC, 1996a and 1996b).

Public information. There is a need to provide information to the public on the quality of the air they breathe. As well as providing up-to-date information, real-time measurement data can be used to alert the public in the event of high pollution episodes which may cause health problems for individuals with respiratory illnesses (episode notification) (AEA Technology, 1996). The specific objective is very important and is now defined in several regulations (EEC, 1996b).

Assessment of the effectiveness of control policies. Monitoring may determine whether current policy on issues such as traffic planning, urban development or emissions controls is effective in controlling air pollution. This information can help the relevant authorities to determine what further action, if any, is required to improve air quality.

Forecasting (short/long-term predictions). Monitoring is required for the development, validation and operational use of pollution forecasting models. The data requirements for model development and validation may be different from those required for operational use of the models. In the case of emergency standards, designed to protect against acute incidents, it is also implied that data should be available on a real-time basis.

Source-effect detection. Monitoring to detect or determine the effect of a particular source has particular siting requirements. There must be sufficient resolution in time and space to distinguish the effect of the source on air quality over advected or background levels. Diffusion modelling is often carried out using sources and meteorological data.

Research needs. Monitoring is needed to answer scientific questions, to evaluate atmospheric processes (photochemical evolution), to study pollution transport phenomena, etc.

Environmental impact assessment. Monitoring is required for the development and validation of multiple-source air pollution models, which predict pollution levels and distributions. These models are used in the preparation of various land-use scenarios for industry, transportation and power generation.

Evaluation of urban climate change. The introduction of pollutants into the air over a city changes the radiation balance, thus affecting the vertical temperature profile and wind field and consequently the dispersion characteristics of air pollutants.

Identification/validation of dispersion models. Air pollution monitoring is required for the identification/ validation of dispersion models, in order to increase the understanding of the behaviour of air pollution sources and sinks, and to develop monitoring networks and control strategies (AEA Technology, 1996).

Air pollution variability

General considerations

Air pollution is a dynamic process which varies in both space and time. The most important factors affecting the space and time variability of pollution concentrations are:

(a) meteorological conditions, notably wind speed and direction, and atmospheric stability;
(b) non-uniform emission rates (irregular distri bution of sources, emission cycles etc.);
(c) local topography and urban terrain which change the wind field, resulting in turbulent flows;
(d) chemical and physical transformations, for ex ample, deposition, coagulation or chemical transformation of suspended particulates.

These factors are interrelated to a certain extent, for example, space heating requirements (and therefore emission rates) depend on temperature and wind. In fact, high correlations between urban air quality and meteorological conditions are usually obtained only when emissions are approximately constant. When appropriate space and time averages are used, the effects of each of the above-mentioned group of factors on air quality can usually be isolated.

Time variability

Concentrations of a pollutant at urban monitoring stations show considerable variability. Variations tend to be largest near emission points, decreasing downwind as the pollutants become well mixed throughout the lower atmosphere. The variations can be classified according to the following time-scales (Munn, 1981):

(a) Instantaneous variations

These are due to fluctuations in wind speed and direction, because the flow in the atmospheric boundary layer is strongly turbulent.

(b) The daily cycle

This behaviour is due to cycles in both emission rates and meteorological conditions. The atmosphere, too, has a well-defined diurnal cycle, nighttime stability alternating with daytime convection. A strong diurnal cycle can be observed when an hourly emission measurement is averaged over many days.

(c) The weekly cycle

On Saturdays and Sundays emissions are lower, and the diurnal cycle is different as compared to other days of the week. In the case of automobile traffic, for example, the morning peak is delayed by an hour or so, and it is not as great on Sundays as on weekdays. As a result, the average concentrations of pollutants are generally lower on weekends than on weekdays, and the diurnal cycles are different.

(d) The annual cycle

Annual cycles in air quality are found in almost every urban area, due to annual cycles in emission rates and meteorological factors.

Space variability

The spatial patterns of air pollution are often classified according to size, as follows: (a) microscale (0–100 m); (b) urban scale (5–50 km); (c) regional scale (100–1000 km); (d) continental, hemispheric and global scales (Munn, 1981).

Each spatial scale is related to different scales of air circulation and characteristic time periods. Hence the urban scale is characterised by local flows (such as sea breeze and synoptic winds) and the regional scale by synoptic conditions, while microscale pollution is influenced by urban terrain.

Correlations

The following kinds of statistical correlations exist in air quality data:

(a) Time correlations (auto-correlations) in a series of measurements of a single pollutant taken at one station.
(b) Cross-correlations between measurements of several pollutants taken at one station.
(c) Spatial correlations between concurrent measurements of a single pollutant taken at two different stations.
(d) Lagged correlations between measurements of a single pollutant taken at two different stations, the time of observations differing by a fixed interval.

Monitoring network systems

Introduction

The first step in designing and implementing any air quality monitoring system is to define its overall objectives, from which the specific data quality objectives (DQO) will eventually be derived. The DQOs are the actual specifications needed to design the study, including the level of uncertainty that the data user is willing to accept (accuracy, precision, completeness, representativeness, comparability). Any organisation is likely to have its own specific objectives for air monitoring, but in every case overall objectives need to be clearly defined. Setting diffuse, overly restrictive or ambitious objectives may result in cost-ineffective programmes with poor data utility. In such circumstances, it will not be possible to make optimal use of the manpower and resources committed to air monitoring. Monitoring and data quality objectives must be clearly specified in order to optimise network design, select appropriate priority pollutants and measurement methods, and to determine the required level of quality assurance/ control and data management (GEMS/AIR I, 1994).

It should be noted that since pollutant concentrations vary in space and time across a given area, it is impossible to measure compliance with air quality standards at all places. Therefore measurements can never be more than a sample from the time-space population. The sum of the sample inaccuracy resulting from the finite number of measurements, and of the inaccuracies in the measurements themselves, gives the overall inaccuracy of the air quality measurement. This overall inaccuracy should not exceed the pre-set inaccuracy of compliance monitoring. It does not matter whether the inaccuracy is in the measurement method or in the sampling strategy (Lahman and Van Der Wiel, 1988).

It is also worth mentioning from the outset that with regard to atmospheric conditions, there is a difference between a typical and a representative location. In an urban area, a typical location in terms of air pollution receives a substantial proportion of its dosage from a few adjacent sources, and is situated over a very non-uniform surface (covered by roads, fences, car parks, buildings, etc.). A representative location, on the other hand, is one that minimises the influences of nearby sources and of surface irregularities. In accepting this distinction and assuming that representative sites are desirable, it is important to ask the question 'representative of what?' and to qualify representativeness in terms of management objectives (EPA, 1995).

An important decision regarding the development of monitoring procedures concerns the frequency of sampling. In brief, the sampling frequency, with regard to both the duration of sampling and the interval between successive measurements, should be selected to provide the information desired for a particular set of survey objectives. Air pollution concentration levels fluctuate; in order to protect potential receptors against air pollution damage, it is generally necessary to measure changing pollution levels, particularly during periods of high concentrations. Although the sampling time does not affect the mean value over a long period, it does affect the range of concentrations recorded. In order to select the optimum sampling time, it is necessary to know which pollutant fluctuation periods are of biological, psychological or physical importance (Rossano and Thielke, 1976). Having fixed the minimum sampling period, one may be faced with a choice of instruments. In general, a faster response instrument is more difficult to operate. It is more expensive and produces more data to handle than a corresponding slower response instrument.

Steps in network design

The basic design of an air quality monitoring system involves a number of steps:

- Selection of pollutants to be monitored.
- Determination of the number, mode (fixed or mobile) and location of sites required.
- Selection of appropriate instrumentation, the analytical techniques and sampling frequencies for satisfactory measurement of pollutant levels and behaviour.
- Development of appropriate data handling and analysis procedures.

A variety of information inputs are necessary to develop a siting plan which will provide answers to these questions. Among the more important of these inputs, in any order of priority, are the following:

- Source inventory-stationary sources, mobile sources, area or point sources.
- Meteorological/climate aspects, in particular with respect to prevalent wind regimes and synoptic conditions giving rise to various pollutant trans port modes and stagnation conditions.
- Geographical/topographical information.
- Land-use data.
- Guideline criteria (defined in legislation, experi ence in other regions/countries, etc.).
- Typical spatial variability of pollutants.
- Location, scale and extent of known air pollu tion effects.
- Modelling predictions.
- Population distribution and density.
- Data from existing or previous networks.
- Data from existing or previous studies/surveys (screening surveys-indicative monitoring, mobile monitoring, random sampling, other short-du ration measurement campaigns).
- Typical costs (EPA, 1995).

Network design systems

There are no hard-and-fast rules for network design, since any decisions made will ultimately be determined by the overall objectives. In practice, the number and distribution of air quality monitoring stations required in any extended network depends on the area to be covered, the spatial variability of the pollutants being measured and the required data usages. If identification or quantification of public health hazards associated with air pollution is an important factor, then the network design should consider the needs of, and information from, epidemiological studies. This may require site or pollutant specific approaches.

National networks may serve a variety of functions, and this is often the case with multi-pollutant measurement programmes. Alternatively, networks can be optimised for specific tasks such as assessing ecosystem exposure or determining compliance with statutory requirements for individual pollutants. In general, the former approach offers financial advantages but clearly cannot be optimised for all network functions.

Two contrasting approaches to network design appropriate over national scales are commonly encountered. The first approach uses a suitably spaced grid siting pattern to provide detailed information on the spatial variability and resulting exposure patterns of pollutants. A more flexible technique involves siting monitoring stations or sampling points at carefully selected representative locations, chosen on the basis of required data usages and known emission-dispersion patterns of the pollutants under study. The latter approach requires considerably fewer sites and is consequently cheaper to implement. However, sites must be carefully selected if measured data are to be meaningful (GEMS/AIR I, 1994).

In urban areas, monitoring is usually undertaken at selected sites. The monitoring sites should be representative of specific location types covering, for example, characteristic central urban, industrial, residential, population-exposure, commercial or kerbside areas. Appropriate site selection must involve consideration of a variety of possible data inputs, including the following:

Overall monitoring objectives. These usually determine the target areas for the study and the number of sites required. Whilst traffic-oriented monitoring may involve kerbside or near roadside sites, health studies may require data from background, suburban and city centre population exposure locations.

Source and emissions. Compilations of emission data can assist substantially in site selection. These will help to identify the most polluted areas, as well as other areas where population or environment impacts may be expected. If background (baseline) concentrations are to be assessed, then monitoring sites should be adequately separated from local pollution sources. By contrast, source-oriented monitoring will often seek out areas where ground-level impacts are at a maximum. If a full emission inventory is not available, alternative statistics such as population distribution, traffic flows and densities and knowledge of principal industrial emission sources may help in estimating likely pollution 'hot spots'.

Existing air quality. Monitoring may already have been undertaken in target areas. If no such studies have been carried out, special screening surveys (indicative measurements) can be designed so as to provide area-wide or local information on pollution problems. These often involve passive samplers and/or mobile monitoring laboratories.

Meteorology/topography. Prevailing meteorological conditions and topographic features strongly influence the dispersion of air pollutants (or, in the case of secondary pollutants, affect their formation in the atmosphere). If used in conjunction with emission data, suitable dispersion models may be used to provide an initial assessment of likely pollution 'hot spots'. Subsequent pollution measurements can also be used to test the models on which predictions are based, providing a basis for prediction in areas where monitoring is not possible.

Model simulations. Results of model simulations can be used to predict pollutant dispersion or deposition patterns, and thus help in site selection. It should be mentioned that these results are of real use only where a reliable inventory is available and the model has been validated.

Other data. These may include demographic, health, population and land-use information and can serve to identify likely effects, mainly health impacts resulting from population exposure to air pollutants.

The site selection process must also take into consideration the spatial distribution of the pollutants within urban environments. For example, CO concentrations are highest at roadside locations, whereas ozone levels have high spatial uniformity, but will be lowest in near-road locations due to scavenging by vehicle NO emissions. For this reason, it may not be possible to optimise measurements for all pollutants at one site location (AEA Technology, 1996; GEMS/AIR I, 1994).

When only three or four stations are planned for a city, a widely used rule of thumb is to select one site in an industrial, another in a commercial and a third in a residential area. When the planned number of monitoring stations exceeds about six, a more objective method for network design is required.

The following network design approaches are common:

1 The land-use approach (Christolis et al., 1992)

In land-use systems, one or more stations are located in different zones of the area. Each zone is homogeneous in terms of type of land use, emission inventory, population density, and topographic and urban terrain characteristics. Quantified estimates of pollution can be provided prior to station selection by air circulation modelling. Usually three categories of stations are considered according to pollution type:

(a) for pollution in the vicinity of a point source (mainly industrial contaminants);
(b) for pollution in the immediate vicinity of heavy traffic (carbon monoxide, lead, etc.);
(c) for pollution occurring on urban, regional or even national scales characterised by weak con centration gradients.

A similar approach is used when distinguishing between two monitored zones:

(a) extended zones such as residential districts or the central districts of cities where stations repre sentative of the whole area should be established;
(b) limited zones in the vicinity of important sources, such as roads with heavy traffic, where high levels of pollution are likely to be measured and many people are affected (EEC, 1985).

2 The statistical approach

The statistical method is based on the analysis of spatial correlations among measurements of existing stations or on determination of the variance of the pollutant parameters (e.g. medians and 98th percentiles) obtained from a pilot survey already made in the area. This design approach is especially applicable to optimal allocation of existing networks in order to assess the minimum number of stations required and obtain the true spatial average value of the parameters to within the accuracy and confidence ranges specified. The statistical techniques used are the distance criterion, the coefficient of geographical variation, structure functions, correlation analysis, the Buell method, cluster analysis, and principal components analysis (Munn, 1981).

3 The grid approach

In the grid system, the area to be considered is classified into districts (cells) and the stations are placed on the intersections of a rectangular grid according to population density, that is 1 km spacing in the city centre area, where concentrations are probably the highest, rising to 2–5 km in suburban areas. This approach is also used for the implementation of a pilot survey by random and short-time measurements, using mobile stations, in order to estimate the spatial distribution of air pollution and to select the fixed station sites.

4 The modelling approach

The basis for the modelling method is that dispersion of pollution can be broadly predicted from a knowledge of emission and meteorological characteristics. It is therefore possible to predict the general features of the concentration field, and to optimise a network to meet any given monitoring objective, for example, to select sites where the concentrations are most likely to be highest in a city. The main drawback of this method is that it requires information on source strength and on the meteorological fields throughout the city. Provided that the forecasts are better than might be expected from chance, this system can help in the appropriate siting of stations.

There are many different kinds of air quality models for a single point source as well as for multisource cases. The most familiar are those of the transfer-diffusion type, sometimes widened to include chemical transformation and deposition terms.

5 Specific measurement techniques

One such technique is that of indicative monitoring (e.g. passive samplers). Here diffusion tubes are densely installed in the area to be surveyed for sampling periods ranging from several days to one to four weeks. This technique, which is very convenient, cost-effective and already widely applied for NO_2, seems to have been tested only for SO_2. It has been used for the redesign of the air quality networks of Paris and Madrid (Atkins *et al.,* 1990). The main drawbacks of this method are that it may not give data in correct time-scale (e.g. annual average rather than the 98th percentile) and includes a relatively high level of uncertainties in the data, for example, uptake rate (AEA Technology, 1995).

Another technique involves the use of mobile-laboratory measurements carried out in real time along streets. This method is convenient for identifying and establishing the distribution of traffic emission sources in urban areas, and for the assessment of sampling sites for short-term air quality standards (Cerutti *et al.,* 1987).

A combination of the above approaches is the best strategy for obtaining more information about the spatial and temporal distribution of pollutants, and for ensuring the most useful siting of monitoring stations.

As more monitoring data become available, and as pollution source distributions or meteorological conditions change, there is a need for periodic reviews of the efficiency of existing air quality networks (EEC, 1996a). Although the optimal monitoring network for a given year may be different from that for another year, the lifespan of a network should be as long as possible in order to achieve continuity of data. Therefore, an optimal network based on an estimated spatial distribution or on data for a year or a season will not be optimal for actual pollution fields during its lifespan (Shinto and Yukio, 1990).

Continuous air monitoring at fixed locations is the most common method, ensuring a time series of data for many years. Mobile sampling, as opposed to the operation of fixed stations, can be a very useful supplement, providing the possibility of evaluating air quality over large geographical areas, checking the accuracy of fixed stations, designing air quality networks and making interpolations between two fixed stations.

Network density

One of the most difficult problems in the design of air monitoring networks concerns the optimal spacing of individual sites. The question facing network designers is 'when is the distance between two sites great enough to justify the siting of another station?'

Once the distribution of concentrations in the area is known, the final choice of a measuring site will be based on the predicted impact assessment of this site. According to EEC Directives for air pollution, three criteria govern the requirements for network design (EEC, 1996b):

(a) The exposure criterion—the integrated exposure encountered by a person relative to the time spent there.
(b) The maximum pollution criterion—measurement stations should be installed at sites where pollution is greatest and humans are likely to be exposed.
(c) The criterion of representativeness—this criterion should be interpreted as the requirement to place stations where human beings are exposed and in such a way that the concentration field is representative of the case under consideration.

The number of sites and stations depends largely on the kinds of pollutants to be measured. The monitoring of photochemical pollutants requires a small number of stations covering a large area; more specifically, ozone stations should be sited in suburban areas at some distance from the busy central district of cities and roads carrying heavy traffic. For SO_2, suspended particulates and black smoke, many stations are required where local emission sources exist.

Empirical guidelines for determining the density of air monitoring networks have been published both in the literature and in USEPA regulations, and an equation for estimating the number of stations has been given, based upon population and the levels of pollution in different zones (EPA, 1984b; EEC, 1996a). For a city with a population larger than 500,000 and where pollution levels are high, a minimum of six to eight stations for SO_2 and TSP (total suspended particulates), and two stations for CO, O_3 and NO_2 has been recommended.

After evaluating monitoring networks for SO_2 and suspended particulates in the EEC countries, one might suggest the following approach to network design: (a) monitoring stations should be sited in all areas where concentrations are likely to exceed 75 per cent of any limit value; (b) the representativeness of local conditions should be interpreted on a scale of 1×1 km; and (c) the maximum interstation distance should be about 4 km in the 'risk' areas (Keddie *et al.*, 1983; McInnes, 1984; Beier *et al.*, 1985).

The range of possible air quality surveillance network configurations is limited mainly by considerations of available resources. In order to arrive at an adequate and acceptable configuration, a compromise must be reached between system objectives and available resources in terms of funds, staff and instruments.

Single-site selection

Concentration gradients in cities are distorted for three reasons: (a) buildings and other obstacles disturb the air flow and generate additional turbulence; (b) multiple heat sources change the local temperature profile, generating turbulent buoyancy flows; and (c) moving vehicles stir the air between buildings on either side of the street—the so-called 'street canyon' effect. Because of the flow-building interaction, a highly complex, three-dimensional, recirculating flow region is established behind the building, and pollutants may become trapped in the cavity wake in high concentrations.

Along a street canyon, stagnation zones may occur near the ground which result in the accumulation of pollutants. If a fumehood or a short chimney on a roof is located on the upwind side of an obstacle, puffs of pollution can reach ground level from time to time with very little dilution.

Sampling devices may disturb the air flow and the concentration fields. Instruments that are particularly susceptible to this kind of effect include the high volume sampler and the dustfall can. In each case, the collection efficiency is a function of wind speed and turbulence intensity.

Over open countryside far from emission sources, wind speeds and concentrations of trace gases usually increase from ground level up to a height of about 10 m; at greater elevations vertical gradients are, most often, insignificant. The magnitude of concentration gradients near the ground depends on the rate of gas uptake by the surface and the atmospheric vertical exchange rate (Munn, 1981).

To assist in the selection of monitoring sites, three general ways of estimating the concentration fields and variabilities will be mentioned, although none is totally satisfactory (EEC, 1996b):

(a) Mathematical modelling

This method can be used in a few idealised cases, such as forecasting the concentrations of pollutants downwind of an expressway that has a generally open exposure. Computational fluid dynamics models provide the possibility of simulating the wind and pollutants concentration fields around complex urban terrains at the microscale (McInnes, 1984).

(b) Wind-tunnel models

A second approach is to test a scale model in a wind tunnel. This method can reveal the velocity field and the concentration patterns, but is time-consuming and costly, requiring, *inter alia,* the simulation of various wind directions and stability categories (Robins and Castro, 1977).

(c) Qualitative appraisal

A third method is to enlist the help of engineers, chemists and micrometeorologists in the selection of sites, relying on their judgement to make the best possible choices according to the stated objectives of the monitoring programme, on the basis of appraisals of the effects of local geometry and pollution sources on air quality.

In selecting the detailed location of sampling points at the microscale level, numerous additional factors need to be considered:

- Height of inlet.
- Manifold design/location.
- Sheltering of inlet by trees, buildings, etc.
- Interfering sources.
- Security (possibility of vandalism).
- Access (accessibility for site visits).
- Electrical power.
- Telephone.
- ‘Visibility’ of the site—public profile.
- Safety of public and operators.
- Planning requirements (AEA Technology, 1995; EEC, 1996b).

According to the USEPA specifications for sampling point selection, the sampling probe should be at a distance of at least 1 m from any vertical or horizontal surface, the flow around it should be unrestricted and a minimum distance from furnaces, fences, trees and other sources or sinks should be ensured. The suggested distance is usually 10–20 m, depending on the pollutant under consideration (EPA, 1984b).

Air quality modelling

Ideally, emission and meteorological data should serve as input to appropriate air quality prediction models for plotting the pollutant patterns expected under various assumed conditions. This has been carried out in a number of situations, and the use of these models, albeit severely limited at present, is to be encouraged. Some of the limitations of formal predictive air pollution models are:

(a) available meteorological data are often in sufficient to permit full exploitation of the ca pabilities of the models;
(b) for primary contaminants, such as carbon mon oxide and sulphur dioxide, existing models for predicting long-term average concentrations are more reliable than those used for predicting short-term conditions;
(c) certain meteorological conditions, such as those involving stability breakup or fumigation, are not easily modelled;
(d) models for the prediction of contaminants formed by atmospheric reactions are in an ear lier stage of development than those used for non-reactive primary contaminants;
(e) complex topography or urban terrains result in strongly turbulent flows, the modelling of which is a very difficult task;
(f) models require fine scale emission inventories for city scale modelling.

Field models that solve the full Navier-Stokes and conservation equations over three-dimensional domains also suffer from some of the above uncertainties and require considerable computer time. They are, however, the most promising, since they include all the physical and chemical aspects of pollution, and they may well become the management tools of tomorrow (Ziomas *et al.*, 1995; AEA Technology, 1995).

Complaint and opinion survey

In an area afflicted by chronic air pollution, there is generally a high incidence of complaints. Many of these complaints are directed at official agencies such as air pollution control, public health and police offices, where records are kept. An insight into the nature and severity of a local problem may be gained by analysing these records in terms of the number, type, time period, nature and location of complaints. Mapping these data may assist in exemplifying the character of the problem.

Since air pollution is a matter of public concern, it is useful to know the degree of awareness as well as the attitudes of the community. Such information may be obtained through public opinion surveys (Rossano, 1986).

A case study of air quality monitoring network design is outlined in Box 4.1. A methodology of network design was recently applied in the Greater Athens Area, based mainly on the 1996 guidelines of the EEC Council Directive on ambient air quality assessment and management.

MONITORING METHODS AND INSTRUMENTS

Classification of methods and instruments

Monitoring networks usually consist of a number of stations spread over the area to be measured. A monitoring station contains a variety of instruments for air pollutant measurements. The instruments may be manual, semi-automatic, automatic, or a combination of these types. Meteorological parameters (e.g. wind speed and direction, temperature, humidity) must also be measured at each station-site. In addition, monitoring networks include laboratories with the dual role of serving as centres from which the monitoring operation is conducted, and of carrying out various laboratory tests and analyses. Usually such centres are responsible for record-keeping and administration (Schneider, 1976; Taylor, 1985a; Harrison and Perry, 1986).

Monitoring instruments

In general terms, a monitoring unit is used to sample and measure several aspects of the ambient air. Taking into account the existing level of automation, monitoring instrumentation can be classified as follows (Rossano and Thielke, 1976):

(a) Manual instruments

The sampling procedure is performed manually. The laboratory analysis of samples may be more or less automated.

(b) Semi-automatic instruments

The sampling procedure is performed automatically by means of a timeswitch combined with a sequential sampler. The laboratory analysis may be more or less automated.

(c) Automatic instruments

Sampling and measurements are performed automatically, without human intervention. Sampling, analysis and signal processing of the measurement are combined in a single device.

Monitoring methods

Air measurement method may be classified as follows:

(a) Chemical methods

These methods involve trapping the gas in a suitable liquid medium (for this reason they are also called wet chemical methods), followed by chemical analysis of the trapped material. The gas to be sampled is drawn, by a pump, through a tube, the downstream end of which is below the surface of the liquid. The gas forms bubbles and is dispersed into liquid which can be analysed later.

(b) Physical methods

These methods involve direct measurement of a physical (usually optical) property either of the pollutant itself or of some product of interaction with another compound.

The physical methods include also the remote sensing techniques. One such technique—based on the differential absorption spectrometry (DOAS) principle—has recently become commercially available. In this technique the concentrations are measured as averages over the path length of an optical beam which may be of the order of a few hundred metres. In principle, a large number of inorganic and organic gases may be measured by DOAS technique.

BOX 4.1 AIR QUALITY MONITORING NETWORK DESIGN IN GREATER ATHENS AREA, GREECE

Athens, a city of about 4 million inhabitants is (like most big cities in the world) facing serious air pollution problems. It is located in an area of complex topography within the Athens basin (~ 450km^2). The city is surrounded by moderated high mountains (up to 1,400 m) forming a channel with only one major opening towards the sea (Saronikos Bay) to the south-west. The mountains act as physical barriers with only small gaps between them. The most important of these gaps is the channel leading to the north-east part of the Attica peninsula, giving the Athens basin access to the Etesians, the system of semi-persistent northerly winds that reduces the likelihood of prolonged pollution episodes. However, the weakening of the background synoptic wind allows the development of local circulation systems, such as sea/land breezes along the axis of the basin (NE to SW) and anabatic/catabolic flows from the surrounding mountains. In such cases the ventilation of the basin is poor, the Planetary Boundary Layer is shallow and the air pollution potential increases. The Greater Athens Area (GAA) gathers about 40 per cent of the total Greek industrial activities and about 50 per cent of the total automatic traffic (Ziomas et al., 1995). Due to its location, local meteorology and dense emissions, the GAA is severely impacted by photochemical smog.

The first automatic air quality monitoring network in the GAA began to operate in 1983 under the auspices of the Ministry of Environment, Planning and Public Works. In 1996, the network was evaluated and redesigned under the programme styled the 'National Ambient Air Quality Monitoring Network'.

The site selection in the GAA was based on:

- the guidelines of the EEC Directive on ambient air quality assessment and management (1996a);
- the evaluation of the previous network and the use of data from large-scale experiments that took place in the area;
- the distribution of population, the meteorology and topography of the GAA, the emission data from industry, traffic, etc.) as inputs in a GIS system;
- the development of the building area in the future years;
- the use of a photochemical model (Urban Airshed Model-UAM).

The following methodology was applied. Two types of stations were considered:

1 Urban background (centre, residential).
2 Street canyon.

With reference to the selection of urban background stations, the study was based on the pollutants NO_x and O_3, due to the fact that the photochemical pollution is the greatest problem in the GAA. It was considered that the stations measuring the pollutants NO_x and O_3 could also serve to measure the pollutants SO_2 and PM_{10}.
As the first stage in the selection of sites for measuring NO_x, the UAM photochemical model was applied for adverse meteorological conditions in the area, and the pollution levels for NO_x estimated. As the second stage, the GAA was divided into three types of sub-areas, based on the mean values of NO_x:

1 Sub-areas where the maximum hourly concentration of the pollutant is between +20 per cent and -20 per cent of the mean value of the concentrations that exceed the value 100 μgr/m3 in the GAA.
2 Sub-areas where the maximum hourly concentration of the pollutant is between the value 100 μgr/m3 and -20 per cent of the mean value of the concentrations that exceed the value 100 μgr/m3 in the GAA.
3 Sub-areas where the maximum hourly concentration of the pollutant is higher than the mean value of the concentrations that exceed the value 100 μgr/m^3, +20 per cent.

As the third stage, the distribution of population density was brought into consideration. The working group estimated the products in each sub-area as concentration of the pollutant×population density, and expressed them as indexes. Hence three types of zones were posited in every sub-area.

1 Zones with high indexes.
2 Zones with medium indexes.
3 Zones with low indexes.

Finally, the zones with the highest indexes in every sub-area were selected, and it was proposed that NO_x should be measured in those zones. The same methodology was applied for O_3. As a result, the number of the urban background stations proposed was 10 in the GAA, with a considered spatial representativeness of 5×5 km.

It was proposed that three street canyon stations should be established. These stations were located at different typical streets in terms of traffic characteristics (density, average speed, composition of the traffic), the particular street's configuration (ratio of height to width) and its orientation.

Source: Christolis *et al.*, 1997

Manual or semi-automatic instruments are generally based on chemical methods, while automatic instruments tend to depend on the physical properties of pollutants for identification and quantification. In the past the operation of most automatic instruments was based on chemical methods. However, certain inherent problems mean that they are not entirely satisfactory for typical field application. They must (for example) be attended frequently, the reagents used are unstable, and the instruments require complex pumping and accurate pumps, which render them bulky and heavy (Rossano and Thielke, 1976).

Data acquisition systems

The flow of air quality data from the monitoring instrumentation to the ultimate data user is effected by the data acquisition system. Approaches to this range from manual techniques to highly automated microprocessor-based—and computer-based—systems (Bryan, 1986). With manual and the semi-automatic instruments the data acquisition is performed manually.

Automatic instruments and meteorological sensors generally produce some form of continuous electrical analogue signal, which varies with the parameter being measured. The final format and the ultimate storage of the data are effected by data acquisition systems. These systems can be classified, according to the level of automation, as follows:

(a) Analogue recording systems

Analogue recording is based on the use of chart recorders which produce an analogue trace proportional to the signal. Charts are periodically removed and manually reduced.

(b) Digital recording, on-site storage

The analogue signals from the analysers are transmitted to an on-site digital data acquisition system (e.g. data logger), which performs an analogue-to-digital conversion, calculates measurements averages and places the averaged values into a storage device (e.g. internal memory of data-logger).

(c) Fully automatic digital systems

The data from an on-site digital acquisition system can be transmitted by means of telephone or radio lines to a central site where the data are recorded and stored into a computer for the subsequent evaluation.

Commands can also be transmitted from the central site to the field stations through the telecommunication system, to initiate zero and span calibration of the monitoring instrument or to control the remote station activities.

Selection of methods and instruments

Introduction

The choice of measurement devices for an ambient quality surveillance (a.q.s.) monitoring programme should be based on the objectives defined (e.g. air pollution warning system, compliance with air quality standards, determination of air pollution trends, etc.) and factors related to the measurement devices (e.g. operational specifications). For example, only automatic devices should be used for the purpose of issuing warnings.

Prior to any selection of monitoring instrumentation for AQS purposes, the number of air pollutants to be measured must be considered. Even though this number is practically unlimited, ambient quality standards have been set for only a few of them. These major, or 'criteria', air pollutants have had proven and well-documented adverse effects on people and their environment at concentrations likely to occur at present (Pitts and

Pitts, 1986). Monitoring must also be conducted for certain other pollutants (e.g. trace metals) generated by local or regional features such as fuel used, specific industrial plants and special activities. Furthermore, several other aspects of polluted air must be measured in order to understand and evaluate the behaviour and fate of the main pollutants.

Air pollutants commonly considered as major are listed below.

- *Gases:* Sulphur dioxide (SO_2), nitrogen dioxide (NO_2), nitrogen monoxide (NO), carbon monoxide (CO), ozone (O_3) and non-methane hydrocarbons (NMHC).
- *Suspended particulates:* Total suspended particulates (TSP), suspended particulates, PM_{10}, black smoke and lead (Pb).

Criteria for instrument selection

After selecting the air pollutants to be measured, the criteria for selecting monitoring instruments must be set. The factors affecting this choice of equipment are as follows:

(a) Funds available

The major part of the capital cost of an AQS system is in the cost of equipment and instruments. The cost of manual and semi-automatic instruments is low in comparison with continuous automatic analysers. On the other hand, labour costs associated with the operation of non-automatic instrumentation are higher than in the case of automatic instruments.

(b) Equivalent of the monitoring methods

In establishing air quality standards the monitoring methods must be clearly specified (Puzak and McElroy, 1987). Of those available, the most accurate serve as reference standards against which other techniques can be compared for determining equivalence. Standards or equivalent methods are necessary with air quality standards.

(c) Sampling frequencies

The greater the frequency of fluctuating pollutant levels, the more numerous are the samples required to assure statistical accuracy. Therefore, the air quality standards also include the factor of sampling frequency. The sampling frequency of an automatic instrument is usually of the order of one minute, but manual and semi-automatic methods require special consideration.

(d) Maintenance staff

The skill requirements for trained staff to operate automatic equipment are high. On the other hand, manual and semi-automatic instruments do not require highly trained personnel, although the subsequent laboratory analysis of samples involves chemists or highly trained technicians.

(e) Operating specifications and housing requirements

Factors such as the temperature and humidity of the station's environment, power requirements and unattended operational periods must be considered when selecting monitoring methods and instrumentation.

(f) Performance specifications

Performance specifications referring mainly to automatic instruments and their generally accepted definitions are as follows (Bryan, 1986):

- *accuracy*—the degree of agreement between a measured value and the true value, usually expressed as a positive or negative percentage of the full scale (see below);
- *precision*—the level of agreement between repeated measurements of the same concentration value, expressed as the average deviation of the data from the mean;
- *noise*—spontaneous deviations from a mean output not caused by input concentration changes;
- *lower detectable limit*—the smallest amount of input concentration that can be detected as the concentration approaches zero, usually twice the noise level at zero concentration;
- *lag time*—the time interval between a step change in the input concentration at the instrument inlet and a reading of 90 per cent of the ultimately recorded concentration;
- *rise time (90 per cent)* —the time interval between initial response and 90 per cent response, after a step increase in inlet concentration;
- *fall time (90 per cent)* —the time interval between initial response and 90 per cent response after a decrease in concentration;
- *zero drift*—the change in instrument output over a stated period, usually 24 hours, of unadjusted continuous operation when the input concentration is zero, usually expressed as a percentage of the full scale;

- *span drift*—the change in instrument output over a stated period, usually 24 hours; of unadjusted continuous operation when input concentration is a stated up-scale value, usually expressed as a percentage of the full scale;
- *interference*—an undesired positive or negative output caused by a substance other than the one being measured;
- *linearity*—the maximum deviation of the mea sured value from the corresponding value on a straight line curve drawn between upper and lower calibration points;
- *full scale*—the maximum measuring limit for a given operational range.

Principles of commonly used monitoring methods

This section deals with the principles of the methods commonly used for measuring major air pollutants.

Sulphur dioxide (SO_2) (Ciccioli and Cecinato, 1992)

(a) Fluorescence (automatic apparatus)

Sulphur dioxide molecules are excited to a higher electronic state using an appropriate light source (zinc or cadmium lamps, at 213.8 and 228.8 nm), and the intensity of the emitted radiation (220 to 400 nm range) is monitored as it returns to the ground state. The instrument employs a pulsed UV source which irradiates the sample gas as it flows continuously through the optical cell. The fluorescence emission is detected by a photomultiplier tube, situated at right angles to the light source to eliminate interference from the excitation beam. The electronic signal-processing device transfers the light energy impinging on the photomultiplier into a voltage, which is in direct proportion to the concentration of SO_2 in the sample stream being analysed.

Lower detectable limit: 0.002 ppm. Response time: 4 min.

(b) Pararosaniline method (manual or semi-automatic apparatus)

Air is drawn through a solution of sodium tetra-chloromercurate (TCM) by means of an impinger or a fritted glass bottle, and sulphur dioxide present in the sample is absorbed by formation of a stable complex. This complex reacts with pararosaniline and formaldehyde to form intensely coloured pararosaniline-methyl-sulphonic-acid. The absorbency of the solution is measured spectrophotometrically in the laboratory. The method is applicable to the measurement of sulphur dioxide in ambient air using sampling periods of up to 24 hours. It is recommended by the International Organization for Standardisation (APHA Intersociety Committee, 1977; EEC, 1980).

Lower detectable limit: 0.010 ppm.

Nitrogen dioxide (NO_2)

(a) Chemiluminescence (automatic apparatus) (Harrison and Perry, 1986)

The principle of this method is based on the detection of light produced from the energy released in the chemical reaction of nitric oxide (NO) with ozone (O_3). Ozone is produced by a generator (e.g. of electrical discharge type), which forms an internal component of the instrument. Air containing NO is drawn into the monitoring instrument, where it is mixed with an excess of O_3. The product of the reaction of NO with O_3 is the electronically excited nitrogen dioxide, NO_2*, a light-emitting species. The NO_2* relaxes by photon emission, and the intensity of the emitted light is proportional to the concentration of NO in the air sample. Only NO is detected directly by the chemiluminescence method. The NO_2 is monitored by convening it to NO by means of a molybdenum converter, which is an internal component of the instrument, at 200°C.

Lower detectable limit: 0.002 ppm. Response time: 3 min.

(b) Griess-Saltzman method (manual apparatus) (Harrison and Perry, 1986; APHA Intersociety Committee, 1977; ISO, 1981a)

Nitrogen dioxide is absorbed by a solution of an azodye forming reagent, by means of an impinger. A pink colour is produced, the intensity of which may be read on a photometer. This is the most commonly used of the wet chemical methods and is recommended by the International Organization for Standardization (WHO).

Lower detectable limit: 0.005 ppm.

Ozone (O_3)

(a) Ultra violet absorption (automatic apparatus)

In a single-cell instrument, air is alternately drawn through the sampling cell directly (first mode), and through a catalytic converter (second mode) which decomposes O_3 without causing loss of other atmospheric constituents. The two-mode cycle eliminates the interferences of other UV-absorbing species (e.g. aromatic hydrocarbons) as the difference in readings between the two modes is solely a function of the O_3 present. Ultra violet light, from a mercury vapour lamp, passes through the cell in each mode and is directed at a light detector photocell where the light intensity is measured for each mode of operation. The electronic part of the instrument calculates the concentration of O_3, using the Lamber-Beer absorption law.

Lower detectable limit: 0.002 ppm. Response time: 1 min.

(b) Neutral KI method (manual apparatus) (APHA Intersociety Committee, 1977)

This method is intended for the manual determination of O_3 and other oxidants in the range 0.01 to 10 ppm. The oxidants are absorbed in a neutral buffered solution of KI liberating I_2, which is measured spectrophotometrically. The procedure is suitable for calibrating other methods in the laboratory.

Carbon monoxide (CO) (EPA, 1984a)

Gas filter correlation nondispersive infra-red spectroscopy (automatic apparatus)
Radiation from an IR source (a simple glowing wire) is directed upon a rotating wheel which contains two IR-transparent gas filter cells. One cell contains nitrogen (N_2 filter), which does not absorb radiation, and the other contains a high concentration of CO (CO filter), which absorbs all the frequencies of the CO absorption spectrum in the incident light. It then passes through the sampling cell where the sample is drawn continuously by means of a pump. In general terms, the CO and the interfering species absorb the N_2-filtered beam, while only the interfering species absorbs the CO-filtered beam (since the CO spectrum of the beam has already been absorbed by the CO filter). This difference in the light intensity of the beams is proportional to the concentration of CO in the sample.

Lower detectable limit: 0.1 ppm. Response time: 30 sec.

Non-methane hydrocarbons

Flame Ionisation Detection, FID (automatic apparatus) (Pitts and Pitts, 1986)
Individual hydrocarbons can be separated and measured using several chemical techniques (e.g. gas chromatography with FID). For routine monitoring purposes, however, such analyses are not feasible due to the time required for each analysis and the need for highly trained personnel. In addition, air quality standards, where they exist, are currently set in terms of total NMHC (non-methane hydrocarbons) rather than individual hydrocarbons, making separation of the components unnecessary from a compliance viewpoint. As a result, several automatic apparatuses for measuring hydrocarbons without separation of the individual components are now in use. Most of them utilise FID, the electrical response of which is linearly proportional to the concentration of hydrocarbons. A feature of such analysers is the ability to discriminate between total and non-methane hydrocarbons in the sample. This is important because, due to its low reactivity, the mainly naturally generated methane is of little significance in photochemical smog formation. The monitors incorporate a scrubber system for the removal of all non-methane hydrocarbons from the sample stream, so that methane is periodically separated and measured independently at selected time intervals.

Therefore, the analysis involves two operational modes, one for monitoring the total hydrocarbons (by passing through the scrubber system) and the other for monitoring methane. The NMHC concentration is then calculated.

Recently, fully automatic gas chromatographers for performing more detailed separation of mixtures of hydrocarbons have become commercially available. One of these is the automatic BTX analyser used to monitor the benzene, toluene and xylene in continuous basis. The detection is achieved by a flame ionisation detector (FID).

Suspended particulate matter (SPM)

(a) High-volume (Hi-Vol) for total suspended particulates (Harrison and Perry, 1986)
A high flow-rate blower (1.1 to 1.7m^3/min) draws the air sample into a covered housing and through a 20×25 cm rectangular glass-fibre filter. The mass of particles collected on the filter is determined by the filter material before and after exposure. The method is also frequently applied to the analysis of aerosol chemical species, which are measured by appropriate methods after extraction from the filters with convenient solvents.

(b) Black smoke method (OECD method) (Harrison and Perry, 1986)
The procedure is performed by drawing a sample of contaminated air through a filter which is held on a clamp, and collecting the suspended particulates on filter paper. The flow rate of the sampler is low (1–1.6 litres/min) and the sampling duration is 24 hours. The darkness of the stained filter is assessed using a reflectometer. This instrument consists of a light source and a photosensitive element mounted together on a measuring head. Light reflected from the filter paper falls on the sensitive element and the current generated is measured. The darker the stain, the lower the intensity of the reflected light. The results can be converted into concentrations of smoke by use of calibration curves.

(c) PM_{10} (particulate matter with an aerodynamic diameter < 10µm)
In the USA, the PM_{10} concentration is the quantity measured to protect human health from effects caused by inhalation of particulate matter. There are commercially available automatic instruments that monitor PM_{10}, utilising measurement by, for instance, b-ray absorption or tapered element oscillating micro-balance techniques. Used with a carefully designed inlet head these instruments can be used to sample the PM_{10} specific size fraction of particulate matter.

Quality assurance (Puzak and McElroy, 1987; Adams and Farewell, 1986; Taylor, 1985b)

Quality assurance may be defined as those operations and procedures that are undertaken to reduce measurement errors and to ensure that the results have a stated high probability of being correct.

Two concepts are involved in quality assurance:

- Quality control—the mechanism used to reduce and maintain random and systematic errors within tolerable limits.
- Quality assessment—procedures that monitor the quality control mechanism and evaluate the quality of the data produced.

A quality assurance programme is necessary in atmospheric pollution management to verify and evaluate the quality of the measurement data. In general, the lower the pollutant concentration or the more complex the sample matrix, the more rigorous the quality assurance must be.

The basic elements under consideration of a well-designed quality assurance programme are:

- general network design;
- specific sampling site selection;
- sampling and analytical methodology;
- probes, collection devices, storage containers and sample additives or preservatives;
- special precautions, such as heat, light, reactivity, combustibility and holding times;
- reference, equivalent or alternative test procedures;
- instrumentation selection and use;
- calibration and standardisation;
- preventive and remedial maintenance;
- replicate sampling;
- blind and spiked samples;
- collocated samplers;
- QC procedures such as intralaboratory and intrafield activities;
- documentation;
- sample custody;
- transportation;
- safety;
- data handling procedures;
- service contracts;
- measurement of precision, accuracy, completeness, representativeness and comparability;
- document control.

Calibration (Woodfin, 1984; ASTM, 1975)

Introduction

One of the principal components of a quality assurance programme is the calibration procedure. Broadly speaking, calibration means the comparison of an instrument's measurement capability with a standard, in order to report, or eliminate by adjustment, any deviation in the accuracy of the instrument.

A prime requirement for calibration is the availability of appropriate and accurate standards. A standard is a gas mixture of known composition, generally containing the substance with respect to which the instrument will be calibrated, and a complementary gas such as pure nitrogen or synthetic air.

The standards may be distinguished according to the method used to determine the concentration, as follows: (a) primary standards, where the concentration is calculated from physical quantities such as mass, volume or pressure; and (b) secondary standards, where the concentration is measured by an analytical method.

Two different systems are used to generate known concentrations of gases: (a) static systems, when an appropriate volume of component gas is added to a known volume of air; and (b) dynamic systems, when a stream of the component gas is blended with a pure air stream to continuously produce a desired concentration.

Static systems

The static systems used in the laboratory are usually containers with flexible or rigid walls that can hold a volume of pure air or nitrogen. A quantity of the component for calibration is added to this container (e.g. by syringe) and filling is completed by means of the chosen complementary gas, until the desired pressure is obtained. This pressure is usually greater than atmospheric pressure, so as to permit easy use of the mixture. Static systems are frequently considered as generating primary standards, since the component concentration is calculated by measurement of physical quantities such as container and syringe volume, pressure and filling temperature (ISO, 1981b).

Dynamic systems

The conditions necessary for dynamic systems are a constant supply of pure air and a supply of the component at flow rates that are in the correct ratio, are accurately known, and can be kept constant over an extended period of time. The supply of the calibration component can be provided by cylinders of accurately known concentration, calibrated by the manufacturer or by permeation tubes.

The procedure using the permeation tube is based on the permeability of plastic materials (e.g. teflon) to gases under pressure, and involves keeping a liquid substance inside a sealed plastic tube. The vapour above the liquid diffuses through the tube wall at a rate dependent on the tube temperature, liquid substance, tube material, wall thickness and surface area. The emitted vapour is diluted into a gas stream which passes over the tube. Permeation tubes may be calibrated by the user if a microanalytical balance is available. The normal calibration procedure is to weigh the tube once a day over a period of several days whilst maintaining the tube at a constant temperature. The concentration of the component can be calculated by the weight loss and flow rate of the diluting gas.

A dynamic system utilising a permeation tube calibrated by the user in the laboratory may be considered as generating a primary standard, since the concentration of the component is computed from the measurement of physical quantities. Dynamic calibration systems using cylinders or permeation tubes are commercially available.

Calibration procedure

Usually in an AQS the following main items of equipment are used for calibration purposes:

(a) In-station calibrators (e.g. a dynamic system utilising permeation tubes) are used to perform zero and span checks. The zero and span checks indicate changes in system performance and demonstrate whether the instrument is operating within control limits (e.g. +/-10 per cent for spans) or not. In-station calibration, which can be automated, must be performed once a day.

(b) Transfer calibrators (calibrators, verified against a primary standard) are used to perform the multipoint calibration. Multipoint calibration involves measuring multiple known inputs over the range of the analysis for checking the accuracy and linearity of the instrument.

Multipoint calibration must be performed whenever in-station calibration indicates changes in system performance that are outside the control limits.

A well-equipped laboratory with calibration devices for accurate flow measurements is vital for the support of an AQS network system.

CONCLUSIONS

The purpose of this chapter is to provide a general view of air quality monitoring, and it focuses on the steps necessary for the design of an effective air quality monitoring network. The first step in designing and implementing such a system is to define its overall objectives. Although any organisation or country is likely to have its own specific objectives for air monitoring, these should be clearly defined in order to optimise network design. A wide variety of information should be collected, and specific choices need to be made: the pollutants to be monitored; the number, mode (fixed or mobile) and location of sites required at national and urban scale; the detailed location of sampling points at the microscale level. It is also very important to select the appropriate monitoring methods and instruments for every pollutant separately. The question of the most suitable instrumentation concerns not only monitoring stations, but also the central monitoring laboratories where the various tests and analyses are carried out. Finally, in order to verify and evaluate the quality of the measurement data, and in order to achieve a high level of performance for any air quality monitoring network, it is necessary to develop a complete quality assurance programme.

REFERENCES

Adams, D.F. and Farewell, S.O. (1986) 'Sampling and Analysis', in A.C.Stern (ed.) *Air Pollution,* VII, New York, USA: Academic Press.

AEA Technology (1995) *Report on Monitoring Network Design for Air Quality Assessment,* UK: National Environmental Technology Centre.

AEA Technology (1996) *Air Quality Monitoring—A Handbook for Local Authorities,* UK: National Environmental Technology Centre.

American Public Health Association (APHA) Intersociety Committee (1977) *Methods of Air Sampling and Analysis,* New York, USA: John Wiley & Company.

American Society for Testing and Materials (ASTM) (1975) *Calibration in Air Monitoring,* USA: Technical Publications No. 598.

Atkins, D.H.F., Lawand, D.V. and Sandalis, R. (1990) *The Distribution of Nitrogen Dioxide in the United Kingdom,* UK: United Kingdom Atomic Energy Authority Report, AERE R 13304.

Beier, R, Gonzalez, P., McInnes, G., Mylle, E., Stevenson, K. and Zierock, K.H. (1985) *A Study of Network Design and Measurement Methods in Member States for the EC Air Quality Directive for Nitrogen Dioxide* (85/203/EEC), Report CR2880 (AP), UK: Warren Spring Laboratory.

Bryan, R.J. (1986) 'Ambient Air Surveillance', in A.C.Stem (ed.) *Air Pollution,* VII, 3rd edition, New York, USA: Academic Press Inc.

Cerutti, C., Sandroni, S., Froussou, M., Assimakopoulos, D.N. and Helmis, C.G. (1987) *Air Quality in the Greater Athens Area,* EUR 12218 EN, Brussels, Belgium: Commission of the European Communities.

Christolis, M., Assimakopoulos, D., Ziomas, G. and Markatos, N.C. (1997) *The Greek National Ambient Air Quality Monitoring Network,* Athens, Greece: National Technical University of Athens.

Christolis, M., Clayton, P., Hecq, P., Payrissat, M. and Petit-Coviaux, F. (1992) *Monitoring Instruction Manual —Volume II: Black Smoke Monitoring,* EUR 14550/II, Brussels, Belgium: ISPRA, Joint Research Centre— Commission of the European Communities.

Ciccioli, P. and Cecinato, A. (1992) 'Advanced Methods for the Evaluation of Atmospheric Pollutants Relevant to Photochemical Smog and Dry Acid Deposition', in J.O.Driagu (ed.) *Gaseous Pollutants,* New York, USA: John Wiley.

EEC (1980) 'Directive No 80/779/EEC', *Official Journal of the European Communities* No L 229, 30 August 1980, pp. 30–48.

EEC (1985) *Council Directive of 7 March 1985 on Air Quality Limit Values for* NO_2, 85/203/EEC, Brussels, Belgium: Commission of the European Communities.

EEC (1996a) 'Council Directive on Ambient Air Quality Assessment and Management' (96/C59/02), *Official Journal of the European Communities* No C59/24, 28 February 1996.

EEC (1996b) 'Directives—Position Papers for Air Quality on Nitrogen Dioxide, Sulphur Dioxide, Lead, PM_{10}', Draft versions.

EPA (1984a) *Performance Test Results and Comparative Data for Designated Reference Methods for Carbon Monoxide,* EPA-600/ 84–83–013, USA: Environmental Protection Agency.

EPA (1984b) *Quality Assurance Handbook for Air Pollution Measurement Systems, Volume I: Principles,* EPA-600/9–76–005, USA: Environmental Protection Agency.

EPA (1995) *Quality Assurance Handbook for Air Pollution Measurements Systems,* USA: Environmental Protection Agency.

GEMS/AIR I (1994) *Quality Assurance in Urban Air Quality Monitoring,* Methodology Review Handbook, Vol. I, Nairobi, Kenya: UNEP/WHO.

Harrison, R.M. and Perry, R. (eds) (1986) *Handbook of Air Pollution Analysis,* London, UK: Chapman and Hall.

Hens, L. and Vojtisek, M. (1997) 'The Establishment of Health and Environmental Standards', in B.Nath, L.Hens, P.Compton and D.Devuyst (eds) *Environmental Management in Practice: Analysis, Implementation and Policy,* Volume I, London, UK: Routledge.

ISO (1981a) *Air Quality Determination of the Mass Concentration of Nitrogen Dioxide in Ambient Air, Modified Griess-Saltzman Method,* ISO/TC 146/sc 3, Dp 6768, Geneva, Switzerland.

ISO (1981b) *Gas Analysis—Preparation of Calibration Gas Mixtures—Static Volumetric Methods,* ISO-6144, Geneva, Switzerland.

Keddie, A.W.C., Lahman, E. and McInnes, G. (1983) *European Community Directive 80/779/EEC: An Assessment of Monitoring Network Design and of the Corresponding Stringency of Annexes I and IV,* Report CR 2455 (AP), UK: Warren Spring Laboratory.

Lahman, E. and Van Der Wiel, H.J. (1988) 'Monitoring and Analytical Methods', personal communication, The Netherlands: RIVM.

McInnes, G. (1984) *Monitoring Network Design and Treatment of Black Smoke and SO_2 Data with Respect to Annex I of the EC Directive 80/779,* EC Workshop, Bilthoven.

Munn, R.E. (1981) *The Design of Air Quality Monitoring Networks,* Hong Kong: Macmillan.

Newill, V.A. (1976) 'Air Quality Standards', in A.C.Stern (ed.) *Air Pollution,* V, New York, USA: Academic Press.

Pitts, B.J.F. and Pitts, J.N. (1986) *Atmospheric Chemistry,* New York, USA: John Wiley & Sons.

Puzak, J.C. and McElroy, F.F. (1987*) The EPA's Role in the Quality Assurance of Ambient Air Pollutant Measurements,* ASTM STP 957, USA: American Society for Testing and Materials.

Robins, A.G. and Castro, I.P. (1977) 'A Wind Tunnel Investigation of Plume Dispersion in the Vicinity of a Surface-mounted Cube', *Atmospheric Environment* 11: 299–293.

Rossano, A.T. (1986) 'Air Pollution Survey', in A.C.Stern (ed.) *Air Pollution,* VII, New York, USA: Academic Press.

Rossano, A.T. and Thielke, J.F. (1976) 'The Design and Operation of Air Quality Surveillance Systems', in *Manual on Urban Air Pollution,* Copenhagen, Denmark: WHO.

Schneider, T. (1976) 'Automatic Air Quality Monitoring Systems', in *Manual on Urban Air Quality Management,* Copenhagen, Denmark: WHO.

Shinto, J. and Yukio, M. (1990) 'Considerations on Air Pollution Monitoring Network Design in the Light of Spatio-temporal Variations of Data', *Atmospheric Environment* 24B, 2: 335–342.

Stern, A.C. (1986) *Air Pollution,* I, New York, USA: Academic Press .

Taylor, J.K. (1985a) 'Principles of Calibration', in *Sampling and Calibration for Atmospheric Measurements,* ASTM 04–957000–17, USA: American Society for Testing and Materials.

Taylor, J.K. (1985b) 'What is Quality Assurance', in *ASTM STP 867,* USA: American Society for Testing and Materials.

UNEP (1991) *Urban Air Pollution,* UNEP/GEMS Environmental Library No. 4, Nairobi, Kenya: UNEP.

WHO (1987a) *Air Quality Guidelines for Europe,* WHO-Regional Publications, European series No 23, Geneva, Switzerland: World Health Organization.

WHO (1987b) *Setting Environmental Standards,* H.W. Koning (ed.), Geneva, Switzerland: World Health Organization.

Woodfin, W.J. (1984) *Gas and Vapour Generating System for Laboratories,* USA: US Department of Health and Human Services.

Ziomas, I., Suppan, P., Rappengluck, B., Balis, D., Tzoumaka, P., Melas, D., Jacobi, G., Papagiannis, A., Fabian, P. and Zerefos, C. (1995) 'A Contribution to the Study of Photochemical Smog in the Greater Athens Area', *Beitr. Phys. Atmosph.* 68, 3: 191–203.

SUGGESTED READING

Atkins, D.H.F. and Quirino, I. (1990) *A Survey of Nitrogen Dioxide in Paris,* EUR 13369 EN, Brussels, Belgium: Commission of the European Communities.

Bartaire, J.G., Bertolaccini, M.A., Fernandez-Patier, R., Houdret, J.L. and Payrissat, M. (1992) *Monitoring Instruction Manual—Volume I: Sulphur Dioxide Monitoring,* Torino, Italy : Environment Institute, ISPRA , Joint Research Centre.

Beier, R., Gonzalez, P., McInnes, G., Mylle, E., Stevenson, K. and Zierock, K.H. (1985) *A Study of Network Design and Measurement Methods in Member States for the EC Air Quality Directive for Nitrogen Dioxide* (85/203/EEC), Report CR 2880 (AP), UK: Warren Spring Laboratory.

GEMS/AIR II (1994) *Primary Standard Calibration Methods and Networks Intercalibrations for Air Quality Monitoring,* Methodology Review Handbook, Vol. II, Nairobi, Kenya: UNEP/WHO.

GEMS/AIR III (1994) *Measurement of Suspended Particulate Matter in Ambient Air*, Methodology Review Handbook, Vol. III, Nairobi, Kenya: UNEP/WHO.

ISO 9004–4 (1993) *Quality Management and Quality System Elements*—Part 4: Guidelines for Quality Improvement, Geneva, Switzerland.

Manns, H., Mucke, H., Nitz, G. and Turowski, E. (1995) *European Intercomparison Workshops on Air Quality Monitoring—Measuring Of SO_2, NO and NO_2*— Vol. 1, Report 7, Berlin, Germany: WHO.

Study of the Redesign of the Paris Air Pollution Monitoring Network (1990) Final Report.

WHO (1976) *Manual on Urban Air Pollution,* Copenhagen, Denmark: World Health Organization.

WHO (1987b) *Setting Environmental Standards,* H.W.Koning (ed.), Geneva, Switzerland: World Health Organization.

SELF-ASSESSMENT QUESTIONS

1 An urban background station is located:
(a) Away from local sources.
(b) Where extreme values of pollutant concentrations are expected.
(c) Within a small distance from a road.

2 A ‘street canyon’ station is located:
(a) In the middle of a road.
(b) At kerbside.
(c) At a certain height of a street building.

3 The statistical approach for network design is used:
(a) When the monitoring network already exists.
(b) When short-term measurements take place.
(c) When air quality models have been developed.

4 The instrument linearity check can be carried out through:
(a) Checking the sampling flows.
(b) Zero and span checks.
(c) Multipoint calibration.

5 Quality assurance activities aim at:
(a) Maintaining the performance of the instrumentation.
(b) Ascertaining the control limits under which the instrument is operating.
(c) Both of the above.

5

NOISE POLLUTION

Causes, effects and control

Marc Van Overmeire, Filip J.R.Verbandt and Robert E.Jonckheere

SUMMARY

Noise is most commonly and most simply defined as unwanted sound. In considering environmental noise, we study how noise affects the health or interferes with the activities of people. We are principally concerned with determining quantitatively the amount and type of noise to which people will be exposed in a given situation. Based upon this, we assess the possible effects on human health and activities. We are also concerned with the potential effectiveness of alternative mitigation measures for reducing noise and its effects upon people.

Considerations in environmental noise assessment may be divided into two categories: those relating to the noise source and those relating to potential receivers. For example, in describing the noise impact characteristics of a source, whether it be road traffic, aircraft, or a stationary source, it is necessary to have a physical description of the sound itself, a description of how the volume varies with time, when the noise occurs and the location of the noise source. The noise environment in which the receiver will be situated must then be assessed and a judgement must be made on the effects of the noise as it relates to sleep disturbance and interference with speech communication, or other human activities. A judgement must then be made as to whether such noise levels are acceptable. All environmental noise problems contain these basic elements.

ACADEMIC OBJECTIVES

The aim of this chapter is to make readers familiar with the terminology used by specialists so that the basic elements of noise can be understood. Readers will be introduced to the various types of sound and the disturbance and damage these may cause. They will be made aware of the difference between an objective noise measurement and the subjective assessment of noise by people. Finally, readers will learn some methods of noise control and laws governing it.

INTRODUCTION

Sounds are an integral part of life, and absolute silence, a very unusual phenomenon, can exist only in anechoic chambers. Sounds are a source of information, for example, the communication we give or receive over the telephone, and a source of pleasure, as with music or birdsong. But sounds can also be unwanted, annoying or even damaging. When they fall into one of these latter categories, they are called noise. The level of disturbance from unwanted sound depends on the sound, the current situation, as well as one's personal response. A dripping tap, loud music or a typewriter may disturb one person but not another. Annoying sound types are those that interfere with other activities, for example, when noise from a machine interferes with speech intelligibility. Damaging sound, on the other hand, can be evaluated more objectively because it results in hearing loss or other physiological problems.

In dealing with the practical aspects of environmental noise, it is not necessary to study the detailed physics of acoustics. Instead, we will draw upon the science of sounds as far as needed to develop an understanding of the methods for dealing with environmental noise.

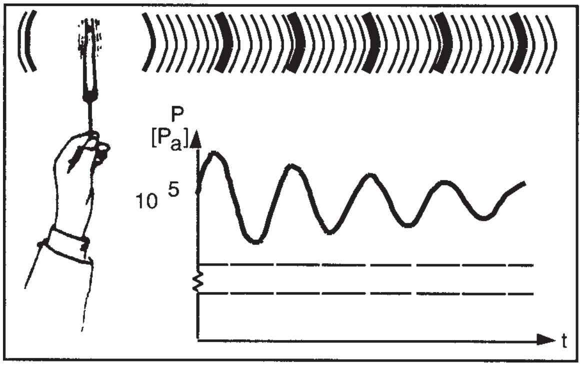


Figure 5. 1 The sound wave emitted by a tuning fork

Source: Brüel and Kjær, 1979a

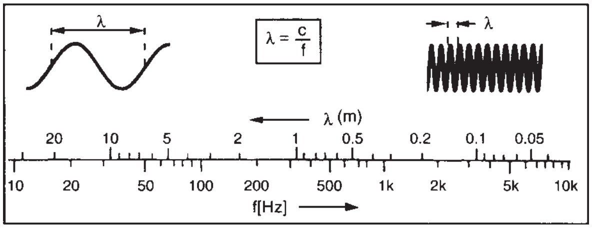


Figure 5.2 Basic relationship between wavelength, frequency for constant sound velocity of 344 m/s
Source: Brüel and Kjær, 1979b

PHYSICAL PROPERTIES OF SOUND

Sound is the sensation we experience when we perceive the vibrations of air particles on our eardrums. The vibrations of these particles give rise to fluctuations in air pressure above and below the prevailing atmospheric pressure (Figure 5.1). The change in air pressure is the physical characteristic of sound to which we most often refer. If sound becomes unpleasant, we call it noise.

Sounds only propagating through a medium (gas, fluid, solid) is a wave propagation phenomenon. Sound waves are generated by a vibrating body or air turbulence. The sound-generating mechanism is called the sound source and the rate at which the disturbance travels through the air is the velocity of sound. It varies with the type and temperature of the medium through which it travels; at a room temperature of 20°C this velocity is 344 m/s.

Frequency is the physical characteristic of a sound that enables us to sense different levels of pitch. It is measured in hertz (Hz), which is the number of cycles of pressure fluctuation per second. The higher the frequency of a sound, the higher the pitch. Figure 5.2 shows the relationship between wavelength, frequency and sound velocity. The amplitude (size) of the pressure fluctuation is known as sound pressure and is expressed in pascal.

For the air pressure fluctuation to be audible, two conditions must be satisfied (in addition to the obvious one that the receiver of the sound should be able to hear). First, the frequency of the fluctuating air pressure (the rate at which the fluctuations occur) should have components within the audible frequency range of the ear (20–20,000 Hz). Second, the amplitude of the pressure fluctuations should be high enough to exceed the threshold of hearing at the frequency involved (May, 1978).

MEASURES OF SOUND

A variety of measuring scales are used to express sound level. It is possible to measure sound objectively by measuring the pressure it adds to atmospheric pressure. Measurements of loudness and noisiness try to take into account the frequency sensitivity of human hearing. The degree of loudness that the human ear registers depends on the frequency of the sound. Loudness measurements, however, use pure tones (a single frequency), while in daily life most sounds are a mixture of tones. Also, different sounds affect people differently. For example, the alternating periods of noisiness and quietness from aircraft taking off and landing may be more disturbing than the more steady noise of an expressway. Different units have been developed to express the variety of sounds to which human beings are exposed. The purpose of disturbance indicators is to be able to arrive at a single number which will reflect the likely annoyance caused by a sound without having to conduct a study of the population's response in every situation (Lawrence, 1989).

Sound pressure level

Sound pressure level is one of the most basic measurements of the level of a sound. Sound pressure is the variation caused by vibrations with regard to atmospheric pressure. The human ear is able to detect variations in pressure from 2×10^{-5} to 200 N/m^2 (Pa). This is a range of about ten million. Instead of using the linear pressure scale, with its unwieldy range of numbers, another, logarithmic, scale has been developed and its unit is the decibel (dB). Most instruments that measure sound pressure level automatically convert the pressure readings to the decibel scale so that it is unnecessary to use the conversion equation. In order to grasp the meaning of a decibel, though, it is useful to understand at a simple level its mathematical and physical basis. It is calculated as follows:

$$L_p = 20 \log_{10} \frac{P_{eff}}{P_0}$$

L_p is the sound pressure level in decibels, P_{eff} is the effective sound pressure, calculated as the root mean square integrated over a period of time of the difference from atmospheric pressure, and P_0 is the reference pressure in air, 2×10^{-5} Pa.

Table 5.1 shows the correspondence of varying sound pressure with different types of sources, as well as the amount of difference between sound interactions.

The instrument used to measure sound pressure levels is called a sound level meter. In this, a microphone receiving the sound produces an alternating voltage proportional to the sound pressure. This voltage is then converted to a sound pressure reading. The meter has a series of band filters in order to measure the sound level of different frequencies.

Loudness

Because the human auditory mechanism reacts differently to different frequencies of sound, the sound pressure level is not useful for showing how loud a sound is when it is perceived by human beings. For example, two tones played at the same decibel level, but one at 100 Hz and the other at 500 Hz, do not seem equally loud: most people will say that the 500 Hz tone is louder.

The loudness of a sound is therefore determined by comparing it with a sound at 1,000 Hz at a given decibel level. Loudness measurements are of pure tone sounds—sounds consisting of a single frequency. Suppose, for example, that we listen first to a sound of 1,000 Hz at 20 dB, and then to a sound of 200 Hz. In order to make the 200 Hz sound seem as loud as the 1,000 Hz sound at 20 dB, the 200 Hz sound must have a sound pressure level of about 27 dB. Thus, the loudness of a sound depends on its frequency as well as on its sound pressure level.

Table 5.1 A sound at the threshold of hearing at 1,000 Hz has a pressure of 2×10^{-5} Pa

(μPa)	*Sound pressure*		*Sound pressure level (dB)*	*Source*
	200,000,000	injurious range	140	Jet engine (25 m)
Threshold of pain			130	Rivet gun
	20,000,000		120	Propeller aircraft (50 m)
			110	Rock drill
	2,000,000	danger range	100	Metalworking shop
			90	Heavy lorry
	200,000		80	Busy street
			70	Private car
	20,000		60	Ordinary conversation (1 m)
		safe range	50	Low conversation (1 m)
	2,000		40	Soft music
			30	Whisper (1 m)
	200		20	Quiet town dwelling
			10	Rustling leaf
Threshold of hearing	20		0	

Source: Milne, 1979:59

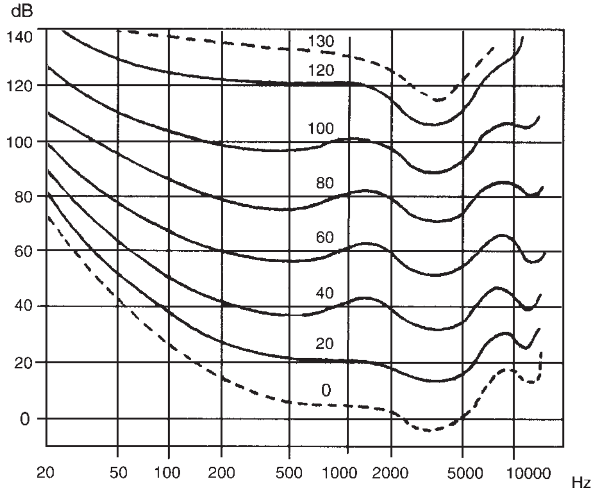
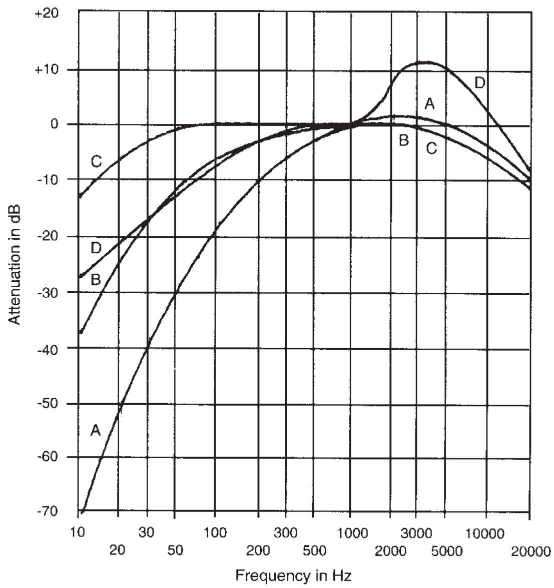


Figure 5.3 Phon curves (lines equal loudness)

Source: May, 1978



octave band, Hz	63	125	250	500	1000	2000	4000	8000
A-correction, dB	–26.2	–16.1	–8.6	–3.2	0	+1.2	+1.0	–1.1

The table shows the corrections factors for the dB(A) scale, where 1,000 Hz is the reference frequency.

Source: May, 1978

Figure 5.4 Sound measurements with A, B, C and D filters.

The loudness level of a tone, in phons, is numerically equal to the SPL of the 1,000 Hz standard. In the above example, then, the 1,000 Hz tone at 20 dB and the 200 Hz tone at 27 dB both have a loudness of 20 phons. Figure 5.3 shows a collection of equal loudness contours, also known as isophon curves. The loudness of a given sound at a particular decibel level can be read from the graph by finding the phon curve that intersects with the frequency and decibel level in question.

Although one might think that a doubling in phon level would result in the tone sounding twice as loud, this is not the case. In fact, an increase of only about 10 phons results in a doubling of loudness. Because it is awkward to calculate how many times louder a tone at 35 dB is than the same tone at 51 dB, a scale called the *sone* scale has been created. The relation between sone (S) and phon (P) is: $S=2^{(P-40)/10}$. A doubling in sone units represents a doubling in loudness, so that 1 sone=40 dB, 2 sones=50 dB, 4 sones=60 dB. In practice, however, the sone is no longer frequently used.

Weighted decibel scales

Another method that has been developed to measure the loudness of sound is to weight the values of sound pressure levels so that they more accurately reflect the loudness of a sound level as perceived by a human being. Four scales have been developed: dB(A), dB(B), dB(C) and dB(D). Within a given scale then, one can say a sound of 50 dB(A) is louder than a sound of 20 dB(A).

The principle behind these scales is that the human ear responds less well to very low and very high frequencies. Some decibels are therefore subtracted from the upper and lower frequencies, while leaving the middle ones about the same. The number of decibels subtracted from a given frequency varies between the scales. Thus the dB(A) scale was derived for sounds that were ‘not loud’, from the 40 phone loudness contour. In the dB(A) scale a tone of 63 Hz will have 26.2 decibels subtracted from its unweighted sound pressure level (Figure 5.4).

The scales can be used for sounds composed of a variety of frequencies. A sound level meter measures the sound level for all the frequency bands, but the signals are put through a filter for the specified scale so that the weighting is done automatically. The decibel levels for each frequency band are then totalled, using a special procedure, to give a single weighted decibel level for a given sound (May, 1978).

The dB(A) scale has proven useful for expressing the loudness of sounds commonly encountered from industry, traffic and building construction. It has also been found that a 10 dB(A) increase results in a doubling of the noise disturbance experienced, leading to more use of the dB(A) scale than the phon or sone, which were introduced in the previous section. Most national and international standards are now expressed in dB(A). The dB(D) scale is used for sounds dominated by high frequency tones such as aircraft noise, where more than 10 dBs are added to the frequencies from 2,000 to 5,000 Hz. PNL dB (perceived noise levels) are also specified in some national standards, and international standards (Smith, 1989).

Noisiness

Loudness measurements use pure tones and a 1/3 octave band of noise. However, most sounds, and particularly noise, are composed of a collection of frequencies. Furthermore, sounds can be masked (hidden) by other sounds whose frequencies include the frequency of the masked sound, and the loudness of a combination of sounds with different frequencies cannot be determined by simply adding together the loudness of the individual sounds. Consequently, a method has been developed that uses bands of frequencies rather than pure tones as a basis for measurement. This method parallels the measurement of loudness, but instead of constructing equal loudness curves for pure tones, curves of equal noisiness use bands of frequencies.

Figure 5.5 shows a plot of equal noisiness curves, also known as noise rating (NR) curves. The x-axis, instead of having the frequencies of pure tones, has the centre frequencies for bands of frequencies. The y-axis is in decibels. Superimposed upon the decibelcentre frequency matrix are the NR curves that are numbered on the right-hand side of the graph. All band centre frequencies with a corresponding decibel level that fall within the same NR curve have the same noisiness. Thus a noise centred at 4,000 Hz and a level of 30 dB and a noise centred at 500 Hz and a level of 39 dB, both have a noise rating of 35.

Because noises may be made up of sounds of more than one frequency band, it is necessary to measure the sound level for each band and then plot the data on a graph of NR curves. The noise rating of the sound is then the highest NR value of the bands of frequencies measured. The sound plotted in Figure 5.5 has an NR value of 83. The 1,000 Hz octave band measurement of 83 dB gives the highest NR, namely 83. Particular NR values are often assigned to various locations. For example, it is recommended that a concert hall should not have an NR above 20, that is the 'noisiness' of the hall when it is empty should not exceed 20.

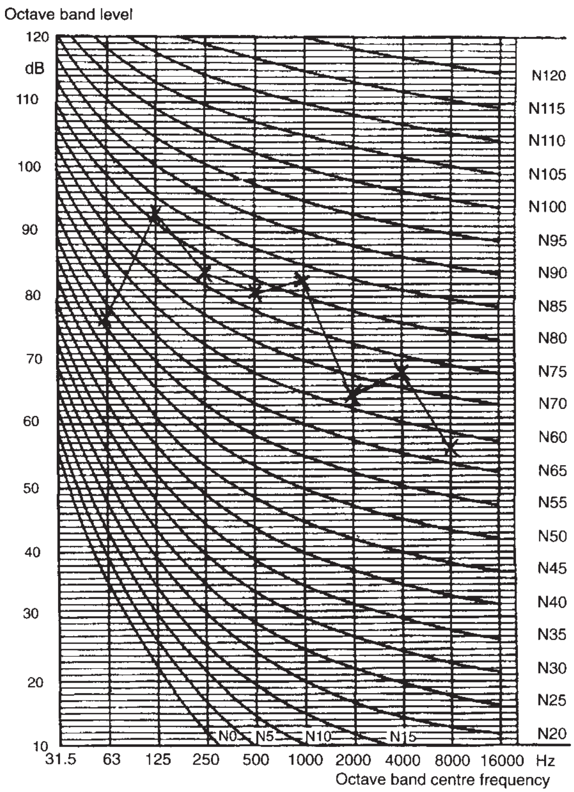


Figure 5.5 Equal noisiness curves. The NR value of the plotted noise is 83.

Noise disturbance as determined by the noise-rating method is useful in that it expresses disturbance caused by a variety of frequencies in a single number. Disturbance caused by noise, though, is not just a function of its level; it is also related to how the noise changes over time. One way of expressing fluctuating sounds is to give the sound level (L_n), usually in dB(A), which is exceeded for a certain proportion of time. L_{10}, for example, is the sound level which is exceeded during 10 per cent of the measuring time, and L_{90} is the sound level which is exceeded during 90 per cent of the measuring time. So L_{10} corresponds to peak sound levels and L_{90} to the background sound level.

Another measure used to express fluctuating levels of sound is the equivalent sound level (L_{eq}). The concept is that the same amount of noisiness occurs from a sound having a high level for a short period, as from a sound with a lower level but persisting long enough for the same amount of energy to be released (May, 1978) as indicated by its formulae (see Box 5.1).

Another way of expressing noise levels has been developed, called the noise pollution level (NPL). It takes into account the amount of variation in the sounds; the greater the variations, the higher the NPL. The NPL can also be used to assess community noise levels with mixed noises (Lawrence, 1989). A good correlation has been found between NPL and the subjective disturbance of a fluctuating sound perceived by human beings. The equations used for calculating L_{eq} and NPL are shown in Box 5.1.

NOISE POLLUTION

Sources and patterns of emission

Sources of noise pollution are usually classified as either stationary or moving. Stationary sources include residential, industrial, construction and demolition sites. Residential noise is assessed either by measuring the noise produced by various domestic appliances or machinery, or in more subjective terms by finding the origins of those sounds that are disturbing (May, 1978). Noise levels have been recorded for some of the more typical industrial noise offenders. Most industrial noise, though, is confined to the facility producing it. In general, an average factory emits about 70–75 dB of noise. Construction noise depends on the particular activity being carried out, such as the breaking of a pavement with jackhammers. Construction noise can reach intermittent peak levels of 110 dB at 3 metres.

Characteristic line noise sources are urban transportation networks and the flight paths of aircraft near airports. Noise from motor vehicles is generated from the engine, exhaust, carburettor inlet, cooling fan, transmission, tyres and aerodynamic effects. Different types of vehicles and different traffic volumes produce different noise levels. Although aircraft noise is a

BOX 5.1 EQUATIONS FOR CALCULATING DIFFERENT TYPES OF NOISE DISTURBANCE

(a) Equivalent sound level (Leq)

$$L_{eq} = 10 \log_{10} \frac{1}{T} (t_1 \times 10^{(L_1/10)} + t_2 \times 10^{(L_2/10)} + \ldots + t_n \times 10^{(L_n/10)})$$

$$T = \sum_i^n t_i$$

where T is the total time and t_n is the time for which the sound level is L_n

(b) Noise pollution level (L_{NP})

$$L_{NP} = L_{eq} + 2.56\,\sigma$$

where L_{eq} is the equivalent sound level as given above and σ is the standard deviation of the instantaneous sound level.

much more localised problem than road and rail transport, it is a significant problem, particularly in urban areas. Newer types of aircraft such as the Airbus series, the Boeing 747, 757 and 767 and the McDonnell Douglas MD-11 are considerably quieter than earlier pure jet aircraft, but the volume of air traffic is increasing and more airports are being built, increasing the noise problem.

In addition to different types of noise sources, different types of patterns of noise generation can be discerned. The categories of patterns are mobility, time and spatial scale. Sources of noise can be either mobile, such as transportation vehicles, or stationary, as mentioned earlier. Noise emissions also vary with the time of day or the day of the week. For instance, road traffic noise may reach peak levels in the morning and the evening, while in recreational areas noise may be highest on the weekends.

There are three special types of noise sources: point, line and surface sources. Line sources mentioned are often composed of a number of moving point sources. Industrial and residential units, when considered individually, are point sources but, because a collection of each type are usually found together, they are treated as surface sources. Construction noise, on the other hand, can be considered as a point source, as it is easy to identify the particular place from which the noise originates.

Generation factors

When investigating the sources of noise and ways of controlling it, one should consider aspects of how it is generated. There are two sets of factors involved in the generation of noise: (1) factors that affect initial emission of noise from the source; and (2) factors that affect the transmission of noise once it is emitted from the source (Berry, 1974). The first set encompasses the physical characteristics of the object generating the noise, for example, the design of an engine. The second set of factors is related to the properties of sound and its transmission. The transmission of noise propagation is affected by geometric spreading—the decrease in noise level as a function of distance from the source; refraction—changes in the speed of sound often due to changes in humidity or wind conditions, air and surface absorption; scattering of sound waves —due to inhomogeneity in the medium, reflection of the sound waves, diffraction; and wall transmission (Berry, 1974:221). It is important to be aware of these effects when trying to control noise emissions. For example, a barrier may be ineffective because the sound is scattered in another direction by air turbulence.

EFFECTS OF NOISE POLLUTION

Health effects

One of the most important effects of noise pollution is hearing damage. The probability of a person suffering damage depends on the level of the noise, its duration, the frequency of its occurrence on a daily basis, the number of years that daily exposure continues, the level of hearing loss one considers as constituting damage and individual susceptibility to this type of injury (Berry, 1974). Noise-induced hearing loss is a result of damage to the hair cells in the cochlea. Ageing is also accompanied by a deterioration in hearing sensitivity, called presbyacusis. It may not be a serious disability in itself, but when combined with noise-induced hearing loss it can be a serious handicap (Martin and Walker, 1978).

Hearing loss can be divided into three phases. In the first phase, the threshold for hearing high frequencies temporarily rises, that is the ear becomes less sensitive to high pitched sounds. Individuals will notice little or no change; a decrease in sensitivity will only appear on an audiogram (hearing test). In the second phase, individuals exposed to intense noise experience a temporary deafness which may last for only a minute or up to several days but eventually disappears. People may also experience a decrease in sensitivity if briefly exposed to very high sound levels such as at a rock concert. This is termed a temporary threshold shift and may not result in permanent damage unless such intense exposure occurs frequently.

In the third phase, a permanent diminution in sensitivity occurs and is characterised by an important loss of sensitivity to sounds of 4,000 Hz. This is called noise-induced hearing loss. As exposure to very intense sounds continues, sensitivity continues to diminish, and the range of frequencies that are difficult to hear broadens (Figure 5.6). It is internationally accepted that a permanent decrease in a person's hearing ability has occurred when his or her threshold of hearing is 25 dB higher than the reference threshold, averaged over the frequencies of 500, 1,000 and 2,000 Hz. Individuals reaching the third phase will always have difficulty hearing high pitched tones, which decreases speech intelligibility and their ability to hear important signals such as a ringing telephone or a car horn. From the third phase on, individuals are conscious of their hearing loss, and it is very important to keep them away from dangerously high noise levels. Further development of hearing loss (apart from presbyacusis) will stop once an individual is not exposed to excessive noise levels. It is best, of course, to prevent individuals from reaching the third phase of hearing loss.

Assuming an exposure of eight hours a day for five days a week, medical-acoustic investigations have shown that 75 dB(A) a day may be considered a safe limit. Levels above 90 dB(A) are dangerous and correspond approximately to an NR value of 85. For higher levels it is necessary to decrease exposure time in order to maintain the same noise dose.

A noise dose is calculated using the following formula:

$$\text{Noise dose} = \sum_i \frac{t_i}{T_i}$$

Ti is the maximum permitted time for a given noise level, and ti is the actual time exposed to the noise level. For 90 dB(A), T is 40 hours a week (eight hours a day, five days a week). In the European Community, the law requires a halving of the exposure time for every 3 dB (ISO standard) increase. Thus the maximum exposure time for 93 dB is 20 hours, and for 96 dB, 10 hours. The maximum allowable noise dose is 1, sometimes expressed as 100 per cent. For example, what is the noise dose a person would receive if exposed to 90 dB(A) for 20 hours and 93 dB(A) for 20 hours? The answer is 1.5, or a 50 per cent overdose; see Table 5.2 for the calculation.

An instrument called a dosimeter is used to measure noise dose directly, as it measures both sound level and duration simultaneously.

Besides hearing loss, other effects of noise on the human physical system include muscular tension, metabolic changes, reduced gastro-intestinal activity,

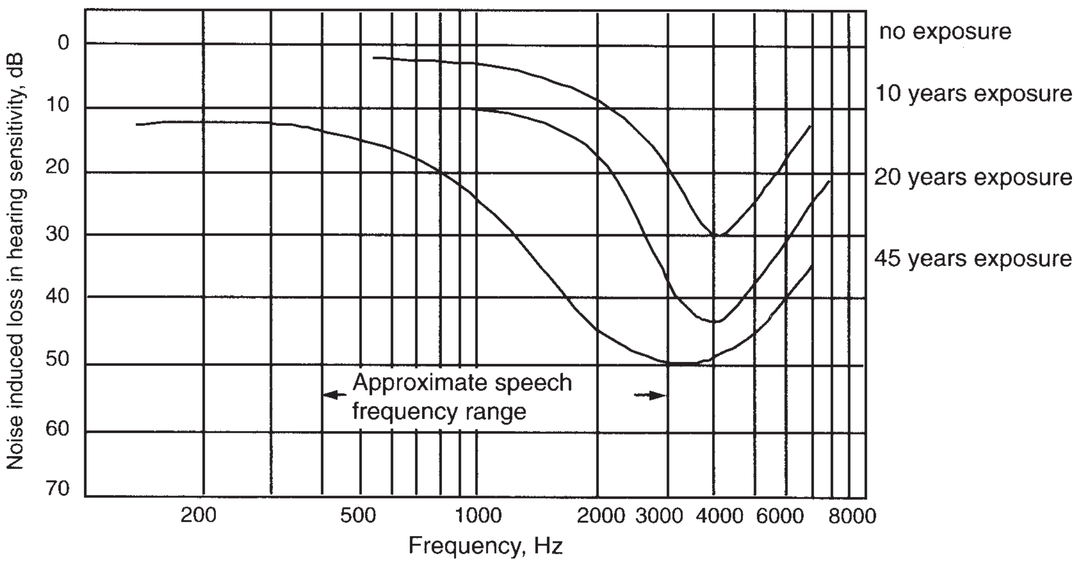


Figure 5.6 Loss in hearing sensitivity due to exposure to noise

Source: May, 1978

Table 5.2 Calculation of a noise dose

	Noise 1 (90 dB)	*Noise 2 (93 dB)*	*Noise dose ($\Sigma\ t_i/T_i$)*
t_i	20	20	
T_i	40	20	
t_i/T_i	0.5	1	1.5

nausea, headaches, tinnitus (ringing or other sensation of noise in the ear), drowsiness and respiratory irregularities, hypertension and isochaemic heart disease (Sound Research Laboratories, 1991; May, 1978). The incidence of these effects depends to a large extent on the individual, and no criteria for 'dangerous' noise levels have been established for them.

Effects on well-being

Besides having serious effects on human health, noise pollution can have negative impacts on psychological, social and economic well-being. Research has been conducted on the effects of noise on sleep, auditory communication, learning and task performance, annoyance and the economy.

Noise does affect the quality of sleep, and those repeatedly disturbed during sleep can suffer both physical and psychological stress such as paranoid delusions, hallucinations, suicidal and homicidal impulses (Berry, 1974). Sleep disturbance depends on a number of variables, such as the characteristics of the noise, fluctuations in noise level, motivation to wake, gender and age.

Clearly, noise interferes with auditory communication and although overall background noise is used as the basis for evaluation, frequency and time distribution also play an important role. The effect of noise on auditory communication is a particularly important consideration in education, where traffic or aircraft noise can seriously disturb students' ability to understand lessons.

The effects of noise on learning and task performance are less clear. The results of studies on these activities are usually complicated and conflicting. Some studies show that there is no effect, some show a detrimental effect, and yet others a positive effect. In fact, in some studies, such as those investigating the effect on two simultaneous tasks, noise is correlated with an increase in performance of one task and a decrease in the other (Lukas, 1978). Although it has not yet been fully investigated, it is expected that the performance of mental tasks would be more seriously affected by noise than that of physical tasks.

The economic effects of noise disturbance have been examined mostly in relation to property values and rent levels, particularly near airports and motorways. It has generally been found that property values are lower in such areas. The economic impacts of noise are an important consideration for urban planners in the siting of roads, highways, railway lines and airports. In the case of airport siting, city governments must appease residents who live in the area of a proposed airport, both because of noise and because of the drop in the values of their homes, which many owners regard as a long-term investment.

NOISE CONTROL

Having established that noise does have an impact on human health and well-being, we now turn our attention to controlling noise pollution. Before it is possible to control noise, it is necessary to analyse the noise, in particular its emission, transmission and immission (Figure 5.7). Although the best approach is to control the production of noise at its source, this is not always possible and so solutions must be found further along the path of the noise. Figure 5.8 shows the production of noise and its transmission paths.

When examining a noise source, one must consider whether the noise is due to air turbulence or to mechanical vibrations. Noise of the former type is produced by ventilators, jet engines, and the intake and exhaust of combustion engines. This kind of noise can be reduced by decreasing airflow velocities

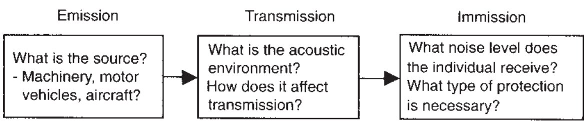


Figure 5.7 Emission, transmission and immission aspects of noise

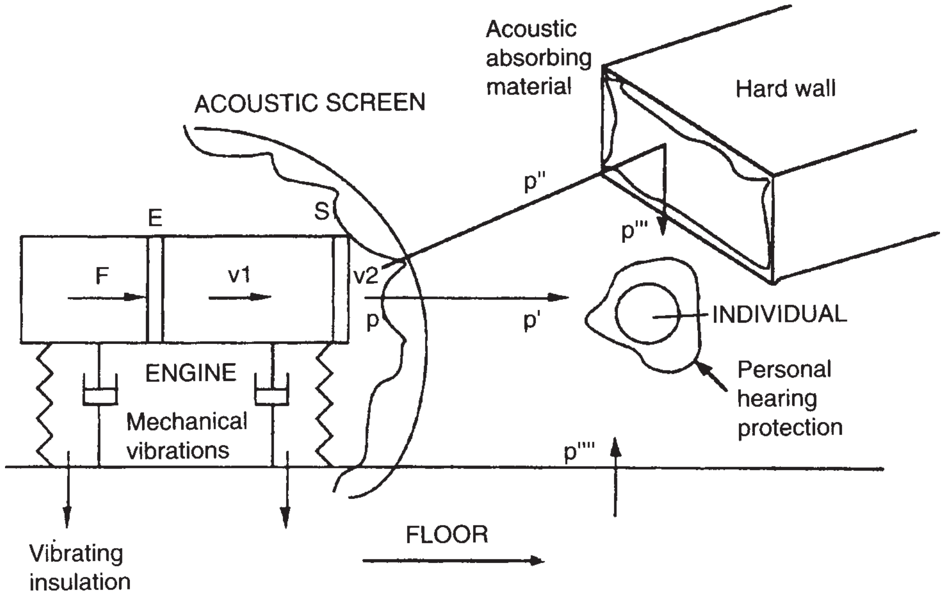


Figure 5.8 Production of noise through mechanical vibration and its transmission paths. A variable force F (sine, periodic, random or impulse) interacts with the structure of the machine at point E, which starts to vibrate with a velocity V . The wave propagates and causes an external surface S to vibrate with a velocity V . This surface will radiate acoustic energy in the form of a sound wave propagating in the air. The sound pressure p will possibly be weakened to value p' before it reaches the human ear. There is also the possibility that wave p?, after reflection on a wall will reach the individual with the sound pressure p?. The individual hears p' and p?. The machine also causes the floor to vibrate, which in turn radiates sound p??.

and duct dimension, the use of turning vanes, avoiding lift fluctuations and steadying combustion. Noise generated by mechanical vibrations can be controlled by isolating or damping the vibrations, eliminating their cause, or by reducing the surface area of the vibrating material.

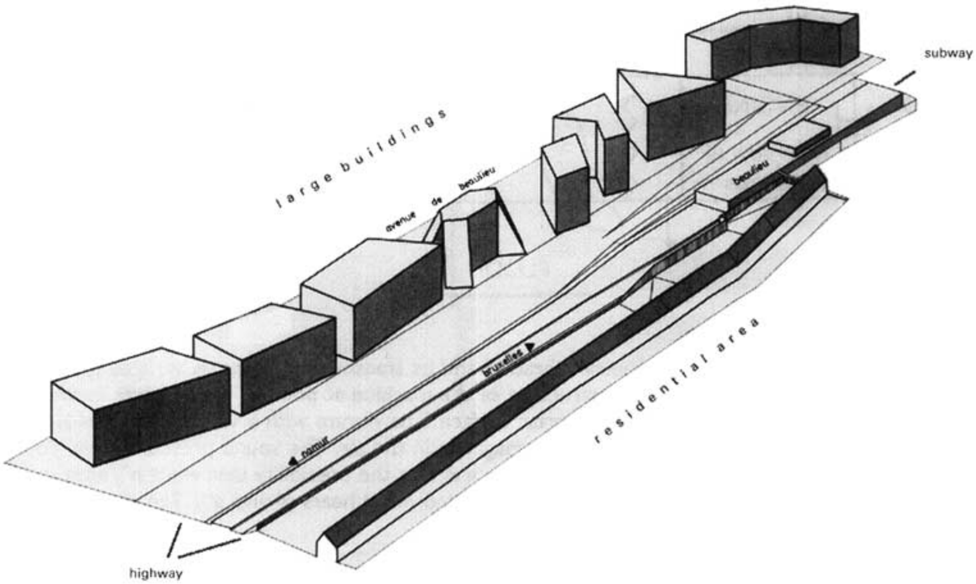
Another method is to change the nature of the vibration including the frequency or force generating the vibration, or the frequency response of the part that vibrates. Such changes can reduce the amplitude of the vibration or shift the frequency of the vibration to a frequency to which humans are less sensitive. With regard to Figure 5.8, the amplitude and frequency spectrum of the force, which is the fundamental cause of the noise, can be modified. Induction and/or damping of the machine structure can be used to weaken the vibration passing from point E to point S. By using the correct shape and material, the external surface, S, will radiate less energy. Finally, insulation material can be placed between the machine and the floor to minimise the transmission of vibrations to the floor, which in turn will start vibrating and radiating noise (p??).

If the noise cannot be completely controlled at the source, the transmission path can be modified by using barriers of sound-absorbing or -reflecting material, and/or by choosing the site so as to minimise the received noise level. Whenever possible, the siting of the noise source should receive careful consideration in order to minimise both noise and the cost of installing barriers. The aim should be to maximise the distance between the source and receiver. However, other factors such as the direction of the prevailing wind must also be considered. Siting is particularly important in the management of noise from highways and airports. When noise control through siting is not possible or not completely effective, noise barriers may be necessary. These include fences, screens, or complete enclosures which block the transmission path. Fences are often used along

BOX 5.2 TRAFFIC NOISE IN THE URBAN ENVIRONMENT

Traffic noise is a severe problem in many industrialised countries. In urban areas, the noise impact can increase due to changes in traffic density, the surface of the roads, reflection on buildings and other obstacles.

In the Beaulieu region of Brussels, the global noise environment in a residential area was drastically changed by the erection of large office buildings. The noise, produced by the traffic on an adjacent elevated highway and the subway, was reflected by the concrete and glass areas of the buildings. This resulted in an important increase of the sound pressure level in the residential area (see figure below).



Geometrical model of the Beaulieu region

Noise management was based on local acoustic measurements and advanced numerical simulation techniques, based on ray tracing (Pleeck et al., 1993).

Various scenarios were analysed in order to find the optimal cost-benefit solution. In the different scenarios the use of barriers, a concrete coverage of the subway and the highway, alternative entrances for the highway, other road surfaces, etc. were introduced and compared.

Some proposed solutions, while very efficient in noise reduction, were too expensive (concrete coverage of highway and subway with an associated cost of nearly US$ 30 million).

The final choice was a modification of the highway entrances in order to shift the car accelerations to excavated roads and to change the road surface by using porous asphalt. This resulted in a decrease of the sound pressure level by almost 4 dBA. This figure was verified afterwards by sound pressure level measurements.

highways and railways. Barriers are usually only partially effective and can be very expensive. In addition, it may not be feasible to use them in enclosed spaces, if the machinery needs large amounts of air for cooling or if a combustion process is involved. Absorption of sound through the use of acoustical absorbing materials is one of the most effective ways of attenuating sound in the transmission path. These materials can be in the form of carpets or tiles. However, it is not effective for direct sound.

Table 5.3 Average values of the attenuation of personal hearing protection (the standard deviation (s) is indicated between brackets)

Type of protection	*125 Hz*	*250 Hz*	*500 Hz*	*1,000 Hz*	*2,000 Hz*	*4,000 Hz*	*8,000 Hz*
Dry cotton wool plugs	2	3	4	8	12	12	9
	(2)	(2)	(3)	(3)	(6)	(4)	(5)
Waxed cotton wool plugs	6	10	12	16	27	32	26
	(7)	(9)	(9)	(8)	(11)	(9)	(9)
Glass wool plugs	7	11	13	17	29	35	31
	(4)	(5)	(4)	(7)	(6)	(7)	(8)
Personalised ear-mould plugs	15	15	16	17	30	41	28
	(7)	(8)	(5)	(5)	(5)	(5)	(7)
V-51 R type plugs	21	21	22	37	32	32	33
	(7)	(9)	(9)	(7)	(5)	(8)	(9)
Foam-seal muffs	8	14	24	34	36	43	31
	(6)	(5)	(6)	(8)	(7)	(8)	(8)
Fluid-seal muffs	13	20	33	35	38	47	41
	(6)	(6)	(6)	(6)	(7)	(8)	(8)
Muffs and safety helmet	14	17	29	32	48	59	54
	(4)	(5)	(4)	(5)	(7)	(9)	(9)

Source: May, 1978

The final option in noise control is to protect the receiver through the use of personal hearing protection and by limiting exposure time. Personal hearing protection includes full and semi-insert earplugs, earmuffs or noise helmets; Table 5.3 shows the effectiveness of the different types.

Full-insert earplugs have the advantage of not requiring any external support, unlike semi-insert earplugs, and thus do not interfere with head covers, masks or goggles. Earmuffs provide better protection, particularly at the higher frequencies, but their effectiveness depends on the materials used. 'Sound proof helmets increase protection at the higher end of the frequency spectrum, but they are not commonly used for hearing protection alone and combine this function with protection against head injury or low temperatures.

It is important that individual hearing protection be comfortable, well-fitted and properly maintained, otherwise the degree of protection may be decreased. Regular checks should be made to ensure that employees are wearing their protection devices.

ACOUSTIC LEGISLATION IN THE EUROPEAN COMMUNITY

Various industrial countries follow the recommendations of the International Organization for Standardization (1975) for the noise protection of people in the workplace. The ISO standard indicates that a noise dose of 100 per cent corresponds with:

8 hours a day, five days a week of less than 90 dBA
4 hours a day, five days a week of less than 93 dBA
2 hours a day, five days a week of less than 96 dBA
1 hour a day, five days a week of less than 99 dBA
0.5 hours a day, five days a week of less than 102 dBA

ISO 1999 also specifies that hearing damage starts to occur from equivalent sound levels of 75 dB(A) for eight hours a day.

According to the EEC directive 86/188 CEE (12 May 1986) the following measures must be taken:

1 If the noise level is higher than 85 dBA but lower than 90 dBA:
   (a) adequate information must be given to personnel concerning:
      - the measures (on the source, environment and individuals) that must be taken to con form to noise control regulations law;

• the danger of noisy environments;
(b) individual hearing protection must be available to all individuals;
(c) hearing tests under medical supervision must be performed every five years.

2 If the noise level is higher than 90 dBA:
(a) information must be given to personnel as to why noise levels are so high;
(b) a task redistribution and hearing protection for personnel is obligatory;
(c) a hearing test under medical supervision must be performed periodically.

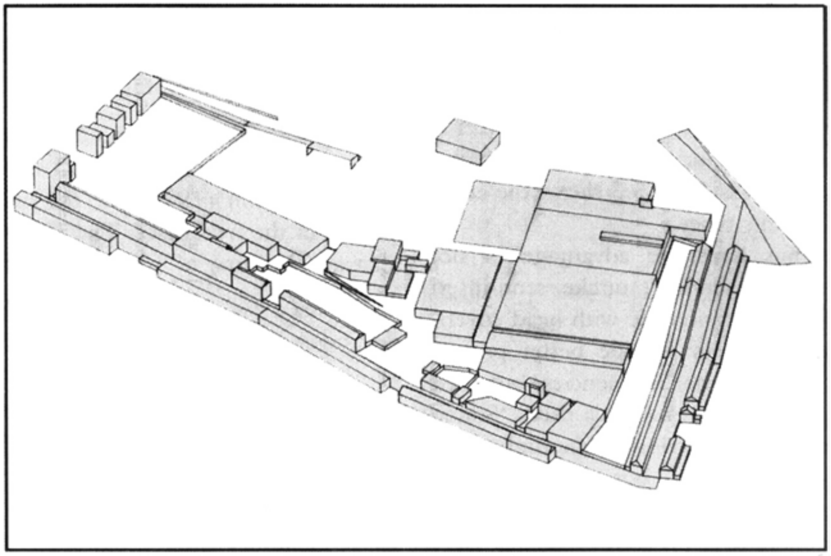
Furthermore:
(a) the maximum allowable peak level is 140 dB;
(b) new machines producing more than 85 dBA must have an acoustic information label;
(c) the directive is not concerned with questions

BOX 5.3 INDUSTRIAL NOISE

Industrial noise is, after traffic noise, the second most important noise problem in the industrialised world. Although recommendations with respect to community noise exist in almost all European countries, the application of and conformity to the above mentioned recommendations will take much time.

Noise management is based on in-situ measurements in enough measurement points and prediction of the sound pressure level as a function of proposed modifications of location, sound power and directivity of the source, the use of screens, sound absorption and isolation.

In this example, we present the study of the noise impact of a large industrial plant, located in the middle of a residential area (see figure below).



Geometrical model of production plant and adjacent urban area

In order to assess the noise impact on the environment, noise measurements of the relevant acoustical parameters were performed. Sound power level measurements of the various sources were also performed to estimate the contribution of each source to the existing sound field.

On the basis of the imposed sound pressure levels (by local regulations: VLAREM 2) and the known sound power levels of the various sources, it was possible to identify the sound sources that were causing problems. Changing their directivity, decreasing their sound power level by replacing them by less noisy sources and improving their sound isolation made it possible to comply with the existing regulations.

of comfort, only hearing loss and secu rity are considered;

(d) the directive specifies the equivalent noise level L_{eq} (dBA) and noise dose measure ments.

Community noise is an increasing problem in today's environment. ISO/R 1996 (ISO, 1971) describes recommendations with respect to community noise. These recommendations are suitable for predicting public reaction to industrial and traffic noise. ISO/R 1996 includes methods for measuring and rating noise in residential, industrial and traffic areas with respect to their interference with rest, working efficiency, social activities and tranquillity. On the basis of ISO/R 1996, national authorities have established adapted noise regulations with respect to communities according to the living habits of the people. Basic criteria take into account time of the day (day, night, evening), type of the district (e.g. rural, suburban, city, industrial), characteristic features of the noise such as impulse, continuous, intermittent or the presence of pure tones.

CONCLUSIONS

Noise management is currently focused on the effects of noise on human health in particular. This is due, in part, to the fact that the measurement of sound pressure levels and their relation to hearing is quite objective. The amount of disturbance caused by noise is much more subjective, and depends on the type of noise and an individual's response to it. In addition, perceived loudness is related not only to the sound pressure level but also to the frequency of the sound, thus making it necessary to develop other units, such as the phon or noise rating to describe loudness.

Noise management should not be restricted to determining whether the sound pressure level is lower than the applicable threshold value. An integrated approach should be adopted, so that the impact of noise is considered in the design of a factory or highway and in site development. The emission, transmission and immission of noise should be kept in mind at all times by designers, operators and public authorities. In industrial occupational situations, once operations have begun, a safety programme should

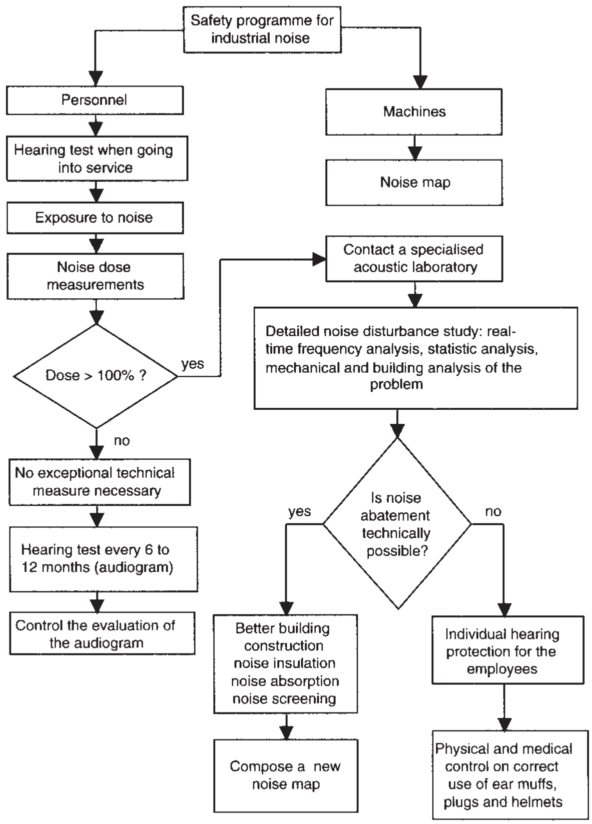


Figure 5.9 Safety programme for industrial noise abatement

be followed (Figure 5.9) in order to ensure the rapid detection of any hearing problems and their causes.

Another aspect of noise is its impact on natural ecosystems. For example, does the noise from highways impose stress on wildlife living in adjacent forests? There have been many complaints from the indigenous people of Ecuador that the noise of helicopters and underground explosions used in oil exploration are having a negative impact on forest animals. Such effects of noise require further investigation, but the solution will always require the reduction of noise at its source, a task for policymakers and equipment designers.

REFERENCES

Berry, B.J.L. (1974) *Land Use, Urban Form and Environmental Quality*, Research Paper No. 155, Chicago, USA: University of Chicago, Department of Geography.

Bruël and Kjaer (1979a) *Sound Level*, Lecture notes no. 311E, Naerum, Denmark: Bruël and Kjaer.

Bruël and Kjaer (1979b) *Frequency of Sound*, Lecture notes no. 321E, Naerum, Denmark: Bruël and Kjaer.

ISO/R 1996 (1971) *Assessment of Noise with Respect to Community Response*, Geneva, Switzerland: ISO.

Lawrence, A. (1989) *Acoustics and the Built Environment*, New York, USA: Elsevier Applied Science.

Lukas, J.S. (1978) 'Noise and Sleep: A Literature Review and a Proposed Criterion for Assessing Effect', in D.N.May (ed.) *Handbook of Noise Assessment*, New York, USA: Van Nostrand Reinhold, pp. 313–334.

Martin, A.M. and Walker, J.G. (1978) 'Occupational Deafness and Hearing Conservation', in D.N.May (ed.) *Handbook of Noise Assessment*, New York, USA: Van Nostrand Reinhold, pp. 251–284.

May, D.N. (ed.) (1978) *Handbook of Noise Assessment*, New York, USA: Van Nostrand Reinhold.

Milne, A. (1979) *Noise Pollution*, North Pomfred, Vermont, USA: David and Charles.

Pleeck, D., Van Overmeire, M. and De Borger, G. (1993) 'Highway Noise in a Critical Built-up Area: Noise Measurement, Improvement Simulations Using an In-situ Tuned Prediction Model', in *Proceedings EURO-NOISE 1993*, Leuven, Belgium, pp. 621–624, 1555–1558.

Smith, M.J.T. (1989) *Aircraft Noise*, Cambridge, UK: Cambridge University Press. Sound Research Laboratories (1991) *Noise Control in Industry*, London, UK: E. & F.N.Spon.

SUGGESTED READING

Berger, E.H., Ward, W.D., Morill, J.C. and Roysta, L.M. (ed.) (1986) *Noise Hearing Conservation Manual*, Akron, OH, USA: American Industrial Hygiene Association.

Franssen, E.A.M., Staatsen, B.A.M., Vrijkotte, T.G.M., Lebret, E. and Passchier-Vermeer, W. (1995) *Noise and Public Health*, Workshop Report No. 441520004, Bilthoven, The Netherlands: National Institute of Public Health and Environmental Protection.

Hassall, J.R. and Zaveri, K. (1988) *Acoustic Noise Measurements*, Glostrup, Denmark: Bruël and Kjaer.

Lawrence, A. (1989) *Acoustics and the Built Environment*, London, UK: Elsevier.

May, D.N. (ed.) (1978) *Handbook of Noise Assessment*, New York, USA: Van Nostrand Reinhold.

Schultz, T. (1982) *Community Noise Rating*, 2nd edition, London, UK: Elsevier.

SELF-ASSESSMENT QUESTIONS

1. What is the unit most commonly used to describe sound level? Why is this unit not useful for describing the loudness of sound as perceived by humans? What units are used instead?
2. What factors need to be considered in the analysis of noise pollution?
3. What is the safety limit of noise exposure for a 40-hour working week? To what dose is this limit equivalent? What is the noise dose a person would receive if exposed to the following noise levels during one week: 90 dB(A) for 16 hours, 93 dB(A) for 4 hours, 96 dB(A) for 10 hours and 99 dB(A) for 10 hours? Is it an acceptable dose?
4. Imagine that the airport authorities of our city want to expand the airport. What might be some of the impacts of the additional noise? What factors should the airport authorities consider with respect to noise, both for its workers and the surrounding community?

PART II

THE SECTORAL APPROACH

INTRODUCTION

Dimitri Devuys

Environmental management will be most effective if it takes the form of a multi-instrument and a multi-target approach. This means that the many instruments for environmental management that were examined in Volume 1, such as environmental planning, environmental reporting, auditing, impact assessment and life-cycle analysis to name only a few, should be used to deal with environmental problems in the different environmental compartments and in the distinct sectors within human societies on this planet. Moreover, several instruments for environmental management need to be adopted in different ways and combinations, depending on the ecosystem in which the environmental problems occur and depending on the specific set of environmental stressors that need to be taken into account.

Part II of Volume 2 examines the following sectors of human societies: the energy sector, the mining sector, the agricultural sector, the fisheries sector, the forestry sector, the industrial sector, the transportation sector and the tourism sector. Looking at possible environmental management practices in each of these different sectors leads us to the 'sectoral approach in environmental management'.

What is so special about the sectors considered in Part II is that while they are all linked to human activities they are also in one way or another linked to natural resources. Sectors such as mining, fisheries and forestry are directly and very closely linked to natural resources. One would expect that people living from the richness brought to us by forests, seas or mines would be very much aware of the need to protect these precious resources, but practice teaches us that this is not true. In certain parts of the world trees are felled and seas are plundered at an unprecedented and unsustainable rate. The health of other sectors, such as industry, transportation or tourism, are also linked to the maintenance and protection of natural resources, although the links might be less obvious. Tourism, for instance, may seem to be a purely human activity. Examining the tourism sector in more detail shows the many links to the environment. An island that is opened up by the traditional tourism industry will undergo tremendous change. Where will hotels be constructed, do we need a bigger airport, what about the availability of drinking water, what happens with the waste produced by the tourists, will certain wildlife species be threatened by daily visits to the forest?

As you will notice from the above questions, all the sectors stand in relation to each other and, therefore, the different chapters in Part II are tightly linked up with each other. To take the example of the island again; bringing in tourists will, for example, touch the transportation sector (visitors will want to move around), the energy sector (western tourists will want to use all kinds of electrical appliances) and the agricultural sector (fields might be converted to hotels and golf courses).

As you read through the different chapters of Part II, you will notice that many authors try to operationalise principles of sustainable development for their specific sectors. This is because environmental management and sustainable development are very closely linked in the sectoral approach.

The relationship between environmental management in the sectoral approach, measures for sustainable development and other approaches to environmental management can be illustrated with an example. Let us consider China. Its population, the largest in the world at 1.1 billion, represents some 21 per cent of the total human species. Thanks to the efficient but much criticised 'one-child family programme', coupled with a number of health and educational programmes, China's population growth rate has been cut in half even as mortality rates continue to fall. The sectoral approach on its own would never have been able to accomplish this very important task of slowing down population growth. The sectoral approach can, however, deal with the increasing energy demand. China's high per capita energy use partly reflects a widespread use of coal at the household level and inefficient combustion. If energy use in China, with its large reserves of coal, ever reaches a level comparable to that in developing nations, the atmospheric consequences could be catastrophic. Environmental management in the energy sector would lead to the search for new technologies and the promotion of the use of renewable energy sources. Also, measures for more energy efficiency can reconcile environmental concerns with economic development for all nations. They can stretch energy supplies, slow climatic changes and buy time to develop alternative energy sources. Sustainable development principles take us one step further. One of the most serious dilemmas is to find ways for poor countries to develop without making the Earth uninhabitable. What would happen if all Chinese citizens started buying cars, televisions, refrigerators, dishwashers, air conditioners and numerous other small appliances like Europeans or North Americans do? Very important ethical questions need to be raised here and it becomes clear that everybody on this planet needs to strive for a more sustainable development in which environmental, social and economic management considerations are brought together.

The management of the transport sector requires trade-offs to be made between three major goals of transport management—growth, equity and environmental sustainability. International trends in emissions of the major noxious air pollutants and several policy instruments might contribute to reducing the impact that transport services have on the quality of the environment. Here it becomes clear again that the sectoral approach does not stand on its own. There is an obvious link with the management and planning of human settlements, as described in the chapter on urban environments in Volume 3 of this text (Chapter 11). One sensible and compelling alternative to technological advances in transportation would be to reduce the need to travel in the first place. Appropriately designed communities could try to put most destinations within walking or bicycle-ride distance. The humanist design principles would allow people of all ages and walks of life to come in daily contact, promoting social cohesion and building stronger bonds between people. By being less dependent on vehicles, life in these communities would also become more energy efficient. As is shown in Chapter 13, urban planning techniques are only one facet of an efficient environmental management of the movement of people and goods.

In the industrial sector manufacturers of the future must consider how to design and produce products in such a way as to make the control of waste and pollution part of their enterprise. The instrument for environmental management known as 'life-cycle analysis' will become important since the industrial sector will need to pay attention to the entire product life cycle, worrying not only about the materials used and created in the course of manufacturing, but also about what happens to a product at the end of its life. Also, the natural world can teach the industrial world quite a bit, and the practices promoted in 'industrial ecology' hold many promises. Similar to a natural ecological system, which is an integrated whole and minimises waste, a fully developed industrial ecology might not necessarily minimise the waste from any specific factory or industrial sector, but should act to minimise the waste produced overall. Waste products of one process might be used beneficially by others.

In the agricultural sector research into 'integrated pest management' is encouraging. Integrated pest management refers to the combination of using hardy plants, crop rotation, tillage practices, biological controls and a minimal amount of pesticides. It draws on the fundamental knowledge of plant and insect biology amassed by botanists and entomologists. For example, several insect attractants (pheromones) have been identified and synthesised. These substances can be used to interfere with the normal reproductive cycle of common pests. Moreover, the success of sustainable agriculture depends on making plants more efficient in converting sunlight, nutrients and water into food and fibre products. Not all people agree that genetic engineering is the right way to try to achieve these goals. Environmental groups raise important ethical and health related questions over, for example, cloning of animals or genetically modified crops that resist pests. Also important is biodiversity and the development of novel diagnostic tools that carry the power of modern science to identify diseases and help farmers manage their crops. For instance, decision support systems are now under development to aid farmers with soil management. Agricultural research will probably yield many new technologies for expanding food production while preserving land, water and genetic diversity. The real trick will be persuading farmers to use them.

The sectoral approach to environmental management can show us one possible way of dealing with today's environmental problems. This approach, in combination with other approaches discussed in this Environmental Management Series, will bring us closer to solving the global environmental, social and economical problems that we are facing.

6

ENERGY, ECONOMY AND THE CO_2 PROBLEM

Hans W.Gottinger and Philip M.L.Barnes

SUMMARY

This chapter explores three essential structural factors of the global energy-economic system—uncertainty, technological shifts and international action—that affect the climate system via CO_2 concentrations in the atmosphere. Each factor has been subjected to in-depth modelling analysis.

The starting point of the analysis is the perception that the CO_2 problem is essentially one of high prospective energy demand, and that there are four major routes towards reducing this growth in global energy use: (1) reduce the population growth rate; (2) reduce per capita energy growth; (3) reduce world economic growth; (4) decouple energy from economic conversion via more efficient energy supply technologies and/or structural shifts in the economy.

Constrained trajectories of CO_2 increases are modelled using a simple carbon cycle model. The change in the curvature of constrained fossil fuel use has important consequences for the growth of new energy technologies, and also induces change in the technological energy infrastructure.

Non-fossil energy use is a composite of various sources (e.g. solar, hydro, nuclear, etc.) with varied potential, sustained production and diffusion rates.

In the international context of the CO_2 problem, we examine how rapidly various groups of nations could or should reduce their use of fossil fuel so that their cumulative CO_2 emissions will not exceed ceiling constraints. This proposal is contrasted with other, more prevalent policy options such as adaptation, mitigation and compensation.

ACADEMIC OBJECTIVES

The aim of this chapter is to help the reader understand the essential structural features of the global energy economy that affect the climatic system and the potential management of energy resources as they relate to the CO_2 problem.

On completion of this chapter the reader should have:

- a broad understanding of the scope of the CO_2 problem and associated constraints in its global context and the range of uncertainties involved;
- the ability to assess, with the help of simple hypothetical models, the speed with which various groups of nations could or should reduce their use of fossil fuels in order to restrict cumulative CO_2 emissions;
- an overview of world fossil fuel use and the potential rates of penetration of non-fossil and renewable energy forms;
- a measure of the steps essential for coping effectively with the CO_2 problem;
- an understanding of the regional aspects of the CO_2 problem and the economic costs, equity issues and interna tional strategies involved.

THE SCOPE OF THE PROBLEM

The focus of this chapter is on addressing and analysing certain specific issues at the interface of energy use and environmental management from a global perspective.

We start from the fact that carbon dioxide (CO_2) in the atmosphere affects the radiation balance of the earth, and that increasing CO_2 concentrations are expected to cause a warmer climate (climate change). Carbon dioxide is relatively transparent to energy as sunlight, but reflects or traps a larger portion of this energy when radiated from the Earth as heat. An important development in climatic studies is the identification of other gases with this same heat trapping property; these include chlorofluoromethanes (major sources: spray cans and other commercial uses), nitrous oxide (major sources: jets, cars and unknown sources) and methane (major sources: anaerobic fermentation in rice fields and natural gas leakages). These gases are often referred to as trace or greenhouse gases (GHG). Uncertainty about future increases in such gases significantly complicates the prediction of future temperatures. The prediction of future temperatures is further complicated by diverse factors. Significant feedback effects are expected to accompany any direct effects of CO_2 heat trapping.

Warming may change cloud, snow and ice cover and alter the Earth's albedo or brightness. Consequently, the reflection and absorption of energy may increase or decrease. Warming may also cause the release of additional CO_2 that is trapped in frozen soils, accentuating the GHG problem. In addition, major uncertain climatic disturbances are foreseen, which are not connected with GHG. Continental drift, fluctuations in the Earth's orbit, solar flux variations and volcanic dust are all natural causes of climatic change. The release of particulates and changes in land use are human activities that may impact local or regional climate.

We assess the suitability of energy resources, alone or in combination, for satisfying these global environmental constraints, and identify reasonable scenarios for environmentally benign energy futures.

This century has witnessed an unprecedented period of growth in energy, economy, population and consumption. By the end of the 1960s, in particular, anxiety was mounting as to whether the world was beginning to reach the ceiling of resource and environmental constraints, or whether it still had time to complete the gradual transition towards a 'steady state', in which population and resources would be in balance. The 'limits to growth' debate (Meadows, 1972) culminated in several studies associated with the Club of Rome. These predicted a catastrophic increase in mortality rates throughout the world beginning in the first decades of the twenty-first century, largely as a result of resource shortages. But even if resources were not limited, global pollution would, a few years or decades later, lead to an even greater catastrophe. The basic thesis was that the longer this catastrophe was postponed by 'technological fixes', the more destructive the final collapse would be.

These studies argued that consumption growth would so pollute the environment as to threaten human health and the life-supporting properties of the biosphere on which human existence depends.

There are three separate issues involved here:

1 General chemical or radiation contamination of the environment (hazardous and nuclear wastes).
2 General deterioration of natural ecosystems (air and water pollution, soil erosion).
3 Changes in the global climate (CO_2 emission and trace gases, ozone depletion).

Potential changes in global climate now pose the most challenging task for the control of global environmental pollution, and we focus particularly on this as a problem of energy production.

Most studies so far have singled out carbon dioxide (CO_2) as the most serious pollutant, with potentially irreversible consequences for global climate. As yet there are no economically feasible control technologies on the horizon for capturing any significant fraction of CO_2 emissions. A prime characteristic of the CO_2 problem is the long time-lag that may elapse between the cause and the identification of significant effects. Another characteristic is the high degree of uncertainty that usually attends predictions of future effects. Such uncertainty is compounded by the fact that CO_2 effects occur at various complex levels.

The first level of effect is the direct physical result of the activity, in this case the actual CO_2 concentration in the atmosphere. The principal evidence for a trend towards increasing CO_2 levels in the atmosphere comes from continuous observations over many (up to thirty) years at various sites.

The second level addresses the partitioning of CO_2 between the atmosphere and other carbon reservoirs. If we look at the carbon cycle, with the exception of the atmosphere, the amount of carbon in each reservoir is somewhat uncertain. It seems that there is a natural circulation of carbon among these different reservoirs, particularly through photosynthesis and oxidation.

This opens a biogeochemical perspective on the carbon cycle. We can take the view that various chemical reservoirs of the earth are comparable to the organs in a human body (Ausubel, 1980). Accordingly, the CO_2 problem emanates from human activities, feeding carbon into circulation faster than the ability of the Earth's organs to digest or metabolise it. This is due largely to the increasing consumption of fossil fuels and less so to the clearing of natural forests. Carbon dioxide emissions from fossil fuels combustion have been growing at the rate of 4.3 per cent per annum since 1860, with the exception of World Wars I and II and the depression years of the 1930s (Rotty, 1977; Flohn, 1989).

In 1860 the release rate was about 0.9 Gt of carbon, in 1900 it was 1.05 Gt, in 1940 it was 1.3 Gt, in 1970 4.2 Gt, and in 1981 it was 5.5 Gt/yr of carbon, with further increases since then. The effect of fuel substitution, largely oil and natural gas for coal, can be seen in the contrast between the rates indicated for 1970 and 1900.

Between 48 and 56 per cent of the CO_2 produced by the combustion of fossil fuels over the past two decades can account for the observed CO_2 increase in the atmosphere. Between 33 and 41 per cent is thought to be taken by the oceans, which suggests that less than 15 per cent must have gone into some other sink. The role of the biosphere is currently uncertain and under research.

Educated guesses tend to assign a lower contribution to forest clearing, perhaps accounting for less than 20 per cent of the fossil fuel emissions, and the potential contribution from this source is much more limited than that from fossil fuel. But that only complicates the search for the 'missing' sink.

The next level of uncertainty is the direct climatic effect of increasing atmospheric CO_2 arising from the increase in surface temperatures—the so-called 'greenhouse effect'. Calculations undertaken with the 'best' available—but still significantly superficial-time-dependent global circulation models (GCM) indicate that a doubling of CO_2 concentrations (say from 300 ppm to 600 ppm) is likely to increase global temperature (?T_S) in the range of 1.5 to 4.5°C. Most models also show how the global temperature increase will be unevenly distributed with probable amplification of up to a factor of 5 in the polar regions.

A major part of the variance in the several estimates of climate models of ?T_S can be explained from the 'built-in' assumptions of the relative importance of various interactive physical processes on climatic feedback mechanisms. Figure 6.1 shows how these interactions occur in a very simplified network, used only for illustrative purposes. These processes could either amplify or dampen by several times the calculations of ?T_S.

Furthermore, other 'greenhouse' trace gases, such as chlorofluoromethanes, methane, carbon monoxide, nitrous oxide and ozone are being added to the atmosphere by human activities; estimates put this additional warming close to that of CO_2 at about 50 per cent.

One fundamental deficiency of all models is the inability to couple the ocean and the atmosphere in

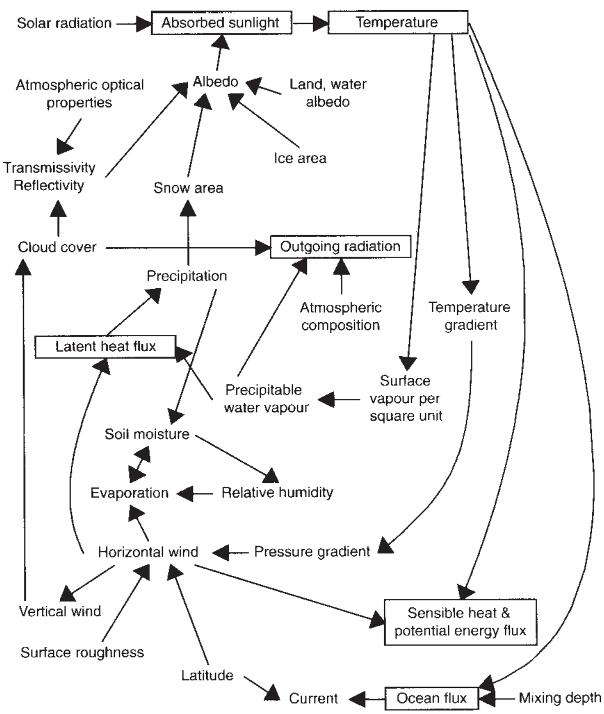


Figure 6.1 The interactions, linkages and feedback loops in the climate system

Source: Kellogg and Schwarz, 1981

ways that describe real energy exchange and division of energy transport between the two. The models hold the ocean's heat transport constant within atmospheric circulation and changing wind stress, which is not realistic. Indeed, the oceans exert tremendous influence on nearly all atmospheric processes and modelling of the fuel interaction is still in the development stage (IPCC, 1990). Despite the uncertainties and probable systematic flaws in any climate model, we should pay attention to a panel conclusion by the US National Academy of Sciences, made public more than fifteen years ago (NAS, 1979) but still valid today.

> We have tried but are unable to find any overlooked or underestimated physical effects that could reduce the currently estimated global warming due to doubling of atmospheric CO_2 to negligible proportions or reverse them altogether.

Can the CO_2 'signal' be filtered out from the 'noise' during the past century? Assuming ?T_S=4.04 In (M/M_o) where M is the carbon dioxide concentration, it can be calculated that the expected ?T_s over the past century is of the order of 0.55°C which lies within the range of natural inter-annual variation in T_s.

The next level of uncertainty concerns the timing and geographical distribution of changes due to the increased global surface temperatures. The last and most important level of uncertainty is the translation of the effects of climatic change and increased CO_2 levels—even if they were certain—into impacts on the environment and human activities in specific regions. This, in effect, closes the chain of causation in the man-climate interaction.

It is possible, however, to identify several major impacts with some confidence. Agriculture is the most prominent. Sea level rise is another major concern. There could be significant changes in the production potential of many areas as a result of temperature, precipitation and variability changes. Shifts in climatic zones and in loci of agricultural activity may create deoptimisation for existing crops and farming activities. Other climate-sensitive areas comprising energy demand, water resources, fisheries, human health, population settlements, tourism and recreation appear to be of lesser significance than food production, but no serious estimates have been made of their magnitude. Second order effects, such as the absence of frost or cold winters, could imply proliferation of pests and pathogens, adding to possible damage. One difficulty in even a crude cost/benefit assessment of these impacts stems directly from speculation about the ability of mankind to anticipate the problems and adapt to them (Schelling, 1988).

Reduction of uncertainty and better, more reliable data on physical impacts and change will be of immense help in using models to calculate the optimal rate of CO_2 emissions. The objective in

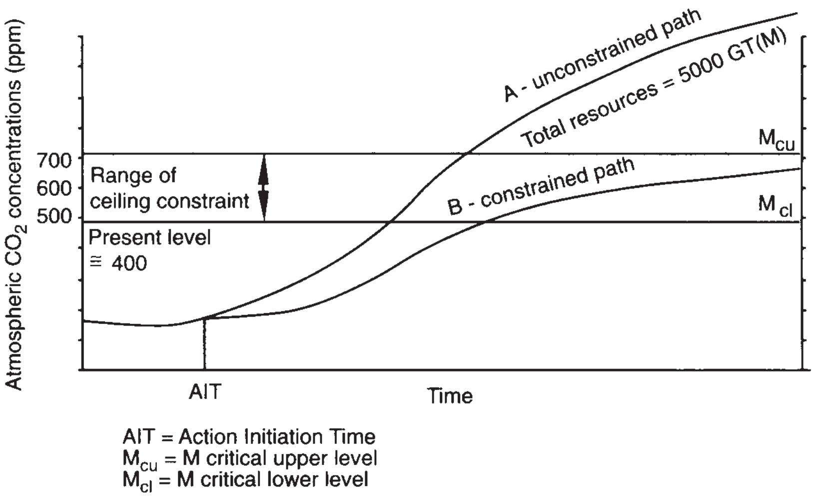


Figure 6.2 Schematic illustration of two CO_2 paths: 'A' and 'B'

energy and resource economics is to determine those time paths of resource use that maximise the discounted present value of the future stream of economic benefits.

In addition to the above constellation of uncertainties in physical climate modelling, carbon cycle models, ecological impacts and economic costs or benefits of the attendant change, there are uncertainties in the social and political domain as well as in mankind's future demand for fossil fuels. But these latter uncertainties about political and social circumstances and fossil fuel demand are of a necessarily different character: they are the outcome of human decisions and hence are more amenable to control. By thinking of CO_2 as a problem in the management of a scarce natural resource—the quality of the global atmosphere—during the transition from a fossil fuel to a non-fossil fuel economy, one is forced to focus on efforts for developing appropriate transition strategies for the management of fossil fuel emissions. The importance of management *per se* is illustrated by the fact that extractable fossil resources are sufficient to increase CO_2 concentrations by a factor of 5 to 8 over the pre-industrial level. If that happens, ?T_S will increase by 6 to 9°C within the next two to three centuries, which would undoubtedly be disastrous. Thus, the world's cumulative use of fossil fuels must be restricted to levels below estimated recoverable resources.

We shall briefly describe the constraining factors that are the 'givens' of our economic modelling. In Figure 6.2, curve A represents the present world energy use, unrestrained by CO_2 considerations.

(1) $M_{cl} \leq M \leq M_{cu}$

(2) $\frac{1M}{M\delta t} \leq km,$ where m is the rate to be determined and k being some fixed parameter.

Curve B is a different world, where CO_2 concentrations satisfy two constraints:

The first constraint to be determined on M, represents a ceiling level that might be approached with 'acceptable' cumulative consequences, but which probably should not be exceeded. In addition, this level should be below any 'threshold' level that heralds catastrophic events.

The second constraint is a rate limitation that depends on the rate of climate change, reflecting in a general fashion the assumption that slow rates of change are more amenable to compensating adjustments in human institutions and in productive activities that depend on climate.

History shows that societies often have failed to recognise the importance of planning for climate changes and have responded differently to climatic perturbations and experienced quite disparate impacts (Flohn, 1985).

Assuming some plausible constraints on the climate system via CO_2 concentrations in the atmosphere we turn to the energy-economic system. Since the CO_2 problem is essentially a problem of high prospective energy demand, it is natural to concentrate our efforts there. There are four routes to reducing this growth in global energy use.

1 Reduction of population growth rate.
2 Reduction of per capita energy growth, with re distribution or without redistribution.
3 Reduction of world economic growth.
4 Decoupling of energy from economic growth through conservation, advances in energy con version, more efficient technologies and/or struc tural shifts in the economy.

The items 1 through to 4 do not constitute free variables that we can adjust, but have a minimum or maximum.

Reasonable variations of those rates have to be discussed in view of the latest global energy demand projections. We proceed as follows. We model mathematically the constrained trajectory of CO_2 increase versus time, and using a simple carbon cycle we find global fossil fuel use. The sensitivity of the fossil fuel scenarios to variations in input parameters is analysed.

The change in the curvature of $F_{fo}{}^{B}$ has important consequences in the growth of new energy technologies (the gap between $F_{nfo}{}^{B}$ and $F_{nfo}{}^{A}$); see Figure 6.3. A slow change in F_{fo} implies a fast change in this gap. This implied rate of change in the technological energy infrastructure is analysed and compared with the present and future industrial capabilities. Some of the discussed 'rates of change' are:

- the market penetration rate;
- start-up rates of new energy technologies;
- the manufacturing capacity expansion rate.

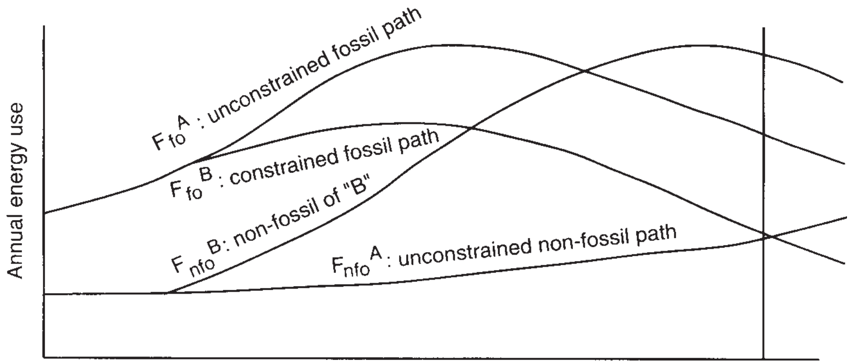


Figure 6.3 Global fossil fuel use scenarios corresponding to CO_2 paths 'A' and 'B'

F_{nfo} in Figure 6.3 is a composite of various sources (e.g. solar, hydro, biomass, nuclear, etc.) with varied potential, sustained production and diffusion rates.

Subsequently, we view the CO_2 problem in its proper international setting.

Recognising that F_{fo}^{B} is the sum of fossil fuel use of various nations while M_{cu} and M_{cl} are globally fixed, and recognising that global climate change will lead to various regional climatic impacts, we will draw a composite hypothetical regional climate scenario based on those (sometimes contradictory) scenarios reported in the literature.

How rapidly could or should various groups of nations reduce their use of fossil fuel so that their CO_2 cumulative emissions shall not exceed those ceiling constraints?

This question and other similar questions posed in the global analysis are approached within the context of a hypothetical multi-regional scheme (pp. 134–135). In the final section, the role of uncertainties in policy models of climate change is examined and implications for policy changes are drawn.

WORLD FOSSIL FUEL USE

Given the CO_2 constraints described in the previous section we will analyse plausible world fossil fuel paths with the aid of a simple carbon cycle model. The model will be used to probe the sensitivity of the results to variations in key parameters that would influence the energy system dynamics. From the economic point of view of fossil fuel use the argument is often advanced that market forces should (and will) play a dominant role in allocating resources, and thus no other mechanism should intervene in this 'optimum' process. According to this view, the 'cost-effectiveness' process is automatically operative, and so there is no reason for constraining exogenously the growth of fossil fuel use.

To put this into the right perspective, note that the resource base for fossil fuels is indeed very large if it includes resources that will cost much more than they do now. About half of the conventional oil and natural gas is recovered inexpensively. The rest comes from a variety of sources associated with higher production costs: poorer fields that require drilling more holes, production from continental shelves, deeper basins and polar regions, etc. All these activities are, moreover, associated with significantly larger environmental impacts. If we assume that cleaning up spills, reclaiming mined-out lands to useful purposes and hazardous waste management are all operations whose cost must be internalised, then these 'dirty' fuels become even more expensive.

In the case of fossil fuels, the standard observation of resource economists—that one never runs out of a resource but simply reaches a point where it is cheaper to use a substitute—is to some extent correct. Unfortunately, very few reliable estimates of fossil fuel resources versus cost curves are available.

It is estimated that remaining and recoverable resources of fossil fuels (Table 6.1) contain nearly 3,900 Gt (of carbon), enough to increase the concentration of CO_2 in the Earth's atmosphere by a factor

Table 6.1 Energy and carbon content of recoverable fossil fuels[1,2]

	Energy content (10^{21} J)	Carbon[3] content (10^{15} J)
Oil		
Reserves 640 × 10^9 bbl	3.9	74
Resources 2,000 × 10^9 bbl	12.2	232
Gas		
Reserves 2,500 Tcf	2.7	38
Resources 9,000 Tcf	9.7	138
Coal		
Reserves 636 Gtce	18.8	440
Resources 5,000 Gtce	147.0	3500
Total	169.0	3870

Notes:
1 Includes only conventional deposits of oil, gas and coal. Excludes heavy crudes, oil shales, tar sands, methane, hydrates, etc.
2 Resources are cumulative, and include reserves. These are estimated recoverable resources, not resources in place. Units 1 bbl=42 gal.=0.159 m^3, 1 Tcf=10^{12} ft^3=2.83×10^{13} m^3; 1 Gtce=10^{15} grams of standard coal equivalent at 7 cal/g.
3 Carbon content/energy content is assumed to be approximately as follows: oil 1.9, gas 1.4, coal 2.4 grams carbon/M.
Sources: Háfele, 1981; Rosenberg, 1989

of four. This means that even a CO_2 limit of 1,000 ppm would restrict the release of fossil carbon well below the 3,900 Gt available in recoverable ores, gas and coal, while a limit of 700 ppm might restrict the release to 1,400–2,200 Gt, depending on the airborne fraction. Table 6.2 summarises the recoverable resources by the type and cost category.

Given an economically ultimate resource base of 5,000 TWyr, an obvious question is: what do plausible production paths look like? Figure 6.4 shows four such curves, each within an integrated area of 5,000 TWyr.

Curves 1 and 2 represent high energy consumption, while Curves 3 and 4 represent low consumption. The difference between Curves 1 versus 2, or 3 versus 4, is whether the difficulties associated with nuclear power will be overcome (Curves 2 and 4). If not, coal will play a major role (Curves 1 and 3). An appropriately weighted integral of the curves in Figure 6.4 represents the projected increase in CO_2 from anthropogenic sources. Figure 6.5 displays the outcome. A CO_2 concentration of 500 ppm(v) will be exceeded in the period 2035–2050, and a 700 ppm(v)

Table 6.2 Summary of economically recoverable additional world resources by cost category[1]

Coal		Oil			Natural gas		
1	2	1	2	3	1	2	3
560	3065	264	200	373	267	141	130

Source: Adapted from Háfele, 1981
Note: 1 Cost categories represent estimates of cost either at or below the stated volume of recoverable resources (in constant $ 1980)
For oil and natural gas: Cat. 1: $15/boe, Cat. 2: $15–20 boe, Cat. 3: $20–25 boe,
For coal: Cat. 1: $ 25/tce, Cat. 2: $ 25–50 tce

will be crossed in 2070–2120 whatever the fossil fuel path will be. Path 1 implies a 2.6pc/yr growth in fossil fuel while Path 4 represents an increase of only about 1.1 pc/yr. By comparison, the historical fossil fuel growth rate over decades has been around 4.3 pc/yr. From this simple exercise and the previous observations, we see that CO_2 concentrations will not be restricted to levels below the ceiling limits indicated earlier (p. 121) by physical depletion or economic factors.

Moreover, although lower fossil fuel growth rates give us more time, they would not eliminate the

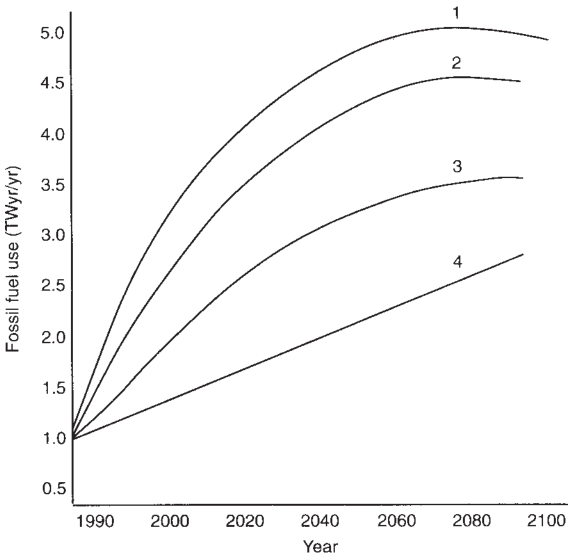


Figure 6.4 Feasible fossil fuel production paths for a total ultimate cumulative consumption of 5,000 TWyr

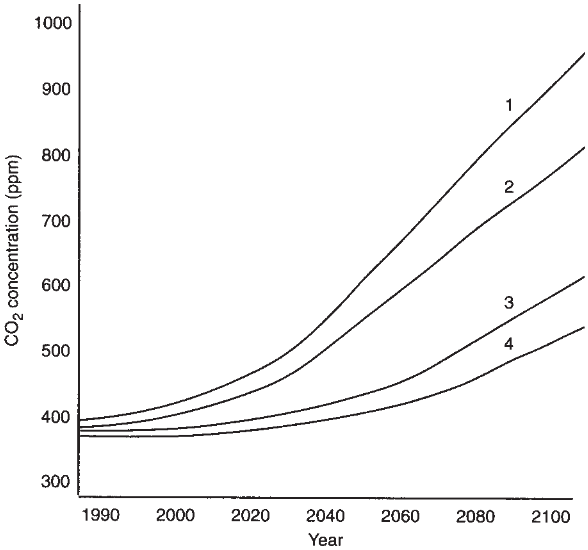


Figure 6.5 Range of projected CO_2 concentration of an ultimate consumption of 5,000 TWyr and production functions (airborne fraction=0.5)

problem. Lower growth rates will satisfy the rate constraint limitation, but not the ceiling. The latter will therefore have to be built into the system response.

A simplified modelling approach

Plans for increased energy use yield various scenarios for increased CO_2 concentrations with time. To capture in simple terms the nature of the problem, we assume that:

1 A minimum realistic level M_{cl} of CO_2 exists, de termined by total integrated demand for fossil fuels cumulated into the future until such time as the world has changed over to energy sources that do not increase CO_2 (principally solar and/or nuclear in various forms). We also assume that if less total fossil fuels are used in the future, there will be considerable social unrest due to unmet demand.
2 A maximum allowable level M_{cu} of CO_2 exists, determined by the limit of tolerable climatic and/ or biological and/or botanical effects.
3 Global fossil energy use can be planned so that the CO_2 Curve F increases smoothly to some value that is acceptable.

The curvature of the CO_2 build-up with time implies that society thinks ahead, and plans to avoid trouble. But do we really have to plan ahead, given the uncertainties, or can we let matters take their normal course, and delay measures until scientific research has finally shown with 'certainty' that we should not exceed a given maximum level of CO_2? In other words, can we behave in such a way that we allow the concentration to increase until it is imperative that a halt is called? In reality it seems that a sudden halt in the use of fossil fuel is practically impossible, as this would imply a complete shift in technological infrastructure away from fossil fuels. Technological breakthroughs notwithstanding, this is an impossible scenario for the world as a whole.

In order to describe the dynamics of carbon accumulation towards the ceiling constraints, we assume that the rate of build-up of CO_2 in the atmosphere is represented by the following equations.

$$M(t)=kF_{fo}(t)+1A(t)-R(t)$$

where dots denote time derivatives

M(t)	is the accumulated mass of carbon in the atmo sphere (in tons) at time t
$F_{fo}(t)$	is the rate of fossil fuel use (TWyr/yr)
k	is the average emission rate of carbon per unit of fuel energy
1A(t)	is the net release rate of carbon from clearing tropi cal forests
R(t)	is the net transfer rate of carbon from the atmo sphere to other reservoirs (sinks).

The objective is to find $F_{fo}(t)$ that conforms to a prescribed asymptotic M_a, but this is hindered by the fact that most terms in the equation are only vaguely known.

Estimates of 1?(t) in the past few years, based on forest inventory statistics with fluxes calculated from a forest clearing rate times estimated carbon stored in particular forests, have shown significant variations.

Currently, the oceans are treated as the sole ultimate sink of carbon, R(t).

Comparing fossil fuel use with the reconstructed and actual atmospheric record suggests an 'airborne fraction' (AF) that according to most calculations with cycle models is between 0.55 and 0.65, depending on the vertical mixing rate within the oceans and the behaviour of the biosphere (Rosenberg *et al.*, 1989).

The type of model we envisage can be sketched as follows. An approximate relationship between the atmospheric mass of carbon M(t) and the fossil fuel burning rate, $F_{fo}(t)$ can be established by:

$$F_{fo}(t) \cong \frac{1}{k \cdot AF} \cdot \dot{M}(t) \qquad (1)$$

and constitutes the starting point for the development of the model.

We shall assume that the carbon added to the atmosphere since ?t_o can be modelled as a logistic function in the form:

$$M(t) = \frac{\beta}{1 + \gamma e^{-\alpha(t - t_0)}} + \lambda \qquad (2)$$

where β, γ, λ, α = constants
t = time
t_o = action initiation time (AIT)

Since $M(t_o) = M_o$ it follows that:

$$\gamma = M_0 - \frac{\beta}{1 + \lambda}$$

A ceiling constraint on CO_2 is imposed:

$$M_{cl} \leq M_c \leq M_{cu}$$

The model developed above attempts to describe a plausible world response to the impending CO_2 problem. The final world trajectory or response function can be thought of as a composite of three stages. The first stage describes the period up to AIT, when the response function follows closely the reference scenarios. Hence, it represents a period during which fossil fuel use is unconstrained by consideration of the CO_2 build-up. However, AIT is not the point when discussions or negotiations to limit CO_2 begin, but rather the time when the resulting policies begin to be effective. The second phase starting at the AIT signals the retardation of fossil fuel growth (because of the CO_2 build-up) and ends with fossil fuel use at its peak. The third phase of the response describes a 'decay' and general 'phase out' of the fossil fuel consumption.

Projections of CO_2 concentrations

Most previous predictions have been based on the assumption of exponential growth in the CO_2 production rate. Table 6.3 shows the CO_2 concentration for exponential growth of 4.3 pc/yr (the historical rate of carbon release), 2 pc/yr (the 1970– 1985 growth rate) and 2.6 pc/yr, the last representing a carbon rich high energy scenario.

A comparison of these three projections suggests the following observations:

1 Doubling of CO_2 (i.e. 600 ppm) would be ex pected by the year 2030 *if* historical growth rates were maintained.
2 Moderating the fuel consumption to 2.6 pc/yr delays the doubling time by 13 years, while fur ther moderating the growth to 2 pc/yr buys only an extra five years. Thus, at most, moderating ex ponential growth will buy us two additional decades.
3 Constraining the trajectory to a 1,000 ppm as ymptote delays the doubling time by three de cades (compared with the 4.3 pc/yr date) and 15 years compared with the corresponding ex ponential path.
4 Constraining the trajectory to 700 ppm produces delays of 45 years and 35 years, respectively.

In other words, there is a marked difference between the constrained and the exponential paths, although there is some overlap in the first few decades.

Table 6.3 Dates by which various CO_2 levels are attained in the constrained and unconstrained scenarios

CO_2 level (ppmv)	*Annual exponential growth (pc/yr)*			*Constrained trajectories asymptotes (ppm)*		
	2.0	*2.6*	*4.3*	*600*	*700*	*1,000*
500	2,023	2,023	2,018	2,050	2,043	2,037
600	2,047	2,042	2,029		2,075	2,057
700	2,063	2,054	2,036			2,074

Energy system constraints and opportunities

We can infer from the previous analysis that in order to meet global environmental constraints, non-fossil energy sources will make a major contribution to world energy supply during the next century. This is hardly surprising to anyone who anticipates a decline in the importance of oil and gas (which presently supply more than 60 per cent of the world's energy needs). However, the recognition that remaining recoverable fossil fuel supplies are really very large— 5,000 TW years or more, that is more than 20 times the total amount consumed so far and enough, at the present rate of consumption, to last another five centuries—makes it less than self-evident that the expected transition to long-term renewable energy resources is in fact considered an urgent necessity among the nations of the world.

It is the anticipated further growth in world energy requirements (as reflected in high and low scenarios) induced by further economic growth of population rich developing countries, coupled with prospective limitations on fossil fuel use well below the physically available resources, that gives rise to the early need for a major growth in non-fossil energy sources. An obvious question arising from this is whether the transition from fossil to non-fossil fuels is likely to be difficult or easy to manage. The details of the transition are complex, and possible answers still appear to be open-ended.

None the less, we are able to analyse a few of the technical, political, strategic and economic constraints which better illuminate the transition. Although no sharp distinctions exist among the various kinds of constraint, it is helpful to group them under several headings.

Technological patterns

Historical data on world energy consumption, when plotted against time as a fractional share of different primary energy sources, follows a very regular pattern. This observation has given rise to the hypothesis that primary energy sources compete for market shares in producing intermediate and final products. Substitution of an existing technology or industrial process to satisfy the needs of a new, emerging technology has been the subject of a large number of theoretical and empirical studies, as evidenced in the journal *Technological Forecasting and Social Change*.

One general finding is that almost all binary substitution processes (e.g. steam engine versus electric power), expressed in fractional terms, follow characteristic S-shaped curves which have been used to forecast future competition between alternative technologies. (This way of looking at technological substitution processes has never been much in favour among economists because they feel it lacks a choice-theoretic or explanatory basis; see Stoneman, 1981.) However, if we can find an endogenous explanation for the way technologies emerge and proliferate, we should be able to parametrise such processes along the line of some logistic substitution model.

Most studies of technological substitution are based on the use of the logistic function, which although not the only S-shaped function is perhaps the most suitable for empirical studies. Another S-shaped function, the Gompertz curve, has also been frequently used, especially to describe population, plant and animal growth.

One of the first such studies showing how technological substitution can be described by an S-shaped curve was the pioneering work of Griliches (1957) on the diffusion of hybrid corn seed in the USA. Griliches showed that hybrid corn replaced traditional corn seed in different States in a very similar way: the S shaped substitution was only displaced by a few years and lasted differing lengths of time in the different States. Following the work of Griliches, Mansfield (1961) developed a model to explain the penetration of an innovation which he suggested was directly proportional to the difference between the expected profit and expected investment associated with the innovation. However, the first systematic attempt at forecasting technological change was made by Fisher and Pry (1970), who extended Mansfield's findings, but only considered fractional shares of a market controlled by two competing technologies.

To handle more than two competing technologies, Marchetti and Nakicenovic (1979) generalised the Fisher-Pry model, since logistic substitution cannot be preserved in all substitution processes. Every given technology undergoes three distinct substitution

BOX 6.1 ENERGY SUPPLY TECHNOLOGIES

The projection energy supply market share technologies can be summarised by the following fundamental model, with f being the total market share:

$$\frac{df}{dt} = (\beta + \alpha f)\ (f_m - f) \qquad (3)$$

$$f\,(t = t_0) = f_0$$

where α, β = constants
f_m = upper limit on market share

The constant α can be interpreted as the index of influenced adoption so that $\alpha f(f_m - f)$ can be thought of as the imitation component. Similarly, β can be interpreted as the index of uninfluenced adoption, so that $\beta(f_m - f)$ can be explained as the innovation component.
The Fisher–Pry model can be obtained from (1) by setting $\beta = 0.0$ and $f_m = 1.0$; that is,

$$\frac{df}{dt} = \alpha f\ (1 - f) \qquad (4)$$

The differential equation has the solution:

$$\left(\frac{f}{1 - f}\right) = \alpha t + \gamma;\ \text{or} \qquad (5)$$

$$f = \frac{1}{1 + \exp\ (-\alpha t - \gamma)} \qquad (5a)$$

This process may be thought of as pure imitation diffusion process.

$$F_{fo}(t) \cong \frac{1}{k \cdot AF} \cdot \dot{M}(t) \qquad (1)$$

$$M(t) = \frac{\beta}{1 + \gamma\, e{-\alpha(t - t_0)}} + \lambda \qquad (2)$$

The table below presents the market penetration rates as well as market penetration times (t_m) of the various energy supply technologies.
Starting with (5), t_m can be calculated as follows.

$$t_m = \left[\ln\left(\frac{0.5}{1 - 0.5}\right) - \ln\left(\frac{0.01}{1 - 0.01}\right)\right]\alpha^{-1}$$

It is clear from the table that characteristic global penetration times are of the order of 100 years, and that the smaller the region or country, the shorter those times will be.

Energy technologies market penetration rates and times

	Technology (%/yr)	*Penetration rates (α)*	*t_m (yrs)*
World primary energy supply	Oil	4.9	94
	Natural gas	4.8	95
	Coal	2.7	170
US primary energy supply	Oil	5.3	86
	Natural gas	4.5	102
	Coal	7.0/4.6*	66/99*
OECD-Europe primary energy supply	Oil	10.0	46
	Natural gas	15.7	29
	Coal	6.9	66

*Note:** The first figure refers to the growth stage, while the second refers to the decline stage.

phases: growth, saturation and decline. They assumed that only one technology was at the saturation phase at any given time, that declining technologies fade away logistically and that new technologies enter the market and grow at logistic rates. The current saturation technology is then left with a residual market share. Once the current saturation technology has reached the stage of logistic decline, the next newer technology enters its saturation phase and the process is repeated until all but the most recent technology are in decline. In order to test this hypothesis, Marchetti and Nakincenovic have applied their model to three different levels of energy system aggregation:

- Primary energy inputs for the world as a whole.
- Primary energy inputs for single nations or re gions.
- Energy subsystems, such as electric utilities.

In total, they used 60 databases to generate 300 examples of 30 different spatial and structural subsets of the world energy system. The quality of it was found to be consistently high in all examples.

The first fact to be observed about the curves is the extreme regularity and slowness of the substitution, which takes about 100 years to grow from 1 per cent to 50 per cent of the market. This length of time, analogous to the time constant of the system, is called the market penetration time. The regularity refers not only to the fact that the rate of penetration remains constant over such very long periods, when so many perturbations are present, but also to the fact that all perturbations are reabsorbed elastically without influencing the trend. Another observation is that different penetration trends have almost the same 'time constant' for all energy sources, except nuclear energy which only achieved 1 per cent share of primary energy in the early 1970s. Its future penetration rate cannot therefore be distilled from historical data, since the initial phases of market penetration will not stabilise on long-term substitution trends for some time yet.

It is prudent to assume from the preceding discussion that new technologies necessary to complement and replace fossil fuels in the future will not be introduced any faster than they were in the past. It should be noted, however, that the dynamics of technological substitution are more complex than depicted here. Since the penetration rate tends to decrease with decreasing rates of energy growth it can be assumed that future penetration rates will be slowed correspondingly. In particular, we need to explain penetration rates endogenously, as part of an interaction of economic factors within an economic model (Gottinger, 1998a).

Penetration of non-fossil and renewable fuels

It has been argued by a number of authorities (e.g. Goldenberg *et al.*, 1988) that global economic goals over the next several decades can be met with scarcely any increase in the present level of global energy use. Particularly when first argued, this finding was in sharp contrast to conventional energy analyses that called for substantial increases in global energy supplies. Conventional projections imply that the scarcity of affordable energy would significantly strain development and that the resulting dependence on fossil fuels would give rise to formidable environmental and security problems.

The energy strategies applicable to achieving a sustainable world are highly relevant to coping with problems of climate change. On the other hand, the present conventional supply-oriented strategies, although placing more emphasis on conservation and efficient use of energy than in the past, are generally not supportive of the goal of sustainable develop ment. For this purpose, attention would need to be focused more strongly on end-use but also on the introduction of genuinely renewable sources of energy.

There are abundant opportunities in both industrialised and developing countries for using energy more efficiently; indeed, it can be cheaper at current energy prices to save an extra unit of energy rather than generate it (Lovins and Lovins, 1991). As discussed earlier, improvement in the efficiency with which energy is used is one obvious way to reduce carbon dioxide emissions. A second is to increase the use of less carbon intensive fuels and of fuels that are renewable. The first approach would reduce the energy required for a unit increase in real GDP, thus increasing the productivity with which energy is used. This does not necessarily imply that less energy would be used overall, but that at the margin every additional unit of energy would be utilised more efficiently than before. It certainly should not imply any reduction in achieving basic social and economic goals in either industrialised or developing countries. But in the long run the achievement of sustainable growth requires a shift to renewable energy resources. The principal renewable sources available are (in declining order of importance) solar, biomass, hydroelectricity and geothermal energy, all of which are currently making a positive contribution to global energy needs. The various solar technologies are in markedly different phases of development, with some, such as solar water heating, passive space conditioning and wind power, already beginning to compete with conventional sources. Others are still the focus of R&D. The fuel cycle associated with the present use of nuclear power is not a renewable one, although there is the potential to achieve a similar effect with fast breeder reactors. Nuclear is one of the principal non-fossil fuel sources available and could be seen as a major element in any strategy to reduce CO_2 levels. However, a lack of public confidence could make it an infeasible choice. In any appreciation of the future role to be played by renewable energy sources, it is important to understand the practicalities involved in realising their potential. Timing, feasible extent of penetration and cost in relation to conventional non-sustainable fuels are of particular relevance and are discussed below. (See Appendix.)

Biomass potential

Biomass energy, for example, fuel wood and agricultural waste, is a form of solar energy that is already in wide use, particularly in the developing world. The proportion of energy demand met from biomass is uncertain but probably ranges from under 2 per cent in the industrialised countries as a whole to 40 per cent, for example, in India. The traditional use of these fuels is declining or, at best, stagnating in many developing countries, generally through a combination of rapidly reducing resources and increasing urbanisation. Biomass is often regarded as a poor people's energy from which the industrialised countries have long since moved away. This makes it difficult for people to take seriously a return to the widespread use of biomass. If this were to mean the reintroduction or retention of inefficient technologies and continued use of wood without replacement, then clearly there would be no benefit at all in terms of reducing CO_2 levels.

However, a biomass economy involving advanced technology, allied to a one for one replacement policy, could have a role to play. In developing countries, gas derived from biomass could be used for cooking at much higher efficiencies than is possible today with wood burning stoves. This gas could also be used in high efficiency gas turbines or other engines. New developments in turbine technology offer the prospect of electricity generation or cogeneration with low cost systems at high thermodynamic efficiency.

There are already a number of projects in existence to convert biomass to liquid fuels, notably in Brazil, where in 1986 27 per cent of total fuel consumption in transportation was supplied by biomass derived fuel and where (as in other LDCs such as China and India) its primary energy share in the 1990s remains between 15 and 20 per cent (Sathaye and Christensen, 1994). Most of the biomass to liquid fuels projects are based on surplus agricultural products and much more needs to be done to set up projects on a self-sustaining basis. A major hindrance to the wider use of modern biomass technologies is the high cost of conversion processes; all existing biomass to motor fuel projects are hopelessly uneconomic at present oil prices. The other major hindrance is the potential conflict over land use.

Overall biomass resources could be very significant. One measure of the size of the resource is that some 100 TW years of chemical energy are stored in plants each year through the photosynthesis process. This is about ten times current global energy needs (Goldenberg *et al.*, 1988). There are three main possibilities for harnessing these resources: through recovery of organic wastes from agricultural and industrial processes, through increased harvesting of biomass from existing forests and through the intensive management of biomass production on energy farms or plantations. The first two may add to CO_2 levels through the burning of material that would otherwise be used to fertilise the soil, for example. Only the route of full replacement of the resource used, allied to efficient technologies, would have a beneficial impact on CO_2 levels.

Solar market penetration

Though not yet fully commercial, perhaps the most promising long-term renewable energy options are offered by photovoltaic technologies that directly convert sunlight into electricity or hydrogen. Other more direct uses of the sun for heating water are already widely used and commercial in a number of countries with appropriate climates. The penetration rates of solar energy technologies will depend to a certain extent on the competitive environment for all energy supplies, while the uncertain cost estimates for most potentially significant solar options limit the procedure that can be used to determine solar market penetration rates.

There are currently numerous market penetration analyses of the solar energy market but some economists (Warkov and Meyer, 1982; Feldman and Wirtshafter, 1980) are sceptical of their value. The criticisms levelled at these models (applicable to any market penetration models) are based on the fact that they are grounded in very simple behavioural assumptions, making them impossible to test. The one claim to legitimacy of these analyses, however, is that of being based on well-developed models of diffusion processes, although their applicability may (as pointed out previously) be limited by the implicit assumption of the pure imitation diffusion process.

Assuming a 1 per cent share for solar energy in the year 2000 in IIASA scenarios (Háfele, 1981) the actual energy share was only 0.6 per cent (High), or 0.15 per cent (Low), taking a market penetration rate consistent with the upper end for other energy sources in the past. This would lead to an f_s (market share of solar in total primary energy) of 6 per cent by 2030 and 18 per cent by 2050. For comparison, note that IIASA predicts f_s of only 1.4 per cent by 2030 (High Scenario), or 1.3 per cent (Low Scenario). Even if we add to solar sources the ‘other’ category in the IIASA scenarios (which include biogas, geothermal and commercial wood use), their combined share by 2030 is only of the order of 3.65 per cent for both scenarios. However, many analysts argue that with adequate government incentives, the rate of market penetration of solar technologies could be far greater than suggested by the application of the historical penetration rates of other energy sources (Okken, 1989).

Wind energy is another form of solar energy which has been harnessed since very early times but which has made a comeback in recent years, helped by new technology based on aeronautics. This has been assisted in the USA in particular by advantageous tax treatment. Under favourable wind regimes electricity from wind generators can compete with fossil fuelled power plants. It is, however, a power source subject to severe fluctuations in availability. There are also limitations to the extent to which it can be used in an integrated power system, and if it is to provide reliable power, storage or back-up facilities must be provided to compensate for the intermittency of the wind. The 1981 IIASA global energy study estimated a global technical potential for wind energy of 3TW (26,000 TWh per year), roughly equal to all the power generation capacity in the world at present. However, there are few signs of it making a significant impact in reducing CO_2 emissions until the economics become much more favourable and there is wider public acceptance of its visual and noise pollution.

Hydro and geothermal

The most widely used renewable resource is hydroelectric energy. The theoretical availability of hydro resources is the potential power that can be developed from the runoff of water from each segment of land surface from some average elevation above sea level. The World Energy Conference 1990 put the theoretical potential from this source at over 15,000 GW which would be equivalent to displacing over 200 million barrels of oil equivalent daily. However, the technically usable potential is usually put within the range 2,200–2,900 GW and, of this, perhaps 30–40 per cent is economic at present. The majority of economic sites have already been exploited in industrialised countries. For example, in Western Europe over 90 per cent of potential capacity is already in place and operating. Most remaining sites are in the former USSR and the developing regions, particularly China, South-East Asia and South America. The economics of hydro development are very site-specific and capital intensive. Costs are also rising in real terms because most of the attractive sites have already been used, but the absence of fuel costs once the development is in place usually makes hydroelectricity the cheapest source of electricity.

Hydro projects usually have a long life and can bring substantial economic benefits, as well as being a source of energy that does not contribute to global warming. Unfortunately, such projects can also create severe ecological and social problems including the opportunity costs of diverting large land areas to hydro facilities, social disruption, water losses associated with the creation of large reservoirs in arid lands, disruption of fish migratory patterns, the spread of waterborne diseases and soil erosion. Nevertheless, many of the potential problems caused by large hydro schemes can be overcome by good planning and management. Although they have clear environmental advantages over fossil fuels that contribute to global warming, in the longer term the pace of hydro development will depend on the availability of capital and the prices of competing fuels seen in relation to local and broader environmental concerns.

There is considerable potential for small-scale hydro schemes, say below 1 MW, in many developed and developing countries. At present, to achieve a given level of electricity output small-scale hydro plants will usually be more expensive unless there are subsidies of one kind or another. They may also be unable to compete with other small-scale sources of power and there are problems of tying them in with existing electricity distribution systems. So, without very specific conditions, small-scale hydro seems unlikely to be more than a very minor contributor.

Geothermal power is important in a few countries lying in volcanic zones. Overall, it could represent a theoretical potential of as much as 1,000 GW, but in practice, with most of the best sites already developed and many practical difficulties to overcome, its potential in global terms is not very significant.

The nuclear option

In the 1970s, when power from nuclear fission was regarded as the key energy source of the future, the size of the uranium reserves was often considered to be a potential constraint. With the recycling of plutonium in breeder reactors seen as a way of overcoming this constraint, together with processing and long-term research into nuclear fusion, nuclear power could be regarded as a renewable resource.

Because nuclear power generates electricity without producing CO_2 it is already making a substantial contribution to reducing greenhouse gases. However, environmental problems associated with nuclear power, particularly its fuel cycle, the disposal of nuclear waste, the proliferation of plutonium and the potential for widespread and very long-term ecological damage have to be weighed against its advantages. Major accidents in power plants and the costs of decommissioning together with a general lack of public confidence have resulted in new orders for nuclear power plants falling dramatically on a global basis. But at the same time there is some revival of interest in nuclear sources as a counter to global warming, and the Brundtland Commission considered it to be one of the few realistic options available for reducing the burning of fossil fuels. New technologies, including ‘fail safe’ reactors and the economic closing of the fuel cycle are under consideration, but the long lead times mean that, even with an immediate revival of interest in nuclear power, little real impact could be felt before the turn of the century.

The future nuclear share envisaged in the work by IIASA is very scenario-specific, given the fixed solar prediction in either the low or high scenarios. But assuming solar meets this challenge, what are the residual burdens that nuclear should bear? Many variants have been analysed, and we shall therefore discuss a representative sample. Taking the year 2000 as possible Action Initiation Time, nuclear power could contribute between 21 and 43 per cent by the year 2030 for the High Scenario with CO_2 limits ranging from 500 to 700 ppm. The Low Scenario, on the other hand, would require 4–28 per cent under the same conditions. In terms of reactors (with a capacity of 1 GWe) the entire scenario range comes to between 500 and 11,000 reactors, which is too wide for useful policy discussion. It may be noted that the world-wide installed nuclear capacity is presently 350 GWe.

In all case studies, only after the year 2025 is the nuclear share markedly affected by the penetration assumed for solar energy. One particular argument, that of electricity growth, seems to speak for the nuclear option. But counter-arguments include cost considerations (Barnes, 1991) and the other nuclear-related issues touched on earlier (Williams, 1990). Currently, the end markets open to nuclear power are mainly in electricity production, and the opportunities there are presented in Table 6.4 showing projected electricity generation and installed capacity in IIASA scenarios (Háfele, 1981).

Nuclear power represented about 5 per cent of both generated and installed capacity in 1975 rising to 17 per cent in 1989, but now stagnating in Europe and North America, (Gottinger, 1996a).

Inter alia, the IIASA scenarios embody a transition to electricity as the reference energy system, with the associated secondary energy carrier possibly being hydrogen. Moreover, since most of the renewable energy sources discussed earlier tend to use electricity as a secondary energy carrier, they compete directly with nuclear for the same market. Hydrogen would be a partner for electricity because it can be stored and transported. But its use as a final energy source would require a significant adaptation of existing infrastructure and thus need time and capital investment.

Projected electricity generation quoted in IIASA scenarios offer constraints only in the near term and involve a variety of assumptions about electrification in various parts of the world. Although the trend of increasing penetration of electricity into end-use markets slows down significantly, especially in industrialised countries, electrification rates could be increased above IIASA assumptions if electricity more rapidly entered the market for space heating, water heating and cooling. Also, opportunities exist in industry if heat pumps penetrate the market more rapidly than hitherto assumed. Another way to accelerate the trend is by increasing the electrification of railways and urban mass transit systems throughout the world and by introducing electric cars more rapidly. However, if the whole transportation sector —which represents about 25 per cent of final energy use—is electrified, then secondary electricity would increase to 11.1 and 7.2 TW yr/yr in high and low scenarios respectively. This new demand for electricity would eliminate the 'oversupply' of electricity created by restrictions on fossil fuel use.

We should not take the implications of such a policy lightly, since they herald structural changes in the way societies around the world are organised. It would also transform our view of electricity which could come to be seen as a primary energy source and the 'standard' or reference against which all substitutions would be measured. In IIASA's words (Háfele, 1981), it would be the 'currency of the future world'.

The other half of future energy structure would be hydrogen. Depending on the extent to which natural gas networks will be constructed over the next few decades, new investment in installations, materials and equipment would be needed to put in place the infrastructure for long-distance transmission and local distribution networks that would ensure a smooth transition to a hydrogen economy.

Interim conclusion

From this tentative analysis it is clear that the following steps are essential for coping effectively with the CO_2 problem:

- A strong global programme to use energy more rationally and efficiently (commonly called 'con servation') and a move towards a self-sustaining world economy.
- Development and utilisation of more non-fossil and renewable sources in the interim.
- Accelerated efforts to shorten the time-lag be tween effective action time and problem resolu tion.

Table 6.4 Global electricity generation

	Base year	*High Scenario*		*Low Scenario*	
	1975	*2000*	*2030*	*2000*	*2030*
Electricity, secondary (TW yr/yr)	0.75	2.1	4.7	1.7	3.0
Installed capacity (GWe)	1600	4390	9845	3550	6320

The envisaged transition comes at a time of considerable social and political dissent over how to solve the energy problem, and when conventional oil and gas are becoming much more costly. It also coincides with rising expectations on the part of the LDC and of former Eastern-bloc countries that are in the middle of a transition to a fundamentally different economic system.

INTERNATIONAL ASPECTS

The discussion in the previous sections has been confined to the global response to impending CO_2-induced climatic change. However, the effects will be extremely varied in their regional impact. For example, some areas of the globe might experience longer growing seasons or more favourable rainfall patterns. Certain highly productive agricultural lands, at present, could be subjected to prolonged or permanent drought. Thus, it is likely that as more is learned about the problem and as more refinements are introduced into the capability of Global Circulation Models (GCMs), some nations will begin to perceive themselves as, on balance, 'winners', while others will come to see themselves as 'losers'. At that point, diverse regional objectives might prove divisive and diminish the prospects for establishing the international co-operation needed to cope effectively with the CO_2 problem.

Climate scenarios

There are many researchers around the world studying the spatial distribution of the future changes in climate, some using numerical modelling like GCMs, others using primarily past warm periods as analogues of the future (Carter, 1992). With the help of these models it is possible to show temperature changes associated with a doubling of atmospheric CO_2. They produce estimates of average changes according to latitude which are relatively small at the equator but rise dramatically at the pole. GCMs also incorporate a hydrological cycle that permits a study of precipitation, evaporation and soil moisture over continent areas.

We can classify the seven regions used in this study into 'wetter' (W) or 'drier' (D) categories based on agreement of at least two sources:

I	NA	US (D); Canada (W)
II	SU/EE	Former Soviet Union, Eastern Europe (D)
III	WE/AUS	W. Europe (D); Japan, Australia, New Zealand (W)
IV	LA	Mexico (W); Rest/Mixed
V	AF/SEA	Western & Eastern Africa & India (W)
VI	ME/NAF	Middle East & Northern Africa (W) except the fertile crescent may be mixed
VII	C/CPA	Southern China (W); Northern (D)

It should be stressed that this exercise of translating global warming into regional patterns is at most a highly tentative speculation and is often based on conflicting sources, but three points need to be borne in mind:

1 Regional climatic scenarios are based on sparse data and often in disagreement.
2 These studies are 'scenarios' and provide only a guide to the patterns of climate change that could occur, not would occur, as the result of global warming. There will be changes in oceanic and cryospheric boundary conditions influencing large-scale circulation patterns which in turn affect regional temperature and precipitation patterns.
3 Even with this level of uncertainty there seems to be a basis for assigning a higher probability to the direction of the change in some regions, such as regions I, V, VI and VII.

Distributive climatic impacts

As the casual speculations begin to suggest, there is no well-defined process for translating global warming into spatial climatic changes. Consequently, the logical translation of these anticipated changes into effects on agriculture and economy, and into societal responses, is even more uncertain. We are essentially multiplying uncertainties. The main point of the previous analysis is that climatic change is likely to have important distributive implications. Some regions will find themselves with more favourable climates, others with less favourable climates. Indeed, distributive issues could well be of great importance among the seven regions and among groups within each region and among the nations of each group. In the context of these general comments we shall explore the impacts on one sector, agriculture, where effects are at once serious and less uncertain.

BOX 6.2 WORLD FOOD PROBLEM

Given present trends, attention to the world food problem is long overdue. The extent of probable disruption and loss of production in major crop-producing areas because of projected climatic changes gives the problem the dimensions of a major world crisis.

In food and agriculture, the rate of growth in production in the developing world averaged about 3 per cent a year in the 1970s. However, the increase in food production failed to keep up with the growth of population in more than half of the developing countries, particularly the poorer ones (Parry, 1990). The undernourished in developing market economies number at least 500 million and that number continues to rise. The immediate situation and outlook have become more precarious than in the past (Kane, 1992). The increase in food and agricultural production in the 1980s and early 1990s has been lagging behind the increase in the 1970s.

World cereal production in the 1980s fell by about 3 per cent below what was achieved in the 1970s. The cereal gap of the developing countries, which is expected to reach 100 million tons of cereal a year, continues to widen.

The rising cost of agricultural imports severely hampers the efforts of many developing countries to increase their food and agricultural production. Especially noticeable is the rapid acceleration of fertiliser prices, one of the most critical inputs. World carry-over stocks of cereals at the end of the season will represent only about 18 per cent of consumption, which is the minimum proportion required for world food security. Inefficiency in the distribution, handling and transportation of food creates logistical difficulties which aggravate the problem.

Framework for world-regions analysis

Even in its rudimentary form, the foregoing analysis of climatic impacts on regional cereal production suggests the potential conflict between putative 'winners' and 'losers' in the climate game (see Box 6.2). In the cases studied, the probable major losers are likely to be the USA, the former USSR and Europe, and major winners China, India, parts of Africa and Latin America. The fate of other regions is less clear. The negative impacts of such cumulative change on the USA and the former USSR are likely to create severe tensions in the global sociopolitical network, leading to future confrontation within existing socio-political subsystems as resources become inadequate. This selective decline of productivity among the major economic and food producing nations *will not* necessarily be offset by a corresponding increase in developing countries, with their lack of industrial infrastructure and transportation and distribution systems. In addition, questions will inevitably arise concerning responsibility for the costs of the adverse climatic impacts. In practice, this means that nations would need to agree on measures to determine the proportional indemnity to be charged to each region or country. But ultimately countries are free to respect or ignore international agreements.

In principle, a solution should be sought, based on global international co-operation and interdependence, which satisfies conditions of equity and fair ness. As an illustration of one possible approach, the following regional scenarios were constructed. The purpose of this analysis is to illustrate varied regional responses to a global decision on restricting CO_2 levels to certain limits. Global analyses tend to imagine everyone doing the same thing at the same time; in reality people do not do so, and this has important consequences. For example, history tells us that countries proceed from a state of relatively little industrialisation to a state of fairly advanced industrialisation in typically 30–40 years or more (Japan and Germany, for instance). Thus, we can easily imagine that a transition to new technological styles might take a century to implement. Even if all major less industrialised countries started now and followed the same track towards new technologies, we foresee almost half a century of transition.

These qualitative arguments have guided our construction of the regional scenarios, which assume that the high energy-using industrialised countries will shift first from fossil to non-fossil energy technologies, with the less industrialised countries allowed to use fossil fuels for a longer time. A fundamental assumption in the following analysis is the creation of an international regime entrusted with setting up CO_2 standards and a mechanism or framework for establishing carbon quota allowances for each region. The stringency built into the energy scenarios was the constraining of world population to 10 billion by 2100, that is 2.2 times the present size. Also built in was the assumption that per capita energy consumption will gradually decline even to negative values, although the less industrialised countries never enjoy a level comparable to the industrialised countries at any time.

A crucial question arises about whether per capita basis or total prospective cumulative energy consumption should be used when setting up the 'carbon quota' system. No penalties were incurred for pre-1980 CO_2 emissions. Using a per capita system also posed the problem of which population levels to use: present levels, stabilised levels or some cumulative measure. However, on average, setting up standards on total cumulative energy consumption allows the industrialised countries (Regions I–III) a 56 per cent greater share of carbon. Table 6.5 reveals very explicitly the political explosiveness of the CO_2 problem. If there is global action to curtail fossil fuel consumption, it would be near to impossible to persuade any of the countries in Regions IV through VII to reduce or even hold their fossil fuel growth at present rates. There has to be sufficient flexibility to allow some equalisation of the almost elevenfold difference between developed and developing regions (on per capita basis).

In our model, the AIT for Regions I–III is 1990, while for Regions IV–VII it is 2010–2020. It should be clear that this 'equity lag' is politically the very minimum. Once regional carbon shares have been allotted, it is then up to each region to decide independently how best to utilise its allocation. This in turn is reflected in the regional choice of fossil fuel mix. We assume for the period up to the AIT that fossil fuel consumption will follow regional scenarios.

Table 6.5 Regional share of CO_2 emissions in 1980

Region	*Share (%)*	*World population (%)*
I	28.7	6.0
II	24.7	9.2
III	28.1	14.2
IV	4.1	8.1
V	4.6	36.0
VI	1.8	3.4
VII	8.0	23.1
Developed countries	81.1	29.4
Developing countries	18.5	70.6

Table 6.6 Share of carbon wealth by region

Region	*Coal*[a]	*Oil & gas*[b]
I	25.7/25.7	20.1/12.4
II	37.0/49.1	21.7/20.6
III	15.5/8.4	5.7/6.2
IV	1.3/0.4	18.4/7.6
V	6.8/2.1	7.5/6.9
VI	0/0	21.2/42.6
VII	13.7/14.6	5.31/3.9

Source: Háfele, 1981; Ausubel, 1980
Notes: [a] First number is based on estimated economically recoverable coal resources; the second is based on total geological resources.
[b] Both numbers are based on total oil and gas economically recoverable. The first also includes contributions from unconventional resources, whereas the second does not.

The need for an international regime is demonstrated in Table 6.6. The North controls 78 per cent to 83 per cent of the world's coal endowment, with the USA and former Soviet Union alone having some 74 per cent. Although China has 14 per cent of the world's coal reserves, such an imbalance in favour of the North looks very suspicious. Perhaps the South has not yet been properly prospected. On the other hand, the remaining oil and gas resources are more evenly distributed on a North-South basis and regions IV to VII contain between 53 per cent and 61 per cent of the world's oil and gas resources.

From these comments we can anticipate the regions likely to be problematic. Regions II (SU/EE) and VII (C/CPA) might not be as co-operative because of their coal wealth, though SU/EE will be negatively affected, which may counterbalance its response. Regions IV and V would not have much bargaining power and WE/AUS (Region III) would be also in a weak position despite its technological capability. ME/NAF (Region VI), though having no known coal resources, would still be in a relatively strong position because of its oil and gas assets. Appreciating the potential complexities and impediments to a global consensus on the CO_2 question, we shall discuss one possible regional scenario construction in the next section.

Adjustment process

Adaptation might be the path of least resistance in responding to expected changes in climate (Schelling, 1988). Qualitatively, it can easily be shown that migration and industrialisation will be the basic implications of any climate change, especially a slow one, as it impacts upon agriculture. Change in agricultural productivity means that the ratio of population density to agricultural productivity will change, and that such changes can be compensated by appropriate changes in population density (migration) or by increasing agriculture or other economic activities (re-education and industrialisation).

Explicit government intervention to change patterns of behaviour could also assist adaptation. For example, zoning laws could be planned to restrict the construction of new buildings in areas prone to flooding because of projected sea level rise, although recent efforts to prevent construction in flood plains have been notably unsuccessful (Lave and Vickland, 1981). Historically, it can be shown that people do not behave in the ways foreseen by advocates of adaptation. A good example of how myopic the adjustments might turn out to be is contained in Ridker's (1981) study of how coastal populations are likely to adjust to a slow but inevitable rise in sea level. In view of the slowness of the change (of the order of a few feet each century), and man's tendency to use a positive discount rate, it is more likely that sea walls and dikes will be built than that people will evacuate affected areas. And once such sea walls are built, it will again appear cheaper to make them a little thicker and higher as sea level continues to rise. Eventually, much of the human race could find itself living below sea level, with the probability rising over the centuries of a catastrophic breach in the dikes. This is an example of how man's normal response to adaptation could eventually be self-defeating.

There are also considerable costs involved in this adaptation process, possibly including major new investment, significant change in production methods, development of new business relations and disruption and dislocation of present human settlements. Adaptation is inherently redistributive, and could therefore be inequitable, especially for developing countries. Since most developing countries lack strong technological infrastructures, they are highly vulnerable, especially to changes in agriculture and water supply, and adaptation could very well accentuate the North-South cleavage.

Economic cost analysis of various strategies

Nordhaus (1977) pioneered investigations into the costs of constraining CO_2 concentrations to predetermined levels via taxation. Using a market allocation model linked to a simplified carbon cycle model, he considered CO_2 emissions as a resource in short supply, allocating them between sectors in such a way as to maximise national income. Losses are discounted at 10 per cent/yr and by minimising economic losses one obtains 'emission paths' that are 'allowable'. However, the model fails to allow for deployment rates of new non-fossil technologies and probably assumes new capacities are well in place when needed. Hence, Nordhaus is able to conclude that it does not pay to curtail carbon dioxide emissions until almost the time when the limit is reached. For example, in the uncontrolled scenario, the doubling of the pre-industrial concentration is achieved by 2040 but abatement measures become necessary only in the period 2010 to 2020. Nordhaus further calculates shadow prices of carbon dioxide. The shadow price indicates by how much the objective function would increase if the constraint were relaxed by one unit. In this case, the objective function is the real income of consumers. With the doubling scenario, the shadow price starts at \$0.14/ton carbon in 1980, rises to \$68/ton by 2040, and reaches a plateau of the order of \$94/ton for the rest of the century.

Nordhaus (1980, 1994) has used control theory to find 'optimal' emission paths in attempting to maximise the discounted expected utility of consumption:

$$\text{Max}W = \int_0^\infty e^{-rt}\, U[C(t)]dt$$

where: $W(t)$ = welfare functions
r = pure rate of social time preference
$C(t)$ = real consumption
$U[C(t)]$ = utility of real consumption

subject to constraints:

$$C(t) = f[F(t)] - h[M(t)]$$
$$\dot{M} = \beta F(t) - \delta M(t)$$

where:
- $F(t)$ = emissions of CO_2
- $M(t)$ = increase in atmospheric concentration of CO_2 from pre-industrial level
- d = β absorption parameter
- β = carbon cycle constraints
- $f[F(t)]$ = consumption
- $h[M(t)]$ = loss in consumption due to CO_2 build-up.

Nordhaus earlier estimates, (h[2×Mo] (doubling)), the economic value of climate change to range from a gain of $301 billion ($1975) to a pessimistic loss of $691 billion. His 'best guess' is a gross loss of *$180 billion*. (Comparative figures have ballooned in Nordhaus, 1994.) Nordhaus decomposes the goods discount rates into social rate of time preference (r) and a growth discount (a g). The choice of a redistributive parameter (a), which is the elasticity of marginal utility of income, is crucial to the results.

Starting from Nordhaus' 'best guess' we can test the sensitivity of the results to his basic assumptions. If the costs of climate change are closer to his 'pessimistic' assumptions, then the current shadow price increases from $38/ton to a steady state of $323/ton. Also, the control rate becomes 43 per cent in 1980 and reaches 100 per cent at the steady state. If the discount rate is reduced from an assumed 13 per cent to about 5 per cent, the full control of CO_2 emissions will have to be imposed. Obviously, such an analysis is highly simplified and questionable in terms of basic assumptions and even approach. Apart from discounting, which we discuss elsewhere (Gottinger, 1991), Nordhaus assumes the cost of CO_2 build-up to be a linear function of concentration, and neglects the non-linearity of the temperature response as well as possible catastrophic discontinuities. He assumes a parameterised abatement function and ignores possible backstop technology. More elaborate computational results based on a range of scenarios are in Nordhaus (1980).

Equity issues

All the previous studies discussed were concerned with deviations from the global mean, but in many cases the redistributive regional costs might be very different. The natural global climatic losses estimated by Kates (1979) at $30 billion per year originate from floods (53 per cent), tropical cyclones (27 per cent) and droughts (15 per cent). What is striking about Kates' study is the great inequity in the spatial distribution of deaths and losses. For example, of the 250,000 people who die each year from natural hazards, 95 per cent are citizens of poorer nations. Also, the annual drought losses in Tanzania are 1.8 per cent of GNP, while the comparable figure for Australia is only 0.1 per cent. Other examples show similar results. On average, in terms of per cent of GNP, *climatic hazards are about 25 times more severe in developing countries* than in developed countries. If the equity issue is at the heart of the CO_2 problem, then the optimal control or other allocative efficiency theories have to be relativised. These theories, emphasising 'efficiency', tend to net out the losses from the gains with no consideration of distributive impacts. More recent calculations show overall changes for the worse and even wide distribution impacts (Fankhauser and Pearce, 1993).

In the meantime, issues of distributional justice have assumed priority in arguments about international order (Spash and d'Arge, 1989). Assuming we can calculate the time paths of annual costs and benefits for each strategy, we are still faced with the difficulty of expressing the results in a way that reflects their relative present day status. This is often done by discounting. But the long time span involved here raises issues of intertemporal fairness: a problem of trade-offs between succeeding human generations. When the present generation evaluates alternative uses of the atmosphere it is making judgements about the welfare of future generations relative to the welfare of the current generation. This judgement often centres on the choice of a 'correct' social rate of discount, which is an extensively debated question in economics.

Another difficulty arises from the redistribution of costs and benefits. A theoretical underpinning of cost/ benefit analysis can only be effected in one direction, from present to future generations. Page (1977) argues that compensation is likely to be only hypothetical and not real, making the whole discounting procedure meaningless on ethical grounds, since actual compensation is unlikely to be paid.

The intertemporal ethical rule states that in balancing risks to human life in the present, in the future equal numbers of lives should receive equal weight. This especially complicates the evaluation process, by making the 'present value' of future human life independent of the time at which it is lived, in contrast to the present value of a bundle of consumption goods.

Such considerations lead some economists, such as d'Arge, Schultze and Brookshire (1982), to reject the traditional cost/benefit analysis when evaluating the CO_2 issue, in favour of one based on ethical beliefs. They find that differing ethical systems imply differing discount rates. They further suggest that one way of postponing a decision to establish an 'optimal' environmental ethic for future generations would be to estimate the present generation's willingness to pay to avoid environmental risks to future generations. It is hoped that such an estimate might embody the present generation's ethical beliefs.

Clearly, then, many argue that resource depletion and the future long-term quality of the environment are not merely problems of market failure, but distributional problems. They are even more difficult than internalising costs, because people and nations will not agree on the distributional criteria for societies. Should the criteria emphasise maximisation of returns to the international community, conservation of natural resources, or the passing on of a stock of environmental wealth per capita at least equal in value to the one that was inherited? In contrast, should a legacy of infrastructure from development and resiliency be passed on which will minimise the 'loss' from climate change (Ausubel, 1980)?

Uncertainty analysis

These high levels of uncertainty have prompted some policy-makers and scientists to adopt a 'wait and see' approach, since no agreement about the 'optimal' strategy is on the horizon. Thus, they caution against any major actions until better data and information are available (Gottinger, 1991; 1998b). It is, however, quite possible that none of these uncertainties will be reduced in the next decade (Ausubel, 1980; Gottinger, 1996b).

In other words, the conventional assumptions about learning over time may be irrelevant in this issue, and it is possible that we will be facing virtually the same dilemma in ten years' time as we face now, with only slightly more reliable information. This is partly because climate change is plagued with the problem of 'indivisibilities' and 'scant sets' (Olson, 1982). These are areas where our knowledge is generally meagre and the stakes are high. They are also areas where we cannot expect to get reliable answers cheaply or quickly, because a decisive experiment entails a policy change, whilst historical experience yields scant information. It is therefore clear that decision-makers must attempt to live with and accommodate the high degree of uncertainty permeating the CO_2 issue, while policy-makers must realise that decisions will have to be made without definitive cost/benefit information. Waiting until the uncertainty is resolved will probably leave an effective decision until too late.

More attention should be directed at resolving or reducing the effects of 'malign uncertainties' (uncertainties that make it impossible to determine whether the outcome of a particular policy will be tolerably good or very bad), as opposed to resolving 'benign uncertainties' (those which make it impossible to tell which of two good strategies will be better) (Burgess and Burgess, 1980). This means that uncertainties, and therefore losses, become asymmetric around the optimal decision and have to be treated accordingly (Morgan and Henrion, 1990).

Institutional barriers

The 'slow' cumulative build-up of CO_2 in the atmosphere poses a series of challenges to the international community and its institutions. The unique blend of political, economic, ethical, legal and scientific issues raised by the CO_2 problem impels existing institutions to break new ground. Indeed, CO_2 induced climate change is a virtual prototype of the sort of problem that is poorly matched to existing human institutions. Its time span is longer than any political leader's career, and the potential effects are enormous, conceivably dwarfing those of normal, man-made technical and social change. This kind of problem presents an almost insurmountable challenge to institutions designed for times when societies were less complex, man's abilities for doing 'good' and 'bad' much more limited, and thinking much more restricted in time and space.

Design of an international regime

Undoubtedly, some form of international regime for the CO_2 issue and many other 'global commons' is not merely desirable but necessary. Design of mechanisms for a CO_2 regime can be partially extracted from studying other mechanisms developed at the regional or international level to deal with analogous situations. The international CO_2 regime envisaged requires much more substantial global agreement, control, management and allocation of resources. Most of the international agreements discussed have failed to reach substantive agreements (beyond agreeing to study the matter further, to monitor and so forth). Even if such agreements were reached, there are real grounds for doubting that official agreements would be seriously enforced.

The relative lack of enforcement machinery internationally necessarily determines the nature of international negotiations. Thus, it is often less expensive and easier to hold meetings and conferences, and even to issue draft 'statements of principles', than it is to do something practical. We already see this in the proliferation of meetings and papers on the CO_2 issue. The difficulties of investing an international CO_2 regime with the requisite responsibilities and authority, and even with legitimate command of coercion, stem directly from the reluctance of nations to yield sovereignty to international bodies, because doing so means losing some control over decisions that directly affect important national interests and domestic constituencies. However, in practice, some delegation of responsibility cannot be avoided and a variety of means are employed by each nation to make it more acceptable.

A serious investigation of 'optimal' strategies should be initiated to reach substantive agreements under which reasonably effective co-ordinated efforts on the CO_2 issue can be organised among sovereign states. Policies should attempt to identify powerful self-interest and incentives to foster an acceptable agreement, which might be a mixture of selfish and social motivations. The fact that diverse opinions and great uncertainties surround the relative redistribution of climate at the regional level might be helpful for achieving early consensus on action. The tactical steps required to initiate an international regime for CO_2 quickly should be carefully considered, with an eye to the requirements of political legitimacy and the dual need for efficiency and internal consistency in decision-making on highly complex issues.

These conclusions lead naturally to the following recommendations:

1 International co-operation is needed in many areas in order to ensure a timely and orderly 'transition' to non-fossil energy systems. Areas of co-operation include:
   - (a) more efficient and rational world-wide energy use and perhaps subsidies for certain end-use efficient technologies;
   - (b) arresting the increasing fuelwood crisis in the LDCs;
   - (c) international R&D in non-fossil energy systems, their development and commercialisation. Attention should be given to avoid 'monism' in energy policy. All feasible options should be pursued;
   - (d) establishing a world energy financing centre with low interest loans to aid LDCs in technology acquisition, adaptation and training.

2 The highest priority should be given to incorporating the CO_2 issue into the world energy policy-making process, and CO_2 should serve as the discriminating factor in the choice of energy options.

3 Investigation should be promptly initiated to help establish an international institution for dealing with the CO_2 issue. Attention should be directed to the following:
   - (a) how best to optimise effective co-ordinated efforts among sovereign states;
   - (b) setting of standards or 'acceptable' CO_2 limits that will serve as policy targets;
   - (c) a system of carbon 'quotas' or allocation schemes;
   - (d) co-ordinated scientific co-operation in data gathering, monitoring and emission regulation;
   - (e) international agreement on enforcement machinery;
   - (f) the needs of LDCs, their ability to substitute, and their developmental objectives.

4 The USA, EU, and the OECD countries should plan during the next decade to take action to limit fossil fuel use. This is required regardless of the CO_2 problem.
5 There is a need to develop guidelines for world coal trade and to reach early agreements towards its 'liberalisation', especially given the signs of the second coming of the 'coal age'.

In conclusion, it is important to recognise the limitations of the analysis. These limitations are imposed by the paucity of relevant data, the multitude of uncertainties pervading all aspects of the CO_2 problem, the lack of an integrative framework and above all the 'social engineering', ethical and value-laden implications of the issues raised.

POLICY AND CLIMATE CHANGE MODELLING

An enhancement of the greenhouse effect may, over time-scales of a century or more, impact human and natural systems through a change in the Earth's climate. This would mean a rise in the Earth's mean surface temperature, changes in global and regional precipitation, and a rise in sea level. There are major uncertainties associated with predictions on this matter—both in the inputs to and components of the physical climate system, and in the resulting impacts on human and natural systems. As a result no clear policy has emerged.

There have been some studies that attempt to characterise these uncertainties and rank their importance with respect to future outcomes (Peck and Teisberg, 1993). These studies have consistently shown that the impacts of climate change, given current knowledge, depend more upon future socio-economic variables (such as population and technological change and social discount rate) than upon the geophysical variables that are the focus of most current global change research.

If socio-economic factors and technological change are the key determinants of the effects of climate change, then making accurate forecasts about the variables that represent them in models of future worlds is important for predictive modelling. While models of socio-economic phenomena based on statistical extrapolations of past behaviour are reasonably good predictors in the very short term, they show poor results and make conspicuously overconfident predictions when it comes to longer time periods. This is partly because they do not seem to capture long-term 'weak' feedback relationships. The 'Limits to Growth' models (Meadows, 1972) epitomise much that is wrong with long-term 'predictive modelling': an inherent inability to characterise and parameterise long-term interactions between the economy, society and environment. Even if we were able to describe qualitatively all the interactions between model variables, we would still need to parameterise the model. A non-linear world with long-term feedbacks is more complicated than the modellers' limited imagination. As shown by Smale (1980), a non-linear dynamic system of more than four variables can show essentially any behaviour.

Model outputs that calculate 'optimal global paths' (Nordhaus, 1994), such as those from the DICE model, must be treated with caution. Integrated models are useful for an overall understanding of the problem. They allow policy analysts to navigate a sea of uncertainties systematically, and may provide prescriptive information for use in research prioritisation and resource allocation.

That the atmospheric concentration of most greenhouse gases has increased since pre-industrial times is, perhaps, the sole irrefutable piece of scientific evidence in our current understanding of climate change. However, the task of determining the sinks and sources of greenhouse gases accurately is made difficult by the global scale of the phenomena in question. Uncertainties in the quantification and prediction of the sources and sinks of CO_2 and methane can arise out of natural variability, lack of measurements and/or measurement error, and incomplete understanding of scientific phenomena.

Sources of greenhouse gases have large natural components and their sinks are controlled by natural processes. In addition, the amount of carbon terrestrial and oceanic reserves is far greater than the amount of atmosphere. The steady fluxes between these natural reserves are very much larger than anthropogenic release. Therefore, small changes in the factors controlling these fluxes can have a large affect on atmospheric concentrations. Conversely, minor perturbations to the parameters in models of biogeochemical cycles can lead to a large deviation in future predictions. Measurements of fluxes from natural sources and sinks are fraught with uncertainties.

A key element of possible policy responses to global climate change is the abatement of emissional CO_2. There is much debate in policy circles about the true cost of carbon abatement. Engineers and 'bottom up' modellers have constructed supply curves for the costs of carbon abatement which show 30–50 per cent reduction in CO_2 emissions at zero cost from unavailed energy efficiency opportunities. Energy economic 'top down' modellers, on the other hand, predict that carbon abatement policies, especially those calling for stringent abatement measures, can be implemented only at very high cost. Most economic approaches to climate abatement attempt to quantify the costs and benefits (damages avoided) of carbon abatement and compare these in search of a rational policy response. Damage from climate change, or more precisely, the benefits of greenhouse abatement, may be the least understood area of climate change research. This gets worse as one tries to quantify catastrophic, low probability, high consequence events. While damage estimates are not expected to achieve a high degree of clarity in the near future, using quantitative damage estimates in scenario analyses may be one way of explicating the range of future outcomes (Cline, 1992). Another approach is the use of optimal control techniques that attempt to provide 'optimal' global economic pathways, primarily following the work of Nordhaus (1994).

CONCLUSIONS

Climate change due to emissions of greenhouse gases is a long-term, global environmental problem. Specific impacts on different regions, and the timing of those impacts, remain uncertain. At the same time it is reasonable to assume that unilateral voluntary action by individual countries to reduce their net emissions of GHGs is unlikely. The fact is that significant reduction of net GHG emissions by a single major net emitter (say, for example, the USA) is unlikely to slow substantially their rate of increase in concentration in the atmosphere: emissions of GHGs world-wide are increasing rapidly with spreading industrialisation. On the other hand, unilateral changes in energy use patterns are widely perceived to have adverse effects on a country's economic growth, on the welfare of consumers and on trade competitiveness. This perception exists in developing (DCs) and industrialised countries (INCs) alike.

REFERENCES

Ausubel, J. (ed.) (1980) *Climatic Constraints and Human Activities,* Conference Proceedings of the International Institute for Applied System Analysis, Oxford, UK: Pergamon Press.

Barnes, P. (1991) *The OIES Review of Energy Costs,* Review Series, Oxford, UK: Oxford Institute for Energy Studies.

Burgess, C. and Burgess, H. (1980) *Energy Decision Making Under Conditions of Uncertainty,* Working Paper No. MIT-EL80–042 WP, Cambridge, Mass., USA: MIT Energy Lab.

Carter, T.R. (1992) *Preliminary Guidelines for Assessing Impacts of Climate Change,* W6II IPCC, Geneva, Switzerland: WMO/UNEP.

Cline, W. (1992) *The Economics of Global Warming,* Washington DC: Institute for International Economics.

d' Arge, R.C., Schulze, W.D. and Brookshire, D.S. (1982) 'Carbon Dioxide and Intergenerational Choice', *American Economic Review* 72: 251–256.

Fankhauser, S. and Pearce, D.W. (1993) *Social Costs of Greenhouse Gas Emissions, Economics of Climate Change,* Paris, France: OECD.

Feldman, St.L. and Wirtschafter, R.M. (1980) *On the Economics of Solar Energy,* Lexington, Mass., USA: Lexington Books, D.C.Heath Co.

Fisher, J.C. and Pry, R.H. (1970) *A Simple Substitution Model of Technological Change,* Report 70-C-215, Technical Information Series, June, Schenectady, NY, USA: General Electric Co., Research & Development Center. See also (1971) *Technological Forecasting and Social Change* 3: 75.

Flohn, H. (1985) *Das Problem der Klimaänderungen,* Darmstadt, Germany: WIBU (in German).

Flohn, H. (1989) *Treibhauseffekt der Atmosphäre: Neue Fakten und Perspektiven* (Greenhouse Effect of the Atmosphere: New Facts and Perspectives), Vortrage No. 379, 27 Sept., Germany: Rheinisch-Westfalische Akademie der Wissenschaften (in German).

Goldenberg, J., Johansson, T.D., Reddy, A.K.W. and Williams, R.H. (1988) *Energy for A Sustainable World,* New Delhi, India: John Wiley and Sons.

Gottinger, H.W. (1991) 'Economic Models of Optimal Energy Use under Global Environmental Constraints' in R.Pethig (ed.) *Co-operation and Conflict in the Global Environment,* Berlin, Germany: Springer-Verlag, pp. 133–185.

Gottinger, H.W. (1996a) 'Greenhouse Gas Economics and Computable General Equilibrium', *Journal of Policy Modelling* (to appear).

Gottinger, H.W. (1996b) 'Choosing Regulatory Options when Environmental Costs are Uncertain', *European Journal Oper. Research* 88: 28–41.

Gottinger, H.W. (1998a) 'An Endogenous Model of Economic Activity and Climate Change', *Metroeconomica* 49, 2: 139–168.

Gottinger, H.W. (1998b) *Global Environmental Economics,* Dordrecht, The Netherlands: Kluwers.

Griliches, Z. (1957) 'Hybrid Corn: An Exploration in the Economics of Technical Change', *Econometrica* 25: 501–522.

Háfele, W. (ed.) (1981) *Energy in a Finite World,* Cambridge, Mass., USA: Ballinger Publishing Co.

International Panel on Climate Change (IPCC) (1990) *Climate Change, The IPCC Scientific Assessment,* Cambridge, UK: Cambridge University Press.

Kane, S. (1992) 'An Empirical Study of the Economic Effects of Climate Change on World Agriculture', *Climatic Change* 21: 17–35.

Kates, R.W. (1979) 'Climate and Society: Lessons from Recent Events' in *Proc. of World Climate Conference,* Geneva, Switzerland: WMO.

Kellogg, W.W. and Schwarz, R. (1981) *Climate Change and Society,* Boulder, Col., USA: Westview Press Inc.

Lave, L.B. and Vickland, K.H. (1989) 'Adjusting to Greenhouse Effects', *Risk Analysis* 9, 3: 283–291.

Lovins, A.B. and Lovins, L.H. (1991) *Least Cost Energy, Solving the* CO_2 *Problem,* Andover, Mass., USA: Brick House Publishing Co.

Mansfield, E. (1961) 'Technical Change and the Rate of Innovation', *Econometrica* 27, 4: 741–766.

Marchetti, C. and Nakicenovic, N. (1979) *The Dynamics of Energy Systems and the Logistic Substitution Model,* RR-79–13, Laxenburg, Austria: IIASA.

Meadows, D. (1972) *The Limits to Growth,* Washington DC: Potomac Assoc.

Morgan, M.G. and Henrion, M. (1990) *Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis,* Cambridge, UK: Cambridge University Press.

National Academy of Sciences (NAS) (1979) *Carbon Dioxide and Climate: A Scientific Assessment,* Washington DC: NAS.

Nordhaus, W.D. (1977) *Strategies for the Control of Carbon Dioxide,* Discussion Paper No. 443, USA: Cowles Foundation, Yale University.

Nordhaus, W.D. (1980) *Thinking about Carbon Dioxide: Theoretical and Empirical Aspects of Optimal Control Strategies,* Discussion Paper No. 565, USA: Cowles Foundation, Yale University.

Nordhaus, W.D. (1994) *Managing the Global Commons: The Economics of Climate Change,* London, UK: MIT Press.

Okken, P.A. (1989) *Climate and Energy,* Dordrecht, The Netherlands: Kluwers.

Olson, M. (1982) 'A Conceptual Framework for Research about the Likelihood of a Greenhouse Effect', in US Department of Energy, *Carbon Dioxide Programme,* Washington DC.

Page, T. (1977) *Conservation and Economic Efficiency,* Baltimore, Md., USA: Johns Hopkins University Press.

Parry, M. (1990) *Climate Change and World Agriculture,* London, UK: Earthscan.

Peck, S.C. and Teisberg, T. (1993) 'CO_2 Emissions Control: Comparing Policy Instruments', *The Energy Journal* 21: 222–230.

Ridker, R.G. (1981) 'Social Responses to the CO_2 Problem', in USDOE Carbon Dioxide Effects: Research and Assessment Program, *Proceedings of the Carbon Dioxide and Climate Research Program Conference,* Washington DC: USDOE, P. 227.

Rosenberg, N.J., Easterling, W.E., Crosson, P.R., Ros, E.V. and Darmstadter, J. (1989) *Greenhouse Warming: Abatement and Adaptation,* Resource for the Future, Proceedings, Washington DC: Johns Hopkins University Press.

Rotty, R.M. (1977) 'Global Carbon Dioxide Production from Fossil Fuels and Cement, AD 1950-AD 2000', in Anderson *et al.* (eds) *The Fate of Fossil Fuel* CO_2 *in the Oceans,* New York, USA: Plenum Press, pp. 167–181.

Sathaye, J. and Christensen, J. (1994) 'Methods for the Economic Evaluation of Greenhouse Gas Mitigation Options', Special Issue, *Energy Policy* 32, 11: 113–249.

Schelling, T. (1988) *Global Environmental Forces,* E-88–10, Nov., Energy and Environmental Policy Center, Harvard University.

Smale, S. (1980) *The Mathematics of Time,* New York, USA: Springer.

Spash, C.L. and d'Arge, R.C. (1989) 'The Greenhouse Effect and Intergenerational Transfers', *Energy Policy* 17: 88–96.

Stoneman, R. (1981) *The Economic Analysis of Technology,* Oxford, UK: Oxford University Press.

Warkov, S. and Meyer, J.W. (1982) *Solar Diffusion and Public Incentives,* Lexington, Mass., USA: Lexington Books, D.C.Heath Co.

Williams, R.H. (1990) 'Low Cost Strategies for Coping with CO_2 Emission Limits', *The Energy Journal* 11, 4: 35–59.

SUGGESTED READING

Congress of the USA, Office of Technology Assessment (1991) *Changing by Degrees: Steps to Reduce Greenhouse Gases,* Washington DC: US Government Printing Office.

Goldenberg, J., Johansson, T.D., Reddy, A.K.W. and Williams, R.H. (1988) *Energy for A Sustainable World*, New Delhi, India: John Wiley and Sons.
Háfele, W. (ed.) (1981) *Energy in a, Finite World*, Cambridge, Mass., USA: Ballinger Publishing Co.
Intergovernmental Panel on Climate Change (IPCC) (1992) *Climate Change 1992*, The Supplementary Report to the IPCC Scientific Assessment, Cambridge, UK: Cambridge University Press.
Rosenberg, N.J., Easterling, W.E., Crosson, P.R., Ros, E.V. and Darmstadter, J. (1989) *Greenhouse Warming: Abatement and Adaptation*, Resource for the Future, Proceedings, Washington DC: Johns Hopkins University Press.
Tester, J.W., Wood, D.O. and Ferraro, N.A. (eds) (1991) *Energy and the Environment in the 21st Century*, Cambridge, Mass., USA: MIT Press.

APPENDIX

Market penetration of non-fossil fuels

Under normal conditions replacement time for industrial equipment could run from 30 to 50 years. We have the non-fossil component of our scenarios plotted as log *y=log*[*f(*1*-f*] where f is the market share of the non-fossil component.

Three logistic curves are computed and the scenarios are of the form

$$\log\ y = \alpha(t - t_0) + \gamma$$

where
α a is the penetration rate
γ is a constant.
γ was calculated on the assumption that the new non-fossil component could be treated as one single aggregate growing from an initial market share of 8.6 per cent (1990). It is clear that such an assumption is invalid and will most probably yield an overestimation of the true potential of non-fossil supplies. For one thing, the hydro fraction has already saturated and is on the decline. Thus, the true growth of the non-fossil component is mainly nuclear, presently making a contribution of the order of 3 per cent of total primary input.

For the High Scenario we describe two curves from our scenarios, namely for AIT of 1990 and the year 2000 along with logistic curves of constant penetration rates of 6.2, 5.2 and 4.2 per cent/yr, representing an accelerated penetration, a historical trend and moderate to slow penetration, respectively (see Table 6.4).

The AIT=1990 curve follows closely a logistic trend with 6.2 per cent/yr until the year 2030, at which time the non-fossil share is on the order of 66 per cent. Afterwards, it moves to a slower trend of about 4.6 per cent/yr. On the other hand, the later AIT curve displays non-logistic behaviour in the early decades before AIT, averaging a penetration rate of about 1.7 per cent/yr. Subsequently, three distinct phases of growth can be discerned: a ‘catch up’ phase, ending about 2035 and characterised by an unprecedented high penetration of about 8.6 per cent/yr; a transition and readjustment phase ending by 2050; a saturation phase marked by slower penetration of 6.5 per cent/yr, but still faster than the accelerated penetration curve.

The Low Scenario, in contrast with the High, is expected to show a reduced penetration rate. But is this the case? The AIT=1990 curve follows very closely the historical substitution trend up to 2040, where the non-fossil fractional share is 63 per cent. Afterwards it follows a saturation rate of about 3.8 per cent/yr.

In the AIT=2000 substitution process four trends are evident. The first, lasting until around the year 2005, represents the unconstrained penetration (i.e. no CO_2 constraint) and shows slow penetration of about 1.4 per cent/yr; the second extends to the year 2035 with a mean penetration of 8.1 per cent/ yr —a very high rate under any circumstances; the third stage is transitional, before the long-term trend defines the fourth phase and closely follows the historical substitution trend.

The previous discussion has been limited to a CO_2 limit of 500 ppm, but how about other limits, say, for the 600 ppm asymptote. The AIT=1990 manifests two trends: 4.6 per cent/yr extending to the year 2035 and 3.6 per cent/yr after that. As before, the AIT=2000 reveals four phases of penetration with annual rates: 1.6 per cent ending by 2000; 5.8 per cent ending by 2035; transition phase ending 2055; and 4.6 per cent. These penetration rates are still high compared either to historical or IIASA projected trends.

SELF-ASSESSMENT QUESTIONS

1 The high degree of uncertainty in predicting the effects of CO_2 emissions is compounded by these effects occurring at various complex levels. All of these levels of effects are significant, but choose from the list below the one that is considered most important.
(a) The actual CO_2 content of the atmosphere.
(b) The relationship between energy use and CO_2 emissions.
(c) The impact of climatic change and increased CO_2 levels on the environment and human activities in specific regions.
(d) The inability of models to couple the ocean and atmosphere in ways that describe the real energy exchange and division of energy transport between the two.

2 In the context of the CO_2 problem, the world's cumulative use of fossil fuels must be restricted to levels below the estimated recoverable resources. What is the key reason for this?
(a) The fuel cost curve becomes increasingly steep as fossil fuel resources are reduced, so that financing the introduction of clean fuel technologies becomes prohibitively costly.
(b) Extractable fossil fuel resources are sufficient to increase CO_2 concentrations by a factor of 5 to 8 over the pre-industrial level.
(c) An accelerated transition to non-fossil fuels can only be achieved through creating the incentive of potential fossil fuel scarcity.
(d) Resources of fossil fuels are inadequate to meet the world's need for energy beyond the end of this century.

3 It is unsafe to assume that new technologies needed as future complements to or replacements for fossil fuels will be introduced faster than in the past. This is because:
(a) new technology is far more complex than that which it replaces;
(b) capital flows required to put the new technologies in place are likely to be inadequate;
(c) Market penetration has been shown to be taking place at a slow and regular pace for almost all energy sources;
(d) Current energy technology has, in the main, not yet stabilised its market penetration to long-term substitution rates.

4 A number of steps are essential for coping efficiently with the CO_2 problem. They include one of the following:
(a) The development and introduction of more non-fossil and renewable sources of energy.
(b) The initiation of a series of North–South dialogues on the environment.
(c) Restraint on private motoring.
(d) The formulation by all countries of national energy plans within a framework agreed at international level.

5 The renewable energy sources such as solar, wind and hydro compete directly with nuclear energy for the same market. This is because:
(a) they are all clean sources of energy;
(b) they use electricity as a secondary carrier;
(c) they are roughly comparable in unit technical cost;
(d) they have a similar pattern and pace of market penetration.

6 A prime characteristic of the CO_2 problem that makes its consequences potentially so hazardous to global climate is:
(a) no economically feasible control technologies are available;
(b) the bulk of emissions comes from the wealthy industrialised countries of the North;
(c) the long time-lags between the cause and identification of significant trends;
(d) the difficulties in obtaining international funding.

7 The CO_2 problem has been described as politically explosive. One of the main reasons for this is:
(a) countries with 29 per cent of the world's population contribute 81 per cent of the CO_2 emissions;
(b) no international agreement on acceptable levels of CO_2 emissions has been reached;
(c) the raising of sea levels as a result of global warming will lead to territorial disputes;
(d) a new world order will have to be introduced if the problem is to be solved.

8 Emissions of CO_2 occur whenever fossil fuel is oxidised. The rate of emission varies among fuels. Which of the following energy sources release CO_2 into the atmosphere?
(a) Nuclear, solar, hydro, geothermal.
(b) Coal, oil, natural gas, biomass.

7

BEST AVAILABLE PRACTICES IN MINING ON LAND OF MINERALS

Alyson C.Warhurst

SUMMARY

Public policy to promote technical change and foster economic efficiency, rather than environmental regulation alone, is more likely to achieve sustained and competitive improvement in the long-term environmental management of our non-renewable natural resources. Three key issues are explored in this chapter.

1 The relationship between production efficiency and environmental performance

There is growing evidence that technical change, stimulated by the 'environmental imperative', is reducing both production and environmental costs to the advantage of those dynamic companies that have the competence and resources to innovate. Such companies include mining enterprises in developing countries, as well as transnational firms, but the evidence is strongest for large new investment projects and greenfield sites. In older ongoing operations, environmental performance correlates closely with production efficiency, and environmental degradation is greatest in operations working with obsolete technology, limited capital and poor human resource management. The development of the technological and managerial capabilities to effect technical change in those organisations would lead to improved efficiencies in the use of energy and chemical reagents and in waste disposal, to higher metal recovery levels and better workplace health and safety. This in turn would result in improved overall environmental management.

2 The economic and environmental limitations of regulation

Currently, the environmental performance of a mining enterprise is more closely related to its capacity to innovate than the regulatory regime within which it operates. Although international standards and stricter environmental regulation may not pose problems for the economics of new mineral projects, there could be major costs and challenges involved for older, and particularly inefficient, ongoing operations. Controlling pollution problems in many of these cases requires costly add-on solutions: water treatment plants, strengthening and rebuilding tailings dams, scrubbers and dust precipitators, etc. Furthermore, in the absence of technological and managerial capabilities, there is no guarantee that such items of pollution control—environmental hardware—will be incorporated or operated effectively in the production process. In some instances such requirements are leading to shut-downs, delays and cancellations as well as reduced competitiveness. When mines and facilities are shut down the clean-up costs are frequently transferred to the public sector, which, particularly in developing countries, has neither the resources nor technical capacity to deal with the problem effectively. In most countries, perhaps with the exception of the USA, the lack of retrospective regulation means the 'pollutee suffers and pays' principle is alive and well. This is not in itself, however, an argument either against regulation *per se* or for the global diffusion of Superfund legislation.

3 The case for an environmental management policy

The implication of this analysis is that to ensure competitive and sustainable environmental management practices in metals productions, governments need to embrace public policy that goes beyond traditional, incremental and punitive environmental regulation. The latter, in the old 'environmental protectionist' mode, tends to treat the symptoms of

environmental mismanagement (i.e. pollution) not the causes (i.e. lack of capital, skills and technology and the absence of the capability to innovate). The challenge will be for governments to ensure that companies operating within their national boundaries remain sufficiently dynamic to be able to afford to clean up when operations cease and to innovate to improve economic efficiency and environmental management in the meantime. Governments need policy tools that enable them to predict 'corporate environmental trajectories' and pick up the warning signs of declining competitiveness and impending mine close-down to ensure sufficient resources are available for the environmental management of mine 'decommissioning'. Policy mechanisms need to be developed to promote technical change and to build up the technological and management capabilities to innovate and manage the acquisition and absorption of clean technology. The privatisation of the state sector and the liberalisation of investment regimes in many developing countries, Angola, Mozambique, Namibia, Botswana, Bolivia, Peru and Chile, with their emerging emphasis on joint ventures and interfirm collaborative arrangements, provides new opportunities for the diffusion of both competitive and environmentally sound best practice in metals production.

Public policy to promote technical change and, complementary to that, to improve economic efficiency, respects the interplay between the environmental and economic factors that constitute a sustainable development approach to the long-term environmental management of our non-renewable natural resources. Environmental regulation at best provides only one element of a public policy for environmental management.

ACADEMIC OBJECTIVES

- To understand the diverse environmental impacts related to the mining of geologically and geographi cally different mineral deposits.
- To understand that there is flexibility regarding the potential to improve the environmental perfor mance of mining operations.
- To understand that different types of firm adopt different types of strategic response to environmental management and that public policies, including environmental regulation, are not the only determi nants of firms' environmental performance.

THE LIMITATIONS OF ENVIRONMENTAL REGULATION AND CHALLENGE FOR PUBLIC POLICY

Regulatory frameworks for safeguarding the quality and availability of land, water and air degraded as a result of mining and mineral processing activities are growing in number and complexity. This has particularly been the case in the major mineral producing countries of North America and Australia as well as Japan and Europe. The norm in environmental regulation is that governments set maximum permissible discharge levels or minimum levels of acceptable environmental quality. Such 'command and control' mechanisms include best available technology standards, clean water and air acts, Superfunds for clean-up and liability determination, and a range of site-specific permitting procedures which tend to be the responsibility of local government within nationally approved regulatory regimes. Such 'command and control' mechanisms tend to rely on administrative agencies and judicial systems for enforcement. Three issues are relevant regarding the appropriateness of industrialised country environmental regulations to reduce environmental degradation and improve environmental management practices in metals production.

First, there is a trend away from a 'pollutee suffers' to 'polluter pays' principle. However, it remains the case that the polluter pays only if discovered and prosecuted, which requires technical skills and a sophisticated judicial system, and that occurs only after the pollution problem has become apparent and has caused potentially irreversible damage. This highlights the tendency of such environmental regulations to deal with the symptoms of environmental mismanagement (pollution) rather than its causes (economic constraints, technical constraints, lack of access to technology or information about better environmental management practices). This can be serious in some instances because once certain types of pollution have been identified, such as acid mine drainage, it is extremely costly and sometimes technically impossible to trace the cause, rectify the problem and prevent its recurrence. Certain environmental controls may only work if incorporated into a project from the outset (e.g. buffer zones to protect against leaks under multi-tonnage leach pads and tailings ponds).

Second, best available technology standards (BAT) may be appropriate at plant start-up, but their specified effluent and emission levels are not necessarily achievable throughout the life of the plant, because technical problems may arise and there may be variations in the quality of concentrate or smelter feed, etc. if supply sources are changed. Moreover, there are serious implications for monitoring. It would also be erroneous for a regulatory authority to assume standards are being met if a preselected item of technology has been installed. Ongoing management and the environmental practices at the plant are also likely to be important determinants of 'best environmental practices'.

Third, related to points one and two above, BAT standards and environmental regulations of the command and control type tend to presume a static technology—a best technology at any one time. This tends to promote incremental add-on controls to respond to evolving regulation rather than to stimulate innovation. This acts as a disincentive to innovate by equipment suppliers, the mining companies and metal producers. Their innovation, which has required substantial R&D resources, may be superseded by some regulatory authority's decision about what constitutes BAT for their particular activity. BAT gives the impression of technology being imposed from outside the firm, not generated from within. The search for profit and cost-savings tends to be a more obvious instigating factor of technical change, and it might be argued that market-based mechanisms, a technology policy that is complemented by a regulatory framework and a good corporate environmental management strategy can better contribute to achieve that aim.

There has been growing interest in the use of market-based mechanisms, whereby the polluter is charged for destructive use by estimating the damage caused. An important justification for the use of market-based incentives is that they allow companies greater freedom to choose how best to attain a given environmental standard (OECD, 1991). By remedying market failures or creating new markets (rather than by substituting government regulations for imperfectly functioning markets), it has been argued that market-based incentives may permit more economically efficient solutions to environmental problems. Two categories of incentives exist (O'Connor, 1991). One set, based on prices, includes a variety of pollution taxes, emission charges, product charges and deposit-refund systems. Another set is quantity-based, and includes tradable pollution rights or marketable pollution permits. The most common of these measures relates to posting bonds up-front for the rehabilitation of mines on closure. There are also discussions taking place about a mercury tax in Brazil and a cyanide tax in the USA. This is standard practice now in Canada and Malaysia. Currently no government has designed a systematic set of incentives for industry to innovate and develop new environmental technology.

There are two further areas where policy approaches can contribute to improved environmental management practices. The first is that of increasing conditionality of private, bilateral and multilateral credit, which frequently requires both prior environmental impact assessment and the use of best-practice environmental control technologies in new mineral projects. A growing number of donor agencies, in Germany, Canada, Finland and Japan, for example, are also concerned with training in environmental management. The second is in the attempts by some governments, particularly Canada, to promote R&D activities (jointly and within industry and academic institutions) to determine toxicity from mining pollution and clean-up solutions. For example, Canada has extensive government-funded R&D programmes to promote the abatement of acid mine drainage and of SO_2 emissions. There is considerable scope for expanding these approaches, as is argued in the next section.

Environmental regulations designed specifically for mining and mineral processing have, until recently, been uncommon in developing countries, although most countries now have in place basic standards for water quality and, less commonly, air quality. A few developing countries have recently adopted extensive regulatory frameworks—sometimes replicas of US models. This, for example, has been the case in Chile and, to a lesser extent, in Brazil. This growing concern about environmental degradation is occurring during a period of rapid liberalisation in developing countries, which finds expression in new policies to promote foreign investment, privatisation schemes and the availability of loan capital. These conditions also influence the regulatory regime of developing countries. Should the developing country pose less onerous environmental burdens on the potential investor to improve the terms of the investment by implying lower compliance costs or a greater assumption by the state of the environmental costs associated with mineral development projects? Should agreements be signed that release new investors from any liability for environmental damage caused by previous mine owners under less-restrictive regulatory regimes? Or will a clear and strict regulatory regime be more likely to facilitate credit flows from increasingly more environmentally conscious lending agencies? Developing countries, desperate for investment in their stricken mineral sectors, will need to determine what the market can stand and how such terms can be structured to reduce to the minimum the risk premium the investor will seek for a given tax or regulatory burden.

It is worth noting that surveys by Johnson and Eggert imply that environmental policy has not been a major factor in determining the investment strategies of international mining companies (Johnson, 1990). However, more recently the industry press has been citing environmental regulations in Canada and Australia as a major factor causing the cancellation and delay of potentially large investment projects and contributing to the shut-down of several mines. For example, in 1989, the Bharat Aluminium Company announced the closure of its bauxite mining project in the Gandhamardham Hills, Orissa State in India, because of strong environmental opposition by the local population. Other projects such as the Phelps Dodge Copper Basin in Yavapai County (USA) have been withdrawn due to delays and excessive costs involved in project approval, while in 1991 the Kennecott Flambeau Mining Company finally received planning permits after 20 years of negotiation for the Grant Copper Mine in Wisconsin, which will operate for only six years.

However, environmental regulation alone is unlikely to solve environmental problems in developing countries due to endemic production inefficiencies. In particular, the approach of state-owned enterprises towards the environment reflects inefficient operating regimes, excess capacity, breakdowns and shut-downs, and poor management procedures, which contribute to worsen the polluting nature of effluents and emissions. Such inefficiencies make it very unlikely that environmental controls will be incorporated effectively.

Production inefficiency is endemic among many mining enterprises in developing countries, and problems of environmental degradation cannot be viewed independently of it. Moreover, obsolete technology is widely used without the modern necessary environmental controls and safeguards. For example, new concentrators and roasting plants tend to be totally computerised. Automatic ore assaying techniques give an extremely accurate picture of the chemical composition of the ore feed which has implications for the fine-tuning of pressure, heat, cooling and specific environmental control systems. This, in turn, will facilitate the accurate prediction and monitoring of emissions. However, where these controls are missing and, in particular, where ore feeds are of variable composition (in terms of the sulphur, lead and arsenic content), emissions also vary with regard to their pollutant content. The inefficient use of energy and poor energy conservation practices also result indirectly in increases of environmental pollution through the excessive burning of fossil fuels. This is particularly the case in poorly lagged roasters and inefficiently operated flotation units and smelters which are very intensive in energy use. It might be further argued that command and control regulatory instruments are unlikely to result in a reduction of pollution since they cannot affect the capacity to implement technical change of a debt-ridden, obsolete and stricken mining enterprise in the developing country context. Such a company might find it preferable to risk not being detected or convicted, to pay a fine or to mask its emission levels, rather than face bankruptcy through investing in radical technical change.

In addition to the problems of inefficient production, there are a range of further reasons why environmental regulations—particularly those of the 'command and control' and incremental 'paper tiger' nature—do not improve environmental management, particularly in developing countries (Panayotou *et al.*, 1990). These are discussed below.

Environmental regulations tend to be of the blanket-type which specify maximum levels of emitted substances, minimum levels of environmental quality and best available technology standards. They do not tend to reflect the propensity of a particular operation to pollute, which in part depends on local site-specific conditions (geology, geography and climate) as well as economic, infrastructure and technology-related constraints. In a desert, tailings dams need not be as highly specified as in rainy climates; dust regulations may need to vary depending on topography, precipitation and prevalent winds; the sub-strata of leach ponds might need to be of different composition, strength and depth, depending on local geology or the existence of an impermeable level of clay. Since developing country regulations are often copied directly from the statute books of the industrialised countries (for example, there are instances in the cases of both Chile and Brazil) whose regulations are adapted to suit their circumstances, they may not be appropriate for the site-specific characteristics of mines in either tropical regions or deserts. They may result in unnecessary and costly adaptations on the one hand, or the lack of necessary control on the other.

Command and control environmental regulations require intensive monitoring to ensure that they are enforced. However, the small and medium mine sector accounts for at least 25 per cent of mineral production in many countries. Although these mines are individually relatively small polluters, collectively they account for a disproportionately large share. These mines are often located high in the Andes or in remote tropical rain forests and are almost impossible to monitor systematically. Indeed, as regulation becomes more sophisticated, such monitoring requires skills and human resources far beyond the technological and managerial capabilities of many developing countries and frequently beyond their budgets. Understanding the diverse range of toxicity and engineering issues behind regulatory aims also poses challenges even in the industrialised countries. The most knowledgeable regulators are often head-hunted by the mining companies.

In a recent interview in Brazil, a spokesperson for one of the companies said that they were requested to monitor themselves and send effluent samples at intervals to an independent laboratory and to report any abnormal results since the state regulatory agency did not have the skills required to monitor the operation itself. Indeed, the skilled people involved in the environmental agencies tend to live and work in capital cities and infrequently travel into the politically dangerous and inhospitable mining regions. Moreover, the enforcement of command and control regulations depends on a system that admonishes with imprisonment and fines. This, in turn, requires a legal structure and judicial system far beyond the capacity of most developing countries. Compliance is also limited since fines are generally a fraction of the costs involved in remedial treatment and abatement technology. They are also only payable if the polluter is detected, and if convicted. Inflation and local currency devaluation which are endemic in the developing country context also eat into the value of such fines. The costs of environmental regulation enforcement are generally hidden from the public eye and regulatory agencies are not generally accountable as such. However, since different site-specific mining contexts often require individual regulation, perhaps for permit approval, this provides opportunities for bribery which is endemic in bureaucracies and industry in many developing countries.

Indeed, the regulatory system does not demand that efforts are made to deal with the cause of environmental pollution once and for all. It simply deals with the symptoms—once they are reported. Even though there is a theoretical threat of mine closures due to non-compliance, most foreign mining companies know that their developing country host can least afford to lose the foreign exchange earnings from their activities. Therefore the risk of closure due to environmental non-compliance of this type is considered relatively low. Pollution rarely produces a one-off disaster—rather it is a constant crisis.

Environmental regulations are often contradicted by other economic and industrial policies. For example, several countries with tropical forests have recently introduced policies aimed at their conservation. At the same time countries such as Brazil, Ecuador and Colombia have parallel economic policies to promote industrial investment, especially by foreign firms, in these remote areas. The example of the government of Ecuador authorising RTZ's mining investment in one of its national parks is one such case—it resulted in the latter company withdrawing to avoid controversy over the issue. Similarly, in Brazil forest conservation policies were in place, in part conditional upon EC and World Bank loans, at the Carajas smelters, which were fuelled by large amounts of charcoal from the neighbouring forests.

Another recently discussed example of potentially contradictory policy issues revolves around the international Basel convention, which restricts the intercountry transportation of toxic wastes. Since certain scrap metals fit theoretically into this category on account of their heavy metal content, this would undoubtedly restrict trade in scrap metal and metal recycling (OECD, 1990). This is considered to contradict many of the new intentions of European and American governments to encourage the recycling of metal-containing materials at the expense of new primary production.

Command and control regulation tends to identify and deal with symptoms (pollution) of environmental mismanagement rather than causes (production inefficiency, human resource constraints, lack of technology and capital). It is also add-on and incremental in nature. Therefore, there is a tendency for it to emphasise end-of-pipe, add-on and capital-intensive solutions (e.g. smelter scrubbers, mine water treatment plants, dust precipitators, etc.) for existing technology and work practices rather than promote alternative environmental management systems or technological innovation. Regulation may also, to a certain extent, presuppose a static technology. If regulation is incremental, technical change may be incremental, involving the addition of numerous new controls at relatively greater cost and with more overall resultant degradation than if a new, more radical change had been introduced in the first place. It may also oblige specific reductions in pollution, regardless of cost or local context. For example, regulation will refer to the chemical composition of an effluent in isolation from how that discharge rate and pattern may be influenced by natural site-specific precipitation, evaporation or soil and geological conditions. In turn, this regulatory approach may get a more unco-operative response from industry that sees the rules always changing and their cost implications increasing. Furthermore, such regulation ignores the human resource elements of sound environmental management by emphasising a specific pollution control technology rather than training, managerial approaches and information diffusion.

TECHNICAL CHANGE AND CORPORATE ENVIRONMENTAL TRAJECTORIES

Enterprise responses to environmental pressures have been characteristically slow, and reflect the regulatory regimes and public climate of either their home country or foreign countries of operation. Their response has also depended on the nature of their operations in terms of first, the mineral involved; second, the level of integration of mining and processing activities; third, the stage in the invest ment and operations cycle which its mineral projects have reached; and fourth, the internal economic and technological dynamism of the company, that is whether it has the financial, technical and managerial capabilities to be an innovator or not.

After a period of using rather 'static' technology, the mining and mineral processing industry is going through a phase of technical change as dynamic companies are innovating by developing new smelting and teaching technologies to escape economic as well as environmental constraints. Rapidly evolving environmental regulatory frameworks in the industrialised countries and the prospects of their application, reinforced by credit conditionality in the developing countries, are stimulating this trend. Changed technological and environmental behaviour in this context is evident particularly in the large North American and Australian mining companies, but is becoming increasingly apparent in developing country-based companies operating in, for example, Chile, Brazil and Ghana. However, it seems to be the new operators and dynamic private companies that are changing their environmental behaviour, while both state-owned enterprises and small-scale mining groups in developing countries continue, with some exceptions, to face constraints regarding their capacity to change environmentally damaging practices.

It is somewhat inevitable that only those companies that are dynamic and have new project development plans are in a position to invest in the R&D required to develop more environmentally sound alternatives, or to raise the capital to acquire them from technology suppliers. None the less, after a long period of only conservative and incremental technical change, alternative process routes for mineral production are being developed which emerge as being more economically efficient as well as environmentally less hazardous. Furthermore, companies are beginning to sell their technologies, preferring to commercialise their innovations to recoup their R&D costs rather than sell obsolete technology and risk shareholders' scorn or retrospective penalties as environmental regulations are increasingly enforced by the developing countries. Some of those companies have pushed technology even beyond the bounds of existing regulations and as a consequence are seeking to increase regulation—particularly on a world-wide scale— because they can meet the requirements and use their new environmentally sound technologies to their competitive advantage.

There is evidence that improving the environmental management of a mining operation may not necessarily be detrimental to economic performance, and in some cases may even be of economic benefit. Furthermore, because environmental regulation is here to stay and bound to become more widely adopted, more stringent and better enforced, who wins in the division of shares in the metals market will not be those companies that avoid environmental control (only later to be forced to internalise the high cost of having done so), but those companies that were ahead of the game, those that played a role in changing the industry's production 'parameters', and those that used their innovative capabilities to their competitive advantage.

The 'environmental trajectories' that different mining enterprises might take in response to environmental and market pressures are categorised in Figure 7.1. This diagram could be a planning tool for both companies and governments. It can help to evaluate the environmental and economic implications of applying different policies on corporate development.

The average mineral enterprise is competitive (i.e. to the left of the threshold of economic competitiveness, X), although to a greater or lesser extent these enterprises produce environmental pollution and to a greater or lesser extent they have internalised the cost of the environmental degradation associated with their metal production, in response to the regulatory regime they are working within. (The threshold of 'environmental competitiveness' for a given regulatory context is also X and company operations in compliance have environmental trajectories in the quadrants below the horizontal axis.) However, market pressures—mainly a real decline in metal prices—combined with their economic inefficiencies, mean some of these companies are going bankrupt (a trajectory towards quadrant B). They will leave a legacy of environmental pollution behind, and as in the cases of COMIBOL (in Bolivia) and Carnon (in the UK), the burden of clean-up will fall on the state and society. Other companies will respond by innovating; moving into quadrant D; building into the new generation of technology both improved economic and environmental efficiencies (protecting themselves in the process from having to undertake relatively more costly add-on, incremental technical change and rehabilitation at later stages in their operation). Indeed, freed from the incumbent costs of retrofitting sunken investments, greenfield plants in particular display new levels of dynamism—the latest best practice technology incorporates both improved economic and environmental efficiencies.

None the less, there exists a growing group of companies which, if obliged to 'add-on' environmental controls in line with new regulations, would have to close down since the cost of the controls and clean-up required would render their operations uneconomic. The environmental trajectories of these companies is towards quadrant C. Currently, such examples are few, and it is difficult to differentiate between purely environmental factors and the range of other reasons as to why a company's cost curve starts increasing. However, as Figure 7.1 shows, that group would be expected to grow in number, since market and regulatory pressures combined will lower the threshold of economic and environmental competitiveness such that the average company will only survive in the new regime if it innovates. Even the previously dynamic companies will need to keep their environmental trajectories moving ahead of the encroaching threshold of economic/environmental competitiveness (X1 and X2).

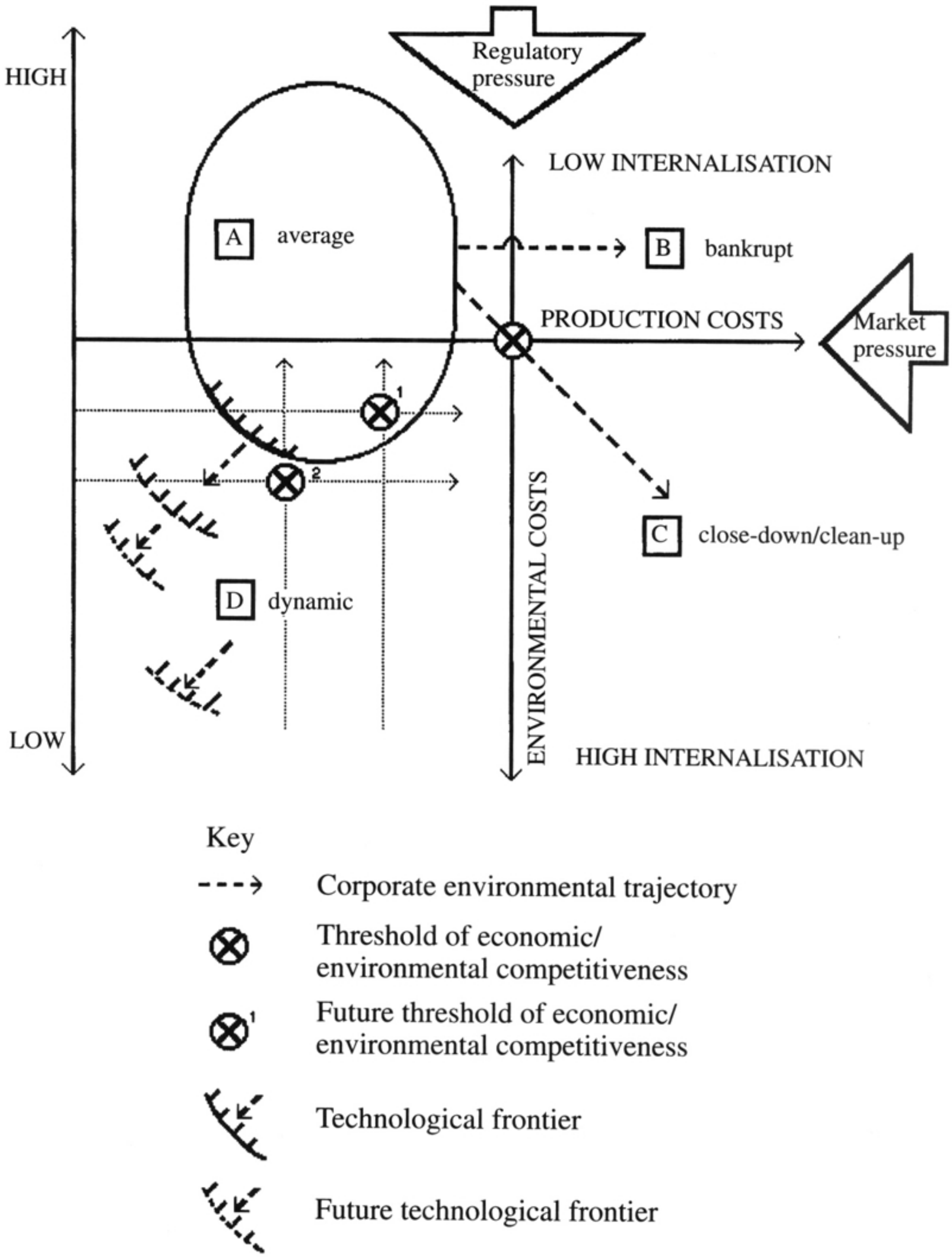


Figure 7.1 Corporate environmental trajectories

Moreover, this implies a serious constraint on the regulatory process for two main reasons, which indeed distinguish mining companies from their manufacturing counterparts. First, an implied closedown due to regulatory burden does not signal the end of environmental degradation. Pollution in metals production is not all end-of-pipe. Rather it heralds a new era—decommissioning, clean-up and rehabilitation all pose significant environmental costs. Second, in very few countries are bankrupt operators liable for the clean-up of their 'sins of the past'. The USA with its 'Superfund' liability laws is an exception. Therefore, by moving the threshold of economic and environmental competitiveness, the overall extent of environmental degradation (particularly that without liability) increases. The policy challenge of the environmental imperative is therefore how to keep firms sufficiently dynamic to be able to afford to clean up their pollution and generate economic wealth.

Cleaning up 'sins of the past' and some environmental effects of current regulation

It is inevitable that those companies with long mining histories and extensive sunken investments in conventional mining and smelting facilities face the greatest technical, and therefore economic, challenges in cleaning up their past facilities and reducing pollution from their ongoing operations. For example, some companies in central and south-west USA have found that dumps from past lead and copper mining operations have now created such serious acid mine drainage and toxic seepage that the government has placed them on its 'Superfund List' which obliges multi-million dollar sums being spent on their clean-up. The government then targets previous owners of the mine, often the richest, making them liable. If a company has already closed down a mine and written off the investment and perhaps is struggling in the current economic climate to manage a new project, it is clear that the costs of such a Superfund indictment, and the legal costs involved in answering it, can be quite crippling. For example, the Smuggler Mountain lead mining site in Colorado has a serious acid mine drainage and toxic seepage problem (EPA, 1989). Its old lead and cadmium mine workings have apparently contaminated soils and groundwater in neighbouring residential areas, requiring a major clean-up operation, the secure repositioning of the toxic waste and the establishment of monitoring mechanisms and pollution controls to prohibit further contamination. The cost of this project as estimated by EPA is currently US$4.2 million. Another Superfund listed mine site is the Silver Bow Creek Site in Butte Area, Montana, which for over one hundred years has been mined for silver, copper, gold and zinc, resulting in severe water and soil contamination and the disruption of local ground and surface drainage water patterns. Currently groundwater is flooding the mine, becoming highly acidified in the process, and it is absorbing high concentrations of iron, manganese, arsenic, lead, cadmium, copper, zinc and sulphate. This toxic seepage is currently threatening Silver Bow Creek, a major river in the region. The clean-up and remedial action is extensive and involves detailed diagnostic analysis and monitoring, water and tailings containment, water treatment and soil treatment. Leading mining companies and individuals, all past owners, are implicated including Atlantic Richfield, AR Montana Corporation and ASARCO. The clean-up cost is estimated to be in the region of several million, to be confirmed once the precise plan of remedial action is determined. Another Superfund-listed old mining site is that of Gregory Tailings in Colorado (EPA, 1986). It was a gold mine, exploited during the late nineteenth century. Waste was placed in inadequate tailings dams and resultant leakages have contaminated local water supplies and soils with acidic waters containing copper, zinc, nickel, cadmium and arsenic. The cost of the clean-up and water treatment has yet to be confirmed, but the strengthening of the tailings dam is estimated at over half a million dollars.

Although there may be economic opportunities associated with such clean-up operations, such as the recovery of extra metal values from acid mine drainage, the commercialisation of innovative water treatment methods or the innovative use of tailings material, these may not be recouped by the mining company itself. Furthermore, companies that previously have had no links to the facility may be nervous of getting involved in case any liability is passed on to them.

This is one reason why some mining companies, which need to clean up their past operations, object to such retrospective regulation and suggest that such restrictions and control threaten their existence. An interesting and illustrative case is that concerning the respective responsibility of both government and industry for the management of mine closure and rehabilitation of old tin, copper and silver mining and smelting operations in Cornwall, UK. The observations that mine closure and rehabilitation were proceeding very inefficiently in the absence of an adequate regulatory framework have been borne out by the recent flooding of polluted acid mine water from the Wheal Jane mine near Truro, Cornwall (Warhurst, 1992). This highly acidic cocktail of dissolved metals, including copper, lead, cadmium, tin and arsenic, entered the Carnon and Fal rivers at a rate of 2–4: million gallons a day and has spread throughout the surrounding estuary and coastal areas. The tourist industry and fisheries have been threatened and local well-water supplies destroyed.

Many have argued that legal responsibility for the mine discharge rests with the current owners, Carnon Consolidated (a management buy-out in 1990 from RTZ, which had bought it seven years previously from Consolidated Goldfields). However, the National Rivers Authority (NRA), the relevant UK regulatory body, publicly admitted that it was aware six months previously that this disaster could happen, but was unable to prevent it since its policy remit did not cover preventative action. The disaster occurred following the withdrawal of a government grant to the Wheal Jane mine which has always required pumping to prevent it from flooding. This meant that plans had to be shelved for turning the site into a golf course and leisure centre to help fund the pumping and maintenance costs at the mine of £100,000 per month. The consequent lack of finance forced Carnon Consolidated to turn off all the pumps on 4 January 1992. Since they had officially abandoned the mine, they denied responsibility for any ensuing flooding and pollution, a claim complicated by the fact that numerous underground shafts from other abandoned mines in the region also provided conduits for the polluted floodwater. The NRA has stated its intention to prosecute Carnon, but loopholes in the Water Resources Act 1991 mean there is no UK regulation to ensure previous mine owners bear the financial liability for clean-up measures (unlike the Superfund legislation in the USA). Abandoned mines are specifically exempted from clean-up liability. Furthermore, the NRA itself has no remit or budget to treat pollution, particularly on this scale.

The NRA has opened an old dam to hold and treat the contaminated water. It is also analysing the potential of biotechnology to assist in the clean-up. For example, certain microbes bred in a slurry of cattle excrement have been shown to be metal-absorbing. Similarly, the creation of wetlands containing plants whose roots absorb metals is another possible long-term solution. However, such solutions are based on piecemeal research being undertaken in this area in different research institutions (e.g. the Colorado School of Mines in the USA, CANMET in Canada, CETEM in Brazil) and none of these techniques has yet been proven commercially. The NRA is approaching the UK government to help pay for the clean-up, which may take decades and could cost over £1 million. The ultimate financial responsibility for mine clean-up in the UK may therefore lie with the taxpayer.

Environmentalists in the UK are pressing for the law to be changed so that companies that abandon mines are liable for any resulting pollution. However, it has been suggested that such a change would make life difficult for British Coal. The NRA has already analysed pollution from rising waters in old coal mine workings which are polluting rivers in South Wales and Yorkshire. They have established that the capital costs of treating the ten worst cases in Yorkshire are estimated at more than £10 million, and indeed as long ago as 1981 the Royal Commission on Coal and the Environment (the Flowers Report) recommended that the costs of remedial action for existing mines abandoned by the National Coal Board are met by central government. However, the tightening of UK laws to ensure that the bills for such pollution are paid by previous owners could have serious implications for the current government's plans to privatise the coal industry.

Similar pollution liability issues are also being faced by governments in developing countries currently engaged in privatising their state mines—Bolivia, Peru and Chile are cases in point. In Peru, for example, the legacy of past pollution from toxic tailings along the river below CENTROMIN's La Oroya and Cerro Pasco facilities was preventing the government from selling those enterprises since the cost of clean-up rendered the investments uneconomic and unattractive to foreign capital. It was therefore agreed that the investment contract for buying these operations would protect the foreign partner from liability for previous environmental damage. They would start with a 'clean slate', as it were, with generous lag-times regarding the introduction of new environmental controls to reduce ongoing pollution. This means that the economic burden for clean-up falls on the state but, in developing countries, that means on society. Where capital is scarce, cleaning up pollution problems from past decades affecting remote rural communities is of low priority.

This poses a policy dilemma. If the government does not waive liability for past pollution, the privatisation schemes, vital to the future of the economy, will not succeed since foreign partners would not be interested, in view of other available investment opportunities in mining. Moreover, international companies are particularly wary of falling prey to new retrospective liability laws and punitive tariffs in their import markets, particularly given the enormity of the clean-up involved. The responsibility for clean-up therefore lies formally with the state. Should loan-conditionality put pressure on the government to clean up using precious capital resources? Should clean-up funds be established and incentives provided to prompt local industry to develop technical solutions? Should aid programmes provide technical assistance and training in clean-up? Should new investors be taxed for old pollution? Clearly the optimal outcome would be for government and donors combined to provide incentive programmes for local firms to seize the commercial opportunities available. Nuñez clearly documents the extensive range of local capabilities which could be harnessed (Nuñez, 1992).

While 'Superfund' in the USA may theoretically be successful, given that one is targeting local investors and traceable companies, it may be more difficult to target and litigate against previous mine owners (prior to nationalisation) in developing countries. Searching out the foreign investors responsible for the many old and abandoned mines may be difficult since most have long since returned home, and local miners have limited resources, which makes the task of determining liability and enforcing clean-up a daunting one.

Companies that are being forced through environmental pressures to deal with pollution problems in their existing operations have been observed to react both defensively and in innovative fashion, depending on the challenges posed, their economic well-being and internal dynamism. For example, depending on the level of enforcement of the regulatory regime, some mining companies, particularly those operating in the developing country context, may prefer to pay financial penalties and fines for affected water and air quality. These may amount to less than the cost and effort involved in remedial action, such as water treatment, and considerably less than the costs involved in innovation or the incorporation of pollution controls. In some instances, as discussed above, the state pays those remedial costs itself and it is clear that it may be subsidising the profit of foreign mining companies at the expense of environmental degradation. Sometimes that trade-off is influenced by the state's absolute dependence on the foreign company as a source of foreign exchange and government revenue.

Indeed, a number of mining companies perceive environmental regulation as imposing a cost burden on their operations that threatens their profitability. They may then enter into negotiations with the state to arrive at a 'stay of execution' or to devise a plan for implementing controls. However, as regulation becomes increasingly strict, backed up by more sophisticated monitoring devices and data processing, companies are being pushed to take remedial action in both the industrialised and developing countries. Data from the USA suggest trends that may be followed elsewhere. According to a US Congressional Report, sulphur dioxide emission controls have resulted in 'substantial capital expenditure' for US copper smelters and increased operating costs due to 'add-on' acid plants (Coppel, 1992). Present levels of environmental control entail capital and operating costs of between 10 and 15 cents per pound of copper. However, the USA has lost substantial smelting capacity. It has been reported that eight out of sixteen smelters operating in the USA in the late 1970s have closed permanently, ‘most because the capital investment to meet regulations was unwarranted given current and anticipated market conditions’ (Office of Technology Assessment, 1988).

Indeed, the evidence from studies in the USA shows that environmental compliance does not distort significantly the economics of new mineral projects, but does place a considerable cost burden on ongoing facilities for either retrofitting or clean-up on mine and plant closure. The US Bureau of Mines has estimated direct ‘environmental’ operating costs for smelting facilities with emission controls. Retrofit capital costs were estimated to be of the order of US $150 million per facility or 5.6 cents per pound of copper produced (US Bureau of Mines, 1990). According to Coppel the overall cost penalty, including capital invested, to the producer for implementing the new smelter and sulphur dioxide capture facilities was estimated to be 7.5 cents per pound after deductions of a 1.3 cents per pound of acid credit. The operating costs for individual smelters ranged from 10–15 cents per pound of copper, and the average operating cost in 1987 was 12.3 cents per pound. Of this amount, 26 per cent or 3.2 cents was calculated by the US Bureau of Mines to be the cost burden of compliance with environmental, health and safety regulations.

It is highly likely that regulation in developing countries will follow a similar pattern. An interesting example is the ALCAN bauxite mine and alumina plant in Jamaica. Foreseeing impending environmental regulation and responding to public concerns in its home country of Canada, ALCAN supported a local university department to develop an innovative solution to the disposal of the red mud sludge from its bauxite mining operations. Previously, the sludge was dumped in a large catchment pond, but toxic seepages into surrounding soils and groundwater had been reported. The university developed a process called red-mud stacking which involved sun-drying of much of the moisture content of the sludge and stacking of the material into much less obtrusive piles, which indeed as bricks may have further use within the plant site (Kelly, 1990). A similar technology was introduced by ALCAN at the Vaudrevil alumina plant in Quebec in 1989 (US Bureau of Mines, 1989).

However, this technology neither seems to offer a solution to the toxic seepage of previously dumped slurries nor is it a solution to pollution *per se*. It appears to replace water pollution by dust pollution which is less stringently regulated against. Moreover, a change in the production process to facilitate the recovery of caustic soda from the ‘mined’ dry mud stacks means that a greater amount of that chemical is discharged than in the previous disposal method. This means that the dust pollution, plus overflows from those parts of the dry mud stacks that become waterlogged during tropical rain showers, may cause a greater toxic hazard than previous low-level seepages. Meanwhile, ALCAN Jamaica’s almost obsolete technology with its recurrent production problems is pushing its environmental trajectory closer to the encroaching threshold of economic and environmental competitiveness.

Dynamic innovators—technical change to improve environmental management

Although some mining companies resist the application of environmental regulation to their existing operations, a growing number of dynamic innovative companies are making new investments in environmental management because they see an evolution towards stricter environmental regulation. Free of the encumbrance of sunken investments in pollutant-producing obsolete technology or with significant resources for R&D and technology acquisition, they have chosen either to develop more environmentally sound alternatives or to select new improved technologies from mining equipment suppliers, who are themselves busy innovating. Increasingly, these new investment projects are incorporating both improved economic and environmental efficiencies into their new production processes, not just in terms of new plant or items of technology, but also through the use of improved environmental management practices. Some examples of these are discussed below in three categories: smelter emissions, gold extraction and waste management.

Smelter emissions

BOX 7.1 SMELTER EMISSIONS: INCO LTD

At one time one of the world's highest cost nickel producers, Inco was until recently the greatest single point source of environmental pollution in North America. This was due to its aged and inefficient reverberatory furnace smelter technology which spewed out excessive tonnages of SO_2 emissions. Inco Ltd had reached the limit of improving the efficiency of this obsolete technology through incremental technical change at the same time as the Ontario Ministry of the Environment began an intensive SO_2 abatement programme to control acid rain. These factors prompted Inco to invest over C$3,000 million in a massive R&D and technological innovation programme (Aitken, 1990).

Under the Canadian acid rain control programme, Inco is required to reduce SO_2 emissions from its Sudbury smelter complex from its current level of 685,000 to 265,000 tonnes per year by 1994: a 60 per cent reduction. To achieve this, Inco plans to spend C$69 million to modernise milling and concentrating operations and C$425 million for smelter SO_2 abatement. The modernisation process will include replacement of its reverberatory furnaces with a new innovative oxygen flash smelter, a new sulphuric acid recovery plant and an additional oxygen plant. By incorporating two of these flash smelters the company plans to reduce emissions by over 100,000 tonnes per year in 1992 and by 1994 to achieve the government target levels of 175,000 tonnes per year. Other environmental benefits include a cleaner, safer work environment.

Inco is now one of the world's lowest cost nickel producers and again, like other dynamic companies that are responding to environmental regulation through innovation, Inco is seeking to recoup R&D costs through an aggressive licensing effort in other copper and nickel processing countries. More than 12 per cent of Inco's capital spending during the last ten years has been for environmental concerns.

BOX 7.2 SMELTER EMISSIONS: KENNECOTT, UTAH, USA

A new smelter project has recently been launched by Kennecott Corporation (RTZ) with the dual aims of setting a new standard for the cleanest smelter world-wide and improved cost efficiencies in processing its ore. Advantages include the capture of 99.9 per cent of sulphur off-gases (previous levels were 93 per cent). Sulphur dioxide emissions will be reduced to a new world best practice level of approximately 200 pounds per hour, less than one-twentieth of the 4,600 pounds per hour permissible level for Utah's current clean-air plan. The investment is US$880 million, resulting in 3,300 new construction jobs and the investment of US$ 480 million in local companies through project development contracts.

The proposed Garfield smelter will expand the concentrate processing capability to the level of mine output (about one million tonnes of copper concentrate per annum) at more than half previous operating costs. It represents the first-time application of oxygen flash technology to the conversion of copper matte to blister (details are based on an excerpt from a Kennecott Corporation press release, 11 March 1992). The two-step copper smelting process consists of smelting furnaces which separate the copper from iron and other impurities in a molten bath, followed by converting furnaces where sulphur is removed from the molten copper. A new technology known as flash converting will then be utilised in the second step of the process at the new smelter. This unique technology was developed by Kennecott in co-operation with Outokumpu Oy, a Finnish company and a leader in the supply of smelting technology. Essentially the new technology eliminates the open-air transfer of molten metal and substitutes a totally enclosed process for producing molten metal. 'Flash converting' has two basic effects: first, it allows for a larger capture of gases than the current open-air process; second, it allows the smelter's primary pollution control device – the acid plant – to operate more efficiently. The smelter will include double-contact acid plant technology.

There will be other environmental benefits from the new smelter as well. Water usage will be reduced by a factor of four through an extensive recycling plan. Pollution prevention, workplace safety and hygiene and waste minimisation will be incorporated into all aspects of the design. In addition, the smelter will generate 85 per cent of its own electrical energy by recovering energy as steam from the furnace gases and emission control equipment. This eliminates the need to burn additional fossil fuel to provide power. The new facility will require only 25 per cent of the electrical power and natural gas now used per tonne of copper produced.

→

The copper refinery's planned modernisation and expansion modifications include major electrical system changes, material handling system improvements and new electro-refining cells. In addition, a new state-of-the-art precious metals refinery will be built. The refinery will be able to process the entire output of the new smelter.

Gold extraction

BOX 7.3 GOLD EXTRACTION: HOMESTAKE'S MCLAUGHLIN GOLD MINE, CALIFORNIA, USA

The McLaughlin Gold Mine, opened in 1988, is perhaps the best example of a new mine and processing facility that has been designed, constructed and operated from the outset within the bounds of probably the world's strictest environmental regime. Environmental efficiency is built into every aspect of the gold mining process, in terms of innovative process design criteria, fail-safe tailings and waste disposal systems and extensive ongoing mine rehabilitation and environmental monitoring systems. The mining operation therefore combines innovative technologies with 'best practices' in environmental management. The most interesting conclusions drawn by the author from site visits and discussions with the firm's environmental officers is that most of these environmental management initiatives have not resulted in any substantial extra cost, and indeed many of these procedures have apparently improved the efficiency of the mine, affecting positively the economics of the overall operation.

For example, before the mining operation began, an extensive environmental impact analysis and survey were undertaken. All plant and animal species were identified and relocated, ready for rehabilitation on the completion of mining operations. The survey also measured in detail prior air, soil and water quality characteristics and flow patterns to provide the baseline for future monitoring programmes. Assaying was undertaken not just of the gold ore, but also of the different types of gangue material and waste, so that waste of different chemical compositions could be mined selectively and dumped in specific combinations to reduce its acid mine drainage generating capacity. Local climate conditions were evaluated to determine the frequency of water spraying needed to reduce dust, and evaporation rates were evaluated to control the water content and flood risk potential of tailings ponds. The tailings ponds themselves are constructed on specially layered impermeable natural and artificial filters, with high banking to prevent overflow and with secondary impermeable collecting ponds in the rare case of flooding.

Unlike other mining projects where rehabilitation is seen as a costly task to be undertaken at the end of a mining operation, often when cash flows are lowest as ore grades decline, at Homestake rehabilitation began immediately and is an ongoing activity. Not only does this serve to spread expenditure more evenly over the life of the mine, but it enables the more efficient utilisation of truck and earth moving capacity as well as of relevant construction personnel. This means that as soon as work piles have reached a certain pre-determined dimension, soils (previously stripped from the mine area and stored) are laid down and revegetation is begun. Although mining has been underway for only three years, extensive areas of overburden and waste have already been successfully revegetated – immediately reducing environmental degradation and negative visual impacts. In addition to these inbuilt environmental control mechanisms, Homestake Mining Company has sophisticated environmental monitoring procedures in place. This means that seepages, emission irregularities, wildlife and vegetation effects can be detected and rectified immediately, which in the long term reduces the risk of expensive shut-downs in operations, costly court cases (e.g. if water toxicity results) and the need for treatment technologies.

Waste treatment

In the minerals industry, considerable waste is produced in the form of overburden, marginal ore dumps, tailings and slags. Much of the toxicity associated with that waste is a direct result of the loss of either expensive chemical reagents or metal values. Currently, public policy has not taken up the challenge to promote and direct R&D in the area of waste reduction and treatment innovations. One interesting area is the application of biotechnology to waste treatment.

BOX 7.4 WATER TREATMENT: HOMESTAKE'S MINE AT LEAD, SOUTH DAKOTA, USA

Homestake Gold Mining Company turned regulatory pressure to clean up a cyanide seepage problem to its advantage. Its own R&D staff developed a proprietary biological technique to treat the effluent, which led to the recovery of local fisheries and water quality in the mine's vicinity at Lead, North Dakota, USA (Crouch, 1990). It is now actively commercialising the technology which can be widely applied at other gold leaching plants.

BOX 7.5 WATER TREATMENT: EXXON'S MINE, LOS BRONCES, CHILE

A mining project in Chile, Los Bronces, was to be expanded into one of the largest open-pit copper mines in the world and consequently required the stripping of very large tonnages of overburden and low grade ore. Before mine development, the Chilean government warned Exxon that it would be imposing financial penalties for the water treatment costs on account of expected acid mine drainage from the overburden of low grade ore dumps into the Mantaro River, the source of Santiago's drinking water. This threat became the economic justification for a bacterial leaching project at the mine. Indeed, the feasibility of this bacterial leaching project was particularly illustrative of the profitability of leaching copper from waste at the same time as prohibiting otherwise naturally occurring pollution (acid mine drainage). Over a billion tonnes of waste and marginal ore below the 0.6 per cent copper cut-off grade are expected to be dumped during the project's lifetime. The waste would have an average grade of 0.25 per cent copper and would therefore contain a lucrative 2.5 million tonnes of metal worth approximately US$3.5 $\times$ 10^9, at current 1985 prices (Warhurst, 1990). The study demonstrated that with a 25 per cent recovery, high quality cathode copper could be produced profitably, at 39 cents/per pound, by recycling mine and dump drainage waters through the dumps over a 20-year period. This was shown to have the double advantage of both extracting extra copper and avoiding government charges for water treatment. At the same time both investment and operating costs were less than two-thirds of estimated costs for a conventional water treatment plant, which would not have had the benefit of generating saleable copper. The Los Bronces mine thus demonstrates the potential economical benefits of building environmental controls into mine development.

In conclusion, these few examples suggest that dynamic companies are not closing down, re-investing elsewhere or exporting pollution to less restrictive developing countries; rather they are adapting to environmental regulatory pressures by innovation and by improving and commercialising their environmental practices at home and abroad.

POLICY IMPLICATIONS FOR MINERAL PRODUCING COUNTRIES

Technical change and the environmental trade-off

Evidence suggests that, at least during the 1980s, environmental policies have not been a major factor in determining where a mining company will target exploration and subsequent investment activities. Geological potential remains of primary importance, which is not to underestimate that in some cases the approval and permitting process is a major cost of compliance. (Johnson (1990) ranked corporate criteria in selecting countries for exploration.) This would suggest that developing countries are not seen as pollution havens and that the industrialised countries' environmental regulations are not stifling new mining investment. Indeed, there are currently several new gold projects in the process of development in California, which has probably the world's strictest environmental regulatory regime. Although environmental policies may not negatively influence the investment activities of dynamic adaptive mining companies, the latter still seek to play a role in determining the detail and focus of relevant legislation so that new regulatory frameworks also reflect, as far as possible, their corporate interests. During preliminary fieldwork by the author in North America and Europe it became evident that this task was an important function of many of the companies' newly appointed environment vice presidents, directors and environmental affairs representatives. For example, the Environmental Vice President of Inco sits on Canada's high level Environment and Economy Committee, and the Environmental Director of Homestake Mining Corporation sits on the Environmental Committee of the American Mining Congress—which works closely with the US Environmental Protection Agency and lobbies government for tariffs on metal imports originating from countries with poor environmental performance.

If one understands the new environmental pressures being placed upon the mining industry in the industrialised countries in the context of hard-won survival following a prolonged period of low metal prices, which gave significant market advantages to their lower cost competitors in the developing countries, then it is possible to understand the recent lobbying by industry alliances, such as the ICME, the Mining Association of Canada and Eurometaux, for industry-wide international environmental standards. Although international standards may not pose too much of a problem for the economics of new mining projects in the developing countries, our analysis suggests there could be major costs incurred by any older ongoing operations. Controlling the latter's pollution problems would in most cases require major water treatment plants, strengthening and rebuilding tailings dams, add-on scrubbers and dust precipitators, etc. The imposition and strict regulation of international environmental standards could make some developing countries' mineral production uneconomic, thus swapping one social cost (environmental pollution) for another (unemployment and poverty, and indeed clean-up, given the absence of liability laws). This is not to dispute the need for improved environmental controls, particularly in the developing countries, but rather to show the complexity of the process by which the underlying power structure of the industry can help to determine the environmental agenda.

Most planned mines (including existing mine expansions) and available reserves are located in the developing countries. Furthermore, after a period of mineral production monopolised by state-owned mining companies (with some exceptions), many developing countries are now embarking upon a phase of liberalisation and have legislated a number of laws and incentives to promote foreign investment. In many cases those investments are being partly financed by credit that is conditional upon good environmental practice and prior environmental impact analysis. The upshot is that this trend in technical change may be to the benefit of the developing countries in that it may enable them to reduce the trade-off between higher environmental costs and lower production costs. This may mean that at least in the case of new mineral projects there may be a wider range of more environmentally sound and economically efficient technologies available to them.

Indeed, new flexible-scale, lower cost, less hazardous hydrometallurgical (leaching) alternatives to conventional smelting may further be to the advantage of developing countries, improving value-added from their mineral production. For example, processing right up to the stage of a final saleable metal product can be undertaken at the mine site—while in conventional process-routes a smelter will require feed from at least ten large mines for full capacity utilisation, and ore may have previously been exported to foreign smelters with consequent loss of by-products and entailing charges for the treatment of pollutant elements.

However, this new prospect of environmental security may have its own costs which require careful analysis. For example, depending how technology transfer agreements are drafted and managed, such new 'environmentally friendly' investment may herald indebtedness, bankruptcy of local equipment suppliers and engineering firms and the loss of employment, etc., reinforced by aid conditionality. On the other hand, smelting, concentration and leaching innovations are being developed by international companies, such as Outokumpu, Mitsubishi, Kennecott, Inco, Cyprus Mines and Homestake, which are adapted to new and prospective regulations in the industrialised countries. These trends may oblige developing countries to export only semi-processed minerals or raw materials for both economic and environmental reasons, reinforced by credit conditionality, new international regulatory agreements and trade tariffs imposed on imports of metals produced not using a predetermined 'best available technology'.

Technology policy for environmental management

Environmental behaviour correlates most closely with a company's capacity to innovate, rather than its size, origin, scale and scope of operations or ownership structure. For example, government policy over time has resulted in a failure by state enterprises to re-invest capital in human resource development, repairs and preventative maintenance, R&D and technology development. Managers became bureaucrats rather than entrepreneurs. This factor, combined with cumulative inefficiencies, a poor waste management strategy leading to metal and reagent losses, and scarce resources means that environmental mismanagement is endemic. It is a structural problem and one not readily solved by recourse to regulation, punitive tariffs or even the simple act of purchasing an environmental control technology. The cases of COMIBOL (Bolivia) and MINEROPERU and CENTROMIN in Peru bear witness to this, as does the case of private companies such as Carnon Consolidated in the UK. However, CODELCO and ENAMI, the state enterprises of Chile, have invested in developing their innovative capabilities both within the industry and through historically close links with local R&D institutions and universities. Although new regulations currently pose a significant technological challenge to these companies, efforts are being made to develop the required human resources and to implement substantial technical change. For example, CODELCO is now at the forefront of metals biotechnology and has made considerable investments in new solvent-extraction/electrowinning technology. In addition, ENAMI is planning to replace its reverberatory furnaces with modern flash-smelting technology at an estimated cost of $300 million, 'largely motivated by the need for environmental improvements' (Coppel, 1992). Environmental degradation from small 'garimpeiro'-type operations is also related to the miners' incapacity to innovate through a lack of access to capital, technology, skills and information. Scale further complicates the choice of optimal low waste, high metals recovery technology. With few exceptions, however, it is the private sector that has so far shown itself to be most innovative and therefore most capable of improving environmental management. In several cases improved environmental management would have been brought about irrespective of regulation due to market pressures to introduce new, more efficient, low-waste technical change.

This is not an argument against regulation but rather to recommend a more sophisticated public policy approach through first, the definition of regulatory goals—something to aim at—and second, an informed technology policy to guide and stimulate those companies along the fastest most efficient route to achieving those goals.

This technology policy could include a detailed technology-transfer strategy, tax relief on R&D and the training of engineers and managers in environmental technology, government grants for collaborative inter-industry and university-industry R&D projects, and information dissemination programmes regarding the moving technological and regulatory frontiers. It requires training for regulators so that they are informed disseminators of information about environmental technology. Finally it requires the provision of new lines of credit—in banks and donor agencies—to promote investment in the development, commercialisation, acquisition and improvement of environmental control technologies.

Another factor that reinforces the need for trained and informed environmental regulators relates to their ability to determine corporate environmental trajectories as a response to regulation. Figure 7.1 provides a planning tool that would enable regulators to plot corporate environmental trajectories against changing thresholds of economic competitiveness and environmental compliance over time. For example, it is suggested that fast-changing 'incremental' regulation would tend to promote add-on incremental technical change—pushing cost-curves upwards rather than down. As a consequence, and evidence of this is slowly emerging, many mining companies, particularly those with large sunken investments, are on a trajectory heading towards close-down and certainly reduced competitiveness. With the growth of market and regulatory pressures that group of companies heading towards bankruptcy, close-down, project delay or cancellation is also likely to grow in number. At the same time, the clean-up problem posed by the closing down of those companies would also be expected to grow in scale and severity, as evidence drawn for Peru and Bolivia demonstrates (Nuñez, 1992; Morgan, 1992). Given the difficulties involved in retrospective clean-up legislation in terms of cost assessment, litigation and technical complexity, it would seem desirable for government to avoid such a scenario. This could be done on the basis of sound prediction and planning, either through imposing a levy on operators as the mine nears exhaustion/abandonment (whichever is sooner) to cover clean-up and rehabilitation, or through promoting tax and other incentives for the required investment in clean-up. For new investments, reclamation bonds or equivalent mechanisms can promote environmental management from the outset, and a carefully planned waste management programme from the start of operations will assist in spreading the costs and reducing potential hazards.

The polluter pays principle only holds if the polluter can survive in order to pay. The capacity to innovate, including capabilities in environmental management, is one key factor in determining a firm's ability to survive, grow and continue to generate wealth from metals production.

The role of technology transfer in an environmental management policy

Technological and managerial capabilities are not required just to deal with new and emerging technologies, they are also vital to an environmental management strategy using existing technology, owing to pervasive inefficiencies. Technology transfer and technology partnership through joint venture arrangements is one way to overcome these constraints, particularly in the developing country context, although such strategic alliances are emerging in all the major mineral-producing countries. Recent examples of collaborative partnerships in innovation include Outokumpu and Kennecott, Outokumpu and CODELCO, Cyprus Mines and Mitsubishi, and Comalco, Marubeni Corporation of Japan and the Chilean power company Endesa.

However, there is a need to broaden the common concept of technology transfer to achieve the desired result of a real transfer of environmental management capability. Traditionally technology transfer has meant a transfer of capital goods, engineering services and equipment designs—the physical items of the investment, complemented by training in the skills and know-how for operating the plant and equipment. Such transfers are often restructured in scope to match the 'step-increments' involved in add-on regulatory-response technical change.

New forms of technology transfer will need to go further to embrace first, the knowledge, expertise and experience required both to operate and manage technical change of an incremental and radical nature and second, the human resource development and organisational changes involved in an overall approach to improve efficiency and environmental management throughout the process route, plant and facility.

Many transnational companies play a major role in the mining industry, usually contributing significant managerial and engineering expertise in joint ventures and subsidiaries. They usually limit their contribution in the light of the costs, capabilities and work involved to fulfil the immediate needs of the specific investment project or physical item of technology transfer.

Empirical research on other sectors demonstrates considerable potential to increase those contributions without adversely affecting corporations' strategic control over 'proprietary' technology (Warhurst, 1991). What is required is a strategy of technology and enterprise targeting and a clear set of technology transfer objectives and financial mechanisms to cover the extra costs involved over and above the investment and basic training budget.

There already exists a developed market and a range of commercial channels through which mine operators can purchase capital goods, engineering services and design specifications; however, the market for knowledge, expertise and experience, and accelerated training programmes is less mature. Bilateral and multilateral agencies and development banks can play a major role in improving this situation. Such an approach was at the heart of the strategy of China's National Offshore Oil Corporation (CNOOC) which targeted specific major oil companies and required them under technology transfer agreements to transfer the capabilities to master selected areas of technology (Warhurst, 1991). Another interesting example is the Zimbabwe Technical Management Training Trust. It was founded by RTZ in 1982 with the aim of training

South African Development Community (SADC) professionals in technical management and leadership. It effectively combines academic and on-the-job training in both home and overseas operations, providing possibilities for accelerated managerial learning by exposure to on-the-job problem-solving situations with experienced colleagues in a range of challenging technical scenarios. Although smaller in scale, this is a similar strategy to that designed by CNOOC to ensure its trainees worked alongside experts in different transnational oil companies in situations that ensured a 'mastery' of the technology rather than the knowledge of how to create it. RTZ absorbs the entire cost of the training and related M.Sc. scholarship programme at the City of London Business School.

It would be quite feasible to build similar in-depth training programmes, concentrating on human resource development in environmental management, into many of the proposed and prospective mineral investment projects throughout the world. Investors and technology suppliers could be selected, in part based upon their proven environmental management competence and their willingness to transfer it. It cannot be over-emphasised that all technology transfer and training has a set of costs for the supplier and these must be covered to ensure optimal results. The danger of not budgeting for these costs would be to resort back to a training programme in operational skills rather than in technology 'mastery' skills. The corporate partners, the government, donor agencies or development banks could assist in finance. However, negotiating power regarding the precise objectives and scope of the programmes would be greater if governments, organisations or firms contributed financially. This paper suggests that it is the capacity to effect technical change, not just the skill to operate an item of environmental control technology, that will ultimately determine the success with which firms build up competence in environmental management.

Moreover, referring back to our case study examples of best practice environmental management, the technical changes introduced illustrate the myriad intangible practices that constitute sound environmental management. Of relevance to technology transfer policy is the fact that it is not the utilisation of one specific technique, but rather a combination of technology, managerial approaches, workplace practices and regulatory and monitoring frameworks that explains the parallel achievements of improving economic efficiency and environmental performance. Most of this capacity would not be secret or proprietary to the firm. It is more a question of knowing what to ask for and then being prepared to pay the cost of the resources, time and effort required for its transfer. This reinforces the need for mining enterprises to build up a range of technological and managerial capabilities as well as workers' skills. This also explains why in the developing country context the simple act of acquiring a new item of technology or an add-on piece of pollution control will not automatically lead to an expected increment of pollution reduction. The transfer of environmental technology does not lead to a transfer of environmental management capabilities unless training and broader human resource development programmes are built into the investment and an overall approach to improve efficiency and good housekeeping at the plant site is adopted. In the developing country context such change is best brought about when stimulated by donor agency requirements supported by credit, government commitment and the participation in the investment project of a partner with proven competence in environmental management elsewhere.

A further public policy implication suggested by the preceding analysis is the need for a two-tier environmental management policy. There is one set of challenges for the ongoing minerals industry, which must encompass the findings above regarding production inefficiency and its environmental consequences and the clean-up requirements on mine shut-down and plant decommissioning. Another set of challenges concerns the policy need to build environmental management into investment and expansion projects from the outset, which requires negotiation at the earliest stage with operators, equipment suppliers and credit sources.

The public policy challenge is, therefore, how to keep firms sufficiently dynamic to be able to afford to clean up their pollution and generate economic wealth through innovation and sustainable environment management practices. The achievement of improved production efficiency and environmental management, particularly in developing countries, will in turn be dependent upon the extent to which far-reaching technology transfer and training clauses are built into the joint ventures and new investment arrangements that characterise the industry, and whether banks, donor organisations and governments demonstrate responsibility by providing the appropriate lines of credit and technical assistance in support of such objectives. Environmental regulation would be one element of that policy and would provide the goal posts for site-specific best practice in environmental management; technology policy to promote technical change through technology transfer and human resource development would lie at its heart.

NOTE

Original text published in R.E Hester and R.M Harrison (eds) (1994) *Mining and its Environmental Impact,* UK The Royal Society of Chemistry.

ACKNOWLEDGEMENTS

The author would like to acknowledge gratefully the kind assistance of Allison Bailey, Gill Partridge and Hilary Webb in the preparation of this article. Support for the research reported here was provided by the John D. and Catherine T.MacArthur Foundation. Parts of this paper were presented in the John M.Olin Distinguished Lectureship Series in Mineral Economics at the Colorado School of Mines in December 1992.

REFERENCES

Aitken, R. (1990) Personal Communication, Inco Ltd.
Coppel, N. (1992) 'World-wide Minerals and Metals Investment and the Environment, 1980–1992', Unpublished Report For RTZ Corporation Plc.
Crouch, D. (1990) Personal Communication, McLaughlin Mine, Calif., USA.
EPA (1986) *Gregory Tailings,* Washington DC, USA: Environmental Protection Agency.
EPA (1989) *Soil Clean-Up Of Smuggler Mountain Site,* Washington DC, USA: Environmental Protection Agency.
Johnson, C.J. (1990) 'Ranking Countries for Mineral Exploration', *Natural Resources Forum* 14, 3: 178–186.
Kelly, D.R. (1990) Personal Communication, Alcan International Ltd, Montreal, Canada.
Morgan, C. (1992) *The Privatisation of State Industries—Guidelines on Environmental Liabilities,* Third Newsletter of Mining and Environment Research Network, December 1992.
Nuñez, A. (1992) 'Environmental Management in a Heterogeneous Mining Industry—The Case of Peru', in *Mining and the Environment: Case Studies from the Americas,* Ottawa, Canada: IDRC, 1997.
O'Connor, D.C. (1991) 'Market Based Incentives', in A.Warhurst (ed.) *Environmental Degradation from Mining and Mineral Processing in Developing Countries: Corporate Responses and National Policies,* discussion document, SPRU, Section 2, Ch. 5, Paris, France: OECD.
OECD (1990) *Environmental Concerns Related to Commodities: Scrap Recovery and Recycling of Non-Ferrous Metals and Environmental Policy Instruments: The 1989 Basel Convention,* 15 October, Paris, France.
OECD (1991) Env*ironmental Policy: How to Apply Economic Instruments,* Paris, France: OECD.
Office Of Technology Assessment (1988) *Copper Technology and Competitiveness,* USA: Congress of the US, p. 16.
Panayotou, T., Leepowpanth, Q. and Intarapravich, D. (1990) *Mining, Environment and Sustainable Land Use: Meeting the Challenge,* Synthesis Paper No. 2, Year-End Conference, Jomtien, 8–9 December, The 1990 TDRI (Thailand Development Research Institute).
US Bureau Of Mines (1989) *Bauxite, Alumina, Aluminium, Annual Report,* Washington DC, USA, p. 6.
US Bureau Of Mines (1990) *Copper Annual Report,* Washington DC, USA, p. 28.
Warhurst, A. (1990) *Employment and Environmental Implications of Metals Biotechnology,* World Employment Programme Research Working Paper, March, WEP 2–22/SP.207, Geneva, Switzerland: International Labour Organisation (ILO).
Warhurst, A. (1991) 'Technology Transfer and the Development of China's Offshore Oil Industry', *World Development* 19, 8: 1055–1073.
Warhurst, A. (1992) *Mining and Environment Research Network Newsletter* 2.

SUGGESTED READING

Warhurst, A. (1993) *Environmental Degradation from Mining and Mineral Processing in Developing Countries: Corporate Responses and National Policies,* Paris, France: OECD Development Centre.

SELF-ASSESSMENT QUESTIONS

1 Indicate which of the subsequent environmental management (EM) and/or policy trends are correct
(a) EM has always been targeted to causes; as a rule it was never targeted to symptoms of pollution.
(b) Best available technology is a management approach with an outspoken dynamic component.
(c) Environmental taxes are outspoken examples of a 'command and control' instrument.
(d) Market-based mechanisms in environmental management work best in combination with command and control and an education-driven policy.

2 Indicate the correct alternative(s): Command and control measures
(a) require intensive monitoring to ensure that they are enforced;
(b) are, for example, standards and environmental taxes;
(c) are older instruments in environmental management than are economic instruments;
(d) tend to deal with symptoms (pollution) rather than with causes (production inefficiency, human resource constraints, lack of technology and capital).

3 Indicate the correct alternative(s): Improving the environmental management of mining operations
(a) may not necessarily be detrimental to economic performance;
(b) can these days take advantage of technical changes which may improve environmental performance;
(c) can be stimulated by a rigid environmental legislation;
(d) can leave behind the intensively contaminated areas of bankrupt companies.

4 Indicate the correct alternative(s): Mining companies that are being forced through governmental pressure to deal with pollution
(a) always react in a defensive way;
(b) are in the medium or long term condemned to bankruptcy;
(c) acquire a competitive advantage as compared to foreign companies;
(d) can react in an innovative way depending, for example, on their internal dynamism;
(e) see this essentially as a cost burden on their operations that threatens their profitability.

5 Indicate the correct alternative(s). Technology transfer
(a) relates uniquely to the transfer of technological hardware, for example, to developing countries;
(b) relates to a combination of technology, managerial approaches, workplace practices and regulatory and monitoring frameworks;
(c) always adversely affects corporations' strategic control over 'proprietary' technology.

8

SUSTAINABLE AGRICULTURE

Kerry Garrett

SUMMARY

Agriculture must meet the needs of the world's current population of 5.8 billion and the huge increase in population anticipated in the next millennium. Any failure to meet this challenge will have catastrophic consequences so the need for a sustainable agriculture is absolutely fundamental to the future of mankind. Interaction between agriculture and the environment is particularly intimate. Agriculture is a manipulation by man of processes in the natural environment and as such creates secondary environmental effects. Some of these effects are highly valued as environmental enhancements whilst others such as pollution of water or air, which may to some extent be inevitable, are nevertheless rated as undesirable. Moreover, agriculture is highly subject to environmental influences especially those concerning climate and the quality of soil, water and air so the interaction between agriculture and the environment is a two-way process. Socio-economic factors also play an extremely important role in both the nature and development of agriculture.

Applying the concept of sustainability to agriculture requires an understanding of these factors but the concept must be given meaning if it is to find practical application rather than providing an abstract theory. The first stage is definition of some key principles that must be applied. These include the requirement for economic efficiency but the application of this principle must be broadly based to consider the natural and socio-economic environment as well as the production process. The second principle concerns conservation of natural resources—both renewable and non-renewable • and further principles relate to flexibility in land use to provide options for the future, and a system of values which must provide the philosophical underpinning to a sustainable agriculture. This approach enables the consideration of various practices which should be incorporated into models of a sustainable agriculture, including the integrated approach to pest and disease management and soil fertility management. The integrated farming approach represents the best current synthesis of these ideas in practical farming systems. It seems likely that this approach can be widely adopted in the developed world but technical and financial constraints continue to limit its wider adoption in many developing regions. It is critically important that such constraints are addressed in development programmes to meet the fundamental human need of a secure and safe food supply.

ACADEMIC OBJECTIVES

This chapter aims to provide the reader with an introduction to the concept of sustainability applied to agriculture. On completion the reader of the chapter should:

- Understand the need for a sustainable agriculture.
- Appreciate the principles of sustainable agriculture.
- Identify key practices in sustainable agriculture.
- Develop ideas on how sustainable agriculture might be promoted in both developed and developing countries.

THE NEED FOR A SUSTAINABLE AGRICULTURE

Nothing is more critical to the future of mankind than the security of world food supply. Whilst malnutrition may prevail because of economic, administrative and political factors in many developing countries, it is widely recognised that food production in the world today is sufficient to meet the world's population of 5.8 billion. Moreover, recent expert analysis (e.g. International Food Policy Research Institute, 1995) affirms that this will continue to be the case for at least one hundred years, in which time the world's population may reach 12 billion. Whilst some of this increase will come from the developing countries, in the medium term at least most of the increase will rely on the two most economically developed regions of the world—North America/Oceania and Europe/former Soviet Union (Dyson, 1996). Achieving this growth in a sustainable manner is probably the most formidable challenge the world will face in the next millennium.

Agriculture involves the manipulation of natural processes to produce desired products. These natural processes include nutrient supply and the population dynamics of plant and animal species, including pest, disease and invasive organisms. In the agro-ecosystem biodiversity is suppressed by management, for example, through cultivation, pest control and soil fertility management to optimise the production of an arable crop. This degree of management intervention represents a significant environmental disturbance within the agro-ecosystem and also in contiguous ecosystems. Moreover, the interaction between agriculture and the environment is a two-way process since climate and other 'external' environmental factors may have major impacts upon production processes. The interactions may also involve feedback loops, for example, ruminant animals and paddy rice production are important sources of the greenhouse gas methane and thereby contribute to environmental change which may in turn impact upon these agricultural systems. When we apply the concept of sustainability to agricultural systems, therefore, we must think beyond the scope of the agro-ecosystem to include all contiguous aspects of the global environment. Indeed, this should include the socio-economic as well as the physical environment.

We should also reflect upon the changing role of agriculture within its socio-economic context. Traditionally the primary aim of agriculture has been the production of food and fibre but in much of the developed world today we have a surplus of traditional agricultural products. Here the agricultural industry is being required to constrain production but at the same time to provide a new range of products—the so-called environmental goods. Some of these, such as recreational opportunity, may have a market value but others, such as clean water or a visually attractive landscape, may have a value to the economy but no direct market value to the farmer. When the public demand such environmental goods, as is increasingly the case, their costs must be met by the farmer if the products are to be achieved through regulation, or by the taxpayer as a subsidy through environmental management payments to the farmer. In much of the developed world, therefore, the agricultural industry has taken on a much wider and more sophisticated role and this trend seems set to continue. This extended role for agriculture must also be incorporated into models for a sustainable agriculture.

The need for a sustainable agriculture is therefore both *fundamental,* being related to its primary role of food production and the need for long-term security of world food supply, and also *imposed* by the democratic will of a public which, in the developed world, increasingly sees the farmer as a custodian of the environment and who is willing to see both the pressure of legislation and the incentive of subsidy applied to achieve long-term benefit in terms of environmental management.

Food choice is another mechanism forcing change in agriculture and driving the industry towards more sustainable models of production. In the developed world the consumer has become empowered by both affluence and information. The consumer is faced by a market in surplus and food prices that are historically low in relation to disposable incomes. At the same time information on food composition and quality, nutrition and health, and on the means of production is more readily available than ever before and the consumer is in a powerful position to make informed choices. Concerns over animal welfare, pesticide residues, etc. now preoccupy an increasing number of consumers and the food processing and retailing industry is requiring the farmer to implement defined systems and methods of production and to install systems of traceability, inspection and accreditation. These link the empowered consumer directly to production methods at farm level and enable public concerns on issues such as animal welfare, pollution and environmental care to impact on farming practice.

At the same time the farming industry in the developed world faces what is virtually an ethical crisis. The industry is no longer based on the production ethic, having developed a capacity to produce potentially huge surpluses of most major commodities. In this situation, and with fewer and fewer people employed directly in farming, the industry has lost much of its former political influence. To some extent farming has also lost public sympathy largely as a consequence of the real and perceived consequences of modernisation—the very process that has assured a plentiful and cheap supply of food—consequences such as pollution, conflicts with conservation and some of the worst aspects of intensive livestock production. The fact that commercial pressures and the distortions of agricultural policy may have driven farmers in these directions does not absolve the agricultural industry from guilt in the eyes of many consumers. There is therefore an urgent need to define a new ethic for agriculture—one that meets the concerns of a diverse range of stake holders (Table 8.1), and which provides agriculture with a vision of the future. The concept of sustainability provides such a vision.

Table 8.1 Stakeholder analysis of agriculture

Stakeholder	*Interest*
Farmer	Fulfilment
	Profit
	Heritage
Customer	Price
	Quality
	Choice
	Security of supply (food industry)
Public	Environmental goods
	Control of development
	Access to countryside
	Production methods
Government	Food security and safety
	Environmental protection
	Rural economy
	Value for money

THE MEANING OF SUSTAINABILITY IN AGRICULTURE

Agricultural sustainability means reconciling and integrating three sets of major objectives—economic objectives, environmental objectives and social objectives. These are potentially in conflict but where conflict prevails farming systems have invariably proved unsustainable.

The term 'agriculture' means literally the husbandry of land. From earliest times man has husbanded land for the production of food. As agriculture became the engine of development in most countries of the world, the husbandry of land for production of food and fibre became an economic rather than a subsistence activity and the primary objective of modern farming systems is the profitable production of food. This is even more so in a society that demands environmental care as a by-product of farming, for it is preferable that this should be paid for from the profit in farming rather than taxation and subsidy. Clearly therefore the primary objective of modern farming is an economic one.

The effect of an unconstrained profit motive, however, is potentially disastrous in an industry that relies on subsidy because it can lead to the production of crops in conditions that are inherently unsuited to the crop in question, overstocking with livestock in sensitive environments or the excessive use of support resources. Nor are such problems confined to the subsidised agriculture of much of the developed world—they can apply even more catastrophically in ill-judged development projects in the world's poorest nations. Clearly the economic objective is not enough—agriculture must also meet key environmental objectives if the goal of sustainability is to be achieved.

The environmental objectives for agriculture are both local and global. Local objectives will be founded first and foremost upon a proper understanding of the local environment—particularly the climatic and edaphic environments. This is essential to informed decisions on land use, soil management and appropriate agricultural practices. Global considerations concern the impact of climate cycles on agriculture and the impact of agriculture on global climate. The objectives will focus on the creation of environmentally benign farming systems which maintain or enhance the productive resource. Once again there are obvious linkages with economic objectives within such systems.

The third set of objectives that must be integrated into this scenario is social in nature. Agriculture is widely seen as the backbone of rural communities. Where agriculture is a principal source of employment, rural depopulation and migration to urban areas may pose major socio-economic and political threats and policy may be directed towards agricultural development and diversification to maintain and develop the vitality of the rural community. In the developed countries agriculture now employs an historically low percentage of the population—generally below 10 per cent—but even at this level the stability of rural communities relies to a large extent on the agricultural industry because it is this that generates a substantial amount of secondary employment whilst at the same time having a long-term commitment to ownership or tenure of the land. Agricultural policy in many developed countries now centres around integrated rural development and seeks to generate rural industries and tourism operating alongside existing farming enterprises. Social objectives must also recognise the importance of the skills and expertise—the knowledge base of which much may have been acquired over many generations and upon which farming depends.

Whilst the integration of economic, environmental and social objectives is the basis of a sustainable model for agriculture, it is also a formidable challenge. Inevitably it involves trade-offs and compromise—but as the later discussion will illustrate, this leads to a stronger rather than a weaker system, in other words one that is more likely to be sustainable.

PRINCIPLES OF SUSTAINABILITY IN AGRICULTURE

The concept of sustainability will have little impact upon agriculture unless it can be developed into a set of workable principles, that is principles that can find practical application rather than being part of a theoretical exercise. Many agricultural systems already embody such principles to a substantial degree—the case study of a Chinese agro-ecosystem provides a particularly good example of a system that has proved sustainable over thousands of years, and there is much to learn from the study of such systems. At the other extreme, examples of catastrophic soil erosion bear witness to the unsustainable nature of farming and land-use systems on formally productive areas. It is easy to see what has proved sustainable or unsustainable with hindsight—it is much more difficult to be definitive about what will prove sustainable in the future. Nevertheless, this must be attempted and the following principles are suggested as a starting point.

Economic efficiency

In a market economy this should be a critically important principle that sustainable systems must embody, for without it the basis for competition and therefore survival is deficient. However, the concept of economic efficiency concerns optimisation rather than maximisation because it must take account of the costs of environmental stewardship as well as those of production. The polluter pays principle is now thoroughly embedded in EU environmental law and in many regions fines for water pollution, for example, can now include the costs of reinstatement of a river or lake following escape of agricultural effluents and such costs can be substantial for most farm businesses. Moreover, many agricultural effluents are much more polluting than domestic sewage effluent so costs of treatment are relatively high. There are, therefore, powerful economic arguments for clean technology in agriculture—systems of production that avoid or minimise the production of polluting by-products rather than reducing their effect by treatment. However, without the implementation of substantial fines for pollution there is little incentive to carry what will generally be the increased costs of a cleaner technology. Likewise the costs of conservation and enhancement will normally be part of the economic equation for sustainable systems.

The development of 'life-cycle analysis' (Azapagic and Clift, 1994) has been a particularly useful way of analysing the totality of environmental costs associated with a production enterprise and there is no doubt that this will help to guide the analysis of economic efficiency as a criterion of sustainability in agricultural systems.

Conservation of natural resources

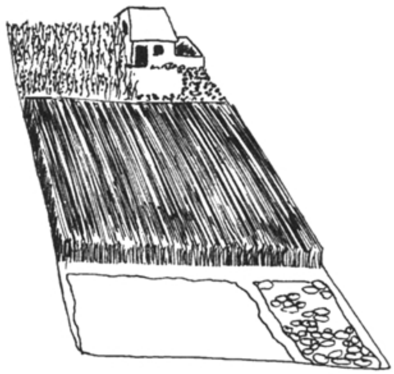
The broad-based approach to be employed in economic analysis must also be applied here, for 'conservation' concerns much more than the wildlife

BOX 8.1 CASE STUDY 1: CHINA

Twenty-two per cent of the world's population lives in China and China is virtually self-sufficient in food. For several thousand years Chinese farmers have practised types of agriculture that are highly intensive but demonstrably sustainable. These systems are extremely conservative in nutrient cycling and are characterised by a very high degree of management. This enables the growth of at least two crops annually and complete use of by-products and crop residues to produce food of good nutritional quality. An idealised farm of 0.1 hectares which might support a family unit is illustrated below and is based on the author's personal observations.

The particular combination of crop types and animals constitutes a highly integrated agro-ecosystem. This integration depends on

1 Deep rooting plants, e.g. mulberry for silkworm/silk production, which capture nutrients relatively deep in the soil profile (background).
2 A pig or hens which consume residues from house and farm.
3 A fish pond which receives nutrient runoff and organic wastes from house and farm and supports fish production from the food chain starting with algal photosynthesis (foreground left).
4 Rice cultivation which receives sludge from the fish pond and nutrient rich irrigation water (middle area).
5 Water plants which capture nutrients from water where nitrogen is also fixed biologically (foreground right).



When this system is studied it is realised that the major nutrient inputs are nitrogen from biological fixation and carbon from photosynthesis. Significant quantities of N and C will also leave the system as carbon dioxide and methane, and in the case of N as ammonia, nitrous oxide and nitrogen gas from denitrification. Otherwise nutrients are conserved through internal recycling of wastes and residues and by scavenging from the soil profile by plants with different rooting depths. This may be supplemented by atmospheric precipitation and occasional flooding.

Many of these traditional agro-ecosystems are now changing rapidly under the influence of industrial development and the availability of supplementary income. This normally involves less attention to recycling, a growing dependence on fertiliser and pesticide inputs, and a reduction in yield if production changes from a two crop to a one crop system. Nutrient losses to the environment may also increase – especially leaching of nitrate and phosphate to drainage water.

resource of the countryside—indeed wildlife conservation is a by-product rather than a primary objective of conservation in agriculture. The three key aspects of conservation in agriculture are:

- maintaining the regenerative capacity of renew able resources;
- shifting from the use of non-renewable to renew able resources;
- phasing the use of non-renewable resources.

Maintaining the renewable resource base implies concern about the biological diversity of agro-ecosystems and the genetic diversity of crop and animal species. As has been pointed out already many agro-ecosystems rely upon suppression of biodiversity to favour the production process but this is not incompatible with biodiversity within areas, for example, the field margin or the hedgerow. Such diversity may favour the pro duction process in the adjacent area, for example, by providing a stable over-wintering habitat for species such as Carabid beetles which are involved in the natural control of agricultural pests such as aphids. There is widespread concern about the increasing reliance of much of world agriculture upon a diminishing range of genetic diversity in the principal crop and animal species. This lack of diversity reduces the adaptive capacity of systems to environmental pressures such as disease or climate change. A dramatic illustration of the value of genetic diversity was provided by the production of new varieties in the so-called green revolution when yield, compositional and disease resistance properties were enhanced in crops such as maize and wheat by introducing genes from wild and ancestral strains of these crops into modern varieties. The protection of soil, air and water are further important aspects of resource conservation but are beyond the scope of this chapter.

Much can also be done to shift from the use of non-renewable to renewable resources. Renewable energy can substitute for fossil energy in many farming activities which require gas, liquid or solid fuels. Recycling of nutrients, water and various materials can often be enhanced through simple measures and improved management and there is much to learn from examples provided by organic farming in this regard. There is also considerable scope for enhancing natural control mechanisms for pests and diseases, thereby reducing dependence on manufactured pesticides.

Obviously such approaches have limits, so it is also important to consider the phasing of use of non-renewable resources to conserve the scarcest. Phosphorus and potassium, for example, are essential nutrients for crop plants but part of the plant's requirement for these nutrients can be replaced by calcium or sodium (Bailey, 1995) which are less expensive and more abundant nutrients. There are numerous examples of this type of approach and throughout this discussion it should be clear that these principles are quite practical as well as being theoretically sound.

Enabling response to change

This principle has already been alluded to in relation to genetic diversity but it applies equally to the issues of land use and enterprise specialisation. Specialised enterprises by definition involve highly selective investment in specialised technology. They may also involve a narrowing of the skills base in the workforce and whilst these features may secure the success of an enterprise in the short term, they are also likely to make it relatively difficult to change and adapt in the medium to longer term, yet support measures for agriculture and the commercial pressures in a very competitive industry can easily foster enterprise specialisation.

The ability to change land use is also a strategically important issue. Concern about the capacity of European and US agriculture to produce a potentially vast surplus of cereals has led to the policy of 'set aside' in which land is taken out of production. In Europe there is now interest in converting both set aside and crop land to woodland and whilst this may be both environmentally and strategically sound it will be important to plan such developments carefully because of their potential to limit flexibility in land use. Events such as the Chernobyl disaster have the potential to take huge areas of productive land out of food production and a rational policy for land use must retain the capacity for flexibility so that strategic adjustments in food supply can be accommodated.

A value system

Finally, it can be suggested that true sustainability for agriculture cannot be achieved without the foundation of a value system—a system that values enterprise and endeavour, that values knowledge and skill, that values the dignity of a farm animal, the integrity of nature and the views of its stakeholders. There can be many philosophies within which these values may be integrated but without them it seems inevitable that the agricultural industry will lack the vision and conviction needed to sustain its long-term development.

PRACTICES IN SUSTAINABLE AGRICULTURE

In agriculture, man manipulates natural systems to favour the production of useful products. Some degree of environmental disturbance is an inevitable consequence and it becomes a highly subjective matter to judge the extent to which this disturbance is acceptable. The judgement will be quite different for those at risk of starvation compared to the affluent consumer preoccupied with choice from an abundant food supply. However, the issue of sustainability is important at both extremes, for both have a stake in the security of food supply and the health of the environment. Moreover, both have a stake in the availability of affordable, safe and nutritious food and this raises the issue of pesticide and fertiliser usage in modern agriculture. It should be understood at the outset that it would be quite impossible for agriculture to meet current world food demand without the massive use of pesticides and fertilisers, let alone meeting the future needs of a growing population. Indeed, the use of both pesticides and fertilisers is essential to sustainable agriculture oil a world scale. This is not to say that it is impossible to produce food without the use of such inputs and indeed high yields can be obtained in 'organic' systems with appropriate management and substitution of inputs—but over half the world's food supply is due to the fertiliser input and without pesticides between 10 and 40 per cent of the world's gross agricultural production would be lost. Moreover, on a regional basis, losses can be even higher than this with devastating consequences. However, the misuse of pesticides, and to a much lesser extent of fertilisers, can also be devastating to the environment so it is essential that these inputs are used properly.

The key to proper management is the use of no more of the input than is needed for the required effect. Unfortunately, the production environment is often unpredictable. Rainfall and drought have major effects on crop growth, but also on leaching of plant nutrients from soils. Heavy rainfall shortly after planting of an annual crop when root systems are still poorly developed can leach nutrients below the rooting zone, whilst under dry conditions nutrients may be unavailable to plant roots and left unutilised in the soil profile after harvest. The farmer may fail to take account of such circumstances or apply extra fertiliser to offset the risk of leaching but in doing so may increase the risk

Table 8.2 Some dangers and benefits in the use of pesticides

Dangers	*Benefits*
Risk of accidental human exposure	Increased and more predictable yield
Residues in food	Animal welfare
Effect on non-target species including natural enemies of the pest	Improved quality and food safety
	Reduced storage losses
Accumulation in the food chain	Enabling improved techniques (e.g. no till/low till farming with reduced energy input and erosion risk)
Development of resistance in pests	

of water pollution. Pesticides also may be applied as an insurance strategy without proper monitoring of pest incidence or in excessive amounts in the misconception that additional benefit may accrue beyond the manufacturer's recommended rate. The role of advice and training is critical to promote the proper use of agro-chemicals, together with surveillance to monitor pesticide residues in food. Unfortunately such systems are poorly developed in many countries. The most important strategy available to the farmer is that of 'integrated management' which seeks to bring together the full range of techniques by which problems such as pests and disease control and the maintenance of soil fertility may be addressed.

Integrated pest management

A variety of approaches may be applied in pest control. These include:

- use of natural enemies—biological control;
- breeding for pest or disease resistance;
- the selective use of pesticides and use of selective pesticides;
- interference with reproduction of the pest;
- environmental manipulation.

In integrated pest management a range of these approaches is combined and applied together with monitoring of pest and disease occurrence (Burn *et al.*, 1987).

There are many successful examples of biological control (Altieri, 1994) which may involve either the introduction of new or exotic organisms which attack the pest, or enhancement of pre-existing populations of such control organisms. Some of the most successful examples are with covered crops where the control agent is confined, but there are serious concerns about the introduction of new or exotic pest control organisms. Enhancement of pre-existing control agents may be more acceptable—but effects may be short-lived because conditions may not be conducive to a high population level irrespective of the food supply. Enhancement works best where the seasonal growth of the natural population of the control agent may be augmented to achieve a more rapid build-up to effective levels for control. The level of natural control agents may also be enhanced by providing suitable adjacent habitats to maintain a reservoir of the control agent or enhance its level by augmenting its food source. This is a beneficial aspect of biodiversity in agriculture and field margins; hedgerows and uncultivated strips in crop land are examples of suitable habitats which can enhance biological control.

The production of disease resistance in crops and animals has made a major contribution to enabling a reduction in pesticide inputs. In many species disease resistance has been associated with undesirable flavour or palatability traits which were eliminated both deliberately and inadvertently by the breeder preoccupied with yield improvements (Conway, 1971). Natural resistance can be reintroduced to modern varieties by incorporating genes from resistant lines by crossing or other gene transfer techniques but attention must be paid to the mechanism of resistance that is conferred. If this is due, for example, to the production of a toxic alkaloid this may pose a risk to the consumer. Nevertheless, disease resistance is the most successful alternative to pesticides.

Modern pesticides are, of course, highly selective compared to those deployed when Rachael Carson published her famous book *Silent Spring* (Carson, 1963). This is the result of public pressure and a massive investment in the development of new pesticides by the agro-chemical industry. Unfortunately the huge costs associated with toxicological and environmental testing for clearance of new pesticides has now reached a stage that is limiting the further development of new products because the payback period under patent protection is inadequate to justify development costs. However, the examples in Table 8.3 illustrate what has been achieved—the best modern synthetic pyrethroids are effective at incredibly low rates of application of the active ingredient and have extremely low toxicity to non-target species.

Chemistry has also provided a sophisticated approach to pest control through mimicking the effect of insect pheromones—volatile hormones released by the female insect to attract males for breeding. Synthetic pheromones may be deployed with traps or by spraying the crop to confuse and disorient the male insect pest. An alternative strategy to disrupt the reproduction of insect pests is to release sterile males and this has been successfully deployed to control and in some cases eradicate certain pests.

Table 8.3 Evolution of insecticides

Compound	*LD_{50}* mg/kg*	*Application rate g/ha*
Organochlorines, e.g. DDT	115	1500
Organophosphates, e.g. Parathion	13	500
Synthetic pyrethroids, e.g. Deltamethrin	5000	10

The most widely practised environmental manipulation limiting pest and disease problems is the strategy of crop rotation, but the use of crop mosaics and other spatial patterns in crop production is widely practised and often effective in reducing rather than eliminating pest and disease problems.

Integrated pest management (IPM) brings together effective combinations of these approaches. It was first deployed on a scientific basis in 1954 for the control of pests of alfalfa in the USA (Conway, 1971) but is extensively deployed today in both crop and livestock systems. The major motivation for the development of IPM has been economic rather than concern for the environment, and the main constraint to its wider application is probably the limited availability of technical advice to the farmer together with commercial pressures designed to maintain the sale of agro-chemicals.

Integrated fertiliser management

The use of mineral fertilisers is a key component of a sustainable agriculture. Fertilisers are employed to improve the fertility of soils with nutrient reserves that are below the optimum for crop production and to replace nutrients removed by a crop. Animal manures also have a valuable role in maintaining soil fertility but unless they are derived from food imported on to the farm they represent an internal recycling of nutrients rather than a net addition of nutrients to the soil. Animal manures have additional benefits, of course, in improving soil structure and water-holding capacity through the addition of organic matter. In addition to these sources of nutrients the biological fixation of nitrogen, especially by legumes such as clover, peas and beans, may represent an adequate input of nitrogen for crop production in many circumstances and may add to soil reserves to support the growth of non-legumes in companion crops or at subsequent stages in a crop rotation. Generally, however, higher yields are obtained with fertiliser nitrogen than with biological fixation.

Integrated fertiliser management (IFM) utilises the optimum combination of mineral fertilisers, organic manures and biological fixation of nitrogen. In some circumstances, it may also utilise crop residues or materials such as sewage sludge although the latter provides a relatively dilute source of nutrients and may have problems due to contamination by heavy metals. IFM also incorporates the use of appropriate crop rotations and cultivation techniques to conserve and enhance soil fertility. The IFM concept was given a powerful endorsement in the UNCED conference in Rio in 1993, and Agenda 21 included these specific objectives:

> not later than the year 2000, to develop and maintain in all countries the integrated plant nutrient approach and to optimise availability of fertiliser and other plant nutrient sources, to develop and make available national and international know-how to farmers, extension agents, planners and policy makers on environmentally sound new and existing technologies and soil fertility management strategies for application in promoting sustainable agriculture.

In the developed world it should be relatively easy to achieve these objectives. There is a need to develop a better understanding of nutrient interactions in crop nutrition and the optimum use of livestock wastes and mineral fertilisers in combination, but it is likely that this can be achieved by research and technology transfer. However, the issues are much less straightforward in many developing countries where the research capability, advice and finance may be lacking and it may be impossible to carry out the long-term research and development work that is needed. Moreover, socio-economic and cultural conditions may mitigate against the development of IFM. Where animal dung is used as a fuel, for example, alternative fuels will be needed together with improved systems for collection and storage of livestock wastes if such wastes are to be integrated with fertiliser management rather than used for energy. It seems likely that enabling the concept of IFM to be put into practice will require the greatest effort in those regions where the need for IFM is also greatest. According to Keatinge (1995) a critical need is the adoption of a systems approach to both conceptualise and address the problem in a development context. It is to be hoped that the development agencies will commit the substantial resources needed for long-term work of this kind.

Integrated farming

In the developed world the use of the systems approach has proved highly effective in elucidating the scientific basis from what is now called integrated farming. Integrated farming is distinguished from both conventional farming and organic farming but incorporates aspects of both (see Table 8.4). It seeks to reconcile objectives for environment and production. Integrated farming demands a high degree of technical management. It can be either extensive or intensive in nature and generally involves a trade-off between reduced yields and reduced input costs. The reduction in inputs may have environmental benefit and the reduction in costs can lead to integrated farming systems being at least as profitable, if not more profitable, than either organic or conventional farming systems. This has been illustrated in many studies and the data in Table 8.5 are based on a comparison of three farming systems in the Netherlands which were studied over a three year period (Vereijken, 1986).

Valuable guidelines for integrated farming systems have been published (El Titi *et al.*, 1993). Data indicating the environmental benefits of such systems are relatively sparse but the enhancement of biodiversity is referred to by Paoletti in the following chapter in this volume and most contemporary studies (e.g. Jordan and Hutcheon, 1993) now incorporate measurements of pesticide and nutrient leaching and in some cases also of soil erosion and provide preliminary indications of reduced pollution risk from integrated farming systems.

Currently the integrated farming approach provides the best practicable means of approaching the ideal of a sustainable agriculture. Moreover, integrated farming can be relatively easily achieved without profound

Table 8.4 Comparison of conventional, integrated and organic approaches in a mixed farming system

Management factor	*Conventional*	*Integrated*	*Organic*
1 *Soil fertility*			
Mineral fertiliser			
Form	Relatively soluble	Relatively soluble	Relatively insoluble
Amount	Economic optimum	Reduced	Reduced
Organic manures	Utilised suboptimally	Recycled efficiently	Recycled efficiently
Green manures	Not utilised	Utilised	Utilised routinely
Crop residues	Generally regarded as waste or exported off farm	Incorporated	Incorporated, composted
Biological N_2 fixation	Used suboptimally, if at all	Utilised efficiently	Utilised efficiently
2 *Weed control*			
Mechanical	Limited use	Extensive use	Exclusive use
Chemical	Extensive use	Minimal use	Not used
Suppressive crop varieties	Rarely used	Extensive use	Extensive use
3 *Crop pest and disease control*			
Resistant varieties	Used if yield uncompromised	Extensively used	Extensively used
Chemical	Economically optimal use	Greatly reduced use	Very limited use only of 'approved substances'
Biological	Used occasionally if economic	Used optimally	Used optimally
4 *Animal health*			
Husbandry conditions	Compromise between production efficiency and welfare	High emphasis on welfare	Favours 'traditional' methods
Disease control	Prophylactic use of certain drugs	Through husbandry first and medication second	As for integrated and favours 'natural' remedies

Table 8.5 Comparative economic results for integrated farming in the Netherlands (in 1,000 guilders) 1982–1984

	Farm type		
	Conventional	*Integrated*	*Organic*
Farm size ha	17	17	22
Labour man/yr	0.5	0.6	1.9
Total returns	127.8	118.2	156.7
Total costs	131.2	118.8	249.9
Net surplus*	–3.4	–0.6	–93.2

Note: * Negative values result from returns without subsidy.

modification of conventional systems. The fact that integrated farming is generally no less profitable than conventional systems, and may occasionally be more profitable, is a powerful incentive for its widespread adoption by conventional farmers. Moreover, the reduced yield of integrated farming systems provides the policy-maker concerned with surplus production in the developed world with a reason to promote this approach through research, advice and policy measures. Properly implemented these measures should enable the further development and widespread adoption of the integrated farming approach.

FOSTERING SUSTAINABLE AGRICULTURE

The concept of integrated farming is probably most advanced in Europe where a coalition of national organisations in six European countries has combined to form the European Initiative for Integrated Farming (see Table 8.6).

Whilst the primary aim of such organisations is to foster the development and adoption of integrated farming they also have a role to present a more favourable image for agriculture, but this is valid when there is evidence of environmental and other benefits. It is important that the public are properly informed about the industry, which is both the source of their

Table 8.6 European organisations promoting the principles of integrated farming

France	FARRE	Forum de l'Agriculture Raisonnée Respectueuse de l'Environnement
Germany	FIP	Fördergemeinschaft Integrierter Pflanzenbau
Great Britain	LEAF	Linking Environment and Farming
Luxembourg	FILL	Fördergemeinschaft Integrierter Landbewirtschaftung
Spain	AGROFUTURO	Gestion Integrada de Cultivos
Sweden	ODLING I BALANS	Rådet för integreated växtodling

BOX 8.2 CASE STUDY 2: ESA SCHEME IN NORTHERN IRELAND

The EU Agri-environment regulation (2078/92) is implemented through regional programmes. In several member states this has involved the designation of environmentally sensitive areas (ESA); in Northern Ireland over 20 per cent of the land area comes under such designations. The aim of the ESA scheme is to encourage environmentally friendly farming practices. The scheme is voluntary and farmers are offered financial incentives to protect and enhance the landscape, wildlife and historic features of their farms. They sign a management agreement with the Department of Agriculture covering a ten-year period but have the right to withdraw after five years. The schemes are designed to take full account of the unique environmental assets of a region and vary substantially between regions. For example, the unique features of the environmentally sensitive area scheme in County Fermanagh include:

- species rich hay meadows;
- species rich limestone grassland;
- wet pasture providing a valuable habitat for breeding waders;
- heather moorland;
- woodland and scrub;
- historic and archaeological sites.

The basic scheme includes:

- restrictions on animal stocking rates;
- limits on the use and timing of fertilisers, herbicides, pesticides and lime;
- prohibition of drainage and reclamation work;
- timing of cultivations to protect nesting birds;
- retention of field boundaries, trees and habitat features with wildlife value;
- agreement to consult the Department of Agriculture regarding various practices.

In addition, more specific prescriptions apply to particular ecosystems and habitat types. The management agreement may also cover a range of enhancement measures. Examples include hedgerow and tree planting, renovation of traditional farm buildings, regeneration of heather and creation of reed beds. Uptake of the scheme has been substantial and research is underway to evaluate the environmental benefits that are being achieved.

A similar scheme was introduced in 1988 for the Mourne Mountains, an area dominated by small beef and sheep farms (average size 18 ha). The objectives were to protect and enhance valuable landscape and wildlife features and to discourage intensification. In this ESA, farmers have received a payment of £30/ha to undertake a wide variety of environmental improvements including planting of 18,000 trees, maintenance of over 350,000 m of field boundaries and sensitive refurbishment of over two thousand traditional farm buildings. The payments made have provided the key incentive needed to enable this work to be undertaken since the cost of the work has been in excess of the ESA payments. The scheme has involved over one thousand farmers (approximately 55 per cent of the total in this area) and research has now shown that participating farmers have developed a significantly more positive attitude to conservation than their non-participating counterparts over the five-year period of this scheme. Moreover, the use of contingent valuation techniques to estimate public perceptions of the value of the ESA scheme reveals a 32 : 1 benefit to cost ratio with aspects such as protection of wildlife habitat being especially highly valued by the public.

The ESA approach provides a particularly effective mechanism for the promotion of the concepts of sustainable environmental management combined with economic farming activity.

food supply and custodian of the countryside, for public confidence and acceptance are key elements to support a sustainable industry. There is evidence that the industry is becoming more proactive in this regard, particularly with the development of open farms where the public are encouraged to experience and learn about farming practice. There is also interest in promoting the products of integrated farming by marketing with a differentiated image although this may not be entirely appropriate when the aim is to promote the widespread adoption of integrated farming rather than to establish a niche market for its products. Perhaps the most significant development is the increasing use of audit trails to enable the customer—both the food processor and the consumer— to be assured about the means of production and to have available a system of traceability in order to validate this. In future agriculture will have to become open to scrutiny by all of its stakeholders and should have nothing to hide when the principles of sustainability have become fully developed and adopted by the farming industry.

REFERENCES

Altieri, M.A. (1994) *Biodiversity and Pest Management in Agro-ecosystems,* London: Food Products Press.

Azapagic, A. and Clift, R. (1994) 'Allocation of environmental burdens by whole system monitoring: the use of linear programming', in G.Huppes and F.Schneider (eds) *Proceedings of the European Workshop on allocation in LCA,* SETAC-Europe and University of Leiden, pp. 54–60.

Bailey, J.S. (1995) 'Sustainable fertiliser use', *Proceedings of the Fertiliser Society* 343: 1–43.

Burn, A.J., Coaker, T.H. and Jepson, P.C. (1987) *Integrated Pest Management,* London, UK: Academic Press.

Carson, R (1963) *Silent Spring,* London, UK: Hamish Hamilton.

Conway, G.R. (1971) 'Better methods of pest control', in W.W.Murdoch (ed.) *Environment, Pollution and Society,* Stanford, Calif.: Sinauer Assoc. Inc.

Dyson, T. (1996) *Population and Food: Global Trends and Future Prospects,* London, UK: Routledge.

El Titi, A., Boller, E.F. and Gendrier, J.P. (1993) 'Integrated production: principles and technical guidelines', *IOBC/WPRS Bulletin* 16, 1: 96.

International Food Policy Research Institute (1995) 'Will the world starve?', *The Economist* 335, 7 918: 63–64.

Jordan, V.W.L. and Hutcheon, J.A. (1993) 'Less intensive integrated farming systems for arable crop production and environmental protection', *Proceedings of the Fertiliser Society* 346: 1–32.

Keatinge, J.D.H. (1995) 'Integrated fertiliser management: the way forward for the third world?' *Proceedings of the Fertiliser Society* 369: 1–23.

Vereijken, P. (1986) 'From conventional to integrated agriculture', *Netherlands Journal of Agricultural Science* 34: 387–393.

SUGGESTED READING

Conway, G.R and Pretty, J.M. (1991) *Unwelcome Harvest, Agriculture and Pollution,* London, UK: Earthscan.

SETAC (1991) *A Technical Framework for Life-Cycle Assessments,* Washington DC, USA: SETAC and SETAC Foundation.

Sustainable Development. The UK Strategy (1992) London, UK: HMSO.

Van Erp, P.J. and Oenema, O. (1993) 'Towards integrated nutrient management', *Proceedings of the Fertiliser Society* 345: 1–32.

Whitby, M. (ed.) (1996) *The European Environment and CAP Reform. Policies and Prospects for Conservation,* London, UK: CAB International.

SELF-ASSESSMENT QUESTIONS

1 Sustainability can best be achieved in agriculture through
(a) greater consumer orientation;
(b) understanding the meaning of sustainability as a visionary concept;
(c) financial support for clean technology.

2 The achievement of sustainability in agriculture
(a) can never be proved;
(b) is a feature of the best current practice in farming;
(c) would secure the future of world food supply.

3 An economic analysis of agriculture should
(a) consider environmental and social as well as production aspects;

(b) focus on costs and benefits;
(c) be based on the 'polluter pays' principle.

4 Integrated pest management involves
(a) use of biological and chemical control strategies;
(b) pest and disease monitoring;
(c) use of all available control strategies in optimum combination.

5 Integrated fertiliser management is
(a) practised widely in developed countries;
(b) dependent on local knowledge of soil fertility and crop nutrient requirements;
(c) the best way to avoid pollution.

6 Integrated farming is:
(a) a combination of integrated pest management and integrated fertiliser management;
(b) a model for agriculture which integrates production and environmental objectives;
(c) a special form of organic farming.

9

TOWARDS A SUSTAINABLE AGRICULTURE

Maurizio G.Paoletti

SUMMARY

The introduction of agriculture between 3,000 and 12,000 years BC enabled humans to reproduce more successfully, thanks to the greater quantities of crops, animals and fibres it made available. But with its benefits agriculture has brought problems, such as the degree of overpopulation emphasised recently by scientists in both developed and underdeveloped countries (see Chapter 1 in Volume 1 of this series). The 'green revolution' of the 1960s and 1970s has also brought benefits (increased yields with new wheat and rice cultivars, self-sufficiency in some underdeveloped countries such as India and China), but at the same time it has brought about more erosion, salinisation, pollution from fertilisers and pesticides, decreased biodiversity and new pest problems. Many crops that are conventionally farmed need a large amount of fossil energy, and the energy efficiency of our crop systems is frequently very low or indeed negative.

Sustainability seems the paradigm for the new millennium in the quest for an integrated, environmentally sound agriculture which could be socially acceptable, economically viable and require fewer external inputs. This depends, however, on better knowledge of the mechanisms of nutrient cycling in relation to biological activity in the soil, water use, sun availability, pest control and plant resistance, interactions among the components, socio-economic and anthropologic patterns, and exchange and knowledge patterns between farmers and consumers. It has been also suggested that local agricultural systems, especially in the tropics, must be better understood and developed before implementation of western prescriptions for food production.

To achieve sustainable agriculture, more emphasis must be given to the biological part of the system, including the soil microbiology and invertebrate activity. Greater emphasis must also be placed on soil invertebrates as bioindicators of input, soil management and contamination.

Genetic engineering together with biotechnology can supply new and powerful tools for agriculture. However, the promotion of new engineered plants, animals and micro-organisms needs a framework of field trials and monitoring protocols to assess the risks clearly. This framework needs to be developed properly in order to avoid possible extensive damage to other biota as well as to the environment generally, to human health and even economics. In using genetically engineered micro-organisms, plants and animals sustainable for agriculture and the environment we must find substitutes for high pesticide use, high tillage and erosion, and high fertiliser addition. We must also promote biodiversity in rural landscapes. In spite of these priorities, most trends are for herbicide resistant crops that can promote higher herbicide use or for *Bacillus thuringiensis* toxins that can control some pests but will also produce pesticide resistant pests.

ACADEMIC OBJECTIVES

This chapter provides an introduction to sustainable agriculture and the concepts that can make it possible. Readers should gain an understanding of how elements of natural ecosystems can be incorporated into agro-ecosystems, farms as complex systems, and finally some of the requirements for sustainable agriculture and its monitoring.

FROM NATURAL ECOSYSTEMS TO AGRO-ECOSYSTEMS—HOW TO DEAL WITH BIODIVERSITY

On the biosphere level, nutrients, energy and biota are associated in different combinations on the planet's surface. It is quite amazing that up to now as few as 1.3–1.8 million species of living organisms have been described, when between 7.3 and 81.4 million species have been forecast as inhabiting the planet, largely in tropical areas (Paoletti and Pimentel, 1992). The discrepancy between the number of species forecast and the number so far described implies shortcomings in our understanding of the planet's structure, composition, diversity and complexity. Moreover, the number of biota on the planet is far greater than, for instance, the number of chemical compounds artificially created by our technology. For example, about 70,000 different chemicals are presently in use in the USA (Newton and Dillingham, 1994), while an estimated 500,000 living species (plants, animals, microrganisms) are present, 95 per cent of which are small organisms (Knutson, 1989).

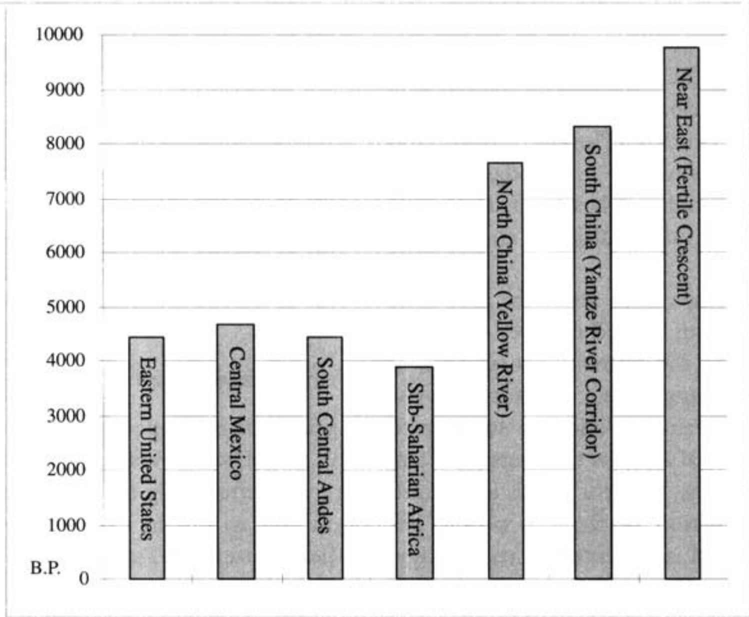
The concern shown by many scientists for the planet's biodiversity is linked to the fact that organic material equivalent to 38.8 per cent of the present net primary production of the planet is currently appropriated entirely by humans (Vitousek *et al.,* 1986). One-third of the global land surface of the planet is still wilderness (48.069×106 km2) (McCroskey and Spalding, 1989) but most of these wilderness areas have humans living in them (Western, 1989).

Around 70 per cent of the earth's ecosystems (50 per cent=agriculture including pastures; 20 per cent =forests) are manipulated to obtain 98 per cent of the food, fibre and wood used by humans (FAO, 1981; Vitousek *et al.*, 1986; Western, 1989; Pimentel *et al.,* 1992). In addition, increasing surface areas are devoted to urbanisation and human settlement.

Although 150 plants are commonly eaten worldwide, just 15 plant species provide more than 90 per cent of the world's food (Pimentel and Pimentel, 1979; Plotkin, 1988) and only three (rice, corn and wheat) produce almost two-thirds of this amount (Raven, 1988). This picture leads us to believe that we know little about the biodiversity of the planet and that we base our food chain on a very small number of plant (and animal) species. This probably dates back to the emergence of agriculture from previous hunting gathering societies. In about eight important centres of plant and animal domestication, the plants cultivated have been developed (Smith, 1995) (Figure 9.1). The domestication of wild plants and animals took place between 10,000 and 14,000 years BP. Most of our current western culture comes from what was domesticated in the Fertile Crescent: wheat, barley, cows, pigs, sheep, goats. Transforming the landscape by increasing monocultures of particular crops has decreased the number and diversity of native alternative plants and animals. In South America, for instance, the large mammals have been destroyed by the new colonising human groups (Martin and Klein, 1984). Despite the concentration on annual crops like corn, wheat and rice in many local societies, alternative plants and animals still have attractions as sustainable resources. But a decline in species number is evident in most developed countries; there are sometimes associated declines in use and knowledge. In spite of the concentration on a small number of biotic resources, some of the practice and knowledge of hunter-gatherers survives.

Apart from the use of bees for honey production, the domestication of small terrestrial invertebrates has not been widely successful. Insects are not culturally accepted as human food in most developed countries, although up to 800 species of insects are in fact eaten by humans, largely in tropical areas (DeFoliart, 1989, 1997). Other terrestrial invertebrates such as scorpions, some earthworms and spiders are also eaten in China and Amazonas. In north-eastern Italy, 56 wild herbs are collected in spring to prepare the folkloristic dish 'pistic'. Although this practice is very old (estimated to date from 3000–4000 BP), only two species in this list of plant species are actually cultivated (Paoletti *et al.,* 1995).

In the Pacific Yap Islands, farmers cultivate over fifty fruit tree species (Falanruw, 1989); in the farm gardens of Indonesia more than five hundred species are cultivated (Michon, 1983). In Swaziland, 220 wild plant species are commonly consumed (Ogle and Grivetti, 1985). A single tribe of Amazonian Indians may use more than one hundred different native species of plants for medicinal purposes alone



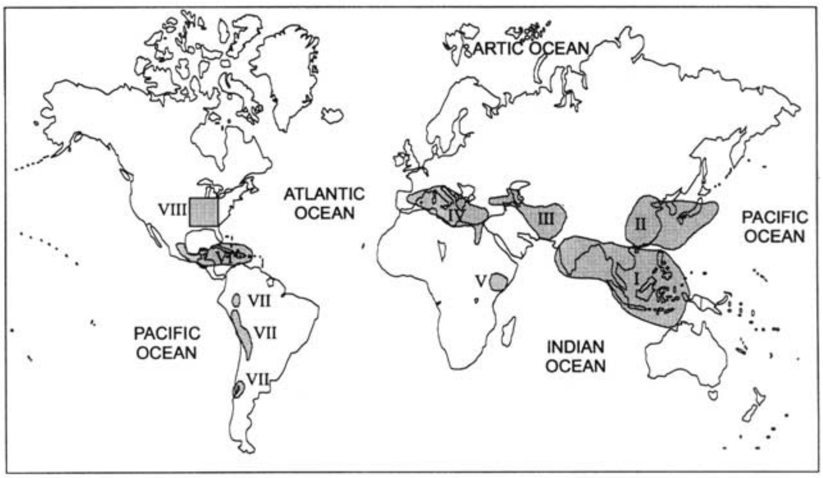


Figure 9.1 The main centres of plant and animal domestications. The bar chart shows the approximate time periods when plants and animals were first domesticated in the seven primary centres of agricultural development. Vavilov's final map, published in 1940, shows seven centres of origin of domesticated plants: I the tropical south Asiatic centre: II the east Asiatic centre; III the south-western Asiatic centre; IV the Mediterranean centre; V the Abyssinian centre; VI the Central American centre; VII the Andean centre (South America); VIII the Eastern USA centre.
Source: Smith, 1995, modified

(Prance, 1982; Plotkin, 1988).). A large number of plants (212) are or have been used by an Amazonian tribe, the Quichuas, in Ecuador (Alarcon Gallegos, 1988). At least ninety tribes have disappeared since the beginning of this century (Plotkin, 1988). And yet the loss of native biodiversity knowledge and use has gone largely unnoticed and has been underevaluated at a time when information is an essential resource for the future.

These reflections prompt us to revise the food base of our planet. The existing limited number of species currently exploited, and the tremendous number of living species available, suggest that we should review the base of our food web and adopt a strategy to develop future agricultural systems that better imitate existing ecological systems. Current trends in western countries are the reduction of farming people per hectare and per unit yield, substituted by technology, fossil energy input and chemicals. At the same time the profitable return of one hour of labour spent in agriculture is less than in other activities such as industry and commerce. Peasantry in most undeveloped (and developed) countries try to move to urbanised areas in the expectation of a better lifestyle. In most cases this urbanisation creates enormous problems of poverty and crime. Some overcrowded countries such as China strictly control movement of villagers from the rural countryside. There is a strong wish to escape from the fields.

Temperate areas

Agriculture in temperate areas has developed extensive and sometimes devastating ways of producing cereals: fire was extensively used to remove previous vegetation and to obtain quick nutrient release from the ashes (Carter and Dale, 1974). In the Mediterranean area and in many temperate European regions, native

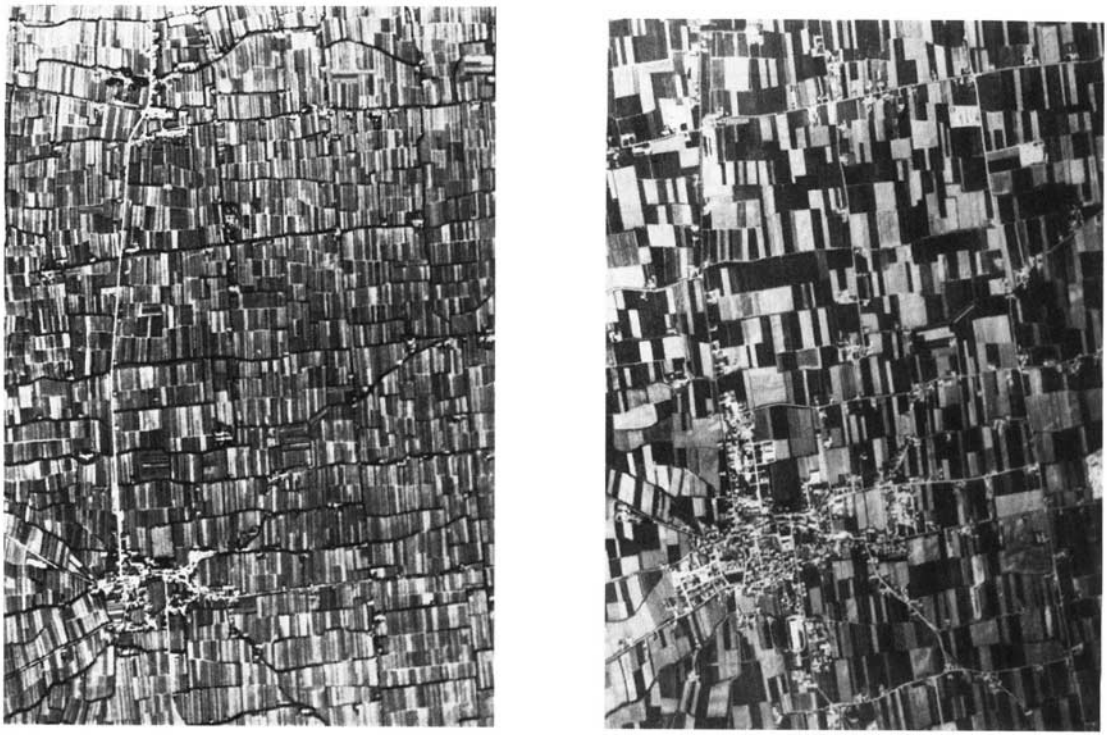


Figure 9.2 Roman centuriation near Riese, Treviso, in north-eastern Italy
The aerial photograph on the left was taken in 1955. On the right is one taken in 1981 showing decreased fragmentation of the landscape in this period.

woodlands have been cleared since previous colonisations. For instance, the Romans resettled (centuriated) the forested areas and farmed (Figure 9.2) most of Italy and part of other European regions 2,000 years ago, clearing the native forest and introducing hedgerows, as a compromise with the original forest. These hedgerows were seen as a limit for property, fences for occasional pasture, paddocks for animals and a reservoir of wood, lumber, foliage and fodder for domestic animals, windbreaks, etc. (Muir and Muir, 1987; Paoletti and Lorenzoni, 1989).

A rural mosaic landscape with hedgerows can introduce a higher species diversity and can also affect the biological control of some crop pests (Paoletti and Lorenzoni, 1989). However, the destruction and removal of hedgerows has been a common practice in all European countries. In Britain, for example, from 1945 to 1970 about 8,000 km^{-y} of hedgerows were destroyed, giving an overall loss of 25 per cent of the total hedgerow length in 25 years (Hooper, 1979).

The hedge maple *(Acer campestris)*, found in the Po Valley and some parts of peninsular Italy, is an example of relict vegetation of the original forests introduced into the agrarian landscape, and is sometimes associated with grape as an alternative to the dry stick supporting the vineyard plants.

Compared with annual crops, arboreal plantations such as peach, cherry, apple and pear trees in pure or mixed orchards and vineyards offer a more stable system of production (for reduced soil tillage, limited erosion, less depletion of soil organic substance and some higher potential for biodiversity). However, if conventionally grown (high input systems) these orchards are associated with a high use of pesticides (herbicides, fungicides, insecticides and acaricides, etc.) which can enter the groundwater, soil and atmosphere and so damage non-target organisms, humans and sometimes the plants/trees themselves (Pimentel *et al.*, 1991; Paoletti and Pimentel, 1992; Pimentel, 1997) (Table 9.1).

The introduction of permanent living mulch under the grapevines and orchards and an integrated or biological system of cultivation (by reducing or abolishing the input of synthetic chemicals) are two important priorities (Figures 9.3 and 9.4) (Edens *et al.*, 1985; Enache and Ilnicki, 1990).

Table 9.1 Mean kg/ha yearly pesticides use in different major crops in the USA

Crop	*Insecticides*	*Herbicides*	*Fungicides*
Corn	0.46	3.6	0
Cotton	2.07	2.1	12.9
Wheat	0.03	0.3	0
Potatoes	3.63	1.45	4.5
Rice	0.19	5.9	0.02
Soybean	0.2	2.3	0.01
Tobacco	5.7	1.7	0.35
Sugar beets	0.22	0.85	0.02
Tomatoes	4.5	0.56	14.12
Sweet corn	2.48	1.65	0.82
Lettuce	0.6	0.6	0.43
Apples	35.3	0.3	17.6
Peaches	7.44	0.9	23.4
Pears	29.41	0.2	4.1
Grapes	6.49	0.9	29
Oranges	30.79	6.1	5.4

Source: Pimentel *et al.*, 1991

In lowland areas, annual crops such as wheat, sugar beet, soybean and corn are frequently grown in monocultures with some herbicides and, depending on the area, a high rate of chemical fertilisers and sometimes liquid manure (except for soybean and alfalfa, which fix nitrogen). These high input cropping systems can often be associated with groundwater pollution and eutrophication of the water bodies. In particular sugar beet production involves a high rate of herbicides and fungicides, and tillage operation for seedbed preparation must be very intensive. All these operations cause environmental problems: environmentally speaking, sugar is bitter.

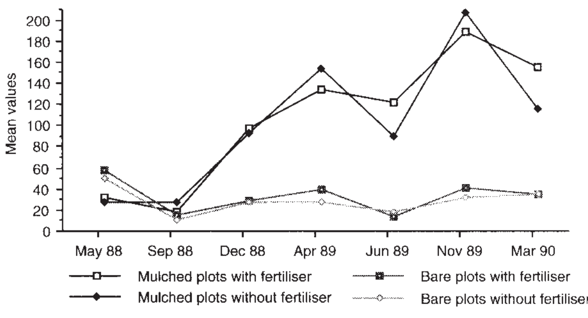


Figure 9.3 Increased invertebrate biomass after the introduction of living mulch *(Trifolium subterraneum)* in a Mediterranean vineyard in Tuscany, Italy

Figure 9.4 Mulched and non-mulched *(Trifolium subterroneum)* vineyard in central Italy. Mean value of mesoinvertebrates found in each sampling date

Mediterranean areas

In the Mediterranean region low precipitation, fire, overgrazing and erosion have dominated many areas, especially in the past and mostly in the southern belt, for instance, Palestine, Southern Italy, Spain and the islands (Carter and Dale, 1974; Golley *et al.*, 1989). Olive *(Oka europaea)*, nut *(Juglans regia)*, fig *(Ficus carica)*, almond *(Prunus communis)*, pistacchio *(Pistacia vera)* and especially citrus-fruit orchards in certain areas are part of a complex three-dimensional system in which trees, shrubs and herbal plants are associated in a multilayer-multicrop pattern. This may be observed in some parts of the volcanic and very fertile areas in southern Italy (in the Agro Nocerino near Naples some relict situations still exist) (Figure 9.5). In many other places in central Italy, for instance, there are mosaic landscapes that have barely changed since the Renaissance (Figure 9.6).

Very ancient compromises between the original forest and annual crops can be seen in some places of the Gallura (Sardinia, Italy) but also in Portugal and Spain, where one can find a thin *Quercus suber* plantation, with cereal (rye) underneath, and the straw after harvest grazed by sheep (Figure 9.7). In such cases erosion and runoff could be minimised, but to my knowledge no experimental data are available. Terraced agriculture is sometimes fundamental to the control of erosion, nutrients, water and microclimate. Extensive terracing operations have been performed in many Mediterranean and tropical areas, but little is known about biological diversity (Carter and Dale, 1974; NRC, 1989).

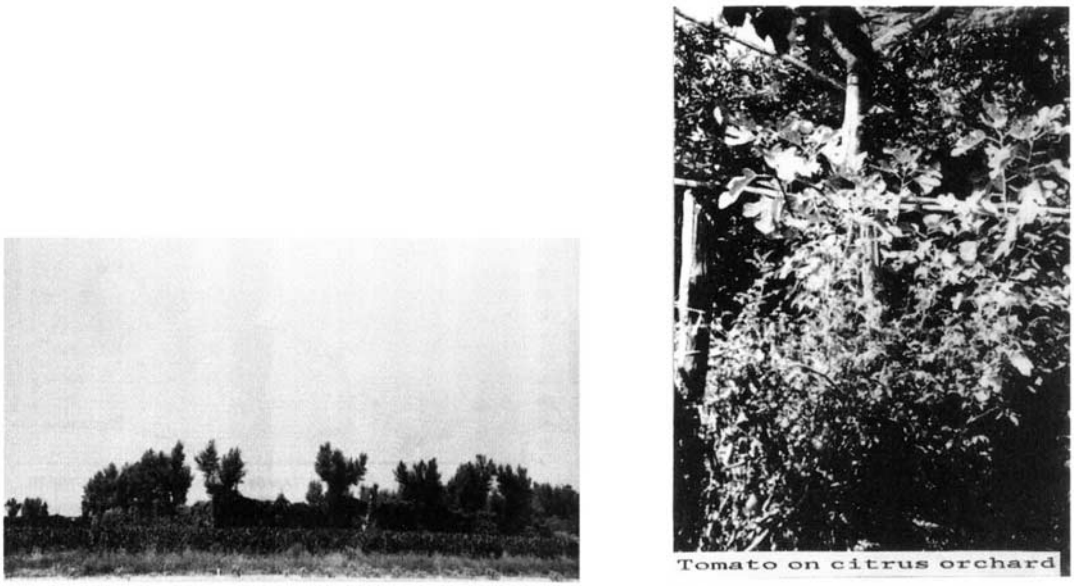


Figure 9.5 Agro Nocerino (Nocera Inferiore, Naples) multilayer vegetation: some relics of a multilayer agriculture

Figure 9.6 Tuscany's Renaissance-like landscape—fresco by Lorenzetti, 1438: 'II buon governo'

Humid tropical areas: the gardens

Only recently S.Gliessman of the University of California at S.Cruz (Gliessman, 1990a), M.Altieri at Berkeley (Altieri and Merrick, 1988), A.Gomez-Pompa at Riverside (Gomez-Pompa *et al.*, 1987) and a few others have underlined some unique structures of traditional farming systems in tropical America. Most traditional forms of agriculture are deeply involved with plant diversity (number of species considered in a multilayer set) and the use of arboreal and weed populations for different purposes including the biological control of pests. Gliessman (1990a) has studied the gardens in Costa Rica which are extremely rich in species numbers (up to 83) and designed in different strata. Hedgerows as an integral part of the landscape are not confined exclusively to temperate areas. Also in the tropics hedgerows have the multipurpose effect of producing biomass, microclimate, lumber, and in general an increased biodiversity within the system (Gliessman, 1990b).

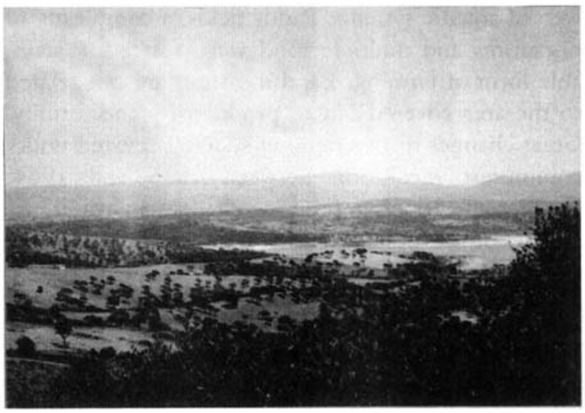


Figure 9.7 Gallura rural landscape in Sardinia, Italy. *Quercus suber* thinly planted, with cereal crops under canopy and occasional flocks of sheep grazing after harvest

Multicrops and alley crops in the traditional forms of agriculture are more the rule than the exception (Gliessman, 1990a). Some archaeological evidence of biodiversity in the Andean regions has been collected (Bray, 1990). Before their basic crops (corn and beans) are ready for harvest, the Tarahumara people in Mexico collect the edible weeds from their fields as a useful integration to their diet (Altieri and Merrick, 1988). Weediness seems a prerequisite for domesticating plants. In fact cultivated plants such as peas, lentils, rye and oats have been evolving as weeds in domesticated crops such as barley and wheat in the Middle East Fertile Crescent (Zohary and Hopf, 1993). Maybe the Andean region, hosting one of the more important areas of domestication (corn, potato, sweet potato, pepper, beans and tomato) has the highest potential, with several undeveloped crops. Over ninety species that are in use locally could be clutivated world-wide if breeding programmes are developed (AA, 1989b; Janick and Simon, 1993). Of 235 Solanum species only one, the potato *(S. tuberosum)*, has been spread world-wide but at least six additional species are edible. Some Solanum

BOX 9.1 THE SAGO CULTURE IN NEW GUINEA, THE ASMAT CASE: A TREE TO PRODUCE BREAD

The Asmat are a Papua group of about 70,000 people. Their territory (c. 30,000 km^2) is located in the south-western part of Irian Jaya (Indonesian New Guinea) and extends from the coast of the Arafura Sea to the interior foothill region. The area's wide alluvial swamps are covered with tropical rain forest and crossed by a dense network of rivers. Of the Asmat population's diet 80 per cent is based on sago, a food obtained from the starchy pith of a palm *(Metroxylon rumphii and M. sagu)*. This grows spontaneously in the humid forest but is also sometimes cultivated, for instance, by the Baroi peoples in Papua New Guinea. In Malaysia other sago species such as Eugeissona sp. *Arenga* sp. *Caryota* sp. are used for the same purpose. To extract the edible part of the sago, Asmat fell the palm, crush the stipe's pith with a kind of adze and then wash it in a trough. The liquid they obtain is filtered and the starch is then collected in a draining pan made out of the plant's leaves. Following Peters 100 g of raw sago contain 27 g water, 71 g carbohidrates, 0.2 g proteins, 0.3 g fibre, 30 mg calcium, 0.7 mg iron, traces of fat, carotene, thyamin and ascorbic acid. There are also discreet quantities of phosphorus and potassium. Although sago is high in calories, its plastic value is limited, and the diet has to be supplemented with the products of fishing, hunting and insect collection. Various insects are also gathered, and one palm grub in particular (Coleopptera, Curculionidae, *Rhynchophorus ferrugineum*). The larvae develop quickly on sago palms. In the Pir-Jimi or 'larvae feast', the palmworm larvae are 'grown' in the fallen trunks of sago palms. After a few weeks' growth they are gathered in large quantities—up to several kilos—to be consumed in the course of the ritual.

In addition, many other insect species are collected from trees, bushes and soil. Sometimes invertebrates support up to 30 per cent of the meat diet requirements otherwise met by fishing and hunting. This is strikingly different from the western model characterised by large ruminants and intensive food production systems. Asmat and other New Guinea peoples, as well as many Amerindian tribes, still live in forested areas, having promoted local diversified resources. In Southern Vietnam, for instance, similar palmworm larvae (collected in bamboo shoots) are eaten by children and villagers, especially near the forested areas where ethnic minorities live (personal observations, October 1996). It is of prime importance to focus on these traditional uses of local biodiversity and mini-livestock in order to maintain and promote the kind of local knowledge that is lost when these minorities come into contact with western civilisation.

species producing fruits *(S. quitoense, S. sessiliflorum)* could be potentially interesting new crops.

East Asia: China

The huge continent of China, with 1.3 billion people, has evolved an interesting and highly efficient agricultural system centred on the Yangtze River and Yellow River basin flooded areas. An aquatic staple crop, rice, has been developed in combination with a set of other aquatic, edible plants—yams (*Discorea),* taro *(Colocasia),* water caltrop *(Trapa bispinosa),* foxnut *(Euryale ferox),* lotus plants *(Nelumbium speciosus)* and water spinach *(Ipomaea. aquatica)* — and aquatic organisms including many fish (at least three herbivores carpas and *Melanobrama*), crabs, shrimps, frogs and turtles (Bray, 1984). The irrigation needed for rice paddy cultivation implies a highly organised and sophisticated bureaucracy and a high degree of discipline. Taxes are paid with rice, and China has developed as an empire using these taxes. But carbohydrates and proteins from paddy fields and ponds are essential in the farmers' diet. The Chinese peasant has developed a very sophisticated strategy to use a plurality of organisms, most of which live in the food web of aquatic systems. Paddy fields, a complexity of organisms and multiple food webs foster a sustainable form of farming. Land measures are not related to the area covered but to productivity and fertility. Great changes to this farming system occurred under communist revolution between 1955 and 1984: changes in cropping systems, changes in water use (reduction of ponds); most parts of the landscape mosaics along the Yangtse River in Hubey have been changed in use (Yu and Baudry, 1998). Instead of large ruminants that need fodder and grassland, ducks, chickens and pigs, which are omnivorous, are reared. Buffalo are used only for ploughing and harrowing, but not for meat or milk production, and at present are low in numbers (Smith, 1995). One highly proteic leguminous plant, soybean, helps in replacing the meat. Soybean is processed to produce the beancurd (tofu) which through its consistency and dietary role takes the place of dairy products. A simple stone mill and chock is easily available to each household for tofu production. Composting, incorporating different vegetal waste materials with animal and human excrements, is much used. Night soils are currently intensively collected for manuring, for instance, the different cabbage varieties in the backyard gardens.

Nutrient tests on soil demonstrate decreasing fertility from backyard garden to rice paddy field to the water fed areas. The distribution and spread of compost follows a precise strategy.

In some rural areas the peasants' sex ratio can reach five males per one female, and this is the visible pressure of limited resources and high population. In some subtropical areas such as Hubey the land per capita is around 800 m^2(!). About 750 million people still live as peasants in the countryside and are compelled not to move to the urbanised areas and the large cities.

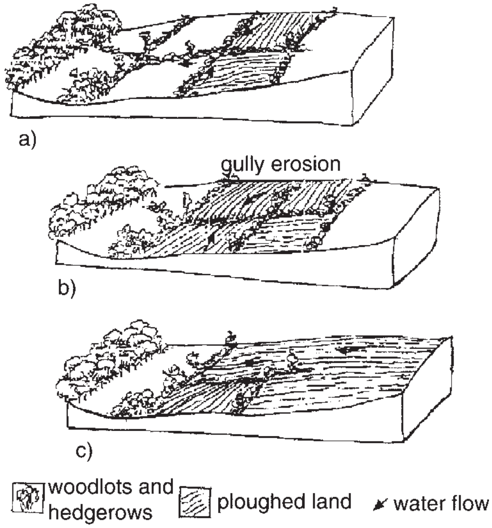


Figure 9.8 Agro-ecosystem structure, evaluation and erosion
Source: Baudry, 1989

THE FARM AS A COMPLEX SYSTEM INSERTED INTO THE LANDSCAPE

In recent times landscape ecology (Forman and Godron, 1986; Paoletti *et al.,* 1989; Bunce *et al.,* 1993) has helped to place the farm unit (agro-ecosystem) in a vast context: the water catchment basin of a river, a regional area, etc. In particular, much of the work in landscape ecology has emphasised continuity and discontinuity of components (physical, animal, plant, etc.). It has also helped to focus on the scale of the different phenomena and the links between structure, function and fate of one landscape. Erosion, for instance, is linked with landscape structure and evolution (Figure 9.8).

Although not new, the idea that the farm is a complex system in many ways similar to an ecosystem

BOX 9.2 AMAZONIAN FORESTED AREAS

There are few references to local crops in the lowland Amazonian landscapes (Posey, 1993) and their possible future. In these areas little or no archaeological evidence exists regarding domestication of plants like yucca *(Manihot esculenta),* tupiro *(Solanum sessiliflorum),* pijiguao *(Boctris gasipaes),* cocurito *(Calathea allouia)* and others (see Hernandez and Leon, 1994). A few crops, such as ocumo *(Xantosoma)* are annuals; most crops of slash and burn agriculture are poliannuals. Sometimes the crops are highly toxic for metabolites (for example, cyanogenic glycosides in yucca, alkaloids in tupiro leaves). Locally evolved technologies have been developed to detoxify the starchy roots of yucca after harvest. Varieties that are known to have low glycoside content, and which consequently require no intensive labour before eating, are not extensively used, despite being demonstrably easier to produce and consume. Among the possible (but not certain) reasons for this permanent use of toxic yucca are defence from parasites; strategies against neighbouring enemies who, because of its toxicity when raw, cannot use toxic yucca as food when assaulting the village; different taste and higher productivity. It should also be stressed that in horticultural communities polyannual plants have been dominant compared with the annuals, and such trees and bushes are abundant. Moreover, the semi-domesticated status of many plants is latent in Amazonia. Use of mini-livestock such as the *Agoutipaca* and *Dasyprocta leporina,* could be developed, as could the use of some insects and iguanas (Paoletti and Bukkens, 1997).

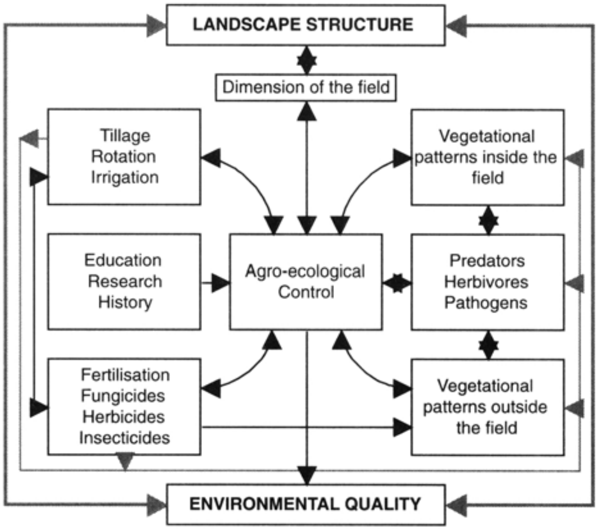


Figure 9.9 Flow chart model of interactions between agro-ecosystem structure and external and internal inputs

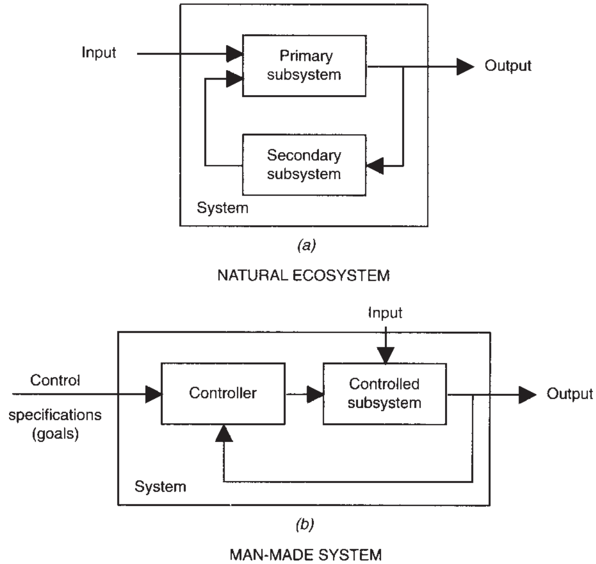


Figure 9.10 Natural ecosystems (a) are controlled by diffuse subsystem feedback in contrast to organisms and man-made systems (b) which have goals or set-points

Source: Odum, 1984

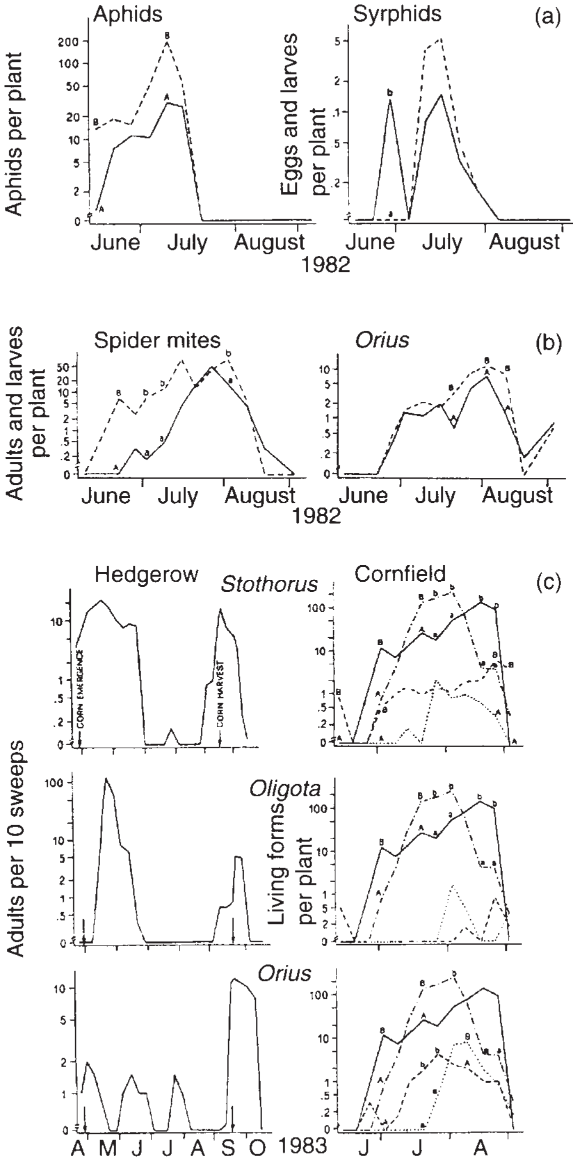
(agro-ecosystem) is not only a question of terminology but implies that subsystems like weeds, spontaneous vegetation, insects such as pollinators, polyphagous predators, detrivores and parasitoids, birds and micro-organisms are linked in a food web with the main crop and the economy of the farm and the landscape structure, water bodies, woodlands and marginal lands (Figure 9.9). Cycling of energy and nutrients in the complex system is also appreciated as part of the proper management of the farm (Edwards, 1989). In this way the introduction of inputs such as most pesticides and chemicals must be decreased and alternative management strategies such as natural inputs, leguminous plants, cover crops, catch crops, rotation, reduced tillage and multicrops adopted (AA, 1989a; Exner *et al.*, 1990; Francis *et al.*, 1990; Janke *et al.*, 1990; El Titi, 1992).

Hedgerows, natural vegetation, weeds and animals in agro-ecosystems

For many years agriculturists have underestimated wild plants and animals as a non-target part of the economy of agricultural activities. Technology, chemicals and energy were applied to normalise problems caused by this bias (Francis *et al.*, 1990). Costs for the environment and consumers and a decreased income and energy efficiency for farming activities have suggested that such a view should be revised (Lockeretz *et al.*, 1981; Poincelot, 1986; Lockeretz, 1990; Madden, 1990; USDA, ICAR, FERRO, 1991) by adopting more sustainable strategies.

Figure 9.10 focuses on the more consistent differences between a given natural ecosystem and an agro-ecosystem. If properly managed, subsystems that interact with the production system's natural vegetation, hedgerows and soil biota can introduce some beneficial effects into the agro-ecosystem and can moderate pest effects. These effects include overwintering refugia for poliphagous and specialised predators, parasitoids and pollinators.

Hedgerows may introduce some pests into the system, for instance, *Brevocoryne brassicae* (Pollard *et al.*, 1974) but various predators are also present and commute efficiently through nearby crop fields (Nazzi *et al.*, 1989; Paoletti and Lorenzoni, 1989). Some beneficial effects of hedgerows have been demonstrated in Italy (Figure 9.11) as well as in other EEC countries (Muir and Muir, 1987; El Titi, 1992; Burel, 1996).



Ground beetles and other predators in the soil increase in number according to the hedgerow microclimate. Detrivores such as earthworms, isopods, diplopods and soil mesofauna can also be affected, increasing the turnover of organic debris on the soil surface near hedgerows. Furthermore, soil microbial biomass (Jenkinson method) is more abundant in the vicinity of hedgerows than in fields, and more abundant in lucerne than in annual crops such as corn; microbial biomass can be correlated with detrivores in soil (unpublished data for north-eastern Italy).

Soil biota, soil activity and processes of fertility

For a long time in the history of agriculture the presence of invertebrates in agricultural soils was regarded unfavourably. Insects and worms in the soil were associated with crop pests, with few exceptions: for instance, Vincenzo Tanara, in 1644, defined agrarian soil fertility: when birds such as ravens, magpies and others are attracted to a freshly ploughed field and scratch on the soil to eat small invertebrates (mostly earthworms) (Tanara, 1731). Charles Darwin, with great scientific authority, helped improve the reputation of soil biota in general and of earthworms in particular (Darwin, 1881). The important role played by the biotic component of the soil (Paul and Clark, 1989; Edwards *et al.*, 1990; Crossley *et al.*, 1991) in the processing of nutrients, plant growth and sustainability of agricultural activities is now recognised. Although this role is still not given much consideration by current agricultural practices and conventional technologies, it is increasingly taken into account by alternative systems such as organic farming (AA, 1989a).

Figure 9.11 Poliphagus predators and herbivores affected by hedgerows in northern Italy. Hedgerows' effects on invertebrate dispersal in a cornfield, S.Dona' di Piave (1982–1983). (a) Aphids and syrphids near the hedgerow (solid line) and 36 m from the hedgerow (dashed line) in the cornfield, (b) Two-spotted spider mites (TSSM) *(Tetranychus urticae)* near the hedgerow and 36 m from it in the cornfield (dashed line); the same for an anthocorid generalist predator, *Orius,* (c) On the left: the presence of some predators of TSSM in the hedgerows. On the right: comparison of TSSM in the cornfield near the hedgerow (solid line) and in a corn monoculture without any spontaneous vegetation (dash and dot line). The predators of TSSM are also reported in the cornfield near the hedgerow (dashed line) and in the corn monoculture without spontaneous vegetation (dotted line). Different capital letters indicate significant differences between sites (P<0.05), different small letters indicate highly significant differences (p<0.01) using Duncan's test.

Source: Paoletti and Lorenzoni, 1989

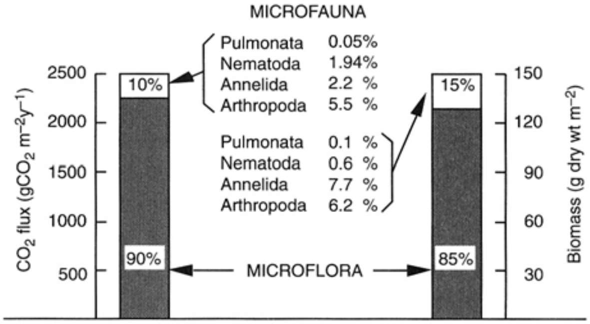


Figure 9.12 Comparative biomass and CO_2 flux for the microflora and faunal decomposers in a deciduous forest litter

Source: Paul and Clark, 1989

The soil living biomass and respiration rate for a forest is reported in Figure 9.12.

Except in a few cases (Crossley *et al.*, 1991), food web structures in soils are very poorly understood. Only recently has emphasis been given to the relationship between mycorrhizae (VAM=vesicular arbuscular mycorrhiza) and soil invertebrates (Rabatin and Stinner, 1988). Also, biotron and rizotron (automatic cameras inserted in the soil) have demonstrated the consistent and close interaction between soil invertebrates and the root terminations (Lussenhop *et al.*, 1991).

Soil is the base and product of a complex of living creatures. Plant roots and living processes in soils are strictly correlated. When chemicals such as pesticides are broadcast in agro-ecosystem soils the effects can be impressive.

SUSTAINABLE AGRICULTURE

Definition of sustainable agricultural systems

It is not easy to define the paradigm of sustainable agriculture. It is perhaps easier to define what it is not.

Many definitions have been tried focusing from different viewpoints. Sometimes definitions are controversial because since 1986–1989, the term ‘sustainable agriculture’ has been adopted by most funding agencies as a prerequisite for project funding; as a result the practical meaning has consistently broadened and has sometimes become confused. For instance, most agriculturists attracted by the achievements of the green revolution equate sustainability with food sufficiency, in which case sustainable agriculture can embrace any means to that end. For economists, sustainability is eminently a matter of efficiency in the use of resources; for environmentalists sustainability means responsibility for the environment and the best use of natural resources (Conway and Barbier, 1990). After the Rio Conference, June 1992, the topic of sustainability became more linked with biodiversity protection (Nath *et al.*, 1996).

Table 9.2 (modified from Francis *et al.*, 1990) shows the historical development of agriculture, and the last step tries to define sustainable post-industrial agriculture. The idea is more a process than a static definition. The process is not associated with pre-industrial agricultural systems, although the agricultural history of a given area provides information for the selection of some strategic ideas and options. New technology and new scientific knowledge of the processes involved are necessary as a basis of any future transformation of conventional, unbalanced farming systems.

The focal challenge for sustainable agriculture is to be economically viable, environmentally sound and socially accepted. In practice it is not easy to satisfy these three requirements. Because most funding agencies have adopted sustainability as their basic reference, the placement of most activities under this heading is misleading. Most current research in agriculture conducted within this nominal framework is irrelevant to this criterion. The problem is to evaluate regionally which farming system is most sustainable and which policies can improve the situation (Francis *et al.*, 1990). The following steps are necessary: a conparticipatory approach to local rural communities; indepth research on alternatives to current trends; in-depth examination of the human food web (the few crops in monoculture and the few large ruminants of basic livestock) and its constraints and destructive effects on biodiversity. In many cases, local communities in the tropics have a better knowledge of the optimal use of their resources; the problem arises when these communities are confronted by western technology and agricultural methods, incorporating alien plants, animals and lifestyle.

Table 9.2 Phases in agricultural evolution and their characteristics

	Institutional development	*Participation*	*Technologies*
1 Hunting/gathering subsistence	Minimal	Entire population	Biological only
2 Participatory, partly commercial	Mostly local, concerned with land use; some markets starting to develop	Dominant portion, increasing with population growth	Mostly biological
3 Participatory, commercial	Generally equal distribution of national, regional and local levels	Dominant portion but decreasing	Well-developed small-scale machinery and other industrial inputs
4 Industrial	Global	Small number actually producing	Industrial inputs substituting for biological
5 Post-industrial (sustainable agriculture)	Blend of local, regional and global institutions based on balance of resource use efficiency and local employment	Stable	Biological controls predominant; systems interactions are intricate and quantified

Table 9.2 Phases in agricultural evolution and their characteristics

Most organic farming is more sustainable than conventional farming based on chemicals and monocultures. However, the market's requirement of unblemished and colourful fruits and vegetables sometimes creates agro-ecosystems stressed by inputs such as 'organic pesticides' and intensive tillage in place of chemical herbicides. The problem for sustainability in the rural landscape encompasses a radical revision of crops, resistance to pests and pathogens, recovery of new crops, and the transformation of the current food web. In addition, in many tropical countries allocation of the land to a few landowners facilitates the promotion of highly subsidised monocultures for exportation and heavy use of chemicals.

In many developed countries (for instance, the Netherlands and Sweden in the EU) policies have been designed to cut pesticide use consistently in agriculture or to reduce fertilisers, tillage operations, monocultures, some polluting crops, to develop crop rotation, etc. Some of these options have also been promoted by the European Union. It is not, however, clear how, when and by whom this desirable transition will be effected (Jordan, 1992; Pimentel, 1997). It is of strategic importance to create policies and to assess the environmental improvements subsequent to the new policy measures (Van Straalen and Krivolutsky, 1996).

How to monitor sustainability

Using biota to produce food, fibre and lumber is the major target of agriculture. The assessment of different agricultural systems using biota (biodiversity tools) is not well developed, but is nevertheless both important and promising. Some recent authors have tried to monitor the sustainability of farming systems using soil organisms as a tool for comparing high versus low input agricultural systems; in other words, to assess simultaneously pesticide residues, tillage, fertilisers, irrigation, rotation and other features such as landscape porosity and landscape evolution. In fact, the evaluation of sustainability seems to be the central item in a new, sound approach to agriculture (and other human activities) especially where the policy adopted has to be evaluated, benefits measured and subsidies to farmers discussed and optimised.

In recent years some research has demonstrated the usefulness of soil biota, especially invertebrates (Foissner, 1987; El Titi and Ipach, 1989; Matthey *et al.*, 1990; Werner and Dindal, 1990; Paoletti and Bressan, 1996). In general, biomass, species number and frequency have been adopted. In the low input sustainable systems, a higher number of species, higher biomass and biodiversity is generally found (Figure 9.13 and Table 9.3). In any case, more

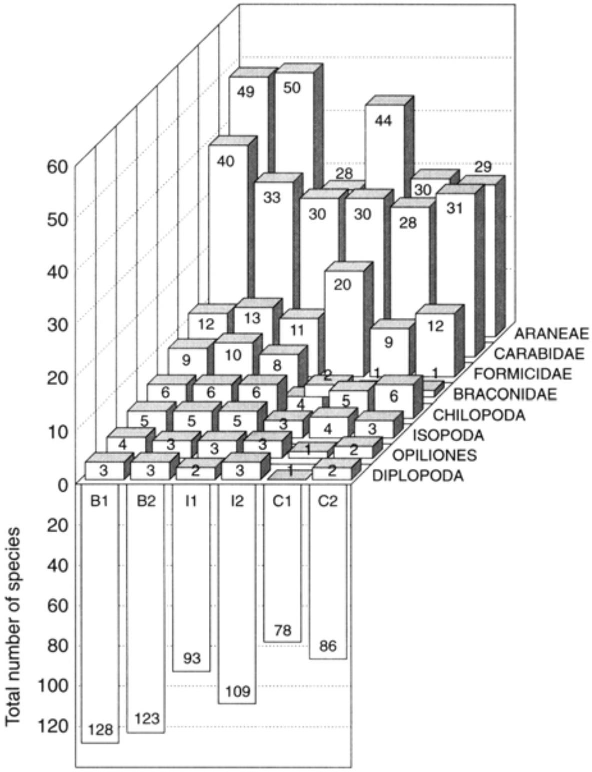


Figure 9.13 Biodiversity assessment of differently farmed orchards. Species number in differently farmed peach orchards (Forti, Italy). B1 and B2 are biological farms; I1 and I2 are integrated pest-management farms; C1 and C2 are conventionally farmed orchards (high input).

research in this field is regarded as necessary to better evaluate local situations.

Designing a sustainable farming system

The introduction of spatial and temporal variability in the conventional monocultural landscape seems to be a consistent trend in increasing sustainability (Table 9.3).

In many tropical areas, for instance, polycultures are the most familiar way of handling crops (Gliessman, 1990a). Terracing and water use was effectively applied in the Incas' very efficient agriculture (AA, 1989a). In this way some part of the original tropical diversity has been maintained in the agro-ecosystems (Paoletti and Pimental, 1992). In the same manner,

Table 9.3 Farming systems which can increment biodiversity in agro-ecosystems

Sustained biodiversity	*Decreased biodiversity*
Hedgerows (1,2)	Wild vegetation removal
Dykes with wild herbage (1,2)	Tubular drainage or vegetation removal
Polyculture (3)	Monoculture
Agro-forestry (3)	Monoculture
Rotation with legumes (4)	Monosuccession
Dead mulch, living mulch (4,5)	Bare soil
Strip crops, ribbon cropping (6)	Conventional cropping
Alley cropping (6)	Monoculture
Minimum, no-tillage, ridge tillage (5,7,8)	Conventional plowing
Mosaic landscape structure porosity (9,10,11)	Landscape simplification, woodland clearance
Organic sustainable farming (4,12)	Intensive input farming
On-farm research (13,14)	Conventional plot research
Organic fertiliser (4,12)	Chemical fertiliser
Biological pest control (15)	Conventional chemical pest control
Plant resistance (15)	Plant susceptibility
Germoplasm diversity (3,16,17)	Standardisation

Notes: (1) Paoletti and Lorenzoni, 1989
(2) Favretto *et al.*, 1991
(3) Altieri and Merrick, 1987
(4) Werner and Dindal, 1990
(5) Stinner and House, 1990
(6) Personal evaluations (Paoletti, 1987–1990)
(7) Exner *et al.*, 1990
(8) House and Rosario-Alzgaray, 1989
(9) Paoletti, 1988
(10) Noss, 1990
(11) Karg, 1989
(12) Matthey *et al.*, 1990
(13) Stinner *et al.*, 1991
(14) Lockeretz, 1987
(15) Pimentel *et al.*, 1991
(16) Lal, 1989
(17) Oldfield and Alcom, 1987

Source: Paoletti and Pimentel 1992

the ancient Romans cleared the forests but, with centuriation, they introduced hedgerows, thereby finding a compromise with the original deciduous primary forest system. In tropical areas, it can sometimes be difficult to see the difference between secondary forest and some gardens, because of the large number of species associated in a multilayer structure (De Jong, 1997). Most subtropical Chinese agriculture is still a polyculture. However, when complexity is increased, and the number of associated species increase, technical and management problems increase correspondingly; most technology has been developed for monocultures rather than polycultures. National policies have also largely promoted monoculture (Pimentel, 1997). In designing efficient, economically and environmentally sustainable farming systems, it is important to integrate natural ecosystems with agroecosystems. However, in view of the ever-increasing demographic pressure, it is not possible to maintain any form of sustainability of resources, including that of agricultural landscapes. Control of the human population can increase sustainability on a large scale by placing less pressure on our natural resources (e.g. the Amazonian basin, but also China, India, Africa and other developed areas, including the EU and USA). Socio-political problems have to be solved before natural forces change the situation for the worse (Pimentel, 1996). It has been argued that through a conparticipatory approach non-governmental organisations (NGOs) might in the future facilitate the implementation of a more sustainable trend, especially in balancing the different needs and options of rich and poor countries.

Genetic engineering and sustainability

Although traditional breeding has (sometimes strongly) modified plant structure and productivity by crossing and selecting taxonomically close species, it is not able to cross disparate organisms. Genetic engineering technologies have made it possible to introduce alien genes into an almost unlimited range of organisms. This makes a big difference. Genetically modified organisms (GMOs), including mico-organisms, animals and crop plants, will soon be brought on to the market; agriculture will become an arena of genetically modified organisms. Over 2,500 field trial releases of different transgenic crop plants have been made in different countries since 1986. The features most commonly engineered are these:

1 Herbicide tolerance (47 per cent).
2 Insect resistance (25 per cent).
3 Altered product quality (20 per cent).
4 Virus resistance (17 per cent).
5 Bacterial and fungal resistance (5 per cent).

About fifty different plants, including the main crops (rice, corn, potato, soybean, wheat, apples, etc.) have been engineered for one or more characteristics. Genetically engineered plants differ from those bred traditionally by humans for several thousands of years because target genetically engineered crops can incorporate genes from micro-organisms, animals and completely different plants. For instance, one of the more often used genes engineered into crops is the *Bacillus thuringiensis toxin*. This biological pesticide has for a long time been used in agriculture as a target pesticide against some lepidopteran pests. BT toxin (different strains) seems specific to some target coleoptera dipteran and lepidopteran pests, with low environmental side effects. In practice these modified crops (especially cotton and corn), which become 'toxic' to some key insect pests, have been criticised because they can quickly induce resistance in the key pest. Moreover, little is known about possible effects on non-key pests, or crop residues in the soil.

The introduction of (GMO) crops resistant to herbicides (for instance, soybean, corn, canola) prompts the use of target herbicides, possibly in higher amounts (like glyphosate). Genes of herbicide resistance can escape from crops to parental wild weeds (the possibility exists for canola in Europe) and possibly cause new weeds to grow. In addition glyphosate is highly toxic to aquatic and terrestrial organisms (Paoletti and Pimentel, 1995, 1996). Modifying plants to resist viruses seems a good idea. However, modifying viral pathogens to improve plant resistance could create new pathogenic viruses (Paoletti and Pimentel, 1996).

Sustainable and desirable use of genetic engineering, and of biotechnology generally, would imply reducing the use of pesticides, reducing erosion, reducing fertiliser application, reducing water needs, improving pest and pathogen resistance without losing nutritional values (Paoletti and Pimentel, 1996). A priority in western countries would be reducing the heavy use of fungicides for crops such as grape, apple and pear. Viewed according to these desiderata, the engineered crops currently released are discouraging.

Strong criticisms are also levelled against certain livestock releases such as the bovine growth hormone (BGH) injected in cows to produce more milk. Genetically engineered bovine growth hormone has the potential to increase milk production in dairy cattle by as much as 40 per cent . Although the FDA reports that the presence of the hormone in milk to be consumed by the public, especially children, is safe, there is a concern about the impacts of this technology on the health of the animals and its indirect effects on humans (antibiotics). Evidence exists that using BGH in cattle increases the chances of bacterial infections, such as mastitis, in dairy cattle. The trend of not labelling engineered food is also highly controversial, especially in the USA, where these technologies were first developed. Some debates have also developed in Europe, for example, when unlabelled soybean from the USA has come into European ports.

REFERENCES

AA (1989a) *Alternative Agriculture,* Washington DC, USA: National Academic Press.

AA (1989b) *Lost Crops of the Incas,* Washington DC, USA: National Academic Press.

Alarcon Gallegos, R. (1988) 'Etnobotanica de los Quichuas de la amazonia Ecuatoriana', *Misc., Antropologica Ecuatoriana,* Museos del Banco Central del Ecuador, Guayaquil, 7: 1–178 (in Spanish).

Altieri, M.A. and Merrick, L.C. (1988) 'Agroecology and in situ conservation of native crop diversity in the Third World', in E.O.Wilson, *Biodiversity,* Washington DC, USA: National Academic Press, pp. 361–369.

Andrews, R.W., Peters, S.E. and Janke, R.R (1990) 'Converting to sustainable farming systems', in C.A.Francis, C.Butler Flora and L.King (eds) *Sustainable Agriculture in Temperate Zones,* New York, USA: J.Wiley & Sons, pp. 281–313.

Baudry, J. (1989) 'Interactions between agricultural and ecological systems at the landscape level', *Agriculture, Ecosystems and Environment* 27: 119–130.

Bray, F. (1984) 'Biological sciences: agriculture', in J.Needham (ed.) *Science and Civilization in China,* Cambridge, UK: Cambridge University Press.

Bray, W. (1990) 'Agricultural renaissance in the Andes', *Nature* 345: 385.

Bunce, R., Ryszkowsky, L. and Paoletti, M.G. (1993) *Landscape Ecology and Agroecosystems,* Boca Raton, Fla., USA: Lewis.

Burel, F. (1996) 'Hedgerows and their role in agricultural landscapes', *Critical Reviews in Plant Sciences* 15, 2: 169–190.

Carter, V.G. and Dale, T. (1974) *Topsoil and Civilization,* Oklahoma, USA: University of Oklahoma Press.

Conway, G.R. and Barbier, E.B. (1990) *After the Green Revolution. Sustainable Agriculture for Development,* London, UK: Earthscan.

Crossley, D.A., Coleman, D.C., Hendrix, P.P., Cheng, W., Wright, D.H. and Beare, M.H. (1991) *Modern Techniques in Soil Ecology,* Amsterdam, The Netherlands: Elsevier.

Darwin, C. (1881) *The Formation of Vegetable Mould through the Action of Worms with Observations of their Habits,* London, UK: Murray.

DeFoliart, G.R, (1989) 'The human use of insects as food and as animal feed', *Bull. Entom. Soc. Am.* 35: 22–35.

DeFoliart, G.R. (1997) 'Overview of the role of edible insects in preserving biodiversity', *Ecology of Food and Nutrition* 36, 2–4: 109–132.

De Jong, W. (1997) 'Developing shifting agriculture and the threat of biodiversity loss', *Agriculture Ecosystems and Environment* 62, 2–3: 187–198.

Edens, T.C., Fridgen, C. and Battenfield, S.L. (1985) *Sustainable Agriculture and Integrated Farming Systems,* East Lansing, Mich., USA: Michigan State University Press.

Edwards, C.A. (1989) 'Sustainable Agriculture Systems', *Agriculture, Ecosystems and Environment* 27: 25–35.

Edwards, C.A., Lal, R, Madden, P., Miller, R.H. and House, G.J. (eds) (1990) *Sustainable Agricultural Systems,* Ankeny, Ia., USA: Soil and Water Conservation Society.

El Titi, A. (1992) 'Integrated farming: an ecological farming approach in European agriculture', *Outlook on Agriculture* 21: 33–39.

El Titi, A. and Ipach, U. (1989) 'Soil fauna in sustainable agriculture: results of an integrated farming system at Lautenbach, F.R.G.', *Agric. Ecosystems Environ.* 27: 561–572.

Enache, A.J. and Ilnicki, R.D. (1990) 'Weed control by subterranean clover *(Trifolium subterraneum)* used as a living mulch', *Weed Technology* 4: 534–538.

Exner, D., Thompson, R. and Thompson, S. (1990) 'Case study: a resource efficient farm with livestock', in C.A.Francis, C.Butler Flora and L.King (eds) *Sustainable Agriculture in Temperate Zones,* New York: J.Wiley & Sons, pp. 263–280.

Falanruw, M.V.C. (1989) 'Nature-intensive agriculture: the food production system of Yap Island', in R.E.Johannes (ed.) *Traditional Ecological Knowledge: A Collection of Essays,* Gland, Switzerland: IUCN, The World Conservation Union, pp. 35–40.

Foissner, W. (1987) 'Soil protozoa; fundamental problems, ecological significance, adaptations in ciliates and testaceans, bioindicators and guide to the literature', in J.O.Corliss and D.J.Patterson (eds) *Progress in Protistology,* Bristol, UK: Biopress Ltd, 2: 69–212.

Food and Agriculture Organization (FAO) (1981) *Production Yearbook,* Vol. 35. Series No. 40, Rome, Italy: United Nations Food and Agriculture Organization.

Forman, R.T.T. and Godron, M. (1986) *Landscape Ecology,* New York, USA: Wiley.

Francis, C.A., Butler Flora, C. and King, L. (1990) *Sustainable Agriculture in Temperate Zones,* New York, USA: J.Wiley & Sons.

Gliessman, S.R. (ed.) (1990a) 'Agroecology: researching the ecological basis for sustainable agriculture', *Ecological Studies* 78: 1–380.

Gliessman, S.R (1990b) 'Understanding the basis of sustainability for agriculture in the tropics: experiences in Latin America', in C.A.Edwards, R.Lal, P.Madden, R.H.Miller and G.J.House (eds) *Sustainable Agricultural Systems,* Ankeny, Ia., USA: Soil and Water Conservation Society, pp. 379–390.

Golley, F.B., Bellot, J. and Odum, E.P. (1989) 'Dynamic analysis of the behaviour of two landscape systems', *Options Mediterranennes* 3: 253–257.

Gomez-Pompa, A., Flores, J.S. and Sosa, V. (1987) 'The 'pet kot' —a man-made tropical forest of the Maya', *Interciencia* 12: 10–15.

Hooper, M.D. (1979) 'Hedgerows and small woodlands', in J.Davison and R.Lloyd (eds) *Conservation and Agriculture,* Chichester, UK, New York, USA: J.Wiley & Sons, pp. 45–57.

Janick, J. and Simon, J.E. (1993) *New Crops,* New York, USA: J.Wiley & Sons.

Janke, R., Pleasant, J. Mt., Peters, S. and Bohlke, M. (1990) *Long Term, Low-input Cropping Systems Research. Sustainable Agriculture: Research and Education in the Field,* Washington DC, USA: NRC/USDA.

Jordan, V.W.L. (ed.) (1992) *Expert Presentations of Future Demands and Perspectives for Good Agricultural Practice,* Brussels, Belgium: Commission of the European Communities, Agriculture.

Knutson, L. (1989) 'On the diversity of nature and the nature of diversity', *Bull. Entomol. Soc. AM.* 35: 7–11.

Lockeretz, W. (1990) 'Major issues confronting sustainable agriculture', in C.A.Francis, C.Butler Flora and L.King (eds) *Sustainable Agriculture in Temperate Zones,* New York, USA: J.Wiley & Sons, pp. 423–438.

Lockeretz, W., Shearer, G. and Kohl, D.H. (1981) 'Organic farming in the Corn Belt', *Science* 211: 540–547.

Lussenhop, J., Fogel, R. and Pregitzer, K. (1991) 'Methods used in rhizosphere ecology. A new dawn for soil biology: video analysis of root-soil-microbial faunal interactions', *Agriculture, Ecosystems and Environment* 34: 235–249.

McCroskey, J.M. and Spalding, H. (1989) 'A reconnaissance-level inventory of the amount of wilderness remaining in the world', *Ambio* 18: 221–227.

Madden, J.P. (1990) 'The economics of sustainable low-input farming systems', in C.A.Francis, C.Butler Flora and L.King (eds) *Sustainable Agriculture in Temperate Zones,* New York, USA: J.Wiley & Sons, pp. 315–342.

Martin, P.S. and Klein, R.G. (1984) *Quaternary Extinctions: A Prehistoric Evolution,* Tucson, Ariz., USA: Arizona University Press.

Matthey, W., Zettel, J. and Bieri, M. (1990) *Invertébres bioindicateurs de la qualité de sols agricoles,* Liebefeld-Bern, Switzerland: Nationalen Forschungsprogrammes 'bonen' (in French).

Michon, G. (1983) 'Village-forest-gardens in West Java', in A.Huxley (ed.) *Plant Research and Agroforestry,* proceedings of a consultative meeting held in Nairobi, 8–15 April 1981, pp. 13–24.

Muir, R. and Muir, N. (1987) *Hedgerows. Their History and Wildlife,* London, UK: Michael Joseph.

Nath, B., Hens, L. and Devuyst, D. (eds) (1996) *Sustainable Development,* Brussels, Belgium: VUB University Press.

Nazzi, F., Paoletti, M.G. and Lorenzoni, G.G. (1989) 'Invertebrate dynamics of soybean agroecosystems encircled or not by hedgerows in Friuli, Italy, First Data ', *Agriculture Ecosystems and Environment* 27: 163–176.

Newton, L.H. and Dillingham, C.K. (1994) *Classic Cases in Environmental Ethics,* Belmont, NY, USA: Wadsworth Publishing Co.

NRC (1989) *Lost Crops of the Incas. Little Known Plants of the Andes with Promise for World-wide Cultivation,* Washington DC, USA: National Academic Press.

Odum, E. (1984) 'Properties of agroecosystems', in R.Lawrance, B.R.Stinner and A.House (eds) *Agricultural Ecosystems,* New York, USA: J.Wiley & Sons, pp. 5–11.

Ogle, B.M. and Grivetti, L.E. (1985) 'Legacy of the Chameleon: edible wild plants in the kingdom of Swaziland, southern Africa. Cultural, ecological, nutritional study. Part IV. Nutritional analysis and conclusions', *Ecol. Food Nutr.* 17: 41–64.

Paoletti, M.G. and Bressan, M. (1996) 'Soil invertebrates as bioindicators of human disturbance', *Critical Reviews in Plant Sciences* 15, 1: 21–62.

Paoletti, M.G. and Bukkens, S. (1997) 'Mini livestock: sustainable use of biodiversity for human food', *Ecology of Food and Nutrition* (s.i.) 36, 2–4: 95–346.

Paoletti, M.G. and Lorenzoni, G.G. (1989) 'Agroecology patterns in northeastern Italy', *Agric. Ecosystems Environ.* 27, 1–4: 139–154.

Paoletti, M.G. and Pimentel, D. (1992) *Biotic Diversity in Agroecosystems,* Amsterdam, The Netherlands: Elsevier.

Paoletti, M.G. and Pimentel, D. (1995) 'The environmental and economic costs of herbicide resistance and host-plant resistance to plant pathogens and insects', *Technological Forecasting and Social Change* 50: 9–23.

Paoletti, M.G. and Pimentel, D. (1996) 'Genetic engineering in agriculture and the environment', *BioScience* 49, 9: 665–673.

Paoletti, M.G., Dreon, A.L. and Lorenzoni, G.G. (1995) 'Edible weeds "pistic" found in W. Friuli, N.E. Italy', *Economic Botany* 49, 1: 26–30.

Paoletti, M.G., Stinner, B.R. and Lorenzoni, G.G. (1989) *Agricultural Ecology and Environment,* Amsterdam, The Netherlands: Elsevier.

Paul, E.A. and Clark, F.E. (1989) *Soil Microbiology and Biochemistry,* New York, USA: Academic Press.
Pimentel, D. (1996) 'Human demography and environmental resources', in B.Nath, L.Hens and D.Devuyst (eds) *Sustainable Development,* Brussels, Belgium: VUB University Press, pp. 111–136.
Pimentel, D. (ed.) (1997) *Techniques for Reducing Pesticide Use: Economic and Environmental Benefits,* New York, USA: J.Wiley & Sons.
Pimentel, D. and Pimentel, M. (1979) *food, Energy and Society,* London, UK: Edward Arnold.
Pimentel, D., McLaughlin, L., Zepp, A., Lakitan, B., Kraus, T., Kleinman, P., Vancini, F., Roach, W.J., Graap, E., Keeto, W.S. and Seling, G. (1991) 'Environmental and economic impacts of reducing US agricultural pesticide use', in CRC *Handbook of Pest Management in Agriculture,* Boca Raton, Fla., USA: CRC Press, pp. 679–718.
Pimentel, D., Stachow, U., Takacs, D.A., Brubaker, H.W., Dumas, A.R, Meaney, J.J., O'Neil, J., Onsi, D.E. and Corzilius, D.B. (1992) 'Conservation in agricultural/forestry systems, biological diversity', *Bioscience* 42: 750–760.
Plotkin, M.J. (1988) 'The outlook for new agricultural and industrial products from the tropics', in E.O.Wilson (ed.) *Biodiversity,* Washington DC, USA: National Academic Press, pp. 106–116.
Poincelot, R. (1986) *Toward a More Sustainable Agriculture,* Westport, Conn., USA: AVI.
Pollard, E., Hooper, M.D. and Moore, N.V. (1974) *Hedges,* London, UK: Collins.
Posey, D.A. (1993) 'The importance of semi-domesticated species in post-contact Amazonia: effects of the Kayapo Indians on the dispersal of flora and fauna', in C.M.Hladick, A.Hladick, O.F.Linares, H.Pegezy, A. Semple and M.Hadley (eds) *Tropical Forests, People and Food,* NY, USA: The Parthenon Publ. Group, pp. 63–71.
Prance, G. (1982) 'The increased importance of ethnobotany and underexploited plants in a changing Amazon', in J.Hemming (ed.) *Change of the Amazon Basin. Vol. 1: Man's Impact on Forest and Rivers,* Manchester, UK: University of Manchester Press, pp. 129–136.
Rabatin, S.C. and Stinner, B.R. (1988) 'Indirect effects of interactions between VAM fungi and soil-inhabiting invertebrates on plant processes', *Agric. Ecosystems Environ.* 24: 135–146.
Raven, P.H. (1988) 'Our diminishing tropical forests', in E.O.Wilson (ed.) *Biodiversity,* Washington DC, USA: National Academic Press, pp. 119–122.
Smith, B.D. (1995) *The Emergence of Agriculture,* New York, USA: Scientific American Library.
Tanara, V. (1644) (1731) *L'economia del cittadino in villa,* 10th edn, Venice, Italy: G.Bortoli (in Italian).
USDA, ICAR, FERRO (1991) *Sustainable Land Use Systems,* Kutztown, Pa., USA: Rodale Institute Research Centre.
Van Straalen, N.M. and Krivolutsky, D.A. (1996) *Bioindicator Systems for Soil Pollution,* NATO ASI Series, Dordrecht, The Netherlands: Kluwer Academic Publishers.
Vitousek, P., Ehrlich, P.R., Ehrlich, A.H. and Matson, P.A. (1986) 'Human appropriation of the products of photosynthesis', *BioScience* 36: 368–373.
Werner, M.R and Dindal, D.L. (1990) 'Effects of conversion to organic agricultural practices on soil biota', *Am. J. Alternative Agriculture* 5: 24–32.
Western, D. (1989) 'Conservation without parks: wildlife in the rural landscape', in D.Western and M.C.Pearl (eds) *Conservation for the Twenty-first Century,* New York, USA: Oxford University Press, pp. 158–165.
Yu Zhen Rong and Baudry, J. (1998) 'Landscape dynamics: the importance of changes', *Critical Reviews of Plant Sciences,* in press.
Zohary, D. and Hopf, M. (1993) *Domestication of Plants in the Old World,* Oxford, UK: Oxford Science Publ. Clarendon Press.

SUGGESTED READING

Paoletti, M.G. and Pimentel, D. (1992) *Biotic Diversity in Agroecosystems,* Amsterdam, The Netherlands: Elsevier.
Pimentel, D. (1997) *Techniques for Reducing Pesticide Use: Economic and Environmental Benefits,* New York, USA: J.Wiley & Sons.
Pimentel, D. and Pimentel, M. (1979) *Food, Energy and Society,* London, UK: Edward Arnold.

SELF-ASSESSMENT QUESTIONS

1 Define a rural landscape.
  (a) An agro-ecosystem.
  (b) A rural watershed basin.
  (c) An agricultural farm.
2 Define agro-ecosystem structure and functions.
  (a) A system that produces biomass of human interest.

(b) A cycle of energy and matter in order to produce economic returns and appropriate economy.
(c) A human ecosystem in which energy and matter cycles are handled together with an appropriate management of natural subsystems.

3 How do you evaluate biodiversity in agro-ecosystems?
(a) By the number of trees, birds and mammals.
(b) By the number of different crops and vertebrates.
(c) By the diversity of the microenvironment and diversity of biota, including invertebrates.

4 How do you measure pollution in agro-ecosystems?
(a) By means of chemical analyses and soil structure.
(b) By means of chemical and biological analyses and bioindicators.
(c) By means of biodiversity and physical measures.

5 What does sustainability in rural landscapes refer to?
(a) Maintaining soil fertility, biodiversity and economic viability over time on the farm.
(b) Increasing productivity and income.
(c) Improving economical and ecological parameters at the farm level.

10

SUSTAINABLE MANAGEMENT OF THE FISHERIES SECTOR

M.J.Peterson

SUMMARY

World fisheries are divided into two main subsectors, namely capture fisheries and fish raising. In each subsector, there are distinct industrial (high technology) and artisinal (low technology) forms of fishing. These are practised by different groups and pose different management problems for fisheries regulators. Fisheries management is a combined discipline, drawing on the insights of fish biology, ecology, economics and sociology to develop policies implemented through national laws and regulations and international agreements. The need for fisheries management and regulation was first perceived at the beginning of the twentieth century, and the answers to regulatory problems sought primarily in fish biology. In succeeding decades, perspectives have broadened as difficulties in fisheries management have indicated the need for a broader approach. On the biological side, single species models were replaced by multispecies and ecological models of fish population dynamics. The use of economic analysis in fisheries management, which began in the 1950s, was supplemented in later decades by paying greater attention to social issues of income distribution in the fisheries sector, class, ethnic and other divisions among fishermen, and distribution of fish to consumers. In all, the discipline of fisheries management has shifted from one based on a combination of single-species and economic analysis to a more fully rounded analysis taking account of interrelationships among fish species and the ecosystems in which they live and the full range of economic, social and political concerns raised in deciding who will undertake a particular form of economic activity and how their products will be distributed to what consumers.

ACADEMIC OBJECTIVES

This chapter has four objectives, each the focus of a separate section. The first familiarises the student with the patterns of work in the fisheries sector by briefly describing the main techniques used in capture fisheries and fish raising. The second summarises the main points of fish biology and ecology relevant to fisheries management. The third outlines the basic principles of fisheries economics to demonstrate why managing open access fisheries is so difficult and to indicate why many analysts are adopting some form of individual or communal property rights in fisheries. The fourth addresses fisheries management, tracing the shift from yield-centred to multidimensional management attempting to integrate the ecological, economic and social dimensions of fisheries.

FISHERIES SUBSECTORS AND TECHNIQUES

Both capture fisheries and fish raising are divided into industrial and artisinal forms of operation. Industrial fisheries involve high levels of capital investment, which is justified by capturing economies of scale. Artisinal fisheries involve use of far simpler equipment, hence lower capital investment, and can be economically viable at a much smaller scale. Industrial operations all produce for distribution to consumers, generally through markets; many artisinal fishermen supply consumers but others fish for subsistence.

Most fisheries still involve capturing fish in oceans, lakes, rivers and streams. Ocean fishing is concentrated in areas where the largest number of fish concentrate: seasonal spawning grounds, seasonal feeding grounds, hydrodynamic boundaries and upwelling areas. Fish raising (the term 'aquaculture' also covers raising underwater plants) involves introducing fish larvae or fry into confined fresh or salt water areas and raising them to maturity under human control. Fish raising has been practised for centuries, expanding rapidly in recent decades to the point of providing 12 per cent of world fish production in 1990 (FAO, 1991:12).

This reflects the late, because difficult, transition from hunter-trapper forms to animal husbandry forms of food production in fishing. The shift has been so recent because the advantages of domestication are not as great with fish as with land animals. Primary nutrient production in the water is lower than on land; the global ocean average is one-fifth of the global land average (Tett, 1977:18). Additionally, the open ocean is less amenable than land to human control, though shallow near coast areas can be penned off. Last, and most importantly, natural fish stocks are self-renewing resources whenever they have favourable habitat, sufficient food and low enough mortality. Until the late nineteenth century, fish stocks in the oceans and larger bodies of fresh water were exposed to few risks of serious decline. The environment was relatively unpolluted, food was generally sufficient, and mortality due to human taking well below levels that threatened a stock's reproductive ability.

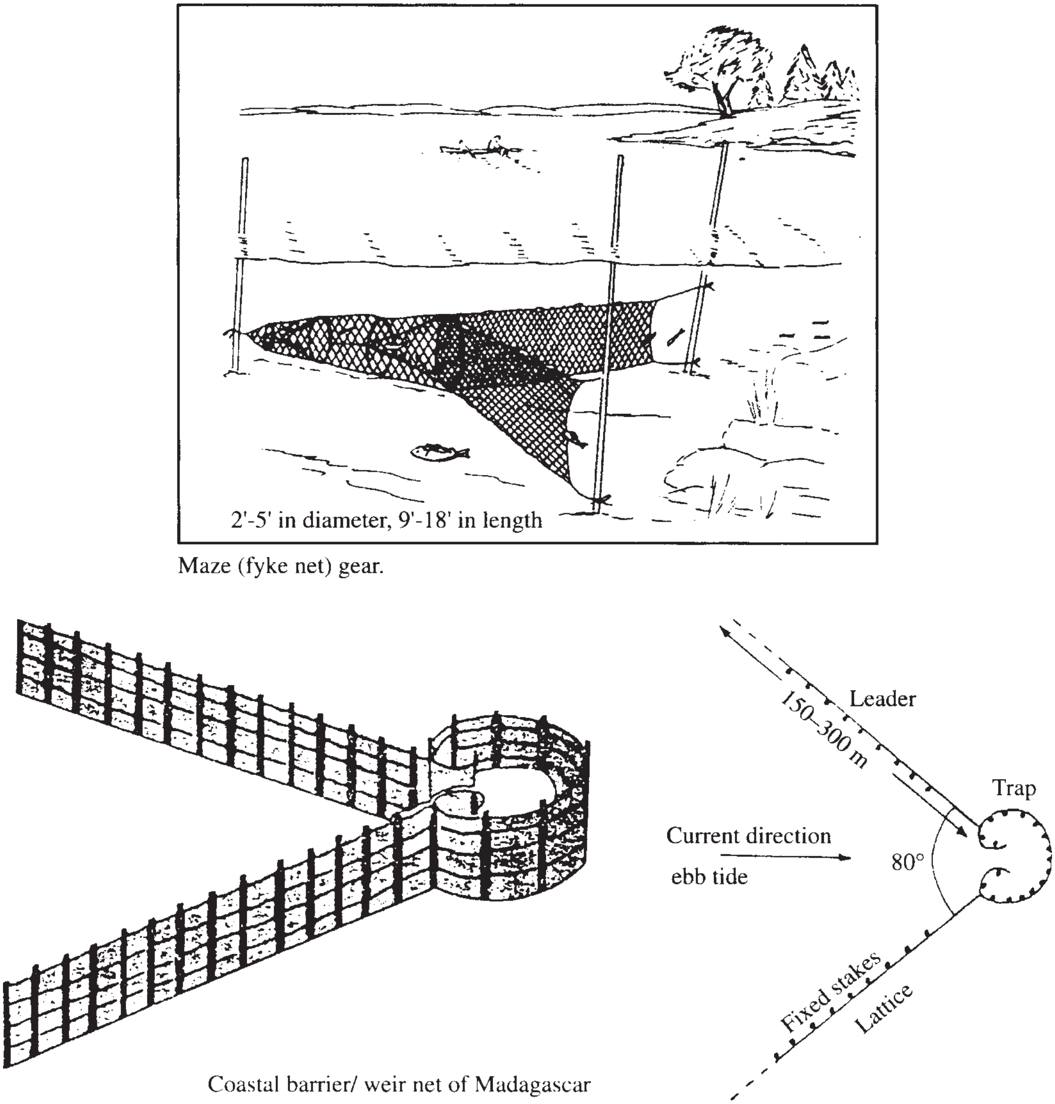
Fish raising has spread for several reasons. First, many capture fisheries have been depleted by overfishing, reducing the advantage of working a self-renewing stock. Second, newly developed technologies permit greater efficiency in fish raising. While unable to compete with capture fisheries efficiency as measured by energy use, labour time and equipment investment, fish raising has become competitive with most forms of land animal husbandry. Because fish are cold-blooded rather than warm-blooded, the portion of caloric intake going to maintain body temperature in land animals can be devoted to growth. Trout, for example, yield 30–40 grams of protein per Meal of feed while cattle yield 2, pigs 6 and chickens 15. Only the intensive raising of broiler chickens—producing five or more slaughter-ready groups a year—now outproduces fish raising on a per hectare basis (Coull, 1993:181). Capital and energy remain limiting factors. Capital is required for acquisition of property on which to base operations and equipment for tending fish. Though some forms of artisinal fish raising use less energy than capture fishing, other artisinal operations and all industrial ones use more (Cunningham *et al.*, 1985:350–351).

Techniques in artisinal capture fisheries

Contemporary artisinal fishermen working from canoes and small boats use hand-cast spears, gorges, hooks, longlines, driftnets and bag nets. Longlines catch fish on individual barbed hooks, driftnets catch fish by the gills, and bag nets by surrounding them. Seventeenth-century Dutch fishermen used longlines to catch North Sea bottom fish and suspended driftnets between boats to form catching areas a kilometre long. Other artisinal forms of fishing involve setting traps or nets on riverbeds or in shallow waters and casting bag nets from the beach. Though these basic forms of gear remain in use today, artisinal fishermen have adopted certain twentieth-century technologies, particularly outboard motors and synthetic ropes.

Techniques in industrial capture fisheries

The industrial techniques of longlining, trawl netting, purse seine netting and driftnetting are all mechanisations and enlargements of artisinal techniques. They were pioneered in England, Scandinavia and Japan where the right combination of accessible fishing grounds, active markets, technological innovation and investment funds existed. Though trawl nets had a long history, their full potential was displayed only in the 1890s after steam engines replaced sails, power winches replaced man-powered capstans, steel warps replaced heavy ropes, and the otter trawl superseded the beam trawl. By 1900 English trawlers had expanded operations from traditional near-shore areas to the Barents Sea and waters off Iceland while Japanese fishermen ventured into the East China and Yellow Seas. Use of industrial techniques expanded rapidly after World War II. Marine diesel engines, first developed in the 1930s, provided more reliable power with less bulky fuel than coal-fired steam engines. Stern trawlers, mid-water trawl nets,



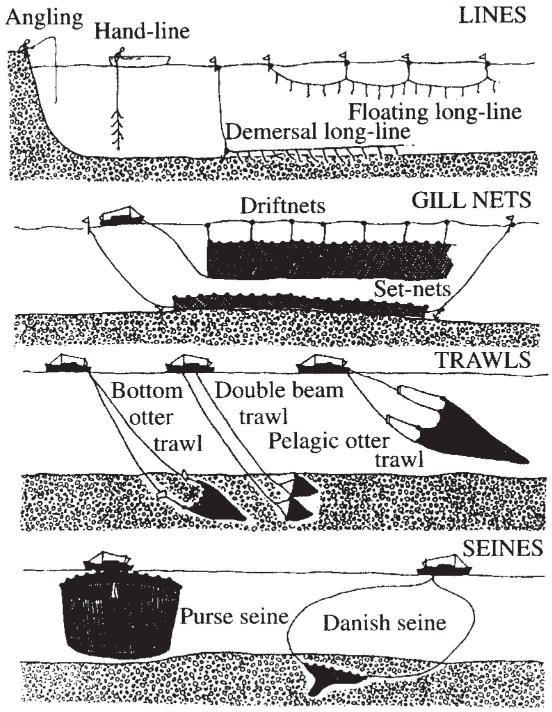
Maze (fyke net) gear.

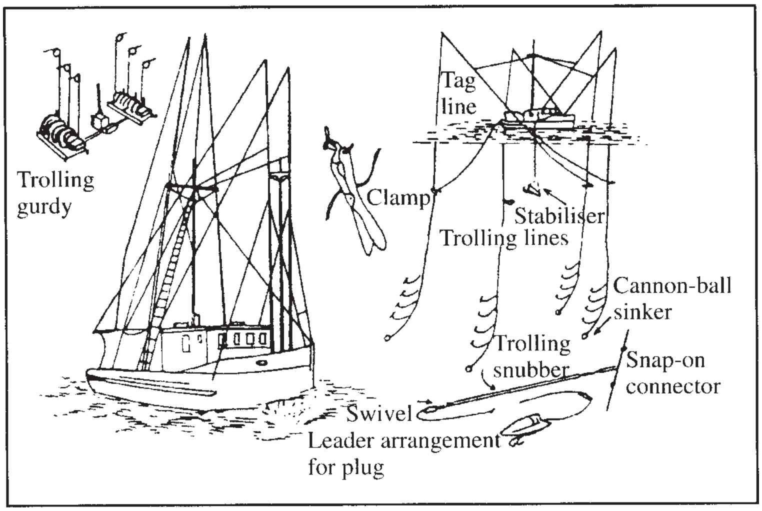
Coastal barrier/ weir net of Madagascar

Figure 10.1 Some traditional fixed fishing gear

synthetic rope, purse seine nets hauled by power blocks, radio navigation aids and echo-locating fish-finding systems permitted catching on far larger scales, justifying investment in larger boats. In the 1950s British and German distant water fleets began using trawler-processors canning or freezing the catch on board, permitting voyages of months rather than the three weeks possible with earlier systems of keeping fish on ice after cleaning them at sea. In the 1960s and 1970s the Japanese and Soviets developed world-ranging fleet operations around large processing ('factory') ships, while the governments of both industrial and developing states promoted growth of industrial fisheries through tax credits, loans and subsidies.

Contemporary industrial fishing activity is divided into middle water and distant water operations. Middle water operations involve locally based fishing





Hook-and-line (trolling gear).

Figure 10.2 Contemporary ocean fishing gear boats equipped using trawl nets, purse seine nets, driftnets, or longlines depending on the species sought. Catch is cleaned on board and kept on ice or under refrigeration until the boat returns to port. Such operations expanded in many parts of the world as economic development programmes came to include building or expanding local fishing capacity, a process reinforced by adoption of the 200-mile exclusive economic zone (EEZ) in the mid-1970s. Locally owned industrial middle water operations can be found today in most coastal countries. Distant water operations employ the same types of gear, but larger boats sailing longer distances to reach the catch. Some employ individual trawler-processor ('trawler-factory') ships that catch, clean and process fish. Others involve fleets consisting of several trawler-processors or several catching ships serving one factory ship plus support ships. In 1970 the Soviet and Bulgarian fleets were among the largest; today the largest distant water fleets are from Japan, South Korea, Taiwan, the EU (mainly Portugal and Spain), China and Poland.

Techniques in artisinal aquaculture

Artisinal ('extensive') aquaculture uses simpler technology and less energy but more labour than its industrial counterpart. Most artisinal operations involve collecting spawn in the wild and transporting it to ponds or paddies where it can be raised away from predators and species that compete for food in its natural habitat. Most artisinal activity relies on species that feed on plankton or natural organic detritus; carp are a preferred freshwater pond species and oysters a preferred saltwater one. East Asia has the longest history of artisinal fish raising; the Chinese *Fish Culture Classic* of 460 BC described several elaborate techniques. A seventh-century prohibition against eating common carp inspired the development of polyculture, simultaneous raising of several species of carp that do not compete for food in the same pond. Pond fertilisation permitted a further increase in yields. Carp were also raised in Europe before AD 1000; methods of collecting and transporting young fish (fry) rather than spawn and of supplemental pond feeding were developed in the medieval period.

Chinese carp ponds are part of an elaborate integrated agricultural system. The larger combined farms of South China maintain a series of ponds for fish at different stages of growth. The ponds are fertilised with pig manure while organic debris collecting at the bottom of the ponds is periodically cleared and spread on crop fields as fertiliser. Some of the crop is then used as fish feed. Fish broods may include as many as five carp species—the mix depending on food availability—as well as some Wuchan bream to control any predators introduced inadvertently during stocking. Normally three broods are introduced to the ponds each year, allowing continuous fish harvesting between late May and early November (Ruddle and Zhong, 1988). In other parts of Asia integrated growing takes the form of alternating carp and rice in paddies. In Bangladesh fish ponds near the coast are used for salt evaporation during the dry season.

Shrimps and prawns have also been cultivated for many centuries. In India and other parts of Asia they are grown in rice paddies after the rice is harvested. In many parts of the world areas of lagoons are fenced off with small barriers and juvenile prawns penned until they grow to harvestable size.

Techniques in industrial aquaculture

Industrial ('intensive') aquaculture relies on technologically advanced and energy-intensive techniques requiring considerable capital investment. Such techniques are most highly developed in Japan, but are rapidly spreading throughout the world. These methods have developed despite their higher initial costs because they offer the possibility of increasing productivity by bringing more of the conditions under which fish spawn survive and grow to maturity under human control, thus producing greater yields.

In general, tropical species are more suited for raising than temperate water ones because their faster growth rates make them ready for market more quickly. For example, carp in South China reach a weight of 1–2 kg within a year. Atlantic salmon are one of the few cold water species raised extensively. Because some species are more tolerant of high stock densities than others, they are now raised all over the world. Rainbow trout were brought into Europe from North America in the 1880s; more recently Pacific salmon were introduced into Chile from the USA and New Zealand, the Japanese oyster to the USA and Europe, and tilapia from Africa to Asia and North America. Better transportation and communications mean such transfers are more frequent today than in the past.

Variations in natural supply and the difficulty of ensuring that individuals of undesired species are not mixed with the spawn or fry to be raised motivated efforts to bring spawning and the earliest stages of fish life under human control by using hatcheries. Doing so requires capturing and maintaining adult fish as they are ready to spawn, and then caring for the offspring until they are past the highly vulnerable larval stage of life. Hormone-controlled spawning and selective breeding for desired characteristics can shorten the growth phase and lead to a more standardised yield. Thus the same incentives that encouraged selective breeding and developing domesticated varieties of land animals also operate in fish raising.

A large fish raising subsector permits specialisation, most commonly between hatchers and growers. With species that hatch best in one set of conditions and mature best in another, patterns of dispersed production reminiscent of some farming develops. Just as seed potatoes are grown in colder areas of Europe while crop potatoes are grown mainly in warmer areas, Japanese oysters are seeded off the north-east coast of Honshu and transferred to the south-west coast for growing. The elvers raised to adulthood on Japanese eel farms are now supplied mainly by breeders on Taiwan. Such specialisation is particularly common in salmon raising because spawning and early life occurs in fresh water while growth to adulthood occurs in salt water.

The best method of raising fish depends on the characteristics of the species involved. Oysters and other shellfish are raised using the suspension techniques, hanging of organisms from ropes or in small cages, first developed in Europe in the late nineteenth century. Suspension revolutionised Japanese oyster culture in the 1920s by permitting a shift from shallows to waters up to 4 metres deep. Using rafts rather than fixed supports permitted moving into even deeper areas. Suspension offers three advantages over using the bottom: an increase in the total area that can be used for raising, an increase in the number of individuals grown in the same space because more than one 'layer' can be tended simultaneously, and more efficient access to food that permits doubling the growth rate. The Japanese now use suspension in raising bivalve molluscs, scallops and abalone as well as oysters.

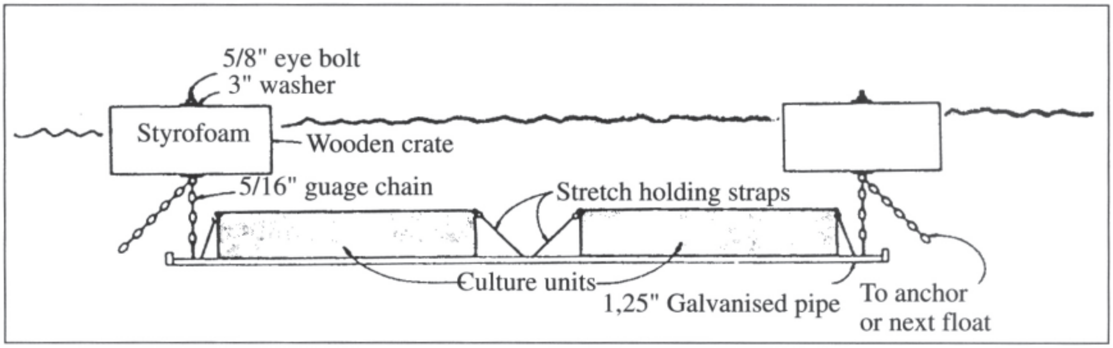
Western advances in stocking during the nineteenth century were initially used to replenish freshwater recreational fisheries. Hatcheries were linked to fish raising during the interwar period when Europeans realised that the recently introduced rainbow trout could stand much higher population densities than native species. Trout raising became particularly well established in Jutland, where they are raised in catchments along small streams or in nearshore saltwater areas; and in France and Italy, where they are raised in mountain valley freshwater 'raceways'.

Raising salmon is complicated by their highly migratory habits. Pacific chum and coho were first raised in Japan and the USA after 1950, and Atlantic salmon in Norway in the mid-1960s. Salmon are spawned in freshwater hatcheries and transferred to saltwater pens for growing. These pens must be located in areas of ample water flow because salmon maintain health only with vigorous swimming. Atlantic salmon production is much higher because they tolerate higher stock densities than the Pacific species.

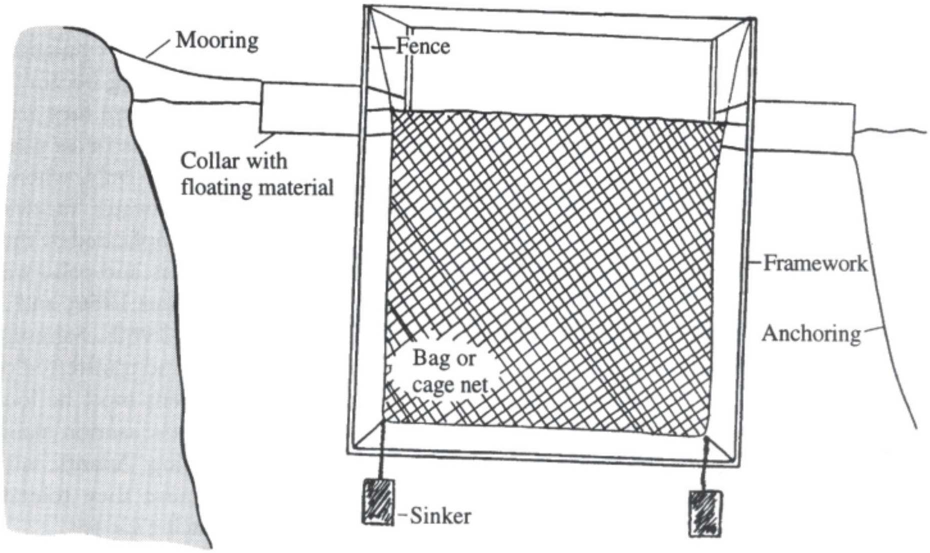
Shrimp are the focus of the greatest expansion of fish raising in recent years. This activity is concentrated in tropical areas because the fastest growing species need water temperatures of at least 26°C. Intensive shrimp raising involves building mechanically aerated ponds in stable, non-acidic soil, securing fry from hatcheries, providing fishmeal because shrimp are carnivores, and flushing the ponds occasionally to restore salinity and clear away faecal matter, excess food and residues of disinfectants, pesticides, fertilisers and water conditioners.

FROM FISH BIOLOGY TO FISH ECOLOGY

Capture fishermen and fish raisers must understand both fish life cycles and the environmental conditions



Cross-sectional construction drawing of a typical bivalve culture raft. (From J. V. Huner and E. E. Brown, *Crustacean and Mollusk Aquaculture in the United States*, Westport, CT: AVI Publishing, 1995.)



Floating fish cage

Figure 10.3 Gear for fish raising

that affect their survival. Though each species has its own particular cycle of spawning, survival to sexual maturity, reproduction and mortality, the main features of life cycles can be understood by classifying them along two main dimensions: the type of water habitat and the type of stock. The most significant habitat distinction is between fresh water and salt water, each of which offers different conditions. Lakes and rivers offer a larger mass of aquatic plants, but most freshwater fish are not herbivores. The photic zones of the oceans where light penetrates offer favourable conditions for the growth of the phytoplankton and zooplankton on which most fish food chains are based. Food for benthos feeding species living near the ocean floor is particularly plentiful in the coastal zones, in upwelling areas and on mid-ocean banks. Both salt and freshwater areas are divided into three main water temperature zones—tropical, temperate and cold—each of which has distinct characteristics that affect fish life cycles. Most tropical species have shorter lives, more spawning periods during the year and faster growth rates than temperate and cold water species. Though overall recruitment is comparable in tropical and other fisheries (Murphy, 1982), the differences in growth rates and periods to sexual maturity (for example, North Sea herring reach maturity in 3–4 years and Arctic cod in 9, while the Peruvian anchovy reaches maturity in 1.25–2) require some modification of the standard stock assessment models developed initially for colder water species (Pauly, 1986). Equally significant in capture fisheries is the type of stock involved. Most wild fish stocks fall into one of four categories: steady-state stocks experiencing relatively limited fluctuation in their numbers; cyclical stocks varying greatly in numbers over predictable cycles of time; irregular stocks fluctuating considerably in no regular pattern; and spasmodic stocks alternating short periods of plenty with longer periods of great scarcity (Caddy and Gulland, 1983). Keeping catch within biologically sustainable limits requires awareness of these differences.

The standard stock assessment models developed in the 1950s focused on steady-state stocks, which still constitute the majority of exploited fish stocks. Two distinct approaches developed. In North America, general production models prevailed. These assume that total biomass is a function of the growth rate plus recruitment (which is assumed to be fairly steady each year) minus catching plus natural mortality, and yield an estimate of safe taking based on aggregate size of the population. In Europe, yield per recruit models prevailed. These derive estimates of safe fishing levels from a series of calculations of how many recruits are added to the stock in a particular year plus their rate of survival in each additional year. The main difference is that the European models take annual fluctuations of recruitment into account. Yet the two models converge on a roughly similar biological optimum, because an abundant stock includes individuals recruited over several years whose simultaneous presence makes stock size more steady than recruitment. Thus both traditions of analysis allowed fisheries managers to proceed on the assumption that year-to-year changes were not too large. Yet it is now understood that even steady-state stocks can fluctuate significantly and that continuing earlier rates of catching during periods of poor recruitment can lead to the same depletion as catching above the sustainable yield. Variations of recruitment to cyclical, irregular and spasmodic stocks do not level out as much, and so cannot be managed in the same way. Any attempt to fish such stocks at a steady annual rate derived from times of high recruitment risks running the species to extinction during the troughs. These characteristics pose a particular challenge to industrial fishermen, whose high capitalisation means they must operate fairly continuously to meet their expenses.

The scientists who developed the standard models of stock assessment were aware of predator-prey relations, competition for food and the impact of oceanic conditions, but assumed that the effects were either steady enough to be treated as constant or small enough to be ignored. By the early 1970s a more holistically oriented group of fish biologists were arguing that the effects were too important to ignore and too variable to be treated as a constant. They proposed various multispecies models for simultaneous analysis of the state of several co-existing stocks. Lack of adequate computational tools inhibited their use at first, but that limitation was overcome as more powerful desktop computers became available in the early 1980s.

Fisheries scientists also became more fully aware of the many ways in which wider environmental conditions—weather fluctuations, shifting currents, loss of habitat, rising pollution levels—affect fish stocks in the 1970s. Particular events, most notably the collapse of the Peruvian anchovy fishery after the El Niño of 1972–1973, inspired concentrated efforts to understand the interactions of ocean conditions and fish stocks more thoroughly. Continued research also yielded greater understanding of how population density affects spawning, recruitment and survival of fish stocks. This new research inspired efforts to develop more realistic models of fish life cycles that took environmental conditions into account. The harm to particular fisheries, most notably in the Great Lakes of North America and in the Black Sea, caused by planned or accidental introduction of non-native species drew even more attention to the importance of ecosystemic balance.

Drawing on concepts developed earlier, many experts began advocating management of fisheries as part of a wider Large Marine Ecosystem. The concept was first endorsed for Southern Ocean fisheries in 1980, but its implications remained unclear until elaboration of the concept of Large Marine Ecosystem (LME) management. LMEs are defined as ecologically coherent areas of ocean and seabed at least 200,000 square km in area. Approximately 95 per cent of world fish production comes from the 49 nearshore LMEs defined by 1991 (Sherman, 1991). An LME approach begins with analysis of the state of target fish stocks, but also takes account of their relation to all other species of marine life (not just other fish) that are their prey, predators or competitors for food, as well as habitat conditions such as pollution, eutrophication levels and currents, and climate or weather-induced shifts of water temperature and turbulence. A few experts have claimed that increased understanding of inter-species relations and fish-habitat connections open up the possibility of 'adaptive management', in which fishing regulations would be used to maintain the relative size of populations or to react to environmentally induced stock fluctuations like the periodic stock reductions off the west coast of Latin America owing to the El Niño or the periods of plentiful food and fish stocks in the North Pacific brought on by the Aleutian low. Some experts now believe that many fish stocks are subject to multi-year or even decadal fluctuations in abundance driven by climatic fluctuations (Rothschild, 1996:20–21). Greater attention to inter-species relations has highlighted the problem of by-catch, hauls of other species caught in the course of many fishing operations. In 1990 by-catch amounted to some 27 million metric tons of fish (Alverson *et al.*, 1994), most of which die after being thrown back into the sea. The proportion of by-catch varies considerably depending on the gear used and the extent to which species commingle, with shrimp trawling routinely amassing far more by-catch than shrimp.

THE ECONOMICS OF FISHERIES

Both capture fishermen and fish raisers must be concerned with maintaining economic viability. This is easier in fish raising and capture fisheries operating under communal or individual property rights regimes, because individual fishermen, fishing groups or enterprises can enjoy the benefits of their own stock management efforts. In the typical open access ocean fishery the incentive of individual fishermen or crews to practise good stock management is weak because the benefits are more likely to accrue to the many anonymous others who are free to take whatever fish are in the sea.

Open access fisheries were viable for so many centuries only because fish-catching technology was sufficiently inefficient, and human demand for fish sufficiently low compared to the reproductive capacity of stocks, that fish were superabundant in the oceans and larger bodies of fresh water. Each fishing boat could catch to capacity and still leave enough for all others to do the same. Fishermen were under no pressure to do all their fishing early in the season, or to invest in additional gear so they could catch faster than competitors. When fish are not superabundant, open access fisheries become economically problematic. The ending of superabundance as industrial fishing techniques spread was apparent in some nearshore areas by 1900 and in many of the middle and distant water grounds by the mid-1970s. The decline of total world fish catch since 1989 indicates that scarcity is now the common condition.

When fish are not superabundant, open access fisheries become hard to manage because the economically optimum and the biologically optimum catch seldom coincide. This results from the different characteristics of the yield and cost curves. The yield curve

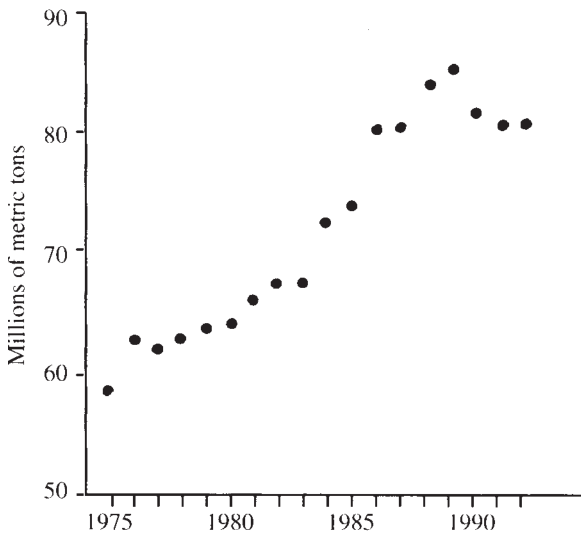


Figure 10.4 Marine fish landings *Source:* FAO, 1993

first rises rapidly from the origin, then more slowly until a high point is reached, and then declines as additional fishing effort removes fish from the stock faster than they can be replaced through recruitment. The cost curve does not have a similar bend; it is far more likely to rise in linear fashion as additional units of effort—boats, gear or more days' work—are added.

This combination of yield and cost functions sets up strong dilemmas for fishermen. Their economic optimum is the point at which the revenue curve most exceeds the cost curve. If prices are steady, the revenue curve resembles the yield curve, climbing sharply with the initial harvest increases, climbing more slowly as the biological optimum is reached, and then declining as fewer fish are caught. Thus the point of greatest revenue above costs lies leftward of the maximum sustainable yield (MSY). However, fishermen find it difficult to restrict effort there—particularly in an open access fishery—because of the profits to be had from additional effort. The marginalist economic analysis—that producers will continue to produce until the marginal cost of an additional unit of effort equals the marginal revenue from the yield of that effort—indicates that any group of fishermen has a strong incentive to catch at or beyond MSY.

The tendency to overshoot MSY is intensified by a backward-bending supply curve. In most enterprises —including fish raising—additional productive effort yields additional supply, holding prices steady or even causing them to fall depending on the level of demand. Capture fisheries are an exception: catches beyond MSY deplete the stock, thereby decreasing the availability of fish. When supply falls, prices tend to rise. The increasing prices inspire continued effort, which only leads to greater depletion, lower stocks, less supply and continued price increases. Though price increases push some consumers out of the market, many fish markets today are supported by enough affluent consumers to sustain the higher prices. In an extreme example, one whole bluefin tuna sells for about $30,000 at Japanese fish auctions, and sushi lovers pay about $75 for two bite-sized pieces (Fairlie *et al.*, 1995:49).

Fisheries economists have argued since the 1950s that the most appropriate policy goal in fishing is the optimum yield (OY) —the level of catch assuring

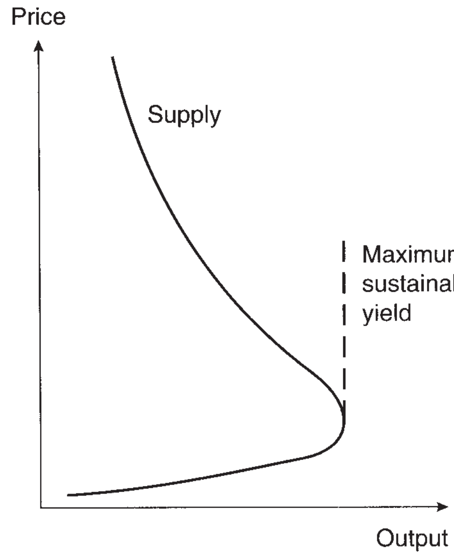


Figure 10.5 Supply and price in capture fisheries

greatest return over costs to the fishing sector. One common definition, derived from the European cohort analysis tradition, defines OY as $F_{0.1}$, a level of taking which yields about 10 per cent less catch than F_{max} (the biological maximum, roughly equivalent to MSY). Fishing to this level requires only about 60–70 per cent of the effort needed to catch at F_{max}, lowering costs considerably.

Yet attaining any optimum has been elusive even on steady-state stocks. Considerable pressure is exerted by those in the fishing sector. Once investment rises beyond low levels or fishermen choose to fish full-time, they seek steady access to catching opportunities. Fish wholesalers and processors also want reasonably steady supplies. This makes it hard for fisheries to accommodate stock fluctuations unless there is room for shifting effort from one target species to another.

Most governments have exacerbated the economic problems of fisheries through ill-considered subsidies, such as paying fees for access to foreign EEZs, providing loans for boats and gear, and implementing price or income supports. The UN Food and Agriculture Organization (FAO) estimated in 1993 that the global fishing fleet was operating at a loss of $22 million annually if operating costs were counted and $54 million if capital costs were also counted (FAO, 1993:20 and 58). As the FAO study noted, the fishing industry is able to secure subsidies in both good and bad times. When stocks are high, fishermen get subsidies for capital investments because governments want to see the plentiful fish stocks exploited. When stocks are low due to natural fluctuation or overexploitation, fishermen secure subsidy of operating costs to maintain jobs and investments. The overall result is an intensification of the economic dynamics that lead to overfishing in the first place.

The main lessons of fisheries economics are clear: there is strong economic incentive to overfish in an open access fishery, and the situation will be made worse by subsidising fishing effort. Yet the cycle is difficult to break. In some countries fishermen form a cohesive group with high access to policy-making because they are concentrated in a few legislative districts (e.g. the USA, eastern Canada, northern Norway). In many countries there is also social consensus that fishing is an important cultural tradition that deserves support (e.g. Japan, Norway). In the 1960s and 1970s many developing countries developed or expanded the indigenous industrial fishing sector to exploit more fully the resources off their own coasts, rather than let outsiders have them.

FROM MSY TO MULTIDIMENSIONAL FISHERIES MANAGEMENT

Management of capture fisheries

Using MSY as the management goal was understandable when fish were abundant and the main problem was developing enough fishing capacity to take full advantage of the renewable resources of the oceans. It is woefully inadequate today. Fish are scarce rather than abundant, new environmental knowledge indicates that the initial models on which MSY was calculated considered too few of the factors causing changes in fish populations, and increasing attention is being paid to the social aspects of fishing.

The social aspects concern two issues: the distribution of fish among consumers and the distribution of opportunities to catch or raise fish. Determining the distribution of fish requires deciding how much fish will go to each of the main end-uses—human consumption, animal feed and industrial uses (including fertiliser), and how much will be available for local or foreign consumers of different income classes and social groups. Determining the distribution of opportunities means deciding on the level of fisheries employment to encourage, the degree of given to different groups of fishermen, and the balance between industrial or artisinal operations. These social issues have become more acute as overfishing and expansion of fish raising have increased conflict among different groups in the fishing sector and as governments have asserted greater control over ocean fisheries.

Governments' responsibility to find solutions has been increased by recent changes in the international law of the sea. In 1945 their authority was limited to inland waters, bays and the 3 nautical miles of territorial sea; the rest of the ocean was high seas open to all. The initial response to declining ocean catches was attempts at multilateral management through establishment of international fisheries commissions charged with drawing up agreed management measures for particular species (e.g. the International Tropical Tuna Commission) or particular parts of the ocean (e.g. the International Commission on Northwest Atlantic Fisheries). Yet the commissions were plagued by foot-dragging over decisions to limit catch or effort, and by failure to enforce many of the measures adopted. A few governments proposed extending national jurisdiction further in the 1950s, but only in the mid-1970s was there sufficient consensus to adopt the EEZ concept extending national fisheries jurisdiction to all areas of ocean within 200 nautical miles of shore.

This change brought some 90 per cent of world ocean fisheries production within the jurisdiction of individual states. The major exceptions concern Southern Ocean stocks, highly migratory stocks and the seaward portion of straddling stocks ranging both landward and seaward of the 200-mile limit. The first are now managed by an international commission established in 1980; the latter two were managed through newly formed or revived international commissions including fishing and coastal states. Yet continued pressure on fish stocks within EEZs as well as dislocations following exclusion of foreign fishermen from those zones soon inspired severe conflict. Latin American and South Pacific countries asserted jurisdiction even over the highly migratory species found within their zones. A number of coastal states attempted to assert control over fishing of straddling stocks occurring outside their zone, claiming it was interfering with their efforts to maintain stocks within their zone. Conflict over straddling stocks became acute in the late 1980s and early 1990s, triggering calls for the negotiations leading to conclusion of the UN Convention on Highly Migratory and Straddling Fish Stocks in mid-1995. It does not permit extension of coastal state jurisdiction, but does commit states to closer co-operation in regional fisheries commissions and permits coastal states to take enforcement measures if flag states decline to act when requested to do so.

Agreement on wider management authority has not been paralleled by consensus on policy goals. A growing coalition of deep ecologists, rural sociologists and other critics of current management techniques present the policy conflict as one between two starkly different visions of future fisheries. In their view, current fisheries management is committed to an industrial model of fishing in which large (often multinational) firms using high technology gear develop a capital-intensive sector catering mainly to wealthy consumers in developed countries, fishmeal producers and industrial users, while ignoring ecological constraints in favour of profits. They offer in response a vision of a labour-intensive, largely artisinal sector of small producers catering more to local moderate and low income consumers and sensitive to ecological constraints. The evident failure of fisheries managers to avoid depletion of stocks lends weight to their critique; the growing popular resistance to many of the implications of economic globalisation makes it possible to form coalitions with groups questioning contemporary policies in other fields.

The stark contrast image has inspired considerable political mobilisation, but is less useful for understanding recent changes in fisheries management thinking. There are significant changes in fisheries management, although they are not going as far or happening as fast as the critics would like. They can be seen first of all in the ongoing experimentation with policies to replace the traditional open access regime in ocean fisheries. Eight possibilities, used singly or in combination, have received the most attention.

Total allowable catch (TAC)

Under a TAC regime, total fishing effort is held to a particular amount (typically weight) of fish caught by all participating boats. When first applied to whaling in the late 1940s and high seas fisheries in the 1960s, fisheries economists and managers underestimated the impact of an important side effect: TACs without suballocations to individual boats or fishermen give a large premium to whomever catches first and fastest. TACs thus intensified the general overinvestment dynamic as fishermen rushed to acquire faster boats, more accurate fish-locating devices, larger nets and more powerful winches to set and haul them. Yet these negative effects soon developed in several fisheries. Depleted fisheries managed through a TAC develop a pulsed activity pattern: intensive fishing and large landings for short periods, alternating with no fishing and no landings during others. Fishermen quickly learned to compensate for downtime by extending their activities to other species or other fishing areas. TAC regimes also proved difficult to monitor and enforce over broad stretches of ocean, because fishermen have a strong incentive to underreport or misreport (specify as a different species) their catch. Fisheries experts admit that such problems are pervasive; perhaps the best-documented instance involves Soviet underreporting of whale catches in the 1960s and 1970s so gross as to cast doubt on population estimates derived from catch data. TACs are used today only in combination with licensing or quotas, where they provide a ceiling on the total number of licences or the total of quotas allocated.

Indirect controls

Indirect (or input) controls limit catches by regulating the season, duration, frequency or efficiency of fishing effort while leaving open the choice of whether to fish. Typical measures include limits on the time (closed seasons) or place (closed areas) of fishing, the size (minimum size requirements) or number (daily, weekly, or per trip catch limits) of fish caught, or choice of gear. Minimum mesh size for trawler or purse seine nets is the most typical gear regulation, but others include maximum length of driftnets, permitting only certain types of nets, or banning trawlers and purse seiners in nearshore areas. Some ecologists and fisheries sociologists advocate relying exclusively on severe input controls not only to protect fish stocks but to promote small-scale fishing operations by reducing the level of capital investment in industrial fisheries and preserving artisinal fisheries, particularly in Third World countries. Yet indirect controls can be hard to enforce, and will not address problems of by-catch in areas where species of different sizes commingle. However, the problems raised by continued investment in more efficient gear and greater awareness of by-catch problems have increased fisheries managers' interest in using indirect controls as supplemental measures in limited entry and quota fisheries.

Limited entry

Limited entry requires fishermen (more commonly, fishing boats) to secure a fishing licence. To succeed in protecting stocks, licensing requires that fisheries managers estimate sustainable catches of fish stocks accurately and distribute licences so that fishing capacity matches them. With fluctuating fish populations, this can be very difficult. Limited entry is easy to implement in an uncrowded fishery; in a crowded one the shift from open access to limited entry creates severe distributional problems because anyone who fails to secure a licence must leave the fishery. Even when fisheries are not crowded, licences can be used to distribute fishing rights among groups of fishermen. In Norway, for example, licensing limits the number of larger vessels and favours fishermen from the isolated communities of the north in line with Norway's overall regional policy. Licensing is seldom cost-effective in fisheries involving large numbers of small boats, yet licensing only larger boats invites evasion by acquiring boats of sizes just below the licensing limits. In the UK, the initial minimum of 40 foot length had to be reduced to 35, and in Iceland the number of small line boats increased enough that their share of the catch went up threefold (Coull, 1993:166). One way to control administrative costs is to delegate licensing of smaller boats to local authorities, as is done in Japan. Success depends, however, on the local authorities having sufficient administrative capacities for such tasks. Even when adequately enforced, limited entry will not solve overfishing problems if the licensed fishermen upgrade equipment; hence the increasing tendency to stipulate boat size, engine power and gear type in licences as well as to combine licensing with other indirect controls.

Individual quota (IQ)

Under an IQ system individual fishermen or, more commonly, boats are allocated quotas of each year's catch; in essence fishermen are allocated property rights in fish. Early individual quota systems expressed quotas by weight, generally using a boat's size and landings in the previous 3–5 years as a guide. Fisheries managers quickly found that this created severe problems when total catch had to be reduced to avoid or reverse overfishing. Expressing quotas as percentages of the total allowable catch allows accommodation of stock fluctuations while still offering fishermen a strong measure of predictability. Individual quotas have several advantages. First, vesting rights to a certain amount of fish avoids the strong competition to fish first that prevails in TAC regimes. Fishermen can moderate their gear investments and distribute their fishing across time to maximise the price they realise in the market; for both reasons their effort is distributed more efficiently. Second, multi-year assurance of a quota creates vested interest in the overall health of the fish stock. However, IQs do not always work as well in practice. Stock assessments are not always accurate, and fishermen still have some incentive to misreport or underreport catch, particularly if the future quota depends on the past fishing record. Enforcement requires considerable administrative capacity: fisheries managers need to be able to keep track of each fisherman's landings. Where fishermen themselves accept the system they will aid enforcement, because quota holders have an incentive to report cheats, but this depends on good relations between fishermen and fisheries managers. Nor do quotas eliminate the wastage of discarding unwanted fish at sea, most of which die soon afterwards. These discards are both by-catch and the result of 'highgrading' —keeping only the larger or more valuable fish of the species sought and dumping the rest back into the sea. Some experts believe that the wastage problem could be solved by allocating quotas of fishing effort rather than of catch, and leaving decisions about which species to catch to the fishermen (e.g. Hannesson and Steinshamn, 1991). Some governments have sought to address the problem with a combination of gear regulations and regulations allowing the sale of by-catch.

Individual transferable quota (ITQ)

ITQs are IQs that can be bought and sold. The transfer provision introduces an element of flexibility into the industry by allowing fishermen to reduce their effort, switch to different species, or leave fishing entirely. The provision for transfer makes ITQs the most market-oriented form of fisheries management, favoured by the neo-liberal economists now providing policy advice in many countries. Critics regard them as tools for favouring corporate fishing interests over others—particularly artisinal fishermen and indigenous peoples—and promoting concentration, as large firms buy out small ones. The extent to which transfers promote concentration can be controlled by limiting the portion of quota that can be held by any one fisherman or company. Accommodating other social interests requires complementary policies.

BOX 10.1 NEW ZEALAND'S ITQS

New Zealand established IQs for offshore fisheries in the 1978 legislation on fishing in the newly proclaimed EEZ. This was a preventive measure, since none of the fisheries were heavily exploited at the time, but it did shift fishing patterns by reserving quotas to wholly New Zealand-owned vessels and permitting no one enterprise to have more than 20 per cent of any TAC. Later the government permitted minority foreign ownership and raised the ceiling to 35 per cent of TAC. IQs were transformed into ITQs in 1986 by permitting the purchase or sale of annual catching rights. A computerised exchange now permits highly effective trading. Quotas are monitored by requiring that fishermen land catches at designated ports and sell only to licensed fish merchants. The 1986 law also extended ITQs to inshore fisheries to address the problem of overcrowding. Inshore fisheries had been managed by a fairly tight limited entry system until 1963, when the government sought to promote fishing by easing the conditions of access. But this easing went too far, and led to overfishing. The government had to suspend issuing permits for many fisheries in 1978 and 1980 and there was thus wide support, particularly from the Federation of New Zealand Fishermen, for using ITQs inshore.

Under the original system, fisheries managers were required to set total catch levels to ensure a maximum sustainable yield, and then distribute this total among current fishermen on the basis of their recent catch records. In extending ITQs to the inshore fisheries, the government also established a quota buy-back programme to encourage exit. Though the government allocated some NZ$45 million (about US$23.6 million) for this, few fishermen participated. Expenses mounted in following years after fisheries managers decided that additional reduction of catch was required and quotas were bought back. The prospect of continued high levels of spending led the government to shift the method of determining quota from weight of fish caught to percentage of the total catch in 1990. This permitted greater flexibility in year-to-year management, but also provoked huge controversies over losses of rights.

The inshore ITQs were also opposed by the Maori, who claimed that their traditional rights to fish had been ignored in the 1986 allocations. They sued, securing a court injunction against further allocations. The government then bought back 10 per cent of the allocations for transfer to the Maori, who accepted this as an interim settlement while continuing their litigation.

Though the ITQ system has protected most stocks, management has been handicapped by lack of sufficient knowledge of stocks and failure to react quickly enough to indications that certain fisheries were still being overfished. Failure to prevent overfishing has been most obvious in the orange roughy fishery, and stems partly from poor data about the stocks and the operation of industry pressure to continue authorising high quotas. This was most apparent in 1993. Though government and fishing industry scientists agreed that the total catch should be set at 3,400–5,900 tonnes, the Minister of Fisheries decided on 14,000. A Greenpeace lawsuit charging the minister with yielding to industry pressure prompted a reduction to 8,000 tonnes, but environmentalists regarded even this as 'too little and too late' (Duncan, 1995:99). The government's inability to devise measures that entirely eliminate highgrading and dumping of by-catch is another significant weakness of the system.

Territorial use rights in fisheries (TURF)

TURF systems involve distributing rights to fish in particular areas of water. The distribution might be done on an individual basis, but most advocates envision distribution on a communal basis—reserving areas for use by inhabitants of a particular district or village. Advocates regard them as particularly useful in the developing countries, where large numbers of traditional fishermen and weak government bureaucracies make licensing, IQs and ITQs unworkable. TURFs are also used in various parts of the industrial world, for instance, in the lobster fishery off Maine. TURFs simplify monitoring and enforcement by creating strong incentives for the individual or community rights holders to police themselves. Unless protected by the government, however, weak TURF holders can have difficulty protecting themselves from more powerful outsiders, such as local or foreign trawler fleets.

Communal management

This is an umbrella term for vesting management of fishing in the hands of local fishing communities. There are several variants in use around the world. In the UK since 1984, most of the national quota has been divided into 'sector quotas' given to each of the Producers' Organisations (associations of boat owners) and monitored by them. While individual boats may seek allocations directly from the government, a considerable portion of the fleet operates under the sector system. More traditional variants, based entirely on local initiative, exist in most countries of the world (Hannesson and Kurien, 1988). Such systems can involve formal allocation of TURFs, but critics of current fisheries management regard even TURFs as too similar to individual property, too likely to be imposed from above without regard for community conditions to work properly, and too likely to result in government distribution of rights to those with the best connections to the elite. They reserve the term 'communal management' for systems of exclusively local management relying mainly on indirect controls, particularly bans on trawling and purse seining. Where national governments do support local initiatives, communal management offers good prospects of avoiding overfishing and overinvestment.

Zoned management

Some governments attempt to balance between different groups of fishermen and types of fishery by dividing their national waters into distinct zones reserved to different sorts of fishing. In eastern Canada, the larger boats, mainly those over 100 feet, are assigned individual quotas and required to fish in the rougher waters off Labrador and eastern Newfoundland. The smaller boats have group quotas for the season, supplemented by weekly or per-trip limits. Indonesia and Malaysia have developed particularly elaborate zoning schemes to allocate opportunity among subsistence and other artisinal fishermen, smaller industrial fishermen and larger industrial fishermen. Actually maintaining the zones depends on considerable patrolling and enforcement, a requirement that challenges the capacities of the most well-supplied governments. Yet zones offer the possibility of applying management schemes more sensitive to the needs of various groups.

The social dimensions of fisheries take distinct form in industrial and developing countries. In most industrial countries, fishermen and those employed in related lines of work form a small part of the population, and their industry a small part of the national economy. However, their concentration in particular localities permits them to wield considerable political clout. They also benefit from the popularity of government assistance for maintenance of traditional ways of life. Consumer interests tend to find little expression because the cost of fish is a relatively small part of household budgets. Yet fishermen's influence is often undercut by divisions between fishermen targeting different species or between smaller and larger operators.

Fisheries in developing countries involve a larger portion of the population, partly because of its predominantly artisinal character and partly because open access rules permit fishing to function as a 'reserve occupation' for people unable to gain employment elsewhere. In many developing countries—Peru, Indonesia, Thailand, Malaysia and India are prominent examples—large numbers of subsistence and artisinal fishermen operate alongside a locally or foreign-owned industrial fleet, with the industrial fishermen often having more influence than the others.

BOX 10.2 INDONESIA'S AND MALAYSIA'S ZONE SYSTEMS

Both the Indonesian and Malaysian governments face the problem of allocating fishing rights among artisinal fishermen—including a large subsistence group—and industrial fishermen—many of whom are foreign nationals. Both governments have sought to address various needs and reduce intergroup conflict by establishing a set of distinct fishing zones where different gear is permitted. The Indonesian system involves four successive bands of sea: shore to 3 miles, 3 to 7 miles, 7 to 12 miles, and beyond 12 miles. Access is defined by a combination of boat size (tonnage), engine size and gear employed, with trawlers permitted only in the outermost zone. In Malaysia, the conflict between artisinal and industrial fishermen is exacerbated by ethnic differences. Malays dominate the artisinal sector, while ethnic Chinese dominate the industrial one. This helped inspire a more complex system of four zones. All waters within 5 miles of shore are reserved to artisinal fishermen using their own boats, those between 5 and 12 miles out are open only to trawlers and purse-seiners less than 40 tonnes, those between 12 and 30 miles are reserved to wholly Malaysian-owned and operated boats larger than 40 tonnes, and waters more than 30 miles out may be fished by Malaysian-owned or foreign boats over 70 tonnes. The government has attempted to exercise additional control over fishing through a licensing system that now extends to all boats and favours smaller craft. National plans for the middle water fleet include limiting it to 646 vessels. Though the zoned systems do show real potential for protecting artisinal fishermen and controlling the size of trawler fleets, limited enforcement capacity has left both vulnerable to violation by licensed boats intruding into other zones and unlicensed boats entering the fishery.

In recent years, fisheries managers have developed the outline of a new multidimensional form of fisheries management. On the biological/ecological side, this involves taking greater account of the full complex of variables affecting fish populations by using 'reference points' rather than a single expected yield. Under this approach, endorsed by the FAO in 1993, the focus of management would shift from catch levels to stock biomass levels. Maximum yield (based on the new ecological models of fish population dynamics) would serve as an outer limit, with actual catches kept below that level. Other reference points would indicate the sizes of spawning biomass at which various management measures—including moratoria, should biomass sink too far—would be initiated immediately (see Caddy, 1995).

On the economic/social side, multidimensional management involves much greater focus on reversing or avoiding overinvestment. Yet there is little agreement on how to do this. Many experts believe that allocation of property rights in fish, fishing effort or fishing areas is the only way to solve the problem. Critics contend that such measures would lead to elimination of all small-scale fishermen and intensification of the industrial forms of fishing they regard as particularly harmful. These economic disagreements feed into and are fed by the social ones. Consensus that the social dimensions of fisheries require attention is not matched by agreement on what policies should be pursued. Competing positions rest more on broad value preferences than analysis of fisheries as such, making it hard to reconcile differences while there is no agreement on which values should guide policy.

Management of fish raising

Fish raising has not faced the difficulties of moving away from an open access system, because development of fixed facilities occurs only after individual or communal property rights have been established over the particular areas used for cultivation. Defining such rights means confronting two distinct questions: the division of areas between aquaculture and other possible uses, such as industrial development, coastal tourism or maintaining natural habitats, and the allocation of opportunities to participate in fish raising. The first issue is becoming more acute, particularly in the tropics, where the massive expansion of shrimp raising for foreign markets has involved taking over natural habitat or areas used for traditional fish raising.

Management of the fish raising subsector is not complicated by a backward-bending supply curve, and so it can rely more heavily on market signals for cues about the amount of fish raising to undertake. Though fish raisers do not have to worry that others' decisions about effort will lead to depletion of fish

BOX 10.3 THE EU COMMON FISHERIES POLICY

The EU was committed to the development of a Common Fisheries Policy (CFP) by the Treaty of Rome, which covered fisheries within the subjects of the Common Agricultural Policy (CAP). Like the CAP, the CFP has the stated goals of rational development of production, ensuring a fair livelihood for fishermen, and stabilising markets. The first CFP, developed in 1966, addressed sectoral structure and established the principle that the waters of each EU member would be open to fishermen from all other members. The EU was involved in setting conservation measures only after the move to 200-mile EEZs forced an overhaul of existing EU and national management institutions. The CFP established in January 1983 for the period 1983–2002 has four main elements:

1 Confiding the determination of TACs and their allocation among member states to the EU Commission.
2 Applying restrictive licensing schemes to protect local fisheries in the 'Irish Box' north-west of the UK.
3 Enforcing thorough at-sea inspection and mandatory reports of catch landed.
4 Centralising authority to deal with international fisheries issues in the EU Commission.

In practice, EU fisheries management has depended mainly on TACs and quotas because of the principle of relative stability. This requires that any measure reducing a member fleet's catch in one fishery must be balanced by measures in others allowing the affected country to maintain its total catch level. As TACs were adopted for additional species, this created very strong political pressures which in many cases led to scientific advice being ignored and the setting of TACs to accommodate inter-member deals.

Portugal and Spain entered the EU in 1986, with a combination of poor home fishing grounds and large fleets. They also needed to find new grounds, as coastal states in Africa and North America limited their operations. This massive addition to the existing overcapacity problem was partly mitigated by an agreement that only half of the fleets would have access to EU members' waters until 2002; in the meantime the EU would assist the other half in securing fishing rights elsewhere.

This added significantly to existing tensions, as catches declined and the CFP goal of reducing EU fishing fleets was largely ignored. By the end of 1991, the EU Commission conceded that the CFP had failed. Massive protest by UK fishermen against foreign (mainly Spanish) trawlers and by Danish, French and German fishermen against severe losses of income due to competition from lower cost fish from Iceland, Norway and Russia in 1992–1993 led to modifications of the system. Fisheries management was shifted to a licensing system effective in January 1995. The EU Commission also began setting multiannual TACs and integrating fisheries management more tightly. This meant, inter *alia,* opening the 'Irish Box' to non-local fishermen and preparing to adopt TACs in the Mediterranean. The EU also adopted a new programme of decommissioning and gear limitation aimed at reducing capacity in round-fish fisheries by 20 per cent, and in flatfish fisheries by 15 per cent, by 1996. The EU has achieved modest reduction of capacity, but conflicts over access remain acute because fishing capacity remains larger than needed for taking sustainable yields.

stocks available to them, they do face the hazard of market fluctuations. For example, the rapid development of Atlantic salmon cultivation in Norway and Scotland during the early 1980s led to oversupply and a sharp decrease in the price of what had been a luxury food. The resulting reduction of profit margins left Norwegian firms vulnerable when the European Community (now European Union) imposed a protective tariff to shelter member state operations from Norwegian competition (Bjorndal, 1990). The industry was just adjusting to the new situation when additional competition from Chile and North America reduced its market share in other countries and created a fresh set of difficulties.

Though not involved in determining total yield, governments do regulate fish raising in various ways. At the least, they define and support the property rights on which it depends. Many governments have also encouraged fish raising by leasing public riverbed or coastal areas at favourable rents, lending money to finance facility construction, subsidising operations, operating hatcheries, or sponsoring research and training programmes. Almost all have used licensing, taxes and regulations to control the type, extent and location of fish raising operations.

Both governments and fish raisers are now devoting more attention to reducing or avoiding the environmental damage caused by fish raising. Artificially concentrating fish affects the immediate surroundings in several ways. Even in carefully monitored operations, excess food is likely to be added through spillage or misestimation of fish appetites. While the excess is often degraded and recycled through natural processes, slow water circulation or low oxygen concentrations in the water can lead to increases in hydrogen sulphite that threaten the whole stock. Use of aerators averts this danger, but adds considerably to energy consumption. High concentrations of fish also create unnaturally large accumulations of faecal matter. The combination of faecal accumulations, excess feed and nutrient leaking from feed pellets before they are eaten can lead to increased eutrophication, with the negative consequences of killing off marine life, increasing populations of pests and hastening the spread of disease. Overstocking (attempting to raise too many fish in a particular place) increases competition for food, producing poorly fed individuals more susceptible to disease. This danger is greater in tropical operations because of warm water and faster growth cycles of fish. Domesticated fish are also in greater danger of exposure to infections from wild animals than are domesticated land animals (OECD, 1989:14). Countering disease through use of antibiotics encourages the evolution of antibiotic-resistant strains of bacteria.

All of these problems have led some observers to prefer traditional fish raising, as much for its lower impact on the environment as for its greater accessibility to the poorer segments of the population. This preference intensifies the 'industrial versus artisinal' controversies. At the same time, managers of industrial fish raising operations are coming to realise that they must tailor the size and timing of their operations much more closely to the natural conditions of the surroundings.

CONCLUSIONS

The broad trend of twentieth-century fisheries management has been a shift away from reliance on linear single-species models of fish population and simple models of fisheries economics, towards multi-species, ecological models of fish population dynamics, more elaborate models of fisheries economics, and greater awareness of the social and political issues raised by allocating fish among consumers and opportunities for fishermen to capture or raise fish. The transition is not yet complete; there is still considerable controversy over the best policies to follow and the best ways of pursuing chosen goals. Yet the current debates do point clearly to one conclusion: no single management regime is appropriate to all fisheries. In capture fisheries, TURFs and communal management work best when the fish are relatively localised and gear is limited in mobility or fixed in place; IQs and ITQs are most appropriate in single-species fisheries, where year-to-year stock fluctuations are not too great; IQs and ITQs applied to multi-species fisheries must be sufficiently flexible to deal with by-catch. Though the notion of indirect control has a general application, the type of indirect control needed in a particular fishery depends on the life cycle of the fish. Similarly, the type of property rights regime that best fosters fish raising depends on local conditions, while the nature of the operation is defined by the species to be raised and the habitat to be maintained. Thus sustainable fisheries management requires a multidisciplinary sensitivity to fish population dynamics, ecology, economics, sociology and politics.

REFERENCES

Alverson, D.L., Freeberg, M.H., Murawski, S.A. and Pope, J.G. (1994) *A Global Assessment of Fisheries Bycatch and Discards,* FAO Fisheries Technical Paper No. 339, Rome, Italy: FAO.

Bjorndal, T. (1990) *The Economics of Salmon Aquaculture,* Oxford, UK: Basil Blackwell.

Caddy, J.F. (1995) *Reference Points for Fisheries Management: Their Potential Application to Straddling and Highly Migratory Resources,* Rome, Italy: FAO.

Caddy, J.F. and Gulland, J.A. (1983) 'Historical Patterns of Fish Stocks', *Marine Policy* 7: 267–278.

Coull, J.R. (1993) *World Fisheries Resources,* London, UK and New York, USA: Routledge.

Cunningham, S., Dunn, M.R. and Whitmarsh, D. (1985) *Fisheries Economics: An Introduction,* London, UK: Mansell.

Duncan, L. (1995) 'Closed Competition: Fish Quotas in New Zealand', *The Ecologist* 25, 2/3: 97–103.

Fairlie, S., Hagler, M. and O'Riordan, B. (1995) 'The Politics of Overfishing', *The Ecologist* 25, 2/3: 46–73.

Food and Agriculture Organization (FAO), Committee on Fisheries (1991) *Environment and Sustainability in Fisheries,* Rome, Italy: FAO.

Food and Agriculture Organization (FAO) (1993) *Marine Fisheries and the Law of the Sea: A Decade of Change,* FAO Fisheries Circular No. 853, Rome, Italy: FAO.

Hannesson, R and Kurien, J. (1988) *Studies in the Role of Fishermen's Organisations in Fisheries Management,* FAO Fisheries Technical Paper, Rome, Italy: FAO.

Hannesson, R. and Steinshamn, S. (1991) 'How to Set Catch Quotas: ConstanEffort or Constant Catch?', *Journal of Environmental Economics and Management* 20, 1: 70–91.

Murphy, G.I. (1982) 'Recruitment of Tropical Fisheries', in D.Pauly and G.I.Murphy (eds) *Theory and Management of Tropical Fisheries,* Cronulla, Australia: CSIRO, pp. 141–148.

Organization for Economic Co-operation and Development (OECD) (1989) *Aquaculture: Developing a New Industry,* Paris, France: OECD.

Pauly, D. (1986) 'Problems of In-Shore Tropical Fisheries: Fishery Research on Tropical Soft-Bottom Communities and the Evolution of its Conceptual Base', *Oceans Yearbook* 6: 29–37.

Rothschild, B.J. (1996) 'How Bountiful are Ocean Fisheries?' *Consequences: The Nature and Implications of Environmental Changes* 2, 1: 15–24.

Ruddle, K. and Zhong, G. (1988) *Integrated Agriculture-Aquaculture in South China: The Dike-Pond System of the Zhujiang Delta,* Cambridge, UK: Cambridge University Press.

Sherman, K. (1991) 'Sustainability of Resources in Large Marine Ecosystems', in K.Sherman, L.M.Alexander and B.D.Gold (eds) *Food Chains, Yields and Management of Large Marine Ecosystems,* Boulder, Col., USA: Westview, pp. 68–90.

Tett, P. (1977) 'Marine Production', in J.Lenihan and W.W.Fletcher (eds) *The Marine Environment,* Glasgow, UK: Blackie, pp. 1–45.

SUGGESTED READING

Burke, W.T. (1994) *The New International Law of Fisheries: 1982 UNCLOS and Beyond,* Oxford, UK: Oxford University Press.

Cushing, D.H. (1995) *Population, Production and Regulation in the Sea,* Cambridge, UK: Cambridge University Press.

Fairlie S. (ed.) (1995) 'Overfishing: Its Causes and Consequences', Special double issue of *The Ecologist* 25, 2/ 3: 41–127.

McGoodwin, J.R. (1990) *Crisis in the World's Fisheries: People, Problems, and Politics,* Stanford, Calif, USA: Stanford University Press.

Meltzer, E. (1994) 'Global Overview of Straddling and Highly Migratory Fish Stocks', *Ocean Development and International Law* 25, 3: 255–344.

Sherman, K., Alexander, L.M. and Gold, B.D. (eds) (1991) *Food Chains, Yields, Models, and Management of Large Marine Ecosystems,* Boulder, Col., USA: Westview Press.

SELF-ASSESSMENT QUESTIONS

1 Fish raising has expanded recently because
(a) governments are now providing large subsidies;
(b) it is more energy-efficient than capture fisheries;
(c) depletion of capture fisheries and high demand for fish make it profitable.

2 MSY is no longer the prime management goal in capture fisheries because it
(a) promotes overinvestment and leaves stocks more vulnerable to natural fluctuation;
(b) conflicts with the investment choices of fishing enterprises;
(c) encourages a level of fishing likely to deplete stocks.

3 Artisinal fish raisers
(a) produce more fish than industrial fish raisers;
(b) use less capital, less energy, and more labour per kilogram of fish produced than industrial fish raisers;
(c) breed fish for others to grow.

4 Tropical fish species
(a) live longer than cold water fish;
(b) swim further than temperate water fish;
(c) spawn more often than temperate water fish.

5 The UN Convention on Highly Migratory and Straddling Fish Stocks
(a) permits coastal states to extend national jurisdiction up to 400 miles from their shores;
(b) endorses concepts of reference point management;
(c) obligates flag states to co-operate more closely with coastal states in regional fisheries commissions.

6 Multidimensional fisheries management
(a) addresses all aspects of fish distribution to consumers;
(b) addresses the ecological, economic and social aspects of managing the fishing sector;
(c) was adopted at the Third UN Conference on the Law of the Sea.

11

SUSTAINABLE FORESTRY MANAGEMENT

Ivan Raev

SUMMARY

Forest ecosystems are among the most important components of the biosphere of Earth and are an essential condition for the existence of humankind. This chapter provides data in brief on the wealth of forests and attempts to outline the exceptional role of forests in the improvement of life on the planet. The main ecological problems in forests are outlined: the destruction of forest resources, especially in tropical regions; air pollution, acid rain and damage to forests; the vulnerability of forests to climatic change caused by the greenhouse effect in the next century; the reduction of biodiversity of forests; forest fires, etc. There are descriptions of the evolution of ideas for sustainable forestry, the theory of the normal forest, the principles of constancy and balanced use, and the principle of sustainable development of forests. The strategy for adaptation of forests towards future climatic change is outlined. Grounds are set out for sustainable management of forests on the basis of entire watershed areas, as a variant of the best possible form of contemporary forestry.

ACADEMIC OBJECTIVES

This chapter aims to define the principal ecological problems in forestry, as well as the potential for their possible solution. It contains the following parts:

- Brief information on forest resources.
- Concern for the future of humankind given the irresponsible use of forests.
- Contemporary concepts of global climatic changes and the state of forest ecosystems.
- Steps for the reduction of vulnerability of forest ecosystems and their adaptation depending on climatic change.
- Formulation of the principles of sustainable development of forest ecosystems.
- The outlining of an ecologically sound system of forest management.
- Differences in ecological problems in forests of different countries.

DEFINITIONS

Forestry is the science of biology and ecology of forests. It is simultaneously a science, business and art of creating, conserving and managing forests and forest lands for the continued use of their resources, material or otherwise (Sharma, 1992).

The concept of a forest comprises the biological community where trees and other tree species prevail (Spurr and Barnes, 1980). However, a contemporary view of forests is broader in scope and it would be more to the point to speak of forest ecosystems, as a forest is a very complex ecological system, of which vegetation is only a part. An ecosystem is a complex of live creatures and their environment (Odum, 1971). Some researchers prefer the term 'biocenosis' (Sukachev and Dylis, 1964), but ecosystem is more commonly used, and the two terms are generally considered to be synonymous. The object of study in forestry is the relationship between forests and humankind.

A forest ecosystem is a complex ecosystem consisting of various trees and other organisms, constituting a given community and its environment.

In other words, forest communities and the environment together make up a forest ecosystem (or forest biocenosis) of organisms and environment in a state of interaction within broad and complex cycles of carbon, water and nutrients.

Each forest ecosystem (biocenosis) consists of the following elements:

- Ecotope: atmosphere; water; soil.
- Biocenosis: vegetation; animals; micro-organisms.

The interrelationship between the separate parts of the ecosystem is a characteristic feature of the conditions of their existence, and can be illustrated schematically as in Figure 11.1 (after Sukachev and Dylis, 1964).

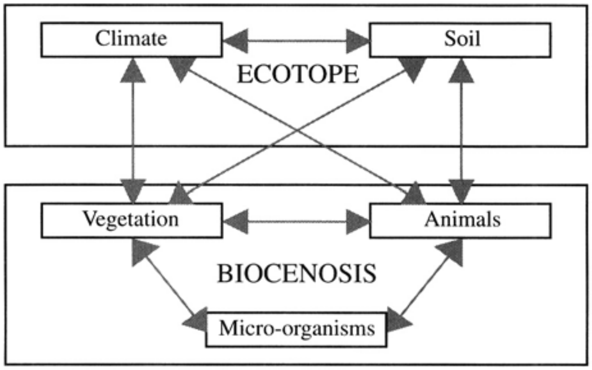


Figure 11.1 Structure and interaction of an ecosystem
Source: Sukachev and Dylis, 1964

Trees, the principal component of a forest ecosystem, fall into two types, namely broad-leaved and coniferous. Broad-leaved tree species are also known as hard nonconiferous or angiosperm species. Coniferous species, mostly evergreen, are also known as softwood or gymnosperm species.

BRIEF DATA ON FORESTS IN THE WORLD

To understand the role of forests as areas of production and accumulation of organic matter on Earth, let us turn to Table 11.1 (after Leith and Whittaker, 1975).

Approximately 47 per cent of the annual biomass is generated in forests covering over one-third of the area of the continents. The hypothesis of the major part played by the world ocean in the production of biological matter has not been confirmed. Oceans and seas constitute over 70 per cent of the surface of the Earth but yield only about 32 per cent of the annual biomass.

However, the processes of accumulation of organic matter in the oceans are insignificant in comparison with the deposition of organic matter on the continents, and in forests particularly. Approximately 92 per cent of the biological matter on Earth is concentrated in forests (1700×10^9 t of the total of 1841×10^9 t on Earth). Only 0.2 per cent of that quantity (3.9×10^9 t) is accumulated in the world ocean.

Timber is an important element of biomass. It constitutes the biomass together with branches, foliage, undergrowth, litter and soil organic material. The greatest stocks of timber are in tropical rain forests (45 per cent of the forests); tropical broadleaf forests account for 15 per cent, boreal 14 per cent, moderate zone broad-leaf forests 12 per cent, moderate zone evergreens 10 per cent, etc.

Clearly forests are the largest source and accumulator of biomass on Earth, which explains their exceptionally important position in the economy of nature.

The distribution of forests on continents is quite uneven. Although about 45 per cent of Africa is covered by forests, only 7.5 per cent of the continent has true forests (Table 11.2, after Sharma, 1992). In Asia and the Pacific only 24 per cent of the territory is forest-covered, while forests in a relatively good state account for 14.9 per cent of the territory, followed by North America (25 per cent), Europe (29.1 per cent), Central and South America (35.7 per cent) and the former USSR (35.8 per cent of the territory). This reflects the differing ecological conditions among the continents, as well as the impacts of human activity on forests.

The differences between continents are considerable, and are related to timber stocks (Table 11.3). Asia and Africa are the poorest in timber; Europe and North America follow. Central and South Africa and in particular the former USSR (with 24.2 per cent) have the greatest shares of the world timber stock.

Table 11.1 Gross primary productivity and biomass of vegetation

Type of ecosystem	Gross production Area $10^6 km^2$	m^2 $g/m^2/y$	Total $10^9/g/y$	Biomass kg/m^2	Total mass $10^9 t$
Oceans and seas	361.0	152	55.0	0.01	3.9
Tropical rain forest	17.0	2200	37.4	45.00	765.0
Tropical broad-leaf forest	7.3	1600	12.0	35.00	260.0
Evergreen moderate zone	5.0	1300	6.5	35.00	175.0
Broad-leaved moderate zone	7.0	1200	8.4	30.00	210.0
Boreal forests	12.0	8000	9.6	20.00	240.0
Rare forests and shrubs	8.5	700	6.0	6.00	50.0
Total forests	56.8		79.9		1700.0
Total continents	149.0	773	115.0	12.30	1837.0
Total (the planet)	510.0	333	170.0	3.600	1841.0

Source: Leith and Whittaker, 1975

Table 11.2 Forest area on Earth

Continents	Forest area thousand. ha	% of the continent	Closed forests thousand. ha	% of the continent
Africa	1,339,077	45	222,419	7.5
Asia and the Pacific	831,646	24	517,068	14.9
Central and South America	1,291,445	63	730,875	35.7
North America	734,353	40	459,356	25.0
Europe	178,353	38	136,652	29.1
Former USSR	929,600	42	791,600	35.8
Total	5,304 597	40	2,857,970	21.6

Source: Sharma, 1992

It should be noted that the valuable coniferous trees are concentrated chiefly in the Northern Hemisphere, particularly in Europe, North America and the former USSR, where coniferous forests account for between 68 and 82 per cent. In Central and South America, as well as in Asia and the Pacific region coniferous forests make up only between 3 and 4 per cent.

World-wide, broad-leaf trees are more important than coniferous trees, both in terms of the area they occupy and of their economic value (for example, the tropical hardwood species).

This schematic description of the state of world forest resources suffices to prove the enormous role played by forests on the planet: they occupy 38 per cent of the continents, are responsible for 69.5 per cent of annual production and accumulate 92.5 per cent of the total biomass of all continents.

As this annual biological production consists chiefly of carbon (C), taken in by green plants in the form of CO_2, it is evident that about 70 per cent of CO_2 taken in from the atmosphere is accumulated in forest vegetation. The total quantities of C in timber account for about 92 per cent of the accumulated organic matter on the continents; forest ecosystems are the main accumulator of carbon and carbon dioxide on Earth and form a powerful buffer against the greenhouse effect.

Through the process of photosynthesis forest vegetation not only takes in CO_2 from the atmosphere, but also gives off oxygen (O_2), creating the conditions vital for the survival of organisms and at the same time exerting a powerful impact on air temperature and the microclimate of the Earth. This pertains especially to young forests, as the processes

Table 11.3 World timber stock

Continents	Total stock M m^3	Stock m^3/ha	Coniferous forests %	Total stock %
Africa	38,790	–	–	14.0
Asia and the Pacific	31,475	–	4	11.3
Central and South America	78,637	–	3	28.3
North America	46,354	113	72	16.7
Europe	15,418	120	68	5.5
Former USSR	66,996	125	82	24.2
Total	277,670			100.0

Source: Sharma, 1992

of releasing oxygen and carbon dioxide in older ones are balanced.

To illustrate the exceptional part played by forests in the establishment of conditions for life on Earth, we list the material uses and ecological advantages forest vegetation offers to humankind:

- The enormous quantity of timber for direct use in the form of building material, firewood, tim ber for chemical and mechanical processing; im proved quality of the air through intake of CO_2 and other noxious gases and the emission of oxy gen, phytoncides, etc.
- Large amounts of non-timber forest products: fruits, mushrooms, herbs, resin, game, fish, etc.
- The establishment of a more suitable climate and microclimate on Earth.
- A consolidating and stabilising role for the con servation of soil wealth of lands; the struggle against water and wind erosion, floods, ava lanches, land-slides, etc.
- An exceptionally favourable role in regulating the water balance of watersheds, the quantities and quality of river waters.
- The preservation of biological diversity on Earth.
- Recreational needs for humanity.

ECOLOGICAL PROBLEMS OF FORESTS

The main ecological problems of forests are these: rapid felling of forest stock and deforestation of forest areas; uncontrolled grazing of livestock; pollution of forest ecosystems; treatment of forests with hazardous substances; aridification and heating of the climate, caused by the greenhouse effect; erosion of forest watersheds; reduction of species variety and stability of ecosystems, forest fires, etc.

Deforestation of forest areas

Deforestation is the principal problem of forestry. According to the World Resources Institute (WRI) (Houghton, 1990) over the period 1850–1980 forests in North Africa and the Middle East have been reduced by 60 per cent; in South Asia by 43 per cent; in tropical Africa by 20 per cent; in Latin America by 19 per cent. Over the same period forest cover on the planet has been reduced by 15 per cent.

The process of deforestation continues to this day. The deforestation rate for tropical forests over the period 1980–1985 averages 11.4 million ha per year (FAO, 1988). The latest studies show that deforestation of tropical forests currently runs at the rate of 17–20 million ha per year (Sharma, 1992). According to Houghton (1990), if the current rate of deforestation of tropical forests continues (1 per cent annually, or more) the expectations are that tropical forests will be entirely destroyed in the twenty-first century.

Among the main reasons for deforestation in tropical forests are agricultural expansion, felling of forests for firewood and uncontrolled grazing.

It is generally considered that 35 per cent of deforestation in Latin America is caused by small-scale farming, and the remaining 65 per cent by turning forests into pasture. The situation in Africa is similar (Browder, 1988; Mahar, 1989). The same reasons apply to 70 per cent of forest destruction in South-East Asia, Indonesia and West Africa (FAO, 1982).

Uncontrolled grazing and the gathering of firewood are together chiefly responsible for deforestation of North Africa, Sahel, the Middle East and South Asia. These are usually arid or semi-arid regions where vegetation is vulnerable to agricultural and livestock expansion. The gathering of firewood is an important factor in deforestation, particularly in Sahel, East Africa, the Himalayas region, plateaux in the Andes, as well as in densely inhabited regions of Central America. This is a critical problem, since almost three billion people, mostly in developing countries, rely on this energy resource (Sharma, 1992).

Air pollution and forests

The problem of rapidly spreading damage to forests and all natural ecosystems in Europe and North America was first raised urgently at the Stockholm conference on the environment in 1972. It was reiterated at the 1992 Rio Earth Summit.

It is considered that a number of pollutants in the air are principally responsible. The most important are sulphuric compounds (SO_x) and nitrogen compounds (NO_x), ozone, chlorine and aromatic substances, heavy metals, etc. which in the form of acids, dust and emissions are carried over large distances by atmospheric fronts. Mixed with water in the atmosphere, they create 'acid rain' with a reaction below pH 5.6 (Israel *et al.,* 1983).

The most important compounds are sulphur and nitrogen. In developed European countries 56 per cent of SO_2 is emitted in power generation, 28 per cent from industrial processes, 16 per cent from transport. Transport is the main source of NO_x pollution (45 per cent), followed by thermal power stations (31 per cent) and industry (24 per cent) (Ulrich and Pankrath, 1983). It is evident from these data that enormous quantities of sulphur emissions are released into the atmosphere above Europe; about a further 30 per cent are deposited beyond European boundaries, so polluting the other continents (Table 11.4).

Forest damage is undoubtedly the result of a variety of factors: weather fluctuations, gradual increase of pests or diseases, hydrological conditions (e.g. lowering of the water table due to drainage or water winning), soil type and texture, thickness of the humus layer and tree species all contribute to a decline in forest vitality. In spite of this, the data on

Table 11.4 Emissions of sulphur dioxide and nitrogen oxides from man-made sources in selected countries in Europe (data for 1990)

Country	SO_2 ($10^3 t_{SO2} a^{-1}$)	Per capita (kg a^{-1})	NO_x ($10^3 t_{NO2} a^{-1}$)	Per capita (kg a^{-1})
Albania	50	–	9	2.8
Austria	98	13.1	209	27.9
Belgium	420	42.3	300	30.2
Bulgaria	1030	114.6	150	16.8
Former Czechoslovakia	2564	177.1	1122	71.3
Denmark	266	52.0	254	49.6
Finland	256	51.5	290	58.3
France	1206	21.5	1742	31.0
Former German Democratic Republic	5242	314.9	750	42.3
Former Federal Republic of Germany	1002	16.6	2707	44.7
Greece	500	50.3	746	75.1
Hungary	1010	95.7	238	22.6
Iceland	6	24.8	12	49.8
Ireland	168	45.2	135	36.3
Italy	2406	42.0	1755	30.6
Luxembourg	10	32.7	15	40.9
The Netherlands	238	16.1	529	35.9
Norway	60	14.2	212	50.1
Poland	3210	83.5	1280	33.3
Portugal	212	20.6	142	13.8
Romania	1800	45.6	390	16.8
Spain	2190	56.7	950	24.2
Sweden	204	24.5	373	23.2
Switzerland	62	9.5	184	28.2
UK	3774	66.3	2690	47.3

Source: UNEP, 1993

Table 11.5 Approximate share of forests in 16 European countries affected by pollution

Country	*% of forests*	*Country*	*% of forests*
Sweden	4	Austria	26
Denmark	4	Poland	26
Norway	5	Czechoslovakia	27
Yugoslavia	10	Finland	35
Hungary	11	Switzerland	36
Great Britain	13	The Netherlands	50
Belgium	18	Luxemburg	52
France	24	Germany (FRG)	52

Source: Orsini-Rosenberg, 1986

damage to forests caused by industrial pollution are clearly understandable in line with acidifying emissions. The combined stress from acid deposition and photochemical oxidants may increase the potential for forest damage. According to the data in Table 11.5, forest damage caused by industrial pollution covers on average 4–52 per cent of the forest area in Europe (Orsini-Rosenberg, 1986; Nossel *et al.,* 1987). The situation is worst, and the impact highest, in Germany, Luxemburg and The Netherlands.

According to research in many countries, coniferous plantations, which often consist of exotic provenances and are managed purely for timber production, are less resistant to industrial pollution than are other forest ecosystems (Smith, 1981; Raev, 1987). This is particularly the case in Central and Western Europe, where over the past centuries local broad-leaved forests were replaced by coniferous monocultures.

Climatic change and forests

Forests are a geographical phenomenon, dependent on a specific climate; they are also a function of climate. It could be said that precisely established vegetation corresponds to various climatic types (Stefanov, 1930). Together with climatic change, changes occur in soil types and plant types on Earth. These changes are slow and gradual and take place throughout the geological history of Earth. However, over the past 100 years the preconditions for anthropogenic climatic change have grown. This is due to the sharp increase of emissions of CO_2 produced by industrial activity.

CO_2 concentration at Mauna Loa, Hawaii has risen from 315ppmv in 1958 to 354 ppmv in 1990. In other words, the annual concentration of CO_2 has increased by 3.7 Gt carbon, or by 0.5 per cent annually. The results of numerous other studies in the cleanest areas of the world, for instance, the Antarctic, where CO_2 concentration in pack ice grew from 2.77 ppmv in 1740 to 3.46 ppmv in 1988, are the same (Neftel *et al.,* 1985; Friedli *et al.*, 1986; Keeling *et al.*, 1989; Boden *et al.*, 1994).

At the world conference at Rio de Janeiro in 1992, the real danger of rapid change in the world's climate from the emergence of the so-called greenhouse effect was outlined. A rise in air temperature was expected on Earth, accompanied in many cases by changes in the quantity of precipitation and distribution of rainfall (IPCC, 1990, 1992). A convention was prepared for signing at Rio which called on all states in the world to maintain present CO_2 emission levels and reduce those levels over the coming decades.

If the trend of the increase of the greenhouse effect continues we have reason to expect drastic climatic changes as early as the twenty-first century. Forest ecosystems, which formed their specific structure and functions in the course of millions of years, will fail to adapt to the sharp changes in ecological conditions and will quickly disintegrate. This raises above all the need for serious research into the vulnerability of forest ecosystems and their adaptation to the conditions of climatic changes.

Reduction of biological diversity

In spite of the fact that tropical rain forests occupy only 7 per cent of the land surface, they contain 90 per cent of biological diversity on Earth (Lovejoy, 1988). According to Reid and Miller (1989), 700 tree species have been found in 10 ha of tropical forest in Borneo, while Raven (1988) counted up to 300 tree species per ha in the Peruvian part of Amazonia. For comparison, in forests in the eastern part of North America the respective species numbers vary between 10 and 20 and coniferous forests in Europe have only between one and five tree species.

If the destruction of tropical forests continues at the current rate towards the mid-twenty-first century we should expect reduction of biological diversity by

25 per cent (Reid and Miller, 1989; World Resources Institute, 1990; Raven, 1988). If the number of species on the planet today stands at around 10 million, this rate of destruction will cause between 15,000 and 50,000 species to be lost annually (Wolf, 1987).

There are several principal reasons in favour of the preservation of biological diversity:

- Utilitarian considerations. Each of the disappearing species could be useful to humankind and the planet. This use is not only in the form of timber, but also a variety of non-material uses (improvement of climate, hydrological regime, soil protection, recreation activities, etc.).
- Aesthetical considerations. Each of the threatened species has a unique form, adding to the variety of nature.
- Moral considerations. Each species has the moral right of existence.
- Ecological considerations. As natural conditions on earth vary, there is a need for enormous variety in existing natural ecosystems (Sharma, 1992).

Creation of coniferous monoculture in the moderate zone

For the last two to three hundred years, a few coniferous tree species, chiefly *Picea abies and Pinus sylvestris*, were substituted for local broad-leaf vegetation in Central and Western Europe. This is how matters reached their present state, where in countries like Germany, Poland, the Czech Republic and others coniferous species make up 70–90 per cent of forests.

This substitution of species was prompted because of the higher volume of obtained coniferous biomass and the easier formation of forests. However, it has been proven over past decades that this substitution was not justified. Research has shown that coniferous species lead to acidification of soils, and with the presence of atmospheric pollution (chiefly SO_x, NO_x and others) they cannot endure acid rain and start dying out on a mass scale.

Today in Europe there has been a return towards resistant local tree species, chiefly broad-leafed, with the view of maintaining mixed forest ecosystems, which are best adapted to the conditions of the natural environment.

Other ecological problems in forests

Other aspects of ecological problems in forests are erosion of forest soil and devastation from forest fires, tourist invasion, etc.

Usually the devastation of forests through excessive felling and uncontrolled grazing runs parallel to intensive erosion of former forest soils up to their full degradation. This is how entire watersheds in the Mediterranean, the Near and Middle East and elsewhere arrived at their present catastrophic condition.

Unfortunately these are not the only problems. In the same regions, known for their arid climate throughout the dry part of the year, there are devastating forest fires. In many cases these are started deliberately, in order to create conditions for primitive livestock breeding. This leads to a systematic loss of valuable forest ecosystems and impoverishment of the people living in these regions.

In areas where tourism is well developed, tourist digression, the degradation of forest ecosystems caused by excessive invasion of small forest plots, can be a problem. Examples of this exist mainly in the Mediterranean part of Southern Europe, where some sea resorts are hugely popular in the summer. In mountainous areas in Europe there is increased concern over erosion and flooding as a result of the increased establishment of ski lanes.

SUSTAINABLE MANAGEMENT OF FORESTS

The sustainable management of forests is based on the concept of forests as natural ecosystems and the directing of human activities towards the preservation of the multifunctionality and resistance of forest ecosystems.

The principle of regular and balanced use of forests

Prior to the arrival of the concept of sustainable development of forests, various technological methods for forest management were proposed with a view to obtaining more timber of higher quality. This led to the emergence of forestry systems, varying according to natural conditions, variety of tree species and economic activities.

The concept of the ‘normal forest’, formulated by German foresters early in the twentieth century, together with the principle of regular and balanced use of forests, was a major step forward in international forestry science. This idea was later connected with the principle of the ‘natural forest’ (Leibungut, 1975).

A requirement for a ‘normal forest’ is a distribution of forests according to age, with a balance of young, middle-aged and old forests. In this way regularity is achieved, and a constant rate in the use of timber, which is the final goal of forestry.

This principle turned out to be very useful, as it took into account the nature of the forest and the constant need of its use. Attempts to reject this principle as ‘reactionary’ in the management of forests led to the rapid exhaustion of all forest reserves and anarchy in forestry in East European forests from 1950 to 1980 (Marinov *et al.,* 1980).

The principle of sustainable development in forests

Over the past decades the view of the forest as a forest ecosystem came to the fore, along with the adoption of the idea of the polyfunctional use of forests.

The principles of sustainable development of forests were formulated in connection with the 10th World Congress in Paris in 1991 (World Forestry Congress, 1991) and the Conference at Rio de Janeiro (1992) and comprise the following main elements: maintaining biodiversity in forests and polyfunctionality of management of forests.

The view prevailed that forests should be seen as a polyfunctional ecosystem, a necessary condition for the existence and survival of human civilisation at the end of the twentieth century, while economic activities in forests should always be directed towards sustainable protection and development of the forest ecosystems.

The protection of biological diversity in forests can be realised through the promotion of a system of protected forests and territories on Earth; through the toleration of local climax types of vegetation (resistant and well adapted to ecological conditions); through specific genetic and selectionist activities, etc. It is important for the aims of sustainable development of forest ecosystems for some species to survive under the conditions of industrial pollution and of unfavourable climatic changes (Botkin and Talbot, 1992).

Polyfunctional management of forests presupposes the utilisation of the timber production function of forests, but also as an environment of enormous biological diversity. It encompasses non-timber products; anti-erosion protection measures and preservation of soil wealth; regulation of the quantities and quality of river waters—the water balance of watersheds; recreation and tourism; improvement of the microclimate; absorption of CO_2 and other harmful gases in the atmosphere, emission of O_2; reduction of the greenhouse effect to the atmosphere, etc.

In June 1993, the principles of sustainable forest management and of forest biodiversity conservation were adopted at the Ministerial Conference on the Protection of Forests in Europe, which took place in Helsinki. Six main criteria and 27 qualitative indicators of sustainable management of forests were adopted (Stanners and Bourdeau, 1995). The ratio of annual timber cut volume to annual wood increment can be used as one such qualitative indicator. For the countries of the Organization for Economic Co-operation and Development (OECD) this indicator varies from 0.39 for Australia to 0.96 for Germany, being in all the cases lower than 1 (UNEP, 1993).

Intensive investigations for developing a suitable system of criteria and indicators of sustainable forest management continue (Ravi Prabhu *et al.,* 1996; Lay-Cheng Tan, 1996; and others). The discussion about what entails sustainable forest management is closely linked with discussions about certification (ecolabelling) of timber from sustainably managed forests, to be introduced in 2000.

The strategy of adaptation of forests to climatic change

The danger of the enhancement of the greenhouse effect in the atmosphere and unfavourable change of the climate on Earth in the next century has been accepted as real in most special studies (Budyko, 1974; IPCC, 1990, 1991).

BOX 11.1 FOREST VEGETATION IN BULGARIA AS AN ACCUMULATOR OF CO_2 FROM THE ATMOSPHERE; VULNERABILITY AND ADAPTATION OF FOREST VEGETATION IN EXPECTATION OF CLIMATIC CHANGES

Absorption of CO_2 by Bulgarian forests

- In period 1987–1993 between 1,435,000 t and 2,364,000 t°C has been absorbed by Bulgarian forests, defined as the difference between the total annual growth of surface timber mass and the quantity of felled timber.
- In period 1955–1990, stocks of surface timber in forests in Bulgaria have grown from 245 to 396 million m^3, while the total area of forests has grown from 3.670 to 3.772 million ha.

Climatic scenario for the assessment of vulnerability and adaptation of forests

- Baseline climatic scenarios for the so-called contemporary climate are determined by means of data from 20 meteorological stations over an area of 111,000 km^2 for the period 1951–1980.
- Three models are used for general circulation of the atmosphere at the doubling of CO_2 in the coming century.
- The supposition is that there will be a rise of temperature from 2.0 to 5.9°C, and for precipitation from ± 50%.

Changes in the 'zones of life' (after Holdridge, 1967)

- Considerable change is expected in the 'zone of life', in particular between 0 to 800 m a.s.l., where approximately 60% of Bulgarian forests are situated. The change is unfavourable.
- Changes are from 'cool temperate moist forest' to 'warm temperate dry forest' and towards 'subtropical dry forest'. As the result of a research project, measures have been devised for increasing adaptation of Bulgarian forests, as well as strategies for enhancing forest potential for absorption of greenhouse gases.

Source: Raev *et al.*, 1995

Concern for the future of forests on a world-wide scale is justified. The expected warming of the atmosphere will create an unnatural environment for the existing forest vegetation, a change in its distribution and quantity of rainfall which might be disastrous for the survival of existing forest ecosystems (Sampson and Hair, 1992; Qureshi, 1992).

Under the Frame Convention of the UN in Climate Change, adopted at Rio de Janeiro (1992), an intensive research programme, sponsored by the USA has been launched, with the participation of over fifty countries. The programme is known as the US Country Study to Address Climate Change.

Box 11.1 is an example of Bulgarian participation in the project, called Bulgaria Country Study to Address Climate Change (1995). Here the problem is addressed in several main aspects, such as inventory of greenhouse gases emission sources; vulnerability and adaptation for Bulgarian forest and agricultural vegetation; alternative energy balances and technology programmes, etc. We will dwell briefly on the results on forest vegetation in Bulgaria, and as these results are of interest for many countries we will present the main measures for adaptation of forests under changed climatic conditions (Raev *et al.*, 1995).

An essential conclusion is that if climatic change occurs in Southern Europe in the course of the next century, most probably its negative manifestation should be expected chiefly in the low forest zone. It is probable that in higher parts of the mountains relatively more favourable conditions will remain. This calls for a differential approach in the future strategy.

At the threshold of possible climatic changes, the strategic task for forestry in the low forest zone, which is the most vulnerable part of forests (in Bulgaria generally below 800 m a.s.l.), should be to adapt forests towards aridisation of the climate and protect the wealth of forests from deteriorating ecological conditions. The most important measures consist of:

- a differential approach in afforestation in sepa rate regions of Bulgaria, given that aridisation is expected to be regionally more intensive in the following rising order: North Bulgaria-South Bul

garia—the Black Sea coast area—the south-west region;
- preservation and expansion of the mixed character of forests in the lower parts of the country, with preference to all species of trees and shrubs that endure a changed climate;
- special attention to windbreaks, which should be made up of more arid resistant species and should cover larger areas.

The potential for preserving more favourable ecological conditions in the higher parts of the mountains in Bulgaria (above 800 m a.s.l.) means that the goals of forestry activities should be different here: preservation of biodiversity in forests, sustainable development of ecosystems, multifunctional utilisation, development of a system of protected areas. The following strategic steps are possible:

- Preserving and expanding the mixed character of mountain forest ecosystems. In this way the requirement for an expansion of biological variety in forests will be met.
- The development of a programme for the establishment and maintenance of multi-storey coniferous and broad-leaved forests of various ages, with the highest ecological potential.
- Management of mountain watersheds should be directed towards multifunctional use: highest biological productivity, combined with a high level of water regulating, protective, recreational and climate forming functions.
- Overall development of a system of protected areas: national parks, reserves, protected forests, etc. through which biodiversity and high productivity, as well as protective functions of the forest, can be maintained.

Sustainable management of forests in watersheds

Sustainable management of the forest area can be applied only within the limits of entire watershed basins (Odum, 1971). Only in this way can the effect of various forms of economic activity be controlled and a high level of multifunctional utilisation of forests be realised.

In view of this it is essential to take the following steps at the planning phase of forestry management of watershed basins:

1. An exhaustive study of natural conditions in the area.
2. Exact data on an inventory of the forest resources in the watershed area.
3. Sustainable planning of forestry initiatives, including:
   (a) steps for technical consolidation, biological reforestation and recultivation of the terrain;
   (b) precedence before natural recovery of forests with indigenous local tree species;
   (c) the application of forestry models for management of ecosystems, which lead to a mixed, multi-storey and multi-age structure of forest ecosystems (i.e. a selective form of management);
   (d) planning the management of game and fish at the watershed;
   (e) steps for the improvement of the state of forests;
   (f) norms for permissible tourist load of the ecosystems;
   (g) fire-fighting structures in the watershed;
   (h) limits for the construction of sports facilities and hydro-technical construction work;
   (i) designating the zones for various forms of commercial activities, depending on the status of the protected area, etc.;
   (j) taking into account the selection of species for afforestation with a view to future climatic changes (Raev, 1989; Raev *et al.,* 1995).

CONCLUSIONS

The exceptional role of forests on Earth is evident from the fact alone that they occupy 38 per cent of the continents, produce 69.5 per cent of the annual production and accumulate 92.5 per cent of the total biomass of all continents.

The most serious ecological problems of forests around the world are these: intensive deforestation of tropical forest regions; air pollution from industrial sources and consequent damage to forests in Europe and North America; vulnerability and adaptive capacity of forests in view of the greenhouse effect and climatic deterioration in the coming century; reduction of biodiversity, particularly in tropical forests; the consequences of the creation of coniferous monoculture; forest fires, etc.

The quest for a 'normal forest' and the principle of regular and balanced use of forests are important stages in the evolution of ideas of sustainable forestry. Today the principle of sustainable development of forests has two main aspects: preservation and stability in biodiversity in forests and multifunctional utilisation of forests. It is necessary to develop a system of criteria and indicators of sustainable forest management to help create a forest policy appropriate to each country. On this basis it is important to build a strategy for forest adaptation to climatic change in the next century, taking local conditions and traditions into account. This strategy should be applied in the sustainable management of forests along watersheds.

REFERENCES

Boden, T.A., Kaiser, D.P., Sepanski, R.J. and Stoss, F.W. (1994) *Trends '93. A Compendium of Data on Global Change,* Oak Ridge, Tenn., USA: Oak Ridge National Laboratory.

Botkin, D.B. and Talbot, L.M. (1992) 'Biological Diversity and Forests', in N.Sharma (ed.) *Managing the World's Forests,* Ia., USA: K/H Publishing Company, pp. 47–74.

Browder, J.O. (1988) 'Public Policy and Deforestation in the Brazilian Amazon', in R.Repetto and M.Gillis (eds) *Public Policies and the Misuse of Forest Resources,* Cambridge, UK: Cambridge University Press.

Budyko, M.I. (1974) *Climatic Changes,* Leningrad, Russia: Gidrometeoizdat.

FAO (1982) *Tropical Forest Resources,* Rome, Italy: Forest Resources Division, Forestry Department.

FAO (1988) *An Interim Report on the State of Forest Resources in the Developing Countries,* Rome, Italy: Forest Resources Division, Forestry Department.

Friedli, H., Lotscher, H., Oeschger, H., Siegentholer, N. and Stouffer, B. (1986) 'Ice Core Record of the 13c/12c Ratio of Atmosphere CO_2 in the Past Two Centuries', *Nature* 324: 237–238.

Holdridge, L. (1967) *Life Zone Ecology,* San Jose, Costa Rica: Tropical Science Center.

Houghton, R.A. (1990) 'The Global Effects of Tropical Deforestation', *Environmental Science and Technology* 24: 414–422.

IPCC (International Panel on Climatic Change) (1990) *The Potential Impact of Climatic Change on Agriculture and Forestry. Likely Impact of Climatic Change,* Report from Working Group II to IPCC, WMO, UNEP.

IPCC (1991) *The IPCC Response Strategies,* WMO, UNEP.

IPCC (1992) *Climatic Change. IPCC 1990 and 1992 Assessments,* WHO, UNEP, Cambridge, UK: Cambridge University Press.

Israel, J.A., Nazarov, I.M., Pressman, A.J., Rovinskii, F.J., Rjaboshapko, A.G. and Filipova, L.M. (1983) *The Acid Rains,* Leningrad, Russia: Gidrometeoizdat.

Keeling, C.D., Bacastow, R.B., Carter, A.F., Piper, S.C., Whort, T.P., Heimann, M., Mook, W.G. and Roeloftzen, H. (1989) 'A Three Dimensional Model of Atmospheric CO_2 Transport Based on Observed Winds: 1. Analysis of Observational Data', *Geophysical Monograph* 55, American Geophysical Union.

Lay-Cheng Tan (ed.) (1996) *Initiatives on Assessing Sustainability: Status and Future Directions,* Summary of the Open Session of the Third IPAP Meeting on Testing Criteria and Indicators for Sustainable Management of Forest, Turrialbe, Costa Rica, 29 February-1 March 1996, Jakarta, Indonesia: CIFOR.

Leibungut, H. (1975) *Wirkungen des Waldes auf die Umvelt des Menschen,* Erlenbach, Zurich and Stuttgart, Germany: Eugen Rentsch Verlag (in German).

Leith, H. and Whittaker, R. (1975) *The Primary Productivity of the Biosphere,* Ecological Studies, 14, Berlin-Heidelberg-New York: Springer.

Lovejoy, T.E. (1988) 'Will Unexpectedly the Top Blow Off?', *Bioscience* 38, 10: 722–726.

Mahar, D.J. (1989) *Government Policies and Deforestation in Brazil's Amazon Region,* Washington, DC, USA: World Bank.

Marinov, M., Procopiev, E., Shipcovenski, D., Raev, I., Jeliaskov, P. and Gateva, R. (1980) *Reproduction and Conservation of Forest Ecosystems,* Sofia, Bulgaria: Zemizdat.

Ministerial Conference on the Protection of Forests in Europe (1993) *Report of the Follow-up of the Strasbourg Resolutions,* Helsinki, 1993, Helsinki, Finland: Ministry of Agriculture and Forestry.

Neftel, A., Moor, E., Oeschger, H. and Stanffer, B. (1985) 'Evidence from Polar Ice Cores for the Increase in Atmospheric CO_2 in the Past Two Centuries', *Nature* 315, 45–47.

Nossel, H., Krieger, H., Schefer, H. and Trost, N. (1987) *Modelling the Dynamic Process of the Dieback for Spruce and Beech,* Krakow, Poland: Workshop on Forest Decline and Reproduction: Regional and Global Consequences.

Odum, E.P. (1971) *Fundamentals of Ecology,* Philadelphia-London-Toronto: W.B.Saunders Company.

Orsini-Rosenberg, H. (1986) 'Die Oekonomische und Oekologische situation in der Oesterreichischen Forstwirtschaft', *Internationaler Holzmarkt* 77, 5: 1–5 (in German).

Qureshi, A. (ed.) (1992) *Forests in a Changing Climate,* Washington, DC, USA: Climate Institute.

Raev, I. (ed.) (1987) *Accumulation of Industrial Pollution on Forest Ecosystems,* Sofia, Bulgaria: N.T.S.

Raev, I. (1989) *Investigations on the Hydrological Role of the Coniferous Forest Ecosystems in Bulgaria*, Sofia, Bulgaria: Forest Research Institute.

Raev, I., Grozev, O., Alexandrov, V., Vassilev, Z., Rosnev, B. and Delkov, A. (1995) *Vulnerability to Climate Change and Adaptation Measures for Bulgarian Forests and Agricultural Vegetation*, Sofia, Bulgaria: Forest Research Institute.

Raven, P. (1988) 'Biological Resources and Global Stability', in S.Kawano, J.H.Connell and T.Hidaake (eds) *Evolution and Co-operation in Biotic Communities*, Tokyo, Japan: University of Tokyo Press.

Ravi Prabhu, Colfer, J.P., Venkateswarlu, P., Tan, L., Soekmadi, R. and Wollenberg, E. (1996) *Testing Criteria and Indicators for the Sustainable Management of Forests, Phase I, Final Report*, Jakarta, Indonesia: CIFOR.

Reid, W.V. and Miller, K.R. (1989) *Keeping Options Alive: The Scientific Basis for Conserving Biodiversity*, Washington, DC, USA: World Resources Institute.

Sampson, R.N. and Hair, D. (eds) (1992) *Forests and Global Change, v. I: Opportunities for Increasing Forest Cover*, Washington, DC, USA: An American Forests Publication.

Sharma, N.P. (ed.) (1992) *Managing the World's Forests*, Ia., USA: Kendal/Hunt Publishing Company.

Spurr, S.H. and Barnes, B.V. (1980) *Forest Ecology*, New York-Brisbane-Toronto: John Wiley and Sons.

Smith, W.H. (1981) *Air Pollution and Forests*, New York, USA: Springer-Verlag.

Stanners, D. and Bourdeau, Ph. (eds) (1995) 'Europe's Environment', *The Dobris Assessment*, Copenhagen, Denmark: European Environment Agency.

Stefanov, B. (1930) 'An Attempt for the Establishment of a Parallel Classification of Climatic and Vegetational Types', *Zbornik na Bulgarska Akademia na, naukite* 26, 1–122.

Sukachev, V. and Dylis, N. (eds) (1964) *Fundamentals of Forest Biogeocoenology*, Moscow, Russia: Science.

Ulrich, B. and Pankrath, J. (eds) (1983) *Effects of the Accumulation of Air Pollution in Forest Ecosystems*, Proceedings of Workshop of Goetingen, W.9., Dordrecht-Boston-London: D.Reidel Publishing Company.

UNEP (1993) *Environmental Data Report*, 1993–1994, Oxford, UK: Blackwell.

Wolf, E.C. (1987) *On the Brink of Extinction: Conserving the Diversity of Life on Earth*, Worldwatch Paper 78, Washington, DC, USA: Worldwatch Institute.

World Forestry Congress (1991) *Proceedings of the 10th World Forestry Congress*, 1991, Paris, France: V.1–9 Revue Forestiere Francaise.

World Resources Institute (WRI) (1990) *World Resources*, New York, USA.

SUGGESTED READING

Dassler, H.-G. (ed.) (1976) *Einfluss von Luftverunreinigungen auf die Vegetation*, Jena, Germany: UEB Gustav Fischer Verlag (in German).

SELF-ASSESSMENT QUESTIONS

1 Which is the main ecological and social problem for forests on a global scale?
 (a) Deforestation of forest regions.
 (b) Global warming and forest decline.
 (c) Climatic changes.

2 Which is the main cause of the greenhouse effect and changes in the climate?
 (a) Increase of SO_2, NO_x and other concentrations.
 (b) Increase of CO_2 emissions.
 (c) Thinning of the ozone layer.

3 Which is the main reason for acid rainfall and the decline of forests in Central Europe?
 (a) The high concentrations of SO_2, NO_x.
 (b) The creation of coniferous monoculture.
 (c) The increase of the concentration of CO_2.

4 Give the best principles for sustainable forestry.
 (a) The principle of regular and balanced use of forests.
 (b) The principle of the normal forest.
 (c) The principle of sustainable development of forests.

5 Which are the main constituents of the principle of sustainable development of forests?
 (a) Preservation of biological diversity in forests.
 (b) Multifunctional management of forests.
 (c) The maximum production of biomass.

6 What is the meaning of multifunctional forestry?
  (a) Optimal production of wood.
  (b) The utilisation of ecological functions of forests.
  (c) Secondary forest products.

12

DEVELOPMENT AND IMPLEMENTATION OF SUSTAINABLE OPTIONS IN INDUSTRY

Giancarlo Barbiroli

SUMMARY

Industry is undergoing great change in the present transformation process, commonly defined as 're-industrialisation'; however, several features and trends of this process are not consistent with sustainable development, notably the persistence of production/consumption features typical of 'open-cycle economic systems', which are in contrast with ecological systems.

This chapter deals with the role of industrial enterprises in long-term sustainable economic development and outlines the various tasks and actions that need to be accomplished to achieve the expected results. The main guidelines to be followed by industries, discussed by means of case studies, are as follows:

1 Adoption of proper criteria for selecting and managing technologies and products in ways enabling the extension of the lifespan of durable goods, the diffusion of cleaner production, the pursuit of new frontiers in the global performance, the intensification of the dematerialising process.
2 Creation and development of eco-compatible materials, chemicals, energy systems, foodstuffs and related technologies, to achieve a real technological pluralism.
3 Establishment of new enterprises suited to 'closed-cycle economic systems' and capable of creating employment opportunities.

In addition, the chapter focuses on steps and methods to assess the different possible technological options and pathways, in order to gather technical and economic data for the optimal choice, both for enterprises and the whole economic system.

In pursuing the aims of sustainable economic systems, industrial companies meet barriers and obstacles, both internal and external, which need several institutional, effective and stable interventions in order to be gradually overcome.

Finally, the efforts in R&D to be made by industry to develop appropriate and economically viable technologies are highlighted in order to illustrate the increasing number of solutions that must be considered a fundamental condition for achieving the goals of a sustainable development; notably, organised structures for technological evolution and transfer—such as the technopolitan sites—will become more and more essential and effective for industries in their reorientation and restructuring.

ACADEMIC OBJECTIVES

- To lay down the fundamental guidelines for manufacturing industries to select and manage technologies and products in a manner consistent with sustainable development.
- To realise how internal economic conditions (enterprises) can be modified to transform environmental constraints into opportunities for competitive innovation.
- To consider the various internal and external barriers in selecting and adopting sustainable production systems to avoid the risk of inefficient and ineffective results.
- To suggest how the increased R&D required both from companies and from institutions should be oriented to develop an increasing number of appropriate and economic solutions.

INTRODUCTION

While facing the issues of the role and responsibility of industry in sustainable development, it is necessary to set out those fundamental points of the new scenarios that inevitably become bench-marks for all decisions and choices.

- The increasing 'unsustainability' of the previous and present economic development is a consequence of several choices and trends in producing and utilising goods: irrational transformation techniques of raw materials and energy sources; production and consumption patterns based on 'open-cycle economic systems', and, consequently, increasing pressure on the natural resources exploitation (both non-renewable and renewable) and on the ecological systems equilibria (flows of any kind of waste and other causes of disruption); a dramatic increase in issues relating to unemployment and quality of life, especially in the large metropolitan areas.
- In the present phase of 're-industrialisation', which began in the 1980s, the features and structures of industry are very different from those of the previous phase of the industrial revolution (more flexible production systems; high and increasing product diversification in place of simplification; reduced importance of distance as an economic constraint in relationships between companies; increasing importance of non-material inputs and of intellectual employment; further intensification of capital requirements in relation to labour) (Barbiroli, 1995), but they are not yet moving in the direction of sustainable development.
- The need for correct use of natural resources, and for a slowing-down in the rate of their use, as well as the conservation of the original environmental balances, suggests that economic systems should be modified from 'open-cycle' to 'closed-cycle'. This radical change entails diffusion of the culture and practice of 'reproducing' in place of 'producing' (Stahel and Jackson, 1993; Baas, 1996).
- Any choices must also take into consideration that those things being sold should be seen in terms of 'functions and utilisation', rather than simply 'goods'. This entails fundamental changes in all economic structures.
- Sustainability can be gradually achieved in the very long term only if industrial and service activities on the one side and consumers on the other profoundly and systematically change their sensitivities and attitudes in a way that truly harmonises the type, quality and quantity of 'sustainable products' with ways of utilising them.
- All changes in the way goods are produced and consumed shall be 'economically viable' in a modern sense (that is, viable both for enterprises and for society as a whole). This should be an aim despite implying a substantial reformulation of the market composition, and of the relative importance of goods and services. This reformulation would be occasioned by the limits to substitutability of both inputs and outputs and consequent changes in cost and price.
- The measures of welfare must be profoundly modified in a way that encompasses all aspects of social and economic life.

Within this context it is possible to identify tasks and actions for manufacturing industries where they come into contact with the idea and culture of sustainability, in compliance with the principles of Agenda 21.

THE STRATEGIC OPTIONS OF INDUSTRY FOR SUSTAINABILITY

The role of industry, in a context of sustainable development, can be characterised by four fundamental routes:

1. Adopting proper criteria for selecting and managing technologies and products in ways that work towards resource conservation.
2. Devising and developing eco-compatible materials, chemicals, energy systems and related technologies to achieve a 'real technological pluralism'.
3. Implementing innovative techniques to enhance availability and quality of foodstuffs and agro-industrial resources.
4. Establishing new enterprises suited to 'closed-cycle economic systems', and capable of creating employment opportunities.

Adopting proper criteria for selecting and managing technologies and products in ways that pursue resource conservation

This area of innovation may be regarded as a fundamental step towards sustainability and may be articulated in four interrelated directions:

1. Optimising the lifespan of durable goods.
2. Spreading cleaner (eco-compatible) production systems.
3. Exploring new frontiers in the overall performance and efficiency of technologies and inputs.
4. Reducing the incidence of materials and energy in the final value of the products (dematerialising of industry).

Each of these is considered and discussed here.

Optimising the lifespan of durable goods

In the last 15 years obvious economic and commercial factors have led manufacturers in industrialised countries to shorten the lifespan of some main products, but this trend runs contrary to resource and environmental concerns. In aiming at the maximal lifespan, all the impacts relating to the product life cycle should be taken into consideration. For instance, extending product lifespans is desirable because it would reduce the volume of waste released into the environment and the quantity of raw materials utilised; equally, the production of new, energy-efficient, durable goods may be advisable from the point of view of slowing the depletion of energy resources (Cooper, 1994).

In some cases of newly available, more energy-efficient appliances, it might be beneficial to replace still-functional old products with new ones. The environmental advantage of such replacements should be ascertained through a life-cycle assessment, comparing the impacts avoided per year of additional life of the product to the reduction in annual use impacts thanks to increased energy efficiency.

However, it is strongly advisable, and not only for environmental reasons, that new products be designed for longer durability. Companies, indeed, should consider the potential effects on their markets of changing consumer attitudes to product life. Durable goods are intended, by definition, to be long-lasting and should not fail prematurely. Durability and repairability are therefore considered key aspects of quality, and companies that develop long-life products will enjoy benefits in terms of their reputation (Barbiroli *et al.*, 1996a).

Furthermore, durability may well prove easier to market than, for instance, recyclability, as the benefits of the former accrue directly to the consumer rather than to mankind as a whole. Improved customer loyalty represents a further potential benefit. Therefore forward-thinking companies have much to gain from exploring potential new market opportunities for products manufactured to last longer. In the case of products specifically designed for long life, it is to be expected that the initial cost of the product will be slightly higher than that of current short-life products, for two reasons. First, durability of products will require the use of higher quality raw materials. This price difference should not, however, be overestimated. Material costs account for a relatively small part of the cost of most products (10–50 per cent, commonly, down to about 1 per cent of the cost of a high technology jet engine, for example) (Stahel, 1986). Second, since the production volume will be reduced, productivity will be lower than for traditional durables and costs will be higher. Longer durability can be obtained through product development/ upgrading, repair, reconditioning and re-use (Coates, 1993; Ayres, 1993).

The upgrading of a durable product's performance and functionality might be achieved simply by replacing only those parts that have become obsolete, so making it unnecessary to discard the whole product and replace it with a new, technologically updated one. In order to allow easy and economical upgrading, product design should take into account the potential technical-qualitative evolution of the product itself. This could be accomplished through modular design, whereby a product is conceived as a combination of several parts (modules), each of which may be easily replaced with others, leaving the general structure of the product unchanged (Owen, 1993; Barbiroli, 1996a).

Computers are a good example of modularly designed products, as in most cases they can be upgraded to higher technological standards simply by replacing single components or adding new parts. Other durable goods (e.g. aircraft, cars) are usually designed as intercompatible systems with distinct functional modules, separating, for instance, structural elements (a car chassis), skin elements (bodywork), wear and tear components (engine) and control components (Stahel, 1992). This makes it possible, for example, to replace obsolete and/or worn components while keeping the same structural parts. Optimising products and their utilisation through further development/upgrading itself encourages innovation. New technology will be enhanced, such as non-destructive testing to measure quality variations over time, process technology for the remanufacturing of components and goods, servicing and product-life extension.

In order to make these transformations possible and advantageous, a specific design approach (eco-design) is needed. This entails devising and manufacturing products whose structure, properties and functions make them suitable to undergo the above processes (Kaldjian, 1994). Such steps call for a new global manufacturing orientation with dramatic changes in the adopted technologies. It should be pointed out again that modular design and manufacturing are particularly suitable, as they allow single components to be separately and easily repaired, replaced or upgraded (Barbiroli *et al.,* 1996a).

Whenever it is unfeasible to increase a product's lifetime, it is advisable at least to extend the use of the materials of which the product is made. This is technically possible through recovery and recycling of the different materials from specially designed products. The use of compatible materials and easy separation of non-compatible materials should be promoted.

As is widely recognised, design in an ideal environmentally oriented hierarchy of resource use should employ all means of enhancing product durability, since this minimises throughput of resources; it also means that fewer products pass into the waste stream, resulting in significant environmental and economic advantages (Morris, 1992).

Even though recycling is widespread in many industrialised countries, and has recently experienced an increase of interest in others because of cheap labour, its economic and technical long-term feasibility is still uncertain (especially for plastics); recycling growth, indeed, is linked to the development of new markets and the modification of design criteria in view of easy recyclability (Raggi, 1992).

Designing for recycling has also been considered recently for various durables such as cars and household appliances. (A number of large disassembly plants are operated by Appliance Recycling Centres of America Inc., in co-operation with companies implementing appliance buy-back programmes (Steuteville, 1992).)

Spreading cleaner (eco-compatible) production systems

Cleaner production means the continuous application of an integrated preventive environmental strategy on processes and products to reduce dissipation of energy and materials, eliminate toxic raw materials and compounds, and to reduce hazardous emissions and wastes.

The widespread adoption of cleaner production requires increased world-wide knowledge of the concept and of ways of achieving it. It is also necessary to support industry, to develop and adopt long-term programmes to change production patterns, and to facilitate the transfer of the most effective of them.

Case studies of cleaner technology development and adoption can still be found in various industrial branches. In the ceramics industry, for example, vitreous raw materials containing fluorine and lead were used for many years in the production of ceramic tiles. These elements, eliminated in the wastewater, posed a serious hazard to public health and contributed to quality deterioration of environmental resources. As tile-producing companies discovered that end-of-pipe removal from wastewaters was more expensive and less efficient, they substituted vitreous materials for new ones, this time fluorine- and lead-free. Such a replacement, however, was carried out by completely modifying the process, from the biscuit covering step down to firing and decoration. It should be emphasised that the changes described represented an opportunity to innovate the entire plant, reduce overall processing time, adopt new, flexible and high-efficiency kilns, and thus increase productivity.

The chemicals industry is also experiencing a period of profound change, as stricter environmental regulations, increasing waste management costs and fewer acceptable disposal systems have combined to make waste reduction, especially through pollution prevention, one of the key features of further development.

When a back-to-basics approach is taken, raw material substitution (with alkanes, even light paraffins and other feedstocks), process redesign (co-products elimination by introducing new catalytic systems), or product changes are the main guidelines. Capital investments are high and lead times typically more than five years, but radically rethinking an entire process may lead to the development of new technologies able to spread to and benefit other productions. Propylene polymerisation through the Spheripol process and substitution of dimethylcarbonate for toxic reagents such as phosgene, dimethyl sulphate and methyl chloride present two good case studies and will be discussed in greater detail.

The Spheripol process is a major step forward in polymerisation technology, in that gas-phase propylene is converted directly into highly isotactic polypropylene particles with a size distribution suitable for conventional processing equipment, such as extruders or injection moulding machines. The process eliminates the need for (a) washing catalyst residues from the resin, as the catalyst used for polymerisation is so active that residues are 2–3 per cent of conventional catalysts; (b) the atactic fraction extraction step, as the isotactic polymer (the one with useful properties) is more than 98 per cent by weight; and (c) the extrusion/palletising step, leading to net energy savings. The process has virtually no effluents and has a better economic balance due to a consistent saving in catalyst, electric power (energy input has been reduced from 650 kWh to 130 kWh per 1,000 kg of polymer) and utilities consumption.

The key to this process is the development of a new family of super-active catalysts, which are now also being introduced into the production of new ethylene-propylene copolymers, with very attractive development opportunities and properties ranging from those of rigid materials to those of very soft elastomers (Barbiroli *et al.*, 1996b).

A second example of diffusion of a technological innovation originally developed to reduce pollution is the case of dimethylcarbonate. This was originally developed as a substitute for highly toxic phosgene in carbonylation reactions. Now it is used for the preparation of diphenylcarbonate, an intermediate for polycarbonates and methylisocyanate; diallylcarbonate, an intermediate for organic glass production (ophthalmic lenses); oxoalcohol carbonates, used as a synthetic base for lubricating oils; aliphatic polycarbonates, used as intermediates for polyurethanes.

The case of biodegradable surface-active agents is also quite significant: in the 1970s and 1980s, the world's large chemical industries were forced to eliminate all persistent surface-active molecules—manufactured for decades—and to first study and then produce molecules that were almost totally biodegradable in water within 48 hours. This radical and rapid change took place with significant progress in the synthesis of new molecules which had hitherto been neglected. This in turn engendered equally radical changes in processes and the relative technologies. The main changes concerned the raw materials used, the phases of synthesis catalysts and process technologies, and also led to improvements in the efficiency of new, biodegradable surface-active agent synthesis processes (Barbiroli *et al.*, 1993).

Moreover, the changes are not limited to the sphere of surface-active agent manufacturers, but have had equally noteworthy effects on other branches of chemistry, with the application of both the knowledge acquired and the new processing technologies to other fields.

Table 12.1 reports, in a concise way, some recent experiences of cleaner production implementation in different activities and countries.

Exploring new frontiers in the overall performance and efficiency of technologies and inputs

One of the social functions of industrial enterprises is to respond to and meet human needs; clearly, industry and related activities are an important factor in the sustainability equation. As a matter of fact, the very definition of eco-efficiency must be regarded as a broad performance ratio to guide entrepreneurs in carrying out their role in modern society; it measures their success in minimising the ecological burden and

Table 12.1 Successful case studies of cleaner production in the world

Type of product	*Industry*	*Features of the cleaner production*	*Advantages*	*Cost savings*	*Capital investment*	*Payback period*
Coke	Knurów Coking Plant, Knurów, Poland	Closed system of dirty water Open system of clean water	Reduction: 91% hydrogen cyanide; 85% hydrogen sulphide; 88% toluene; 88% benzene; 88% xylene	$166,000	$20,000	1 month
Chemicals	Rhone Poulenc Chemicals, Leeds, UK	Longer product runs to reduce the number of wash outs required Reduction in use of washing aids Appropriate sequence of washouts	Reduction: from 18 to 6 kg of COD/ton of product	£51,000	£10,000	< 3 months
Sugar milling and refining	Central Azucarera, Don Pedro Luzon, The Philippines	Reduction in water use and wastewater generation/waste segregations to dry handling of fly ash, recycling bagasse, recycling clarified water Good housekeeping	Reduction: wastewater from 18,000 to 1,500 m^3/day; lead sub-acetate from 4.5 to 2.5 kg/day	P990,000	P500,000	9 months
Paper	M/S Ashoka Pulp and Paper Mills, India	Separate water storage tank and pump to ensure a constant water pressure at the edge-cutting nozzles New technology to reduce fibres in the pulp	Reduction: paper breakage and fibre loss by 0.5 tons/day Pollution load: COD 800 kg/day TSS 600 kg/day	$118,000	$25,000	< 3 months
Metal plating and finishing	Robbins Company, Massachusetts, USA	Closed loop system to purify and recycle water (hydrogen peroxide destroys cyanide; acid and caustic adjust pH level; particulate filters remove solids; carbon filters remove organic compounds; ion exchange resin remove salt and metals)	Reduction: water consumption from 500,000 to 700 gallons per week; chemical savings 82%; toxic sludge replaced by metal recovery	$117,000	$240,000	2 years
Steel pipe coating	Ferrum Steelworks, Poland	Shot blasting Successive deposition of: powdered epoxide resins; copolymers as a combining agent; ethylene plastics as the outside layer	Reduction of air emissions: dust from 140 to 33 kg/yr; CO from 323 to 31 kg/yr; NO_2 from 2,492 to 235 kg/yr; Phenol from 125 to 0 kg/yr; Benzenapirene from 0.1 to 0 kg/yr; Aliphatic hydrocarbons from 41,125 to 0 kg/yr; Aromatic hydrocarbons from 6,396 to 0 kg/yr	$3,500,000	$6,200,000	1.75 years

Source: UNEP, 1995

simultaneously maximising the economic value of their investment. Eco-efficiency is not an absolute measure, but evolves as a function of innovation, custom values and economic policy instruments. For this reason it is important and useful for each company to evaluate the manifold aspects of each production process, and of the inputs used, in order to adopt a strategy to maximise its performance (effectiveness and efficiency).

In recent research (Barbiroli, 1996a) it has been highlighted that the aspects to consider and measure in a production process number at least twelve:

- Materials cycle efficiency.
- Energy cycle efficiency.
- Process overall environmental efficiency.
- Final product environmental efficiency.
- Energy cycle environmental efficiency.
- Product absolute quality efficiency.
- Product constant quality efficiency.
- Equipment static operating efficiency.
- Equipment dynamic operating efficiency.
- Product mix variability efficiency.
- Product volume efficiency.
- Input efficiency.

One can observe that all selected aspects must be considered fundamental for the success in a modern sense of a production process, especially since they contribute to connoting not only a traditional enterprise orientation but also a socio-economic one (above all, those concerning energy, materials, environmental and quality efficiency).

Of course, each of the above listed aspects of performance may be applied at both a technical and an economic level, which are closely interlinked. Table 12.2 gives a detailed description of how these twelve aspects can be defined and measured (Barbiroli, 1996a).

A few fields of production in some countries have recently seen important advances, entailing the full redesign of the production processes. This has been done by introducing new technologies and combining the production factors with modern criteria to achieve a high increase in energy yields (reducing their unit impact by as much as over 50 per cent with respect to previous processes); the elimination or drastic reduction of jointly produced effluents; a reduction in processing times; better phase integration; increased productivity; improved performance; and often greater flexibility and high diversification of the products in question.

This is happening, for example, in the steel- and iron-making industry. Here, major changes have been taking place, with the introduction of innovative technologies in the melting stage: one of these, which is replacing the traditional blast furnace, is the 'melting reduction' process, whereby iron ore in a molten state is reduced to molten iron, which is then directly converted into steel. There are also innovations in the casting stage, such as 'near-net-shape casting' techniques, where molten steel is cast directly into very thin slabs ('thin-slab casting') or even into thin sheets ('strip casting'), or the 'spray casting' processes, where a liquid steel stream is atomised and deposited as a thin layer on a substrate, obtaining different final shapes depending on the shape and movement of the substrate surface (Barbiroli, 1993, 1996b).

The main advantages of these techniques are the production of clean, defect-free materials, with properties comparable to those of forged materials (thus eliminating the forging stage), increased size tolerance, considerable savings of the metal used and high energy efficiency (which is the main factor stimulating overall process innovations).

The world's most advanced cement plants (Japan, USA, Austria, Germany) have radically changed their production systems over the last eight to ten years, involving every stage and type of equipment, to the point that they have already entered a 'new technological generation'. Redesigned mixers, dryers, precalcinators and rotary kilns in wet processes have together nearly halved the unit consumption of thermal and electrical energy, and productivity has greatly improved, to the point that investments are paid back in two to three years.

The most important revolution in this field is the introduction of the fluidised-bed system, whose advantages are unit fuel costs reduced by 30 per cent, fewer nitrogen oxides produced, the possibility of using different types of coal, fully automated operations, increased plant lifespan, reduced maintenance, less space occupied by equipment, high flexibility and high qualitative diversification of the concretes obtained.

Table 12.2 Structure of the various indicators for measuring the manifold aspects of efficiency of production processes

	Technical efficiency		*Economic efficiency*			
1. Materials cycle efficiency	Quantity of materials actually transformed and included in the products (tonnes)	(0–100)[a]	Additional cost for materials due to the actual conversion rates (total cost for materials × rate of non-utilised materials)	+	Cost for upgrading the materials not utilised in the process	(100–0)[b]
	Quantity of the original materials introduced in the process (tonnes)		Value of the materials actually included in the products (total costs for materials × conversion rate)	+	Value of the materials included in the by-products	
2. Energy cycle efficiency	Quantity of energy actually utilised in the various phases of the process (MJ)	(0–100)[b]	Additional cost for the energy due to the actual conversion rates (total cost for energy × rate of non-utilised energy)	+	Costs for managing and controlling the energy cycle	(100–0)[b]
	Quantity of the original energy sources and forms introduced in the process (MJ)		Value of the energy actually utilised in the process (total costs for energy × conversion rate)			
3. Process overall environmental efficiency	Total quantity of original and intermediate materials and compounds (potentially polluting) not released into the environment (tonnes)	(0–100)[b]	Total costs for reducing the dissipative potential of the original and intermediate materials and compounds (potentially polluting) used in the process and not transformed into products			(100–0)[a]
	Total quantity of the original and intermediate materials and compounds (potentially polluting) not transformed into products (tonnes)		Value of the materials actually included in the products (total costs for materials × conversion rate)			
4. Final product environmental efficiency	Quantities of non-dissipatable materials into the environment present in the products (natural and anthropic metabolism) (tonnes)	(0–100)[b]	Total costs for reducing dissipatable materials present in the products			(100–0)[b]
	Total quantities of materials present in the products (tonnes)		Value of the materials actually included in the product (total costs for materials × conversion rate)			

Table 12.2 (Continued)

5.	Energy cycle Environmental efficiency	(Total quantities of polluting effluents not released into the environment during the energy cycle (tonnes)) / (Total maximum (theoretical) quantities of producible polluting effluents in the energy cycle of the process (tonne))	(0–100)	(Total costs for minimising the dissipative potential of polluting effluents in the energy cycle) / (Value of the energy actually utilised in the process (total costs for energy × conversion rate))	(100–0)
6.	Product absolute quality efficiency	Global performance indices	(0–100)[c]	[(Unit production cost for the most valuable product / Corresponding highest performance index) – (Unit production cost for the least valuable product / Corresponding lowest performance index)] / (Weighted means for cost/performance index ratio of the product mix)	(100–0)
7.	Product constant quality efficiency	(Maximum detected interval for the performance indices, over time – Absolute sequential mean difference $\frac{\sum_i \lvert \Delta x dt \rvert}{n-1}$) / (Maximum detected interval for the performance indices, over time (weighted mean of the single values for the various products, measured using the totality of the products or a statistical sample))	(0–100)	(Total costs for maintaining the highest constancy of quality properties) / (Total increased commercial value of the products with high constancy of quality properties (obtained by increasing the sale price and/or the quality of the products to be sold))	(100–0)
8.	Equipment static operating efficiency	(Total potential working time of the equipment (hours) – Total break time for set-up and other causes for the consolidated product mix (hours)) / (Total potential working time of the equipment (hours)[d])	(0–100)	(Additional amortisation share due to breaks for set-up and other causes (maintenance) for the consolidated product mix) / (Amortisation share in the average unit production cost for the consolidated product mix)	(100–0)

Table 12.2 (Continued)

9. Equipment dynamic operating efficiency	(Total potential working time of the equipment (hours) – Total break time for set-up and other causes after the introduction of new products (or articles) without modifying the process (hours)) / Total potential working time of the equipment (hours)[d]	(0–100)	Amortisation share in the average unit production cost for the new product mix / Amortisation share in the average unit production cost for the consolidated product mix	(100–0)
10. Product mix variability efficiency	Number of new products (or articles) obtained by differently combining the inputs without modifying the structure of the process / Number of products (or articles) usually obtained in the process	(0–100)	Average unit production cost for the new product mix obtained by differently combining the inputs, without modifying the structure of the process / Average unit production cost for the consolidated product mix	(100–0)
11. Product volume efficiency	Quantity of products sold / Quantity of maximum producible products	(0–100)	(Maximum obtainable products value – Actual sold products value) / Maximum obtainable products value	(100–0)
12. Input efficiency	Optimal total lead time (hours or minutes) per unit of product, after having optimised all the inputs used in the various phases of the process (elaborated by means of the linear programming technique-simplex method) / Actual total lead time (hours or minutes) per unit of production, detected under normal operating conditions	(100–0)	(Actual unit production cost – Optimal unit production cost) / Optimal unit production cost	(0–100)

Notes: [a] Addition of the single values for the various materials and products
[b] Addition of the single values for the various products
[c] Weighted mean for the various products
[d] Addition of the single times for the various process units

For years, glass has been considered a static, declining material; the use of new melting, casting and layering techniques has made this industry dynamic, especially in relation to the use of glass in food and drink containers, where it has long suffered competition from aluminium, plastics and paperboard-backed materials.

Technologies that have so far contributed most to innovations in terms of increased overall efficiency of the resources used—and will contribute even more in the future—are lasers, optical fibres, sensors, continuous casting, vacuum furnaces, fluidised-bed furnaces, fluidised-bed reactors, heat-recovery furnaces, direct contact heat exchanges and membrane separators.

Reducing the incidence of materials and energy in the final value of the products (dematerialising of industry)

A consequence of the adoption of sustainable production processes and products, and the pursuit of new frontiers of global performance by modifying business functions, are high added value products: that is, products in which the relative contribution of technology and non-material inputs to the final value is higher and higher. It means that obtaining a monetary unit of income requires a lower percentage of physical resources.

Among the business functions of R&D, design, manufacturing, management, marketing and technical assistance, manufacturing has traditionally been pre-eminent.

Since companies are increasingly shifting their focus to the upstream and downstream functions, research and design are assuming an increasingly significant role in meeting the diversified and advanced needs of consumers and society. Similarly, improved technical assistance has made the role of properties/ performances of products more important. All these changes have jointly contributed to significantly increasing the non-material components of the final product (or service), leading to further steps towards dematerialising of industry and economy.

The car sector is a particularly significant example of this trend. Cars now are technologically very different from those produced some decades ago; costs have shifted towards the design stage, which is increasingly expected to satisfy safety and functionality requirements.

Meanwhile in the field of data storage and transmission, dramatic improvements in information density (that is, the quantity of stored or transmitted information per unit of material resource) have led to greater dominance of the software component over hardware.

The highest added values can be obtained by improving or maintaining components that are part of complex systems, such as railways and telecommunications systems, or by reconditioning goods where utilisation value and economic value are very different, such as depreciated computer systems. The lowest added value will normally result from the recycling of raw materials, or from product-life extension activities using manufacturing technology such as the retreading of tyres using pressure and high temperatures.

Devising and developing eco-compatible materials, chemicals, energy systems and related technologies to achieve a real technological pluralism

The commitment of industry to make available an increasing number of man-made materials, compounds, energy systems and related technologies must be considered a necessity, because most environmental and resource problems arise from a lack of alternatives able to give more advantages than disadvantages in varying economic and social conditions, and able to contribute to achieving sustainability. These are separately considered and discussed here.

The creation of an increasing number of man-made materials and chemicals must be viewed from the standpoint of 'steady state' between extracted resources and artificially produced materials, which is one of the central points of sustainable development, although it is one of the most difficult to achieve. The increasing specification of materials has reduced the quantities of each type sold. Features that were once exceptional (a minimum level of technical performance, reliable quality, keeping to deadlines) have rapidly become commonplace. Materials requirements are becoming more and more differentiated and specific (Bomsel and Roos, 1990).

The superalloy blades of turbines for aeronautics engines need to withstand severe mechanical and thermal stresses (temperatures above 1000°C). Steels used in the chemical or nuclear industry or in the building of off-shore platforms have to resist corrosion. That resistance may either be intrinsic or provided by a protective coating: large parts of nuclear reactors are made of standard carbon steel covered in stainless steel or Inconel; platform components are coated with offshore paints or use cathodic aluminium-zinc protection, sacrificing metal lost to the sea.

The transport industry, especially aeronautics, needs light, high performance materials. Aluminium and magnesium alloys are extensively used in planes and also in cars (though not as widely as some producers had expected) because it makes them lighter. The breakthrough in composite materials used in aeronautics has happened for the same reason: the specific resistance of unidirectional carbon or aramid reinforced composites is several times higher than the resistance of standard light alloys. Aluminium producers have approached this problem by experimenting with aluminium-lithium alloys which offer both density and mechanical characteristics gains (around 10 per cent for both).

In the high performance field some consumers demand high purity metals: for electronic components, optoelectronic equipment, special electrical generators (used in torpedoes, rockets, etc.). Some mass-producing industries also have high performance requirements. Car makers, for instance, have increased their consumption of coated steels (corrosion resistant), high resistance steels (lightweight) or low alloyed steels (cutting costs and reducing the risk of alloy metal shortages). Aluminium or tin rolling mill operators have to reduce the thickness of products used in making cans without increasing the frequency of faults and thus increasing the can maker's rejects. The standards of protection to be met by these packaging materials are becoming ever higher.

Increased use of man-made materials, from polymers (especially advanced ones, such as metal-matrix composites) to neoceramics and others can be achieved by increasing substitution and (especially technological) research, provided that they simplify processing, reduce processing times and the amount of energy required during the entire life cycle, and are recyclable. Substitution may involve a material only, but more usually involves both a material and a process. In car industry foundries, for example, instead of talking about competition between sand cast iron and aluminium it is more appropriate to speak of competition between sand cast iron and pressure die cast aluminium or between powder metallurgy and lost foam foundry for both metals.

The development of small mechanical parts made by powder metallurgy is linked to both the evolution of sintering materials (granulometry, purity, etc.) and that of shaping processes. From a chemical point of view, though, the alloys used today are very similar to those traditionally used to make the same parts. In many cases there are economic reasons behind the choice of substitutes. The price of conventional materials is often their major advantage (for automobile bodies, the cost ratio between steel and plastic is 1:2). But the costs of transformation and utilisation often favour newer materials. These costs also foster a reduction in the number of parts and thus cut assembly costs. A reduction in weight will allow fuel consumption cuts.

However, the use of new structural materials implies a change in manufacturing methods (organic matrix composites in aeronautics, for instance, are closer to chemistry or even textiles than metallurgy), and hence possible hold-ups. Generally speaking, only profound changes in the methods of production of the whole system will enable new materials to be substituted (for instance, increased differentiation between final products, smaller-sized series). Competition between materials results from technical-economic choices in the whole product chain.

The effective or potential use of alternative materials in downstream products (composites in aeronautics or plastics in the car industry) leads metal suppliers to intensify their innovative efforts in order to meet competition. By furthering the development of new materials, customers stimulate all their suppliers.

However, not all materials are capable of substitution. In some cases particular chemical elements are indispensable: nitrogen, potassium and phosphorus are essential to fertilisers. In others, substitution is technically feasible but will remain too costly on a long-term basis: the performance-price ratio of some materials for the time being is unequalled. Examples are numerous. The mechanical resistance of steels makes them indispensable in construction, heavy mechanics and electromechanics. In spite of research on composite armour plating, involving metal, synthetic fibres and ceramics, steel is not threatened in this application. Superalloys seem to be virtually irreplaceable in jet engines (in spite of research on ceramics and ceramic-ceramic composites) and in the vapour generator tubes of nuclear plants. The electronics components industry uses small quantities of metals but these are irreplaceable (aluminium, silver, gold, molybdenum, tungsten, etc.). There are numerous alternatives to lead in car batteries, but they are all too expensive (cadmium-nickel, lithium, etc.). Plastic or composite tubes for the petroleum industry remain too expensive in a depressed oil market. On the other hand refiners are interested in plastic tubing: its good resistance to corrosion reduces the cost of tubing (steel itself is cheaper but requires additional coating) and maintenance. However, some technical problems have yet to be solved (heat resistance, difficulty in assembling parts).

Price fluctuations—as well as other factors—often speed up the substitution process. Substitution of copper by aluminium for the transport and distribution of electricity in France was favoured because of the relative stability of aluminium prices, in spite of the low weight/conductivity ratio. Substitutions are not irreversible: in European car engine manufacturing, aluminium is facing the revival of cast iron: the latter has become technically very competitive (thanks to thin wall moulding) and its price is less volatile than that of aluminium. Substitution is therefore not permanently binding, but a free strategic choice available to the consumer.

To sum up: specially developed materials are increasingly replacing commodities. Such materials may be described as 'evolutionary'. A material is evolutionary if its features and the nature of the services linked to its sale are defined in partnership with customers, adapted to their needs and even, in some cases, specific to a particular customer.

There has long been a demand for evolutionary materials in high-tech industries. What is new is the fact that this demand is now spreading to mass-production industries such as the car industry, construction and packaging (Barbiroli, 1996b).

As far as new energy resources and systems are concerned, we can emphasize that industry has a broad field of activity, both in developing new systems and in conserving energy in the several applications. The energy alternatives based on non-traditional and huge resources have not undergone any significant progress since the 1970s, at least not to the point of making them economically competitive and thus available for large-scale commercial distribution. The difficulties encountered everywhere and the low prices of oil served to slow down, if not discourage, the use of enormous financial resources by research establishments and industries throughout the world.

Only nuclear fusion is attracting adequate financial resources to achieve controllability within a few years, and even then only in a very few countries. The mid-1980s saw a sharp slow-down in the construction of nuclear fission power plants; since then, the major industrial countries have aimed largely at achieving control of fusion, which uses raw materials that are very abundant in nature (deuterium and lithium) and which does not cause environmental disruption in any phase of the cycle (Jackson, 1989).

In addition to this technological trajectory, which can in any case satisfy a basic, centralised energy demand, some technologies in the field of solar energy appear to be assuming economic significance. These include those that combine the concentration of solar radiation (especially through photovoltaic systems) with the conversion of water to hydrogen, which can become a real medium-scale energy vector capable of contributing to the diversification of energy systems and their decentralisation (Winter, 1991). Following a complete cost/benefit/risk analysis, the solar energy-electrical energy and/or electric energy-hydrogen circuits appear to be feasible and advantageous, especially after progress has been made in certain technical aspects crucial to success in this field (e.g. hydrogen storing). Solar energy must also be advantageously captured and converted in order to expand the resources and energy systems available for development, with photovoltaic technologies and with others related to the use of biomass (Kühne and Aulich, 1992).

Water resources, for the simultaneous production of electricity and water, are among those that have been neglected. Indeed, the potential water resources for this purpose are still very large (Sims, 1991) and the proportion of them currently exploited is low in many parts of the world (the exploitation rate is high only in Europe). The main reason for this is that many potentially utilisable water resources are located far from the centres of demand for electrical power, and the high-voltage transmission systems used throughout the world until a few years ago did not make it viable to carry electricity in a useful manner for more than 300–350 km. In addition, since the advent of nuclear fission energy in the 1960s, funding and interest have been very much oriented towards this form of energy. We should point out that as early as the 1960s, projects were developed to harness water in large basins in various parts of the globe (North America, South-East Asia, South America, Australia, United Arab Emirates), although they were not put into practice. Thus these and other projects (such as in Russia) may only be realised within the framework of an economic policy that gives priority to water resources, once very high voltage transport systems have been developed (as high as 1,000 kV), and once the social and environmental impact assessment has given positive results.

It should also be pointed out that just as there is great hydroelectric potential in large and very large basins, there is also considerable potential in small and medium-sized basins, although this is universally ignored and considered non-economical (Wilson, 1991). However, it should immediately be emphasised that the concept of 'economic' that has been used until now in relation to solutions for large plants of every kind falls within the context of economies of scale, and the tendency towards centralisation; these concepts are undergoing dramatic change in a world that requires decentralised, flexible solutions, and where the criteria for judgements of economic convenience must therefore be radically reconsidered. Consequently, with cost/benefit assessments, plants with small and medium potential will inevitably become more important.

There are several reasons for improving the field of synthetic liquid and gaseous fuels, even given the need to eliminate them as combustibles: diversification of the sources and areas, stabilisation of the markets, improvement in the environmental quality of fuels, value enhancement of original resources, creation of industries and service activities for the equipment and technical assistance, job creation, improvement in the technological balance, advances in scientific and technological research into hydrocarbons, catalysis and related fields.

However, one must also consider the high production cost of synthetic fuels, if only at the beginning of their widespread development: between US$35 and US$53 per barrel of oil equivalent for liquid hydrocarbons extracted from oil shale and tar sand, between US$36 and US$46 per barrel of oil equivalent for surface synthetic coal gas, between US$25 and US$37 per barrel of oil equivalent for underground coal gas (Barbiroli and Mazzaracchio, 1995). Consequently, their diffusion on a large scale needs a higher level of prices for crude oil, to reduce the price difference.

Also, developing technologies for converting many sources of biomass into fuels will increase the pluralism in those fields and thus decentralise the systems for conversion/use of the energy obtained. This will have significant advantages in terms of the diversification/appropriate use, thus helping to create a balance between sources and systems of energy.

As far as the real prospects for developing systems that use other renewable energy sources are concerned—wind energy, geothermal energy, ocean wave energy—objective consideration of the limits imposed by their location, form of origin and the potential environmental disturbance caused by their use, leads us to feel that advantageous, economic results can only be achieved on a limited scale. However, this conclusion must not lead us to give up on the possibility of progress and results which could form the basis for interesting economic applications.

The production of energy using non-traditional systems, such as Magnetic-Hydro-Dynamic (MHD), represents a solution that might produce electrical energy at lower costs than traditional thermoelectric systems, thanks both to the high yield (over 60 per cent) that can be achieved by a plant based on this system, and to the economy of the fuel. Also, fewer combustion products are released into the atmosphere; even thermal release per unit of electrical energy generated can be reduced by a factor of between 1.2 and 2. Plants with an MHD system can be powered by any type of fossil fuel, from natural gas to coal with a high level of sulphur and ash.

Each solution has specific features making it suitable for differentiated, appropriate applications. Combined-cycle plants with gas-turbine benefit from numerous international applications, which qualify their reliability. Gas MHD plants are currently in the stage of demonstrative development.

Implementing innovative techniques to enhance availability and quality of foodstuffs and agro-industrial resources

The availability of increased quantity and quality of agro-food as well as agro-industrial resources has become a critical point for human life, especially as a consequence of the world population increase and the productivity limits of traditional cultivation techniques.

Biotechnology seems to be one of the major commitments for industries, public authorities and institutions for the near future (Goodman *et al.*, 1987), and can contribute greatly to sustainable development (Sasson, 1988; Christi, 1993), as long as the possible solutions do not create problems for health or the environment.

The focus now is on developing the range and scale of production technologies for many of the traditional foods through better understanding and control of fermentation processes, on utilising the well-established food-grade micro-organisms in fermenting edible substrates previously employed and, to a lesser extent, on finding new, safe microorganisms for protein enrichment of lower grade substrates. The latter approach has led to a totally new, highly successful, mucoprotein food, Quorn (meat alternative), which in spite of utilising a liquid substrate, presents a shining example for solid substrate fermentations to emulate. Moreover, a process for upgrading lignocellulosic matter from edible plant residues has been developed using an edible microfungus traditionally employed in some solid-state fermentations (a solid substrate slurry fermentation system is employed for the new process) (Mannion, 1992). Production of microbial biomass rich in proteins in the era of food shortages has recently gained importance. By using biochemical reactors, it is now possible to explore microbial growth for the acceleration of protein biosynthesis. Proteins of higher fungi could be used for food.

One of the most important fields of development in biotechnological applications to foods is that of novel yeast strains. Although current yeast strains have been selected for their desirable properties, there is still the potential to improve their performance in particular processes by the use of recombinant DNA technology, once it has been demonstrated that genetic manipulations of whatever form do not have drawbacks. The use of biotechnological processes for the utilisation of spent yeast also seems quite interesting.

A wide field of activity is that of *protein engineering* as a source of new enzyme uses in the food industry. By understanding how amino acids interact to form three-dimensional structures, biotechnologists can design polymers with desired functional properties. This has enormous potential for the food industry. Modification of existing structures is already in hand, although there is a long way to go before the design and construction of new proteins for the food industry becomes commonplace (Lambert and Joos, 1989).

The potential for immunological methods of detecting bacterial contamination of foods, and the development and use of rapid diagnostic methods in the food industry, are equally interesting.

Using biological routes to produce food flavours instead of applying organic chemistry may be considered as an alternative. The two approaches used to adapt biological systems to flavour technology are the use of whole cells and the use of isolated enzymes. One area of current activity is the production of natural cheese flavourings both as part of the cheesemaking process and for other purposes.

Areas that would benefit from manipulations using enzyme technology include food preservation, fat replacers, microwave cooking and frying. There seems to be a promising future for plant cell culture technology. Commercial processes for pigments and enzymes are likely to come into operation over the next few years.

Moreover, microbiologically derived oils are opening new paths. Apart from one or two exceptions, few processes have reached commercial realisation for the production of bulk oils because of the inability of biotechnology to compete with the low costs of agricultural seed-oil production. Where biotechnological processes may be able to compete is in the production of speciality oils providing nutritionally important polyunsaturated fatty acids. One process that has achieved commercial success is the production of oil of Javanicus from the mould Mucor javanicus. This mould produces an oil content of 20–25 per cent with 15–22 per cent gamma linoleic acid, the important

BOX 12.1 CASE STUDY 1: NESTLÉ AMSTERDAM

The Nestlé Amsterdam production plant consists of three separate units: flavour plant, soup plant and the packaging department. The main process in flavour production is the acidic hydrolysis of vegetable protein. The major waste streams generated at Nestlé are the flavour plant residue, wastewater and mixed solid waste.

After the pre-assessment, these three streams were selected as priority areas. During their assessment, 12 viable prevention options were identified. The implementation of four of them started within 16 months of establishing a university–industry co-operation. Feasibility studies were started for six other options. The other options proved not to be feasible in the short term.

The most successful options for Nestlé were enzymatic hydrolysis, decanting, on-site re-use of vacuum ring water, reprocessing of first rinse water from evaporators, substitution of evaporator cleansing agent and off-site recycling of industrial waste components.

The table summarises the major environmental and economic benefits for each of the chosen options.

Prevention options and results at Nestlé Amsterdam

Option	*Environmental effects*	*Economic effects*
1. Enzymatic hydrolysis: use of enzymes for disclosure and (partial) hydrolysis of vegetable protein	About 50% reduction in use of adjunct materials (hydrochloric acid, soda and energy) and reduction of salt load of wastewater and residue	Savings on adjunct material purchases, effluent charges and waste disposal costs
2. Decanting: replacement of cloth filtration unit for salt filtration by a centrifugal decanter	Improved separation between salt and product, which results in improved product yield Reduction in rinse water usage	Payback period approximately 1 year
3. On-site re-use of pump water: feed vacuum pump wastewater from soup plant into wet scrubbers of flavour plant	Reduction of water intake by 118,000 m^3/yr	Payback period within 2 months
4. Reprocessing first rinse water from evaporators: reclaim flavour from first rinse water from evaporators	Reduction of water pollution with about 485 population units	Reduction of effluent charges with Dfl 32,000/yr, and improved product yield
5. Substitution of evaporator cleansing agent: substitution of cleansing agent for evaporator	Reduction of water pollution with about 127 population equivalent units	Dfl 11,000/yr, savings on operational expenditures
6. Off-site recycling of industrial waste components: segregated collection of paper, plastic and product spills for off-site recycling	In total 98,000 kg/yr, industrial waste suited for off-site recycling	Dfl 29,500/yr, savings on waste transportation and disposal cost

Source: van Berkel, 1995

fatty acid in evening primrose oil, which is used to treat a number of ailments, including eczema. In comparison evening primrose oil contains 8–9 per cent gamma linoleic acid.

The application of biotechnologies to agro-industrial production seems to be effective, in quantitative as well as qualitative terms (Pelsy, 1988, 1989).

With regard to renewable resources, some words must be devoted to agro-industrial resources (wood, plants for cellulose, natural fibres, rubber, etc.) which are increasingly scarce and whose intensive exploitation negatively contributes to environmental disruptions. The transformation industry and the primary sector must co-operate to enhance yearly non-wood fibre plants which give good yields and economic advantages (Madden and French, 1989).

The growing demand for paper places pressures on the global environment and opens the way for competitive alternatives to the well-established wood fibre feedstocks. At present, non-wood pulp represents only 7 per cent of the world's pulp production, but for 90 per cent of its history paper has been made almost exclusively from non-wood plant fibres. About 70 per cent of this non-wood pulp production occurs in China and India, where soft wood is in short supply and where the collection of non-wood material is economic, thanks to low labour costs. This resource, then, is greatly under-utilised.

Non-wood material suitable for pulp production can be broadly classified into three groups:

1 Agricultural residue, such as straw, bagasse and cotton linters; or plants, such as kenaf and sorghum fibre capable of intensive mechanised agriculture.
2 Material capable of producing pulps with special characteristics making them uniquely suited to the manufacture of speciality paper products, such as tea bags, bank notes, cigarette and filter papers. This group includes abaca and flax. These materials are generally grown for their fibres and although they sell at premium prices, their production and collection is labour intensive and viable only where labour costs are low.
3 Plant material available as an extensive indigenous resource, such as bamboo (India and China), esparto grass (North Africa) and certain reeds (China and Europe).

Many of the hundreds of thousands of fibrous plants have been tried for pulp and papermaking, and technically many have yielded a product with some desirable properties. However, the mere fact that a plant fibre can be converted into pulp and paper is no indication that it can be grown, collected, stored and processed in a viable way. Necessary characteristics are ample supply of the raw material; availability at the pulp mill all year; capability of storage without excessive deterioration; geographic concentration; moderate collection and transport costs; high yield of good quality fibre; low cost of conversion to pulp; and sufficient demand for the product at a price that will ensure profitability.

Kenaf and sorghum (fibre) represent two examples of alternative raw materials that can be produced in relevant amounts in developed or developing countries, mainly because their production cost has only now made them competitive with other irrigated crops such as cotton, maize, sunflower and others.

It must be stressed that these and other non-wood fibre plants may be utilised in industry as a source of cellulose, and at the same time as fuel feedstocks for chemical synthesis: this makes these plants an attractive prospect for major development.

Establishing new enterprises suited to 'closed-cycle economic systems' and capable of creating employment opportunities

A 'closed-cycle economic system' needs special activities to be established and developed, able to perform specific functions connected with lifespan optimisation (development, upgrading, re-use, reconditioning, repair, recycling of products), management of the flow of energy and materials, technical assistance to different forms of cleaner production, and to create direct and indirect employment opportunities. Such objectives might be accomplished by new industrial and service enterprises, the features of which are foreseen to be somewhat different from those of the present economic systems.

The main characteristics of the new activities are smaller-scale, labour- and skill-intensive work units (Stahel and Jackson, 1993).

Mobile goods such as ships, cars and aircraft with easily exchangeable components are ideally suited to small workshops which can be located according to prevailing needs, conditions and demand. Buildings and other immobile systems with exchangeable components require both on-site intervention and workshop activities. Immobile systems such as sewers or railway tracks are ideal candidates for mobile reconditioning units that can perform repairs and renovations in situ.

It must be emphasised that components that become obsolescent because of leaps in technology will best be recycled to recover the base materials, rather than reconditioned.

This new approach to economic systems will lead to a so-called 'service economy' that is not the same as the traditional 'service sector' but is based on the concept of replacing production and manufacturing

BOX 12.2 CASE STUDY 2: YANTAI SECOND DISTILLERY

Yantai Second Distillery is a medium-sized, state-owned enterprise in the coastal city of Yantai, People's Republic of China. It is engaged in the production of grain liquors and sweet potato wines. The plant employs 510 persons and has a production capacity of 5,000 tonnes of alcohol per year. The production equipment dates from 1982–1986. The single largest environmental problem in the distillery is the generation of 14 tonnes of distillers grain per tonne of alcohol. Since 1986, an anaerobic wastewater treatment facility has been in operation in order to treat the 50,000–60,000 mg/litre of COD in the distillers grains. At present the resulting methane gas is co-fired in the coal-fired boiler. In order to identify prevention opportunities, a plant-level assessment team conducted an assessment for the distillery during 1994 under the guidance of the Yantai Environmental Protection Bureau, the Ministry of Light Industry and the Chinese Research Academy of Environmental Sciences.

During the pre-assessment it was found that the alcohol plant is the largest source of wastewater. In addition, several obvious good housekeeping opportunities could be detected in the bottling department. The implementation of these options gave rise to an annual benefit of over 500,000 Yuan at an investment of 12,500 Yuan. This encouraged the team to proceed with the deeper assessment of the alcohol plant. A number of low cost technology optimisation options had been identified which could be implemented at no extra cost during the annual overhaul period (summer 1994). This created another annual saving of just over 5 million Yuan. In addition, three high cost technology replacement options were identified and evaluated in detail in terms of their technical, economic and environmental aspects. The results obtained at Yantai Second Distillery are summarised in the table.

Prevention options and results at Yantai Second Distillery (China)

Option	*Technical content*	*Environmental effects*	*Economic effects*
1. Good housekeeping in bottling department	Repair of leaks and proper instruction of workers in order to avoid excess filling of the bottles	4.5 tonnes/year, production increase	Investment: 12,000 Yuan Annual benefit: 53,000 Yuan
2. Equipment repair and optimisation in alcohol plant	Optimisation of distillation reflux utilisation Revamp of steam and condensate pipes Reuse of distillation tower condensate for fermenter disinfection Recovery of fermenter washout water	Potatoes input reduced from 3.81 to 3.09 tonnes/tonne alcohol Production increase of 1,200 tonnes/year	Investment: implemented at no extra cost during yearly overhaul Annual benefit: 5,040,000 Yuan
3. Differential distillation	Replace existing multitower distillation with differential distillation system	Steam conservation of 40% Water conservation of 30%	Investment: 3,454,000 Yuan Payback period: 1.5 years.
4. Continuous fermentation	Replace batch-operated fermenters with continuous fermentation system in order to improve efficiency and eliminate wastewater from fermenter washout and disinfection	At least 1% production efficiency improvement Elimination of 5,000 tonnes/year wash water	Investment: 480,000 Yuan Payback period: 3.1 years
5. Boiler replacement	Replacement of the boiler Continuous use of methane gas for power generation instead of co-firing	Coal conservation of 635 tonnes annually Electricity conservation of 750,000 kWh/year	Investment: 1,231,000 Yuan Payback period: 4.4 years

Source: van Berkel, 1995

for operational service, selling 'functions and utilisations' instead of goods (Stahel, 1992).

STEPS AND METHODS FOR ASSESSING AND SELECTING SUSTAINABLE OPTIONS

As stated above, enterprises are requested continuously and intensively to develop sophisticated ideas in manufacturing and managing materials, chemicals and derivatives, energy systems and products. Such intense commitment requires a great deal of thought on the part of entrepreneurs and managers as to the best methods for conceiving and realising these ideas.

The ability of a company to deploy the new features and trends of a production activity depends upon some basic characteristics of the company, notably:

- nature of the company's industrial process;
- size and structure of the company;
- attitudes affecting operations of the company;
- information available to the company;
- assets available for developing and implementing new technologies.

The main steps in choosing the most suitable technologies, both economically and environmentally, are (1) conducting an auditing process; (2) conducting a technology assessment.

The auditing process includes several components:

- Evaluation of opportunities to implement technologies in the specific field of action of the enterprise, consistent with the targets connected with environment, resources, labour, etc.
- Evaluation of products delivered to the market for their potential impact on man, environment and lifestyle.
- Involvement of the labour force as an active participant, within renewed principles of managing total quality systems.
- Training of managers and auditors.
- Knowledge of applicable laws, regulations, and of other governmental inspection and enforcement approaches.
- Environmental expertise and awareness.

A systematic method for the execution of a technology assessment consists of:

- a tool for identifying the most realistic options;
- a procedure (organisation of the necessary activities to develop and implement opportunities);
- a guidance and supervision system (external service to stimulate and check the team responsible for the assessment).

The assessment is usually a planned procedure with the purpose of identifying the best way to contrast current trends, either by simply reducing or by eliminating the generation of waste and emissions from a process, or by adopting totally new production processes which give higher outcomes at the various levels of performance (productivity, quality, variety, etc.).

The assessment serves three basic functions:

1 Analysis of the environmental burden of the production processes and of its causes.
2 Inventory and evaluation of improvement options for the production processes.
3 Integration of the feasible options into the processes and into the structure of the company.

A strategic management of process innovations enables the enterprise to achieve environmental improvements in key areas such as resource conservation, toxic substances and waste elimination.

The central point of a technology assessment should be to examine all the technical and economic features and implications of the production process. This evaluation procedure consists of a 'source identification', a 'cause evaluation' and an 'option generation'.

'Source identification' begins with the drafting of a list of unit operations, with their associated material inputs and outputs and transformations (production generally comprises a number of such unit operations). Clearly it is important to identify the right level of detail during the subdivision of the production process into unit operations. By connecting the individual unit operations in the form of a block diagram, one can prepare the process flow diagram. For each unit operation, material inputs are placed at the left side of the diagram and material outputs at the right side.

An essential step is to check the process flow diagram. There must be a perfect balance between inflows and outflows, so that inputs shall have related outputs (product or waste) and vice versa. Each unit operation can be the source of various waste streams and, therefore, the completed process flow diagram should be used to check all unit operations for waste generation and thereby compile the list of all waste sources.

'Cause evaluation' is carried out to evaluate all material flows, to quantify the volume and composition of all material flows which could result in a mass balance for all individual unit operations or for the entire company; even if it is hard to compile reliable mass balances for each of the constituents of the input and output material flows.

To understand the cause of waste generation, one needs to gather data concerning the impacts of product specifications, materials, technological factors (process design, equipment, lay-out), operating practices and waste handling procedures on the volume and/or composition of the process wastes and emissions.

'Option generation' is carried out to develop alternative options for eliminating or controlling the causes of waste generations. To this end, five general prevention techniques can be used. These are production modification, input substitution, technology modification, good housekeeping, and on-site re-use and recycling. All five possible causes of waste generation can be dealt with by using a particular prevention technique (Figure 12.1). Information from the cause evaluation is used for the identification of the most appropriate technique, and information from the source identification for targeting the source unit operation. The scheme of the whole procedure is described in Figure 12.2, with the main points of the five phases (van Berkel, 1995).

Finally, the importance of reliable economic data —both internal and external—must be highlighted,

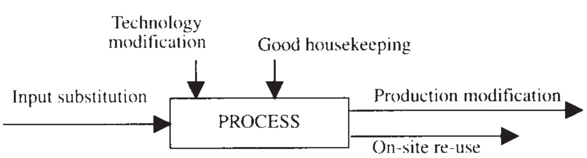


Figure 12.1 Option generation: application of the general approaches

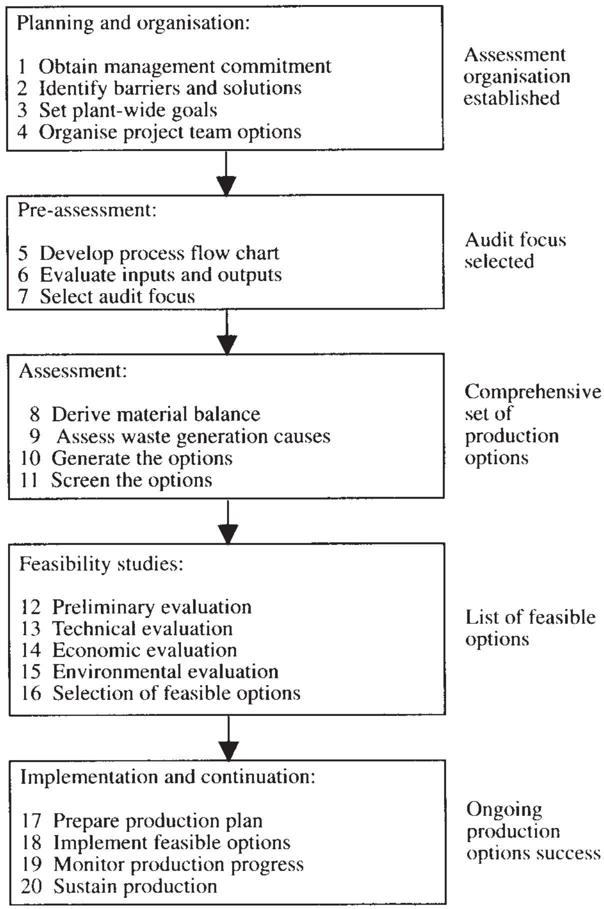


Figure 12.2 Scheme to select and assess production options

especially in dealing with ex-ante estimates and analyses.

APPROPRIATE TECHNOLOGIES FOR DEVELOPED AND DEVELOPING COUNTRIES

In the field of cleaner production, one must consider that different technologies may be developed and adopted and that each one leads to different overall results, with reference to eco-efficiency on the one hand, and to size of companies and employment on the other.

In order to identify the solutions most appropriate to different economic and social situations and to achieve the different goals of companies and governments, we must consider and define four types of cleaner technologies:

- *Type 1—Business-driven technologies:* fairly sophisticated production systems, improving production quality and/or efficiency, improving competitiveness, reducing costs. Such technologies improve eco-efficiency within overall performance improvement and are highly beneficial.
- *Type 2—Clean technologies:* fairly sophisticated production systems, developed and adopted for the primary purpose of improving environmental performance, they are marginally beneficial (e.g. waterless printing).
- *Type 3—Appropriate technologies:* fairly simple production systems that improve environmental performance, but are adopted primarily for economic development purposes.
- *Type 4— 'Low fruit' technologies:* fairly simple production systems that modify existing ones to improve environmental performance (e.g. waste heat recovery/recycling with special furnaces in aluminium smelting).

The first two types of technologies seem to be more suited to large companies, at least in some main fields of application, whereas Types 3 and 4, because of their small scale and low capital requirements, can be autonomously set up and more easily adopted in developing countries, so avoiding the need for their import from industrial countries.

The advantages of adopting Type 1 of cleaner production technology may be seen in the case of an electroplating factory that eliminates the need for pollution control and cuts down its operating expenses through improvement of its operations (Lin, 1994). In a brass plating factory workpieces pass through a three-stage process, starting with a soak bath to remove grease and soil, followed by electroplating proper in a plating liquor tank and finally a rinse. During the rinsing stage large amounts of polluted wastewater are generated. Traditional control options such as pollution removal through precipitation and sedimentation result in shifting pollution from one medium into another (from water to sludge). In a cleaner production approach, several production modifications can significantly reduce pollution: avoiding unnecessary high concentrations of chemicals in the plating tank, leaving workpieces to drop longer before they are moved to the rinsing stage, blowing them off with compressed air, rinsing in a counter-current multistage arrangement (the rinsing water is re-used in the preparation of the next bath), or introducing a drag-out tank before the rinsing tank. Not only do these measures control pollution, they also contribute to cost savings on raw materials.

Type 2, far more common in the real world, makes an overall reduction in production costs possible by reducing the pollution control cost with little or no increase in the manufacturing cost. An example of this would be the replacement of cyanide-containing bath chemicals with non-cyanide substitutes. While the latter are usually more costly and pollution control is still necessary, the total of manufacturing cost and pollution control cost is less than that required for a system using cyanide. In Table 12.3 examples of the relative production cost associated with Types 1 and 2 of cleaner production are illustrated.

Many production processes are currently characterised by low input efficiencies, especially in terms of energy, materials and environmental efficiencies. The introduction of corrective measures on existing processes can lead to limited improvements of both aspects mentioned, while meeting significant investment, management and maintenance costs.

Of course, the different types of technologies that may be adopted as a function of sustainable

Table 12.3 Economic performance of different types of cleaner production

	Cleaner production		*End-of-pipe*
	Type 1	*Type 2*	*treatment*
Sales (10^3 $)	1,300	1,300	1,300
Cost of production			
Manufacturing cost	950	1,050	1,000
Environmental cost	0	30	100
Total	950	1,080	1,100
Profit	350 (75% increase in profit)	220 (10% increase in profit)	200

Source: Lin, 1994

development will perform differently under different economic and social conditions. This means that the new production options cannot be implemented with principles having a general validity; they must be specifically tailored in order to work.

The identification of the most suitable production branches and technologies must take into consideration the major variables of national as well as local situations, in terms of territorial, economic, social, environmental, industrial, structural and employment conditions.

After this fundamental step, co-operation and even economic integration among different countries and areas should be established, in order to make it easier to achieve national and local targets. New forms of co-operation, particularly those between developed and developing countries, should be primarily aimed at creating effective infrastructures and conditions, able to prompt and support the adoption and development of the most appropriate options.

Some crucial points for helping the developing countries to move towards an appropriate route are funding, financing, awareness, access to information, independence of choice, and scientific, technical and training upgrading of human resources. All of these points must reach a sufficient level of effectiveness, otherwise no stable advancement will be achieved in these countries.

In the last five to six years UNEP has carried out specific programmes in several countries (China, India, the Philippines, Brazil, Indonesia, Tunisia, Thailand, Costa Rica, Egypt and others), and in various production fields (distilleries, textile processing, synthetic dyes, pulp and paper making, sugar cane refining, electroplating, soap manufacturing, battery manufacturing, leather tanning, etc.). The aim of these programmes has been to experiment with new cleaner production techniques, to demonstrate their overall effectiveness/efficiency and to ascertain definitively their sustainability.

BARRIERS TO THE DIFFUSION OF SUSTAINABLE OPTIONS

The profound transformation that is necessary to pursue sustainability meets several barriers, at different levels, which may be classified as follows:

- Economic barriers (resource price and availability; availability and cost of funds; inadequate investment planning; *ad hoc* investment criteria).
- Technical barriers (technology limitations; limited access to technical information; technology gap; limited in-house R&D, design, maintenance facilities).
- Organisational barriers (non-involvement of employees; decision-making powers; high turnover of technical staff; lack of recognition).
- Attitudinal barriers (lack of good housekeeping culture; resistance to change; lack of leadership; lack of effective supervision; job insecurity; fear of failure).
- Systemic barriers (inadequate and ineffective management systems; poor record-keeping and reporting; lack of systems for professional upgrading of employees; *ad hoc* production planning).
- Governmental barriers (lack of industrial policy; lack of institutional support; economic incentives for end-of-pipe technologies).
- Demand barriers (change-resistant consumption attitudes; strong resistance from the traditional pollution control industry).

Any one of these barriers alone can prevent companies from implementing cleaner production and related innovations, and so joint efforts must be made to create global conditions (Barbiroli, 1996a), with different responsibilities.

- International institutions should adopt specific guidelines on the main aspects of sustainable development, to be followed by the private and public operators involved in economic and social activities taking practical steps towards achieving sustainability. In this sense, Agenda 21 must be integrated with more detailed and operative indications (for instance, the price level of raw materials).
- Governments and local authorities should provide, along with general economic conditions (industrial policy, intersectorial and territorial programmes, effective criteria for funding and financing, removing economic incentives for end-of-pipe technologies, adopting appropriate incentives and adequate laws and regulations), effective infrastructures (support for

R&D institutions, facilities, professional upgrading systems) and dissemination programmes (information to consumers).

- Industrial systems should adopt strategic long-term programmes for implementing new production technologies, adequate management systems, professional upgrading of employees and the involvement of employees.
- Financial systems should provide adequate funding, devoted to selected investment.

Special efforts should be made by all concerned to increase the number and performance of feasible alternative solutions.

THE EFFORTS IN R&D TO SUPPORT TECHNOLOGICAL EVOLUTION

The innovation process that enterprises must undertake to be consistent with the aims of sustainability is definitely very complex and expensive. It entails, first, developing new technologies, second, attaining a 'pre-economic' stage, and, third, selecting and implementing the most advantageous solution. The process includes R&D activities, design, preengineering and engineering investment, implementation, re-organisation, skill upgrading, marketing, and technical assistance for the new products.

Of course, R&D is still the basis for any type of change, even if it is only a part of the innovative process (on average, less than 50 per cent of the overall cost). R&D is currently being given a great deal of attention all over the world, but it is not yet well-developed or specialised enough to provide the knowledge required for achieving a sustainable economy. Car manufacturers provide a recent example of expensive and ineffective research: after the first (1973) and second (1979) oil crises, large car manufacturers have invested huge amounts of money (tens of billions of dollars) to increase the energy and environmental efficiency of engines, with limited results (improvements of few percentage points); the same huge investment could have been devoted to designing totally new engines and/or new transportation systems more appropriate to sustainable development.

R&D may be carried out at both internal (company) and external level. In the major industrial countries, companies, especially large ones, are generally devoting increasing effort and resources to setting up and developing new technological systems. Even so, they might also need the benefit of research elsewhere. Hence the links between all research centres must be strengthened. To this purpose, the 'technopolitan sites' seem to be the most suitable solution for orienting the innovative process towards an 'environmental, resource and employment qualification'. These centres must be considered the most effective structures, above all for small and medium-sized firms, because they do not have their own R&D centres, even if they participate directly in establishing and managing the 'Technological Poles', the 'Business Innovation Centres' (BICs), the 'Scientific Parks', the 'Entrepreneurship Incubators', the 'Industrial-University Liaison Centres' (Formica, 1991; Gibson, 1992); of course they must contribute to give directions and inputs in order to get the expected outcomes and information for innovation.

The improvement and dissemination of direct information for innovation is also a necessity for the less developed countries, once they start an appropriate development phase in which the specific features of the different situations must be considered. Therefore, great efforts have to be made in order to make possible the achievement of increasing cultural, scientific, technical and managerial autonomy and capacity, even if it will be more easily achieved within a global context, where international organisations are requested to act as 'starters' and 'accelerators' for scientific, development and innovation improvement (Barbiroli, 1996c). Otherwise, neither development *tout court* nor sustainable development can take place in the LDCs, and this goes against the principles of Agenda 21.

TECHNOLOGICAL PLURALISM VERSUS PRODUCT DIVERSIFICATION

Since the 1980s, the trends in all branches of the manufacturing industries are towards increasing product diversification, which can be obtained by means of flexible manufacturing systems. But these trends cannot be automatically considered consistent with a sustainable development and the interest of consumers. This is especially true if, to achieve a wider range of quality diversified products, great investments for R&D, design, equipment, management, marketing and technical assistance are required from companies, and these lead to increasing production costs without a corresponding benefit (quality and/ or productivity increase) for consumers.

This negative result concerning the cost/benefit ratio of production trends contrasts with the need for a real 'technological diversification' (plurality of alternatives to achieve equivalent outcome) of the ways to produce goods to enable appropriate choices to be made for achieving a maximisation of benefits, as compared to costs, for companies, consumers and overall society (environment, resources, territory, employment, lifestyle standards). For instance, in the hydrocarbons sector no real alternatives have been developed in the last two decades; in the electricity field, no real industrial alternative to thermoelectric and nuclear fission power plants has emerged.

This does not mean that some product diversification within each process cannot be considered advantageous both for companies and for consumers, particularly if the price/performance ratios of the products are equivalent (the ratios can be obtained by relating each price to a global performance index for each good, elaborated as proposed by Barbiroli, 1989). Indeed, at present, high price differences for slight or even illusory quality differences among the same line and type of products (automobiles, electric appliances, clothes, shoes, foodstuffs, etc.) may be reduced; this is mainly for production reasons (additional cost for flexible manufacturing, advertising, technical assistance, etc.) and/or demand reasons (subjective quality factors, status symbol and market distortions).

Of course, if one considers the aims and requirements of sustainable economic systems, it is nearly impossible to set, for each good, the 'optimal' degree of diversification both for companies and for consumers. On the other hand, it is fairly easy to measure and assess the objective quality/ performance characteristics of goods.

Otherwise, whereas in all the industrial countries the average level of global quality/performance of highly diversified products cannot be considered higher than it is (and was) of low diversified products (if only in a few cases), the price for the same products as related to the yearly per-capita average revenue has significantly changed, in a way that suggests they have a variable economic meaning over time.

On the one side, excessive, ingenious and possibly marketing-driven diversification/flexibility often alters the real needs of consumers. This requires additional resources, funds and management, which increase production costs and take funds away from other priorities, such as the modification of all production criteria to be consistent with sustainability (type and characteristics of products, cleaner production systems, etc.). On the other side, the price increases— above all those of basic goods— reduce the buying power of citizens and change their attitudes towards goods and services.

This situation, in conclusion, prevents new principles for sustainability being examined and implemented, and above all it stops the development of an increasing number of technically and economically viable alternative production solutions in the same fields of activity, of which the most appropriate might be selected and implemented, in different environmental, economic, territorial and social situations.

DISCUSSION

Production industries can play an outstanding role in the quality of life and in pursuing and building up sustainable economic systems, both on the side of resource conservation/enhancement and of employment (finally in the quality of life). But they can only do so if they profoundly modify the principles that all their choices are based on, notably, types and performance of products and technologies, oriented towards a rational utilisation of natural resources, leading to a dramatic slow down of non-renewable resources and to a corresponding dramatic increase of renewable resources (notably foodstuffs) if the world population continues to increase. This modification can be achieved only by adopting effective long-range plans and a modern entrepreneurial culture oriented to pursue, at the same time, the typical goals of companies and to satisfy the needs of societies, as a whole, in a 'sustainable perspective'.

This ambitious target is neither easy nor inexpensive for all parties concerned, and will definitely meet severe barriers of different nature, which usually interact to reduce the possibility of moving towards sustainability.

Great efforts will jointly be made by all private and state enterprises, international organisations, governments, local authorities and financial systems to overcome any kinds of barriers, of course with different roles and responsibilities; suitable economic policy instruments will be set and implemented to create the general conditions to induce companies to adopt fully the above stated culture and principles.

Of course, political and social stability in all countries is the basic condition to induce manufacturing industries to invest by introducing innovative principles, but economic advantage—both in a short- and long-term perspective—still remains the core condition. Actually, since any trend modification makes uncertain any choice and investment, international organisations and national governments should adopt multilateral integrated agreements, setting the fields for selected investment in the different countries (renewable and non-renewable resources, environment, manufacturing branches, technology transfer, services, information, etc.) and the relative priorities. These agreements should be regarded as innovative tools for implementing effective forms of economic, technological and scientific co-operation, where enterprises could find new perspectives of single and combined activity.

International organisations can play an important role in the field of emission standards and of indicators of sustainability, in order to set the upper and lower limits to several manifestations of human life.

At the national and local level both governments and local authorities should identify and propose the guidelines, that is to say the branches that seem to be appropriate to achieve the local 'sustainable goals' and, on the contrary, those that seem to be incompatible. To this purpose, the intersectorial and territorial programmes seem to be the most suitable economic policy instrument.

Laws and regulations have to be considered as a forceful way to persuade companies to modify their directions, but recent examples can lead to positive evaluations about the necessity of their adoption in cases where remarkable drawbacks are engendered by production industries (the laws concerning biodegradability of surface active agents in the 1960s, in all industrial countries, have been implemented and effective in less than three years).

The public companies should more and more assume a leading and guiding role, in the sense that they should become models of technological and managerial innovation, in the fields considered strategic for sustainable development, and their performance should be useful to all private companies —especially the small and medium-sized ones—and to the whole economic system, directly and indirectly. Moreover, they can contribute significantly to R&D activities, on their own or together with private companies, by managing specialised centres, as has been stressed earlier.

The financial and fiscal tools should be addressed to the most innovative and 'society-oriented' companies, to support them in achieving their own goals. All of these economic policy instruments can be highly effective especially within adequate forms of relationships between industry and governments/institutions; the old-style confrontational approach, which has existed—and still exists—in some countries, is unable to induce the desired transformation and an appropriate creativity.

Moreover, it must be highlighted that all transformations needed for achieving a sustainable development are inevitably bound to be in conflict with the fundamental environmental directions in the present phase of economic development, notably, the widely consolidated 'end-of-pipe' technologies, which have become a great business for many large companies.

Finally, specific attention must be given to small and medium-sized enterprises all over the world, mainly because they lack an autonomous capacity for developing new technologies and implementing suited innovations. For SME the presence and activity of 'technopolitan sites' assume a vital importance for their innovative capacity and their competitiveness.

REFERENCES

Ayres, R.U. (1993) 'Industrial metabolism', in T.Jackson (ed.) *Clean Production Strategies,* Boca Raton, Fla., USA: Lewis, pp. 165–188.

Baas, L.W. (1996) 'An integrated approach to cleaner production', in K.B.Misra (ed.) *Clean Production. Environmental and Economic Perspectives,* Berlin, Germany: Springer Verlag, pp. 211–229.

Barbiroli, G. (1989) 'The elaboration of global performance indices for evaluating durable goods', *Technovation* 9: 83–100.

Barbiroli, G. (1993) *Il fattore energia come volano di mutamento tecnologico. Verso nuovi criteri di efficienza globale nei processi produttivi,* Milano, Italy: Angeli (in Italian).

Barbiroli, G. (1995) 'Measuring technological dynamics and structural change, their interrelationships and their effects', *Structural Change and Economic Dynamics* 6, 3: 377–396.

Barbiroli, G. (1996a) 'New indicators for measuring the manifold aspects of technical and economic efficiency of production processes and technologies', *Technovation* 16, 7: 341–356.

Barbiroli, G. (1996b) 'Strategic technological pathways for sustainable development', in K.B.Misra (ed.) *Clean Production. Environmental and Economic Perspectives,* Berlin, Germany: Springer-Verlag, pp. 121–169.

Barbiroli, G. (1996c) 'Guiding the evolution of technologies in pursuit of sustainability and the cost of change', in *Proceedings 3rd ICEP,* Budapest, Hungary, April, pp. 61–69.

Barbiroli, G., Fiorini, M., Mazzaracchio, P. and Raggi, A. (1993) 'Eco-compatible technologies: criteria for their assessment as strategic patterns', in *Proceedings Second International Conference on Environmental Pollution,* ICEP 2, Barcelona, Spain, September-October, pp. 489–496.

Barbiroli, G. and Mazzaracchio, P. (1995) 'Synthetic fuel technologies as strategic pathways', *Energy Sources* 17, 5: 595–604.

Barbiroli, G., Raggi, A., Fiorini, M. and Mazzaracchio, P. (1996a) 'The contribution of eco-compatible technologies to resource conservation and upgrading', in *Proceedings 3rd ICEP,* Budapest, Hungary, April, pp. 476–484.

Barbiroli, G., Raggi, A. and Fiorini, M. (1996b) 'Resource scarcity as a stimulus for advantageous innovations', in K.B.Misra (ed.) *Clean Production. Environmental and Economic Perspectives,* Berlin, Germany: Springer-Verlag, pp. 479–499.

Bomsel, O. and Roos, P. (1990) 'From commodities to evolutionary materials', *Resources Policy* 16, 2: 104–111.

Christi, Y. (1993) 'Protein enhancement of foods', *Biotechnology Advances* 11, 3: 467.

Coates, J.F. (1993) 'Three Rs: recycle, reclaim, remanufacture', *Research-Technology Management* 36: 6–7.

Cooper, T. (1994) 'The durability of consumer durables', *Business Strategy and the Environment* 3: 23–30.

Formica, P. (1991) *Tecnopoli, luoghi e sentieri dell'innovazione,* Torino, Italy: ISEDI.

Gibson, D.K. (ed.) (1992) *The Technopolis Phenomenon. Smart Cities, fast Systems, Global Networks,* Lamham, Md., USA: Rowman & Littlefield.

Goodman D., Sorj, B. and Wilkinson, J. (1987) *From Farming to Biotechnology: A Theory of Agroindustrial Development,* New York, USA: Blackwell.

Jackson, T. (1989) 'Is fusion feasible? An assessment of the methodology and policy implications', *Energy Policy* 17, 4: 407–412.

Kaldjian, P. (1994) 'Environmental results through industrial design', *World Wastes* 37: 15–19.

Kühne, H.-M and Aulich, H. (1992) 'Solar energy systems: assessment of present and future potential', *Energy Policy* 20, 9: 847–860.

Lambert, B. and Joos, H. (1989) 'Fundamental aspects of rhizobacterial plant growth promotion research', *Trends in Biotechnology* 7, 8: 215–219.

Lin, C.M. (1994) 'Clean production in small and medium-sized industries', *Industry and Environment* October-December: 68–70.

Madden, K.N. and French, J. (1989) 'Papermaking properties of some Australian non-wood fibres', *Appita* 42, 6: 433–437.

Mannion, A.M. (1992) 'Sustainable development and biotechnology', *Environmental Conservation* 19, 4: 297–306.

Morris, D. (1992) 'Getting the most from our resources', *Economic Development Review* 10: 8–12.

Owen, J.V. (1993) 'Environmentally conscious manufacturing', *Manufacturing Engineering* 111: 44–55.

Pelsy, G. (1988) 'Les applications des biotechnologies dans la filiere agro-alimentaire', *Annales des Mines* 10–11: 61–71.

Pelsy, G. (ed.) (1989) 'Agricultural biotechnology. Opportunities for international development. Synthesis report executive summary', *International Service for National Agricultural Research,* Canberra, Australia: World Bank, Australian Centre for International Development Assistance Bureau.

Raggi, A. (1992) 'Design and manufacturing of durable goods in view of recycling', *Forum Ware* 20: 24–29.

Sasson, A. (1988) *Biotechnologies and Development,* Paris, France: UNESCO (CTA).

Sims, G.P. (1991) 'Hydroelectric energy', *Energy Policy* 9, 8: 776–786.

Stahel, W.R. (1986) 'Product life as a variable', *Sci. Public Policy* 13, 4: 185.

Stahel, W.R. (1992) 'Waste prevention through an alternative product design and new technical and commercial strategies: three case studies', in *Proceedings 6th International Solid Wastes Congress and Exhibition,* Madrid, Spain, vol. 1, pp. 1–13.

Stahel, W.R. and Jackson, T. (1993) 'Optimal utilisation and durability', in T.Jackson (ed.) *Clean Production Strategies,* Boca Raton, Fla., USA: Lewis, pp. 261–294.

Steuteville, R. (1992) 'Appliance recycling takes off', *BioCycle* 33: 48–49.
UNEP (1995) *Industry and Environment: Cleaner Production* vol. 2.
van Berkel, R. (1995) 'Introduction to cleaner production assessments with applications in the food processing industry', *Industry and Environment* January-March: 8–15.
Wilson, E.M. (1991) 'Small-scale hydroelectricity', *Energy Policy* 19, 8: 787–791.
Winter, C.J. (1991) 'Solar hydrogen energy trade', *Energy Policy* 19, 5: 494–502.

SUGGESTED READING

Berry, R.J. (1993) *Environmental Dilemmas. Ethics and Decisions,* London, UK: Chapman & Hall.
Chechile, R.A. and Carlisle, S. (eds) (1991) *Environmental Decision Making. A Multidisciplinary Perspective,* New York, USA: Van Nostrand Reinhold.
Dalzell, J.M. (ed.) (1994) *Food Industry and the Environment,* Blackie Academic and Professional, London, UK: Chapman and Hall.
Faulkner, W. and Senker, J. (1995) *Knowledge Frontiers,* Oxford, UK: Clarendon Press.
Friedman, F.B. (1993) *Practical Guide to Environmental Management,* Washington, DC, USA: Environmental Law Institute.
Huising, D. (1989) *Cleaner Technologies through Process Modifications, Material Substitutions and Ecologically based Ethical Values,* Industry and Environment, vol. 12, no. 1, Paris, France: UNEP.
Jackson, T. (ed.) (1993) *Clean Production Strategies,* Boca Raton, Fla., USA: Lewis Publishers.
Kirkwood, C. and Longley, A.J. (eds) (1995) *Clean Technology and the Environment,* Blackie Academic and Professional, London, UK: Chapman and Hall.
Misra, K.B. (ed.) (1996) *Clean Production. Environmental and Economic Perspectives,* Berlin, Germany: Springer-Verlag.
Nussbaum, M.C. and Sen, A. (eds) (1993) *The Quality of Life,* Oxford, UK: Clarendon Press.

SELF-ASSESSMENT QUESTIONS

1 How can industry contribute to achieving resource conservation?
   (a) By modifying the existing production processes by increasing materials and energy transformation yields.
   (b) By implementing and spreading cleaner production systems, selecting appropriate production systems and expanding the lifespan of durable goods
   (c) By replacing renewable resources for non-renewable in production processes.
2 How can 'life-extension' of goods be primarily achieved?
   (a) By increasing material resistance.
   (b) By redesigning and manufacturing durables in a way so as to enable their further development/upgrading, re-use, repair, reconditioning and recycling.
   (c) By increasing their prices and changing consumers' attitudes.
3 What are the most important aspects of efficiency a company must achieve?
   (a) Environmental efficiency, exclusive of other aspects of efficiency.
   (b) Input and quality efficiency.
   (c) Several aspects of performance, including environmental, input, quality, materials, energy efficiencies.
4 How can 'dematerialising' of goods and of industry be achieved?
   (a) By simply reducing the weight of products.
   (b) By increasing the incidence of non-material goods (know-how, software, intellectual skills, etc.).
   (c) By reducing the importance of manufacturing and increasing that of service activities.
5 What should one regard as 'real technological pluralism'?
   (a) A wide number of economically equivalent alternatives for producing the same category of goods so as to increase the potential replaceability.
   (b) A wide product diversification in the same manufacturing branch.
   (c) A large number of companies adopting the same technologies.
6 Where should efforts for R&D be mainly addressed?
   (a) To increase productivity
   (b) To improve product quality/performance and diversification.
   (c) To widen the applied knowledge in a way so as to form the basis for appropriate and economically viable innovations and to increase labour intensity.
7 What are the most remarkable internal barriers to the adoption of cleaner production by companies?

(a) Technical, organisational and attitudinal barriers.
(b) Economic and demand barriers.
(c) Governmental barriers.

8 What should be the most desirable features of technologies for helping the less developed countries achieve their own sustainable development?
(a) Business-driven and fairly sophisticated technologies.
(b) High labour intensive and appropriate technologies.
(c) High capital and resource intensive technologies.

13

TRANSPORT AND THE ENVIRONMENT

David A.Hensher, Paul Hooper, Ross Robinson and Sophia Everett

SUMMARY

The objective of this chapter is to highlight the environmental impacts associated with the provision of transport services necessary for the movement of passengers and goods. The management of the transport sector requires trade-offs to be made between three major goals of transport management—growth, equity and environmental sustainability. To place the discussion on ways of reducing the environmental impacts of transport activity in context, we review international trends in emissions of the major noxious air pollutants; we then take a close look at policy instruments that might contribute to reducing the impact of transport services on the quality of the environment.

ACADEMIC OBJECTIVES

The aim is to gain familiarity with the contribution of the transport sector to the environmental problems of air quality, global warming and energy consumption, and ways in which the negative impacts can be reduced. On completion of this chapter readers should be able to appreciate the challenges facing the transport sector in coming to grips with mounting environmental problems. They should also have an understanding of the potential of specific policy instruments for assisting in the environmental management of the transport sector. Case studies are provided for urban passenger transport and global warming, shipping and water pollution, aviation and emissions.

INTRODUCTION

The community views transport in the modern era with mixed feelings. It provides the technological means to facilitate movement of passengers and goods, but is also at the centre of the growing concern about environmental degradation in the form of air pollution, global warming, noise and reduced safety. Combined with traffic congestion in major conurbations—at ports, airports and on roads —the transport sector has been identified as a major contributor to the ills of twentieth-century society. Roads in particular, which provide the infrastructure for moving cars and trucks, have come under increasing criticism on environmental grounds. The question has been raised of whether they are the servants of technology rather than offering a positive opportunity to mould the environment.

Transport systems provide the mechanistic infrastructure used to facilitate movement; they also have a broader and socially valuable role in contributing to the economic, social and environmental fabric of civilisation. What would a nation look like and how would it function in the absence of streets and roads, airports and ports? Transport systems today are key elements in the global economic system, components in a society's amenity infrastructure, and settings of mounting economic, environmental and social challenges.

The emphasis on outcomes (e.g. accessibility, clean air, safety) rather than means (e.g. cars, buses, trains, trucks, planes, ships) is more important today than ever before, since the traditional transport emphasis fails to accommodate institutionally the growing set of ways of 'moving' information and people's contributions to a nation's activity. Setting constraints on particular forms of transport to achieve desirable outcomes must be evaluated within the broader set of ways to satisfy opportunities offered by 'high-tech' and 'high-touch' industries, in contrast to the imposition of a priori beliefs that only improvements in particular modes of transport will 'solve' the ills of society.

Community concern and trade-offs

Many environmental problems that are both real and sensitive community issues stem from the use of transport infrastructure by passenger and freight vehicles. Automobiles, aeroplanes and trucks are major sources of local pollutants, such as lead, carbon monoxide and noise. Traffic congestion exacerbates these problems, and also imposes direct economic and health costs on users and non-users in the form of wasted time and money, stress and other illnesses. Transport systems also make a significant contribution to global warming through emissions of carbon dioxide and other greenhouse gases.

As the observed level of wealth in society increases and the demand for transport use increases, society faces special challenges in containing and reversing trends in a wide range of negative impacts made by greater mobility. Communities are not expressing blanket concern about transport *per se;* rather they are concerned about specific issues (principally motorways, toll roads and the location of seaports and airports) and about the 'failure' of government to do something about the harmful effects of these specific investments. Governments, however, face the difficult trade-off between what they know are appropriate policies to help curb the desire for mobility (especially by a range of pricing instruments) while lacking the political will to execute such policies. The emphasis on physical incentives/disincentives to achieve change relative to financial opportunities continues to be a major constraint on containing the environmental costs of transport systems.

While recognising the concern over adverse environmental effects of transport, its many positive features should not be overlooked. The challenge is to manage the benefits of transportation better, so that the broad set of environmental impacts are reduced to acceptable levels, while ensuring acceptable outcomes in terms of economic performance and equity. Civic pride embellished in the design of transport systems must be given centre stage in the deliberations.

Emphasis of the chapter

Transportation should be evaluated within a broad set of transport management goals—growth, equity and environment—and attention paid to the impact of incentive- and disincentive-based policy instruments, with respect to topical evaluative criteria such as air quality, traffic noise, traffic congestion, global warming and accidents. The chapter emphasises key environmental issues by concentrating on four performance criteria linked with the environment: *air pollution, global warming, energy consumption and traffic congestion;* however, we recognise the importance of water pollution with a case study in the maritime sector. Additional case studies on urban passenger transport and airports emphasise greenhouse gas emissions and air quality.

A major objective is to *illustrate* the debate through a set of reasoned arguments based on better/best practice information, transmitted in a readable form. Historical and contemporary evidence can assist provided it has been properly interpreted. If readers reconsider their own knowledge base and interpretations of the debate on the role of transportation to a nation, then the chapter has achieved its purpose. The distinction between rationality and advocacy is often cited as a distinguishing attribute of alternative views on the role of transportation; however, rational debate needs advocates.

OPPORTUNITIES FOR ENVIRONMENTAL MANAGEMENT —AN INTEGRATING FRAMEWORK FOR PLANNING TRANSPORT SYSTEMS

We know a great deal about the nations of the next 35 years because they are essentially the nations of the present. Much of what we see and are likely to see in the future is driven by inherited geography, topography and climate, and by the highly specialised economic and political functions that they have acquired. Transport facilities are intertwined with the culture and character of nations. But the utilitarian role must also be fulfilled. Designing aesthetically pleasing transport systems is as much a challenge to a diversifying culture as seeking solutions to its evergrowing desire for better accessibility, lifestyle and environmental sustainability.

The urban area is a major focus of environmental management, and illustrates the potential of pursuing environmental management through an integrated framework. Typical of many recent empirical simulations studies, Roy *et al.* (1995) simulated the relationships between urban residential density, job decentralisation and transport energy consumption when new housing is added as outward urban expansion or infill (redevelopment) within an existing urban area. They find for large Australian cities that strong infill can produce energy savings in commuting up to 17 per cent higher than the best sprawl scenarios, so long as the infill policies are accompanied by significant improvements to the level of service in public transport. However, the degree of infill required would take many years to occur, especially given that much of the housing stock has been constructed during the last 40 years and is generally in good condition. Furthermore, the energy advantage of infill (which is also a greenhouse advantage, since greenhouse gas emission changes are almost directly proportional to energy consumed) shrinks to 2–3 per cent when workplace location choice occurs in the nearest sub-centre.

Roy *et al.* (1995) conclude that transport demand management policies and job matching to disperse sub-centres in residential areas will have a greater impact on reducing greenhouse gases in the next 10–20 years than will infill policies (which could take anywhere from 40 to 100 years to have a noticeable impact). They also suggest that if these sub-centres contain ancillary services, public transport stops and other public transport nodes such as stops for circumferential express buses, market forces will automatically increase housing densities in areas surrounding these sub-centres. This will create a natural equilibrium in levels of infill without the need for intrusive land-use control. This is the basis of the idea of the urban village. Studies in Adelaide have shown that urban housing consolidation and infill does not result in significant increases in population density; its role has been, at best, to arrest the trend towards lower population density in the inner suburban areas.

The opportunity for transit corridor retrofit is also real. Travel densities that support public transport can be produced from low density residential activity, provided we allow for a wider range of more flexible forms of public transport such as hail-n-ride bus services using both mini-buses and conventional sized buses. Limiting public transport to very rigid traditional forms of transport such as rail and scheduled route bus services is not helping the rejuvenation potential of public transport as an environmentally appealing alternative to the automobile.

Other considerations affecting location choice, and thus commuting times, such as high job turnover, high residential relocation costs and employment heterogeneity in multi-worker households have been suggested by Small and Song (1992) as reasons why households seek accessibility to an array of possible future jobs rather than just to their current employment. Wachs *et al.* (1993) tracked the differences over six years between home and work location among 30,000 employees of a large health care provider in Southern California. They found that work trip length had in general not grown over the six years, but the growth of the workforce had contributed more to the growth in local traffic congestion than had a lengthening of the work trip over time. This implies that strategies for reducing vehicle kilometres should reconsider the predominant interest in commuting activity and give more emphasis to non-commuting travel as vehicle kilometres increase.

Securing higher residential densities regardless of distance from the core of an urban area appears on balance to reduce automobile kilometres travelled, but only if accompanied by travel pricing policies designed to make the car less attractive and by complementary improvements in public transport. The increase in density near rail stations and bus routes once it is combined with road pricing has been recognised for many years; the constraint is the weakness of political will to implement serious road pricing. Increasingly urban and regional simulation studies are finding that a comparison between dispersed-growth and contained-growth scenarios finds no clear winning scenario in terms of emissions. Concentration of travel in the centres leaves the peripheral areas less congested and therefore people travel farther in these areas.

The anthropological invariance view of travel behaviour is very appealing. When combined with the residential density effect (and pricing of automobile use) we begin to see niche opportunities for public transport—train *and* bus—throughout the urban area.

EMERGING DIRECTIONS

Out of a heritage evolve current trends and speculations as to which ones are likely to dominate the patterns of national evolution over the next 35 years and beyond. Accurate prediction of the distant future is, as history all too often shows us, a risky business. Suggestions that some appealing infill policies will require 40–100 years to have a noticeable impact may disappoint some, and encourage others to doubt the ability of public transport (as a beneficiary of infill) to help bring about desirable change; maybe the 'solution' is to redefine the period of time in the future for which the current generation should be responsible, and to propose the next 100 years as a candidate. Questions on the agenda of the 1990s, such as global warming, local air pollution, energy consumption levels, 'sprawling' cities and loss of amenities may well be handled best by technologies that some currently criticise, or even regard (like cars and roads) as inimical to society, while favouring technologies such as railways. Yet railways have been described by some road advocates as:

> a technology belonging to the nineteenth century [which] are about as efficient when compared to roads as waterways and canals…are when compared to railways. But they enjoy a special place in the affections of many otherwise sensible individuals.
>
> (Ross Swan, Editorial in *World Highways,* April 1995:7)

Two major international trends need to be set out relating to the performance criteria of interest. The first trend is the continuing rise in demand for transport services. Reductions in trade barriers and the globalisation of business have led to record levels of international trade and the pace of this growth is unlikely to abate in the next two decades. This generates increasing flows of goods and business people and coupled with a booming business in international tourism, shipping operators, seaports, airlines and airports are planning for expansion. However, we draw particular attention to the trends in ownership and usage of automobiles.

For example, between 1970–1971 and 1989–1990, total passenger vehicle registrations in Australia grew by an average of 3.5 per cent per annum (a low of 1.25 per cent in 1985–1986 and a high of 5.6 per cent in 1972–1973), equivalent to 0.39 vehicles per head in 1970–1971 and 0.44 vehicles per head in 1988–1989. This trend is likely to continue until automobile ownership levels off when, as is happening in the USA, it approaches 0.6–0.8 personal vehicles per adult. The increase in ownership by females is most noticeable. Europe is witnessing a similar trend. Korea has witnessed a 25 per cent annual increase in car ownership since 1982, growing from 1 million to 8 million vehicles in a population in 1995 of 70 million people. In the last ten years, the number of passenger vehicles in China has increased by an average of 15 per cent per annum (rising as high as 30 per cent per annum in developed coastal regions). There are currently 9.93 million cars and 30 million Chinese people licensed to drive a car. As many societies approach a value of 0.8 vehicles per head (assuming 20 per cent non-adults), it is likely, given the strong evidence that distance travelled per passenger vehicle has shown remarkably little variability over time, that per capita use will also level off. The growth in recent years in the leasing of automobiles in the rental car industry to a growing market of short-term visitors to cities together with the growing use of taxis adds further automobile traffic to the system.

Since 1971, average annual kilometres per passenger vehicle in many countries have varied between 10,000 and 16,000. Total time spent travelling per person also has shown remarkable regularity across countries and time periods (confirming the presence of a constant travel time budget). As improvements in transport infrastructure and service levels occur, households and firms relocate to take advantage of other benefits of location while preserving the mean and variance of travel times throughout the urban area (Marchetti, 1992). Any future improvements in road infrastructure that increase average speeds would tend to increase annual distance travelled without affecting travel times in any noticeable way.

The second trend is the very noticeable reduction in total noxious air pollutant emissions in the western world. However, in emerging economies there are worrisome prospects of increases in greenhouse gas and all other emissions.

Stringent new requirements for emissions of hydrocarbons, carbon monoxide and nitrogen oxides from new automobiles (and trucks) have been introduced in many countries. Energy consumption per vehicle kilometre travelled is declining in western economies although increases in the growth of automobiles and total vehicle kilometres result in a net increase averaging 3.36 per cent per annum. Thus the absolute reductions in emissions are even more impressive when we see the growth in vehicle use. The benefit of emission control legislation is evident. Disturbingly, however, efforts in countries such as Australia and the USA amount to very little internationally in view of the reduction world-wide in carbon monoxide and hydrocarbon emissions being negated about ten years from now by the projected growth in countries where emission controls are minimal (Table 13.1).

This contrasts with the continuing increase in greenhouse gas emissions, primarily CO_2. More than three-quarters of the carbon dioxide emitted from all transport sources in Australia and the USA comes from automobile and truck fuel (Table 13.2). Changes in CO_2 emissions are highly correlated with changes in automobile fuel efficiency and vehicle use, strongly hinting at the major benefits available from improvements in the fuel efficiency of automobiles and reduction in vehicle use. Dobes (1995) has compared greenhouse gas emissions in Australia in 1900 and the year 2000 and concludes that

> Within the limits of long-term historical comparisons and availability of data, it may be concluded that use of the internal combustion engine itself has not contributed disproportionately to greenhouse gas emissions in the transport sector. The equally qualified corollary is that an economy of size similar to that of today would not have generated a significantly lower quantity of greenhouse gases had the motor car not replaced animals and steam from 1900.
>
> (Dobes, 1995:19)

The automobile is the dominating form of motorised transport in terms of the major emissions. For

Table 13.1 Global trends in motor vehicle emissions

Year	*Car*	*Light trucks*	*Motorcycles*	*Heavy trucks*
Carbon monoxide (tons/year)				
1990	223,357,376	1,260,248	8,168,139	7,793,019
1995	217,043,366	1,410,969	9,209,773	8,301,436
2000	183,131,401	1,623,464	10,464,134	9,374,474
2010	97,559,141	2,077,267	14,166,965	12,214,989
Hydrocarbons (tons/year)				
1990	30,025,462	506,570	5,568,461	1,818,987
1995	26,309,692	529,987	6,387,750	1,830,407
2000	23,314,293	607,874	7,075,987	2,038,046
2010	22,084,536	798,924	8,227,297	2,637,515
Nitrogen oxides (tons/year)				
1990	11,049,831	1,995,856	481,970	14,654,156
1995	10,651,242	2,205,343	550,760	13,459,297
2000	8,387,873	2,517,257	619,217	15,054,203
2010	5,996,606	3,113,332	782,274	18,752,930
Carbon dioxide (tons/year)				
1990	2,140,563,394	648,810,244	115,235,655	1,095,306,335
1995	2,326,778,635	714,188,146	131,007,340	1,272,857,434
2000	2,287,475,047	764,561,988	147,578,254	1,468,158,497
2010	2,588,738,693	802,074,961	190,301,058	1,934,317,592

Source: Walsh, 1993

Table 13.2 Estimated emission levels by transport mode and transport proportion of total emissions, 1992–1993

a. Australia ('000 tonnes)

Trace gas	*Road*	*Rail*	*Air*	*Sea*	*Total transport*	*Total emissions energy use*	*Transport as a % of total emissions*
Carbon dioxide	53815	1602	8618	4132	68167	288353	23.6
Carbon monoxide	3073	14	80	124	3291	4470	73.6
Nitrogen oxides	307	40	33	91	471	1276	37.1
NMVOCs	402	3	3	38	446	628	70.9

b. USA (millions of short tons)

Trace gas	*Road*	*Rail*	*Air*	*Other off highway*	*Sea*	*Total transport*	*Transport as a % of total emissions*
Carbon dioxide	311.3	9.1 (est.)	54.3	in road	21.5 (est.)	396.2	31.8
Carbon monoxide	59.99	0.12	1.02	12.88	0.06	74.07	77.4
Nitrogen oxides	7.44	0.95	0.15	2.04	0.18	10.76	44.5
NMVOCs	6.09	0.04	0.20	1.91	0.04	8.28	35.6
PM-10	0.20	0.05	0.05	0.27	0.03	0.59	1.3
Sulphur dioxide	0.44	0.07	0.01	0	0.2	0.72	3.3

Source: USA—*National Air Quality and Emissions Trend Report 1993*, Office of Air Quality Planning and Standards, EPA, October 1994;
Australia—*Greenhouse Gas Emissions from Australian Transport*, BTCE Report 88
Notes: USA—all highway vehicles
Australia—passenger cars

example, in Australia (Table 13.2a) we see that 78.9 per cent of all carbon dioxide, 93.3 per cent of all carbon monoxide (smog), 65 per cent of nitrogen oxides and 90.3 per cent of NMVOCs in the transport sector is produced by cars and trucks. Importantly, less than one-quarter of all carbon dioxide emitted derives from the transport sector compared to nearly three-quarters of carbon monoxide. In OECD Europe (OECD, 1991), in 1991 the estimate of transport's share of nitrogen oxide, carbon monoxide and carbon dioxide is 60 per cent (49–76 per cent across countries), 78 per cent (71–86 per cent) and 21 per cent respectively. Similar findings for carbon monoxide and nitrogen oxide exist in the USA (Table 13.2b) although the VOC contributions are much higher as a percentage of all sources of VOC.

Table 13.3 shows the world-wide growth in cars, light-duty trucks (LDTs), heavy-duty trucks (HDTs) and motorcycles (MCs), the dominant road vehicle in many fast developing economies; it clearly indicates the potential increase, if no effort is made to curb them, of air pollutants. And further interesting indications emerge: economies in the rapidly industrialising Asian-Pacific have annual vehicle growth rates at least 50 per cent higher than the USA and Europe, signalling the challenges the world faces in containing the growth of the automobile and the truck. The CO_2 emissions per capita for selected countries (Table 13.4) show the much higher rates for wealthier nations and signal the direction in which emerging economies will be heading without appropriate actions to stem the massive increases in carbon dioxide. Most emerging and relatively poor economies currently exhibit rates per capita less than 1 tonne in contrast to an average of 3.34 for OECD countries. The challenges facing these countries to reduce emissions are huge, but might benefit from the lessons of policy instruments implemented in the mature economies such as emissions controls on automobiles and trucks. The Australian evidence (Figure 13.1) is very encouraging.

Table 13.3 Annual road vehicle growth rates

	USA	EC	EFTA	EE	OECD Pacific	RICA	RICB	ROW
1989–1995								
Cars	1.5	2.5	2.0	3.0	4.0	7.0	7.0	3.5
LDT	2.0	2.5	2.0	3.0	4.0	7.0	7.0	3.5
MC	0.0	0.0	0.0	3.0	−1.0	7.0	7.0	3.5
HDT	2.0	2.5	2.5	3.5	3.0	5.0	5.0	3.5
1995–2000								
Cars	1.5	2.5	2.0	3.5	3.0	6.0	6.0	2.5
LDT	2.0	2.5	2.0	3.5	3.0	6.0	6.0	2.5
MC	0.0	0.0	0.0	3.0	−1.0	5.0	5.0	3.5
HDT	2.0	2.5	2.5	3.5	2.5	4.0	4.0	3.5
2000–2010								
Cars	1.5	2.0	2.0	3.5	2.0	5.0	5.0	3.0
LDT	1.5	2.0	2.0	3.0	2.0	5.0	5.0	3.0
MC	0.0	0.0	0.0	2.0	0.0	5.0	5.0	3.0
HDT	2.0	2.5	2.0	3.0	2.0	4.0	4.0	3.5

Source: Walsh, 1993

Notes: EC=European Community; EFTA=European Free Trade Association; EE=Eastern Europe and the republics of the former USSR; OECD Pacific=OECD countries of the Pacific including Japan, Australia and New Zealand; RICA=rapidly industrialising countries that have taken steps to introduce state-of-the-art pollution control on cars (Brazil, Chile, Hong Kong, Mexico, Singapore, South Korea and Taiwan); RICB=rapidly industrialising countries with minimal pollution control programmes (Indonesia, Malaysia, Philippines, Thailand); ROW=rest of the world including much of Africa, Asia and Latin America; LDT=light-duty truck; HDT=heavy-duty truck; MC=motorcycle

Table 13.4 Carbon dioxide emissions for selected countries, 1990

Country	Population (millions)	Emissions (m tonnes)	Transport emissions per capita (tonnes)	Contributions of transport to emissions (%)
Australia	17.1	288.4	3.87	23.2
Belgium	10.0	122.1	3.64	29.7
Bulgaria	8.8	81.5	1.34	14.5
Canada	26.5	452.7	5.46	32.0
Denmark	5.1	55.5	2.92	27.0
Finland	5.0	60.6	3.01	24.8
France	56.4	350.5	2.07	33.3
Hungary	10.6	75.2	0.99	13.8
India	850	563	0.08	11.7
Indonesia	178	118.2	0.17	26.0
Italy	57.7	371.8	1.76	27.3
Japan	124	1074.7	1.73	19.9
New Zealand	3.4	26.2	3.02	39.3
Nigeria	115.5	71.6	0.05	8.7
Philippines	61.5	38.4	0.22	35.2
Poland	38.2	473.5	0.91	7.3
Sri Lanka	17.0	3.5	0.12	60.0
Sweden	8.6	56.3	2.84	43.2
Switzerland	6.7	45.9	2.52	36.8
Turkey	56.1	182.5	0.39	11.8
UK	57.4	564	2.10	21.4
USA	250	5224	6.11	29.2
OECD	839	10,300	3.34	27.2
World	5,292	22,700	0.83	19.3

Source: BTCE, 1995

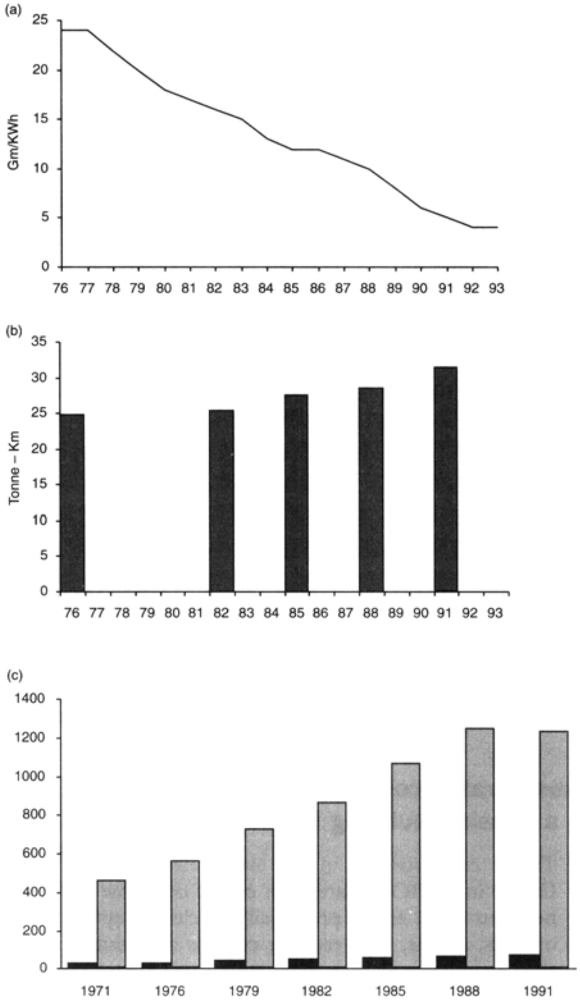


Figure 13.1 a) Truck emissions: nitrous oxides (gm/kWh); b) Truck fuel efficiency (tonne km per litre);
c) Improvements in truck productivity (thousands of tonne km per year per vehicle)

Source: Australasian Transport News, 1994

MANAGING THE ENVIRONMENTAL EFFECTS OF TRANSPORT

The impacts of automobiles and trucks on air quality, on peace and quiet, on scenic value and on the ecology concern many in the community. An extreme view is that these modes impose a net cost to society and only a very small minority of people are willing to reverse the long-standing and pronounced shift towards personal mobility represented by automobile ownership. However, there is little if any economic evidence to justify draconian measures in reducing the use of automobiles and trucks sufficiently to 'solve' environmental problems (Small, 1991). The desired responses would have to be either a switch to public transport and/or a reduction in travel. Given that only 8 per cent of all urban passenger movements in Australia are by public transport, a switch away from the automobile of just 1 percentage point (for example) represents an increase of around 10 per cent in public transport patronage. Public transport systems would therefore require substantial adjustment after even very slight reductions in car use.

Fortunately there are plenty of alternative possibilities—technological and behavioural—that could greatly reduce the problems without much affecting overall mobility. For example, the evidence that a small fraction of the cars is causing a disproportionately high fraction of air emissions suggests that a greater effort to improve the inspection and maintenance of the highly sophisticated pollution control devices on cars and trucks would probably greatly reduce this particular environmental impact. This may not satisfy the proponents of alternative means of movement such as public transport, yet it clearly answers their concern about the major adverse environmental impact of the car. Even so, the challenge for reducing greenhouse gas emissions remains, although a recent study undertaken by the Institute of Transport Studies (unpublished) supports increasing controls on automobile technology, combined with pricing from the set of possibilities (increased fuel excise, carbon tax, congestion pricing and parking pricing) and the promotion of alternative work practices.

Apogee Research (1994) has reviewed the literature on transport control measures (TCMs) to identify their effectiveness in reducing regional emissions. They conclude that pricing has the strongest impact on reducing emissions of mobile sources, with land-use planning, telecommuting and compressed work weeks having high potential, although the evidence for the latter is currently speculative. TCMs such as high-occupancy vehicle (HOV) lanes, incident management, employer trip reduction, transit improvements, signal timing, area-wide ride-sharing, park and ride parking facilities, bicycle/pedestrian facilities and buy-back of older cars have a very weak impact on mobile source emissions. Many of these poorly performing TCMs have been recognised for some time as being cosmetic in impact and no substitutes for real impacting policies. The recent increase in off-street quality parking at key suburban rail stations has encouraged switching of modal access from bus to car, with negligible impact on linehaul modal switching. Buy-back of older vehicles, a policy being promoted by some governments, forces purchase of younger vehicles. These vehicles are more fuel efficient, cleaner and have lower operating costs, tending to encourage increases in vehicle kilometres in line with the idea of a constant expenditure budget of transport (approximately 16 per cent of gross expenditure when automobile capital is included).

Cambridge Systematics (1994) evaluated the effects of a large number of land-use and TCM strategies on commuting behaviour in the Los Angeles Metropolitan Area. These strategies were classified as financial incentives (which included transit subsidy, employee parking subsidy, carpool/vanpool subsidy), flexible work schedules (including flexible work hours, telecommuting programme, compressed work week programme) and assistance programmes (including employer-based matching programmes, guaranteed ride home). They concluded that if the aim is to reduce the drive-alone modal share then:

> A successful travel demand management strategy should be built around a core of financial incentives, regardless of the land use and urban design characteristics of a particular site.
>
> (Cambridge Systematics, 1994:41)

A recent assessment of the health costs of road vehicle emissions in Australia suggests that the air quality of Australian cities tends to be relatively unpolluted compared with cities in the USA and Europe, and that there is currently no evidence that fine particles are above safe levels. Air quality monitoring data based on sites around Australia demonstrate that concentrations of most air pollutants are now at ‘acceptable levels’, for which the evidence indicates no health risk. The pollutants NO_x, CO and SO_2 do not currently exceed acceptable levels and are not expected to do so in the foreseeable future. Atmospheric lead levels have been excessive, but with the reduction in use of leaded fuel, future atmospheric lead concentrations for Melbourne are estimated to remain below the existing acceptable level of 1.5 μg/m^3 as well as the proposed tighter level of 1.0 μg/m^3. Ozone at levels above the current one-hour standard of 0.12 ppm, which is the threshold for definite health risk, occurred in Sydney and Melbourne an average of three days over the period 1989–1993. In Victoria this figure rose to 17 days when ozone exceeded the 0.08 ppm one-hour standard (the range of uncertain health risk). Recent evidence in the USA, reported at a conference on the full social costs and benefits of transport, suggests that within 10 km of road, the nitrogen oxides ‘eat’ the ozone and hence eliminate a major health risk attributable to automobile use. The chemical with the greatest potential risk is PM-10, referred to as ‘road dust’ or ‘fugitive dust’. There is, however, great uncertainty as to its real health risk (McCubbin and Deluchi, 1995) and the extent to which the contribution of cars is confused by wind erosion, salt spray, power plants and other sources that produce particulates (even common house dust).

Road traffic congestion and pricing —a missing building block

> Traffic congestion is not a new phenomenon. Julius Caesar in 45 BC declared the centre of Rome off limits between 6 am and 4 pm to all vehicles except those of officials, priests, high-ranking citizens and visitors.
>
> (Dobes, 1995:1)

Road traffic congestion, more than any other single item, provides a daily reminder of all of the inefficiencies in the current transport system. Traffic congestion adds billions of dollars to road travel costs. Calls for improved public transport are a common response; calls for congestion pricing are rare (although increasingly being heard); calls for discouraging road construction as a contribution to reducing traffic congestion by encouraging the use of public transport, and assisting the move towards a more compact urban area, are on the increase. It is true that urban mobility in many cities throughout the world is being ‘strangled’ by the large amounts of time wasted in traffic jams, draining the urban economy. It has also been said that 'congestion is the sign of a healthy urban economy—what is lacking however is the presence of organised congestion'. Congestion pricing does not eliminate congestion —rather it ensures that the level is the outcome of efficient prices. The supply-side response of more roads is not an efficient or sensible 'solution' in the presence of distortionary pricing which is way out of line with the full set of effects of such pricing. This does not preclude the development of strategic road links built for reasons other than congestion mitigation, such as links to major transport hubs like seaports and airports.

We must continue to make the case for appropriate charges (as distinct from taxes) which reflect the real cost of resources consumed in travel. Congestion pricing is arguably the only policy that will make a noticeable difference in peak congestion levels in the world's most congested cities. Other policies can create real and substantial benefits (see below), but cannot do much to reduce the most severe congestion. There is so much latent demand for car travel at peak periods and during the shoulder periods that whatever capacity we can feasibly expect to build, or which can be freed up by enticing a few drivers off the road, will quickly be filled by people who are currently deterred solely by congestion. This is a well-documented empirical reality known as the 'fundamental law of traffic congestion' (Downs, 1962, 1992).

Neutze (1995), commenting on the relationship between roads and urban patterns says:

> I believe that if you correctly price roads, you will increase the extent to which the investment will cause movement of employment to the outer parts of cities. If you price them correctly, the areas where the price will be high will be in the inner urban areas because that is where the road costs and land costs are high. Land is scarce, therefore it is expensive to provide roads just as to provide buildings in those areas. That will discourage the use of roads in urban areas and that is one of the reasons why, even with optimal investment, you will have and should have high levels of congestion in places where land prices are high.

Implementation of congestion pricing involves recognition of the following issues:

- Congestion pricing would result in a net benefit to change their behaviour.
- Congestion pricing would cause some motorists to society.
- Congestion pricing is technically feasible.
- Institutional issues are complex but can be resolved.
- All income groups can come out ahead given an appropriate distribution of revenues.
- Some motorists will lose.
- Congestion pricing would reduce air pollution and save energy.
- The political feasibility of congestion pricing is uncertain.
- Evaluation of early projects is crucial.
- An incremental approach is appropriate.

Efficient pricing, however, is a necessary but not sufficient condition for a socially desirable outcome. There must be a role for other policy instruments such as physical planning. The limits to pricing as a planning tool are vividly illustrated in a UK House of Commons Transport Committee hearing in which the expert witness (Goodwin, 1995) said:

> there is the intriguing test of intuitive common sense. It is noticeable that there are some transport policies that nobody suggests should be determined by 'willingness-to-pay'. An example is the division of road space between vehicles and pedestrians. It would be possible to say that the relative width of sidewalk and carriageway should be determined by the amounts that pedestrians and vehicles are willing to contribute, or even more specifically that pedestrian-actuated traffic signals should require the insertion of a coin. The logic in one sense is similar to that of road pricing, but it does not command serious consideration. Nor does there exist (as far as I know) an underground of hard-line road prices biding their time until the moment is right to implement pedestrian charging with push-chair supplements and a penalty for elderly slow walker.

Efficient pricing signals and physical planning ordinances should be viewed as being as much potential complements as they are potential substitutes. The 'dark green' end of the environmental spectrum has tended to treat physical planning (constraints) as an alternative, at least partially, to 'failed' pricing. Pricing, however, differs from physical planning in one important aspect—it provides money. Under the new realism banner, eloquently documented by Goodwin *et al.* (1991), it is argued that the huge revenue sums raised from any change in road user prices should in part at least be allocated in a way consistent with the preferences of both society and transport users.

Allowing for both economic reasoning and political reality, the 'rule of three' is actively promoted in a number of countries. The road space initially released by congestion pricing can be used as follows: one-third reclaimed for environmental improvement, including pedestrian and non-transport uses, one-third used for extra traffic for which the reduction in congestion would be important. For example, use the revenue to favour buses, delivery trucks, emergency vehicles and disabled travellers. A final one-third would have the effect of reducing congestion delays for all remaining traffic. Maintaining this benefit will require a combination of pricing and non-pricing instruments to offset the tendency for traffic growth to eliminate the achieved speed increase.

Urban public transport

Opportunities to make better use of public transport exist in all urban areas. Some of these opportunities, however, involve a choice between bus-based and rail-based systems and a recognition that rail systems have the inherent advantage of a dedicated track, while bus systems offer greater flexibility and lower cost.

Roads are used by public transport; indeed they are arguably the most flexible form of infrastructure in accommodating mass public transport, and are capable of assisting public transport in adapting to changing levels of traffic density for relatively low cost. To be specific, buses can be interacted with roads in a low density mode (i.e. buses mixing with all other traffic); as demand for public transport increases buses can be given dedicated road space (possibly in the interim mixing with high-occupancy automobiles and taxis). As traffic densities increase even more, buses can take on the characteristic of linked vehicles (which are called trains) and operate over sections of the infrastructure under a single control unit. The provision of opportunities to expand the role of buses and bus systems (or bus-trains) is greatly enhanced where motorway-level infrastructure is in place, since it is most likely to provide the required alignment essential for public transport to accommodate changing traffic densities. There must, however, be a cultural change within the planning community in recognising the important role of road infrastructure in public transport provision and to promote such capacity specialisation in the future. Unfortunately mixing buses and cars in high-occupancy vehicle (HOV) lanes is not a marketable strategy, no matter how sensible it may be on other criteria. One would like to imagine that in a world of institutional reform centred on outcomes, modal-planning would be replaced with outcome-planning which allows motorways to become busways and, at very high levels of traffic density, railways.

This ability to accommodate flexible densities efficiently and effectively is not a trait of fixed-track rail systems, simply because the latter cannot be used for other forms of transport (e.g. cars and trucks). A common track as offered by a road is the most efficient form of infrastructure technology for accommodating changing traffic densities. Combined with efficient pricing, it will ensure that it is efficiently utilised and will not succumb to the indivisibility constraint of rail track.

To illustrate the value of bus systems with dedicated road infrastructure for the linehaul component of service, the Adelaide O-Bahn should be re-visited. Chapman (1992) undertook a post-implementation social cost-benefit analysis of the economic impact of the O-Bahn system in Adelaide. Chapman concludes by saying that

> Adelaide's O-Bahn Busway…has been one of the relatively few public transport projects that can be considered to have in any way contributed to the economic welfare of the community. It has been extremely popular with commuters, initial ridership projections having been exceeded. Some very large travel time savings have been provided, and commuters clearly appreciate the combination of limited stops, high capacity, smooth ride and congestion free travel offered by a dedicated right-of-way, and the high frequency, flexibility and through service into suburban areas offered by a conventional bus system.

Furthermore, Chapman says (p. 99) that

> in a city of Adelaide's size and urban density it has proven to be a much more effective and economic public transport service than conventional heavy or light rail systems. It was constructed at approximately half the cost of a comparable rail-based system and is one of the few public transport

systems in this era of automobile dependency that has been able to attract (and retain) passengers.

Taking the general transit patronage decline and population growth in the corridor into account, the net overall impact of the Busway is a patronage level approximately 53 per cent higher on a daily basis than otherwise would exist.

Managing the environment means paying one's transport way

A number of studies throughout Europe, Canada and the USA have consistently shown that transport users generally do not pay enough user taxes and charges to cover their external costs. In a review of five major studies, Gomez-Ibanez (1995) concludes that public transport users do not pay their way largely because the fares they pay are not sufficient to cover the capital and operating costs, not because they generate significant amounts of pollution and other social costs. For automobile users, by contrast, government capital and operating expenses constitute only about 20 per cent of total external costs. Among the external costs, parking accounts for about 20 per cent, air pollution about 20 per cent, accidents about 20 per cent and energy security about 20 per cent. These figures are approximations, but they do highlight where the externalities exist. Some of the international evidence is summarised in Table 13.5. It should be interpreted with great caution since our

Table 13.5 Estimates of external costs and subsidies for typical urban passenger trips

Cents/passenger km (cents US, 1994)	*Germany car*	*Germany train*	*Aust. car*	*Aust. bus*	*Aust. train*	*USA car*	*USA bus*	*USA train*
Government:								
Capital			1.64 (*)	10.5 (*)	13.9 (*)	0.25–1.4	0.18–4.4	8.75
Operating and maintenance						0.0–2.2	27–33	19.0
Other govt (police, fire, etc.)						0.18–1.1	0.07–0.16	0.06
Subtotal			1.64	10.5	13.9	0.4–4.7	27.3–37.6	27.8
Societal:								
Congestion			2.8	0.01	0.0	0.25–9.7	2.3	0.0
Air pollution	2.4	0.38	2.0	2.4	2.2	0.6–4.7	1–2.8	0.9–3.2
Noise pollution	0.24	0.06	0.9	0.2	1.0	0.06–0.5	0.03–0.3	0.13
Water pollution			0.1	0	0	0.06–7.5	0.06	
Solid waste						0.13	0.0	
Accidents	1.7	0.18	1.7	0.1	0	0.88–2.1	0.43–1.4	0.38
Energy			0.8	0.1	0	0.44–3.2	0.56–1.8	0.25–0.8
Parking			0.5	0	0	0.5–6.8		
Other						0.008–5.2	0.25	
Subtotal	4.34	0.62	8.8	2.81	3.2	2.9–39.7	4.63–8.91	1.66–4.51
User payments:			16 (**)	8 (**)	7 (**)			
Fares, tolls						0.0	8.8–11.9	8.8
Taxes and charges	2.1					0.4	0.0	0.0
Subtotal	2.1		16	8	7	0.44–1.3	8.8–11.9	8.8
Net subsidy	2.24		–5.56	5.31	10.1	2.6–43.1	23–34.6	21–24

Source: Gomez-Ibanez, 1995

Notes: Germany and Spain data are from the European Federation for Transport and the Environment; the USA data are from the World Resources Institute, the National Defence Council and Todd Litman (an independent consultant). Australian data is sourced from Austroads (1994)

(*) = sum of capital, operating and maintenance costs

(**) includes operations, ownership and fares knowledge of costing many of the items is both immature and often subject to huge variations caused by the context in which transport services are provided. The higher estimates in the range for the USA are for urban peak trips. The Australian evidence for the automobile suggests that there is overpayment for automobile use, which might be queried in urban congestion contexts. The message is simple— there is much scope for correcting the (under) pricing of externalities via a mix of pricing and technological change to eliminate/reduce such external impacts. There is also a need for much more research into identifying the variation and sources of variation in each of the unit cost items.

PLACING TRANSPORT IN A POLICY CONTEXT

The community should rightfully raise questions about suitable policy instruments and strategies designed to produce outcomes reflecting the goals of transport management, and which have high levels of positive impact in terms of the performance criteria that define success. There are many policy instruments available; too many to identify and describe explicitly. What is useful is a tabular summary (Table 13.6) of the degree of impact of a selected set of potential instruments in terms of the performance criteria chosen as measures of success.

These policy instruments are representative of the broad types of action worthy of consideration. The directional indication of impact reflects what we believe will be the likely degree of influence of a policy, within a range of application which we believe might realistically be introduced, given constraints such as political feasibility and widespread community support. Disagreement on the impact is almost certain, because of the complex system-wide interaction, but that is accepted; the primary objective is to encourage debate.

EMISSIONS FROM COMMERCIAL AIRCRAFT—A CHALLENGE IN INTERNATIONAL MANAGEMENT

> Somewhere on earth, every second of every day and night, 45 commercial aircraft commence their takeoff runs.… There are freeways in the sky, and crowding them are machines with un-equalled rates of converting kerosene into persistent pollutants.
>
> (Thomas, 1996)

The focus of this chapter so far has been on the management of transport in an urban context, but there are significant challenges in other parts of the transport system. For example, the growth in international airline traffic has been spectacular, particularly since the introduction of wide-bodied aircraft into commercial airline fleets since the 1970s, and as competition and productivity improvements have kept the real cost of airline travel falling. The combination of low fares and higher incomes has resulted in a world airline market in 1994 equivalent to one billion passengers a year travelling an average of 2,049 kilometres (McDonnell Douglas, 1995). In terms of the triad of growth, equity and environment objectives, it is necessary to question whether there has been undue emphasis on the growth dimension. However, this has to be put into perspective. By far the greatest amount of travel is undertaken using some surface transport mode and aircraft account for no more than 3 per cent of all CO_2 emissions, and probably less than this for NO_x emissions, according to the Intergovernmental Panel on Climate Change (a part of the World Meteorological Organization) (see Table 13.2). Nevertheless, commercial aviation raises problematic issues of management. There is scope for improving air quality at the local airport level, but problems of aircraft emissions in the upper atmosphere and issues of aircraft design require a multilateral approach to management and policy. The example of aviation illustrates some challenges of a different kind to those discussed above.

At the regional level, the issue of air quality can be traced directly to particular airports. However, large airports are located near major urban populations and their contribution to total emissions is very minor. In the case of Schiphol Airport (Amsterdam), for example, emissions from aircraft fell below 5 per cent of the region's total beyond a distance of 3 kilometres from the site. Once beyond the airport boundary, the movement of aircraft makes only a small impact on total emissions (Kinhill, 1991a). At the airport, of course, aircraft generate the largest share of all types of emissions, in taking off and

Table 13.6 Summary of potential impact of policy instruments (within the likely achievable range and likely behavioural responses over next 40 years)

Degree of impact on measure of success (blank = no or negligible effect, ? = unsure)	*Air pollution*	*Global warming*	*Energy consumed*	*Traffic congestion*
Congestion pricing (mix of charges and taxes)	+++	++	+++	+++
Increase parking charges (CBD, regional centres)	+++	++	++	+++
Parking rationing/restrictions in CBD	++	+	+	++
Toll road charges (selective major routes)	+	?	?	+
Restrictive automobile access to CBD	++	++	++	++
Reduce sales tax on new autos – skew/eliminate	+?	+?	+?	
Increase vehicle registration charges (by age, weight, fuel)	+	+	+	
Limits on company car provision and use	?	?	?	?
Limit maximum age of vehicles	+	+	+	
Carbon tax (linked to alternative fuels)	+++	+++	+++	+
Fuel excise by fuel type: increase/exemptions	+++	+++	+++	+
Tradable permits	+	+	+	
Fee-based compulsory emissions checks	+++	++	++	
Price rebates/discounts on alt. fuelled vehicles (end-use impacts only)	+++	+++	+++	?
Govt purchase and scrap high emitters	++	+	+	
Alternative fuels – electric vehicles (end-use impacts only)	+++	++	++	?
Alternative fuels – LPG , CNG, diesel oil (end-use impacts only)	++	++	++	?
Automobile engine/transmission technology improvements – conventional fuels	++++	++++	++++	
Automobile vehicle design (weight, drag)	++++	++++	++++	
Intelligent transport systems	?	?	?	++
Route guidance	?	?	?	++
Traveller information systems	?	?	?	++
Urban form and density (physical planning, dwelling mix)				
Legislative actions (zoning etc.)	+?	?	?	+
Compact cities with regional nodes	?	?	?	+
Work-related incentive strategies:				
Ride-sharing and employer incentives	+	+	+	
Telecommuting	+++	+++	+++	+++
Compressed work week (time use)	++	++	++	+++
Non-motorised options – bicycle, walk	+	+	+	
New public transport – light rail	+?	+?	+?	
New public transport – bus priority systems	++	++	++	
Public transport – park-n-ride/kiss-n-ride	?	?	?	?
Existing public transport – reduce fares				
Existing public transport – increase levels of service	+	+	+	+

landing, during taxiing, during maintenance and in refuelling. However, an airport is the hub of a considerable amount of activity, with arriving and departing passengers, cargo movements and a sizeable workforce. All of this results in additional emissions from motor vehicles. Other contributions to emissions can arise if aircraft dump fuel in flight and when fire safety training exercises are conducted.

Collectively, the airport-related emissions include hydrocarbons, nitrogen oxides, carbon monoxide, suspended particulate matter, sulphur dioxide, ozone and lead. The difficulty in devising policy responses at the regional level is to identify management measures that will result in changes in the behaviour of airlines, maintenance operations, and individuals as they travel to and from the airport. Taking the

BOX 13.1 CASE STUDY 1: URBAN PASSENGER TRANSPORT POLICIES AND IMPACT ON CO_2 CHANGES IN PERTH, WEST AUSTRALIA, 1993–2003

Using a model system of the household sector developed for Perth, West Australia, we have evaluated the impact that four policy instruments might have over the period 1993–2003 in respect of carbon dioxide emissions and total end user cost. To understand how we identify the impact of a policy instrument, consider a fuel tax increase. The imposition of an increase in tax on automobile fuel, via its impact on unit operating cost (c/v km), has an immediate and direct influence on: (1) the use of each vehicle for particular trips such as the commuter trip (i.e. mode choice, which includes both a switch to public transport and vehicle substitution from within the household's vehicle park); (2) a change in the timing of the commuter journey to reduce the increased costs associated with traffic congestion; and hence (3) a change in the overall and non-commuting use of each automobile available to a household. It also directly affects the household's choice of types of automobiles from the set of conventional fuel, electric and alternative-fuel vehicles (the last two vehicle fuel types introduced in any year, under a reasonable expectation of availability). The indirect impacts include a change in residential location via the change in modal and spatial accessibility to work opportunities, and a change in the number of vehicles in a household (given the increased operating costs). Changes in residential location may further affect the total use of each automobile, as well as the mix of urban (commuting and non-commuting) and non-urban kilometres. The adjustment in commuter travel may also affect non-commuting car use if a vehicle previously used for commuting is released for use by another non-working member of the household. Some adjustment in the loss rate of automobiles will also occur.

The adjustments in vehicle, travel and location choices at the household level translate at the aggregate level into a new set of equilibrium levels for traffic congestion (broadly measured by the ratio of travel time to distance travelled), residential densities, total kilometres of travel by automobiles and various forms of public transport, fuel consumed and greenhouse gas emissions. The table below summarises the four illustrative policy instrument changes to be evaluated *separately*. All changes commence in 1996. The cost items are calculated in constant dollars ($93), but are converted into present values at a real discount rate of 8 per cent per annum for all dollar-based costs.

Illustrative policy instrument evaluative strategies

Policy instrument	*Acronym for table*	*Units*	*Range of assessment*
Sales tax on new automobiles	Stax	% of wholesale price	10, 20, 30
Increase in fuel excise on fossil fuels	FEx	cents/litre	60, 80, 100
Change in fuel efficiency (litres/100 km)	FEff	% decrease	–5, –15, –25
Existing public transport fares	PT fares	% change	–25, –5, +20

Note: The conversion of a carbon tax from cents/kg to cents/litre is as follows: for petrol = 5, 10, 15, 20 and 25 cents/kg is equivalent to 3.1, 6.3, 9.2, 12.3 and 15.4 cents per litre in $94 assuming a retail price of 70 cents a litre. The equivalent diesel prices are 3.7, 7.3, 11.0, 14.7 and 18.5 cent per litre

The impact of various policy instruments in Perth, West Australia, on CO_2 and end user cost, 1993–2003

FEx c/litre	*FEx CO_2 % change*	*TEUCpv $ pa*	*Stax %*	*CO_2 %change*	*TEUCpv*
60	–5.46	758 m	10	3.94	–310 m
80	–14	1.37 bn	20	3.69	33 m
100	–21.6	1.82 bn	30	3.52	387 m
FEff %	*CO_2 % change*	*TEUCpv*	*PT fares*	*CO_2 % change*	*TEUCpv*
5	2.66	–100 m	25% inc.	4.1	115 m
15	.41	–310 m	25% dec.	3.45	–140 m
25	–5.45	–540 m	50% dec.	3.06	–340 m

Within the ranges evaluated, improvements in vehicle fuel efficiency have the greatest impact on reducing CO_2, as well as reducing the cost to the end user. In contrast, a fuel excise reduces CO_2 but also has the effect of increasing the cost to the end user. The sales tax policy has been evaluated at 10 per cent, a level lower than the current 20 per cent for the majority of new passenger vehicles (excluding luxury vehicles), as well as at a higher level (30 per cent). We see that CO_2 improves relative to the 'do nothing' effect in each year up to 2003, but it does not better the 1993 level. The cost to the end user decreases for the lower sales tax, but increases at a disproportionately higher rate for the larger sales tax. A public transport fare decrease of 50 per cent has a small but noticeable impact on changes in CO_2, decreasing it by 3.06 per cent, while reducing total end user cost by $340 m (in $93). There is a consequent 2.65 per cent reduction in the car's modal share for commuting and a reduction in total vkm of less than 1 per cent overall.

In evaluating these policies one has to consider some of the broader implications. For example, the fuel efficiency policy adds traffic to the road system by encouraging an increase in vehicle kilometres travelled. While this may be attractive for individuals with latent travel demand, it nevertheless may be undesirable from an environmental point of view. Possibly a congestion charge might accompany this policy.

passengers first, it needs to be understood that airport planners are concerned with issues of ground transport access. A consequence of planning major airports without provision for public transport systems is a reliance on motor vehicles. The contribution of this vehicular traffic to total airport-related emissions of carbon monoxide and NO_x is non-trivial. Also, there is scope for the airport operator and the urban planner to influence this component of the emissions. Policy responses range from the construction of heavy or light rail access, improved bus services, increases in parking charges for motor vehicles, and other traffic calming measures in the region of the airport.

On the airport itself, it is possible to take actions that can achieve results. For instance, ancillary vehicles that operate on the airport itself can be converted to alternative sources of energy with lower emissions. Also, power can be supplied by terminal operators to parked aircraft to maintain air conditioning and services. Aircraft engines are designed to be at their most efficient when cruising; taxiing is the least efficient phase of a flight and it goes without saying that airport operators can play a role in the reduction of hydrocarbons, carbon monoxide and nitrogen oxides emitted on their premises. For example, it has been found that the construction of a parallel runway (rather than having runways crossing) can reduce taxiing and queuing times.

It is during the take-off and climb-out that an aircraft generates most of its nitrogen oxides. Airport managers might take the view that there is little scope for them to exert influence on the airlines, or for the airlines themselves even, to reduce this impact. The take-off phase leads us to consider a more difficult question about emissions from aircraft en route. The evidence is that as aircraft climb through the troposphere, that part of the atmosphere that extends up to 9 kilometres from the earth's surface, their emissions contribute to ozone production. Turboprop aircraft stay in this region, but jet aircraft tend to cruise in the tropopause, the next layer of atmosphere up to 15 kilometres above the surface (depending on the geography and season). Some jet aircraft do reach into the stratosphere, but this is the operating domain of supersonic aircraft.

More research is needed into the effects of emissions, but it does appear that pollutants stay longer in the upper atmosphere. Nitrogen oxides released into the upper atmosphere are said to have 30 times more impact on global warming than similar emissions at ground level. In the upper atmosphere, the emissions result in the depletion of ozone; one estimate is that aviation contributes more than 10 per cent of global greenhouse gas warming (Thomas, 1996).

Airframe manufacturers and engine producers are conscious of the fragile financial state of the world's airlines, and ways are being tested of reducing fuel consumption and, consequently, emissions. For example, aerodynamics can be improved through advanced wing design, through weight reduction, and through attention to materials used in the 'skin' of aircraft. Newer jet turbine engines, according to Airbus Industries, will halve the nitrogen oxide emissions of an aircraft by the year 2010. Though there are gains to be made in engine technology, a basic problem has been that an engine designed for maximum fuel efficiency also produces high amounts of NO_x. Another approach is to connect jet turbines to propellers to achieve greater fuel efficiency at lower altitudes. On-board sensors and instrument landing systems remove some of the concerns about flying at lower altitudes in prop-jet aircraft in poor weather conditions. Airlines can choose not to fly their jets in the stratosphere, but to do this they have to sacrifice some fuel efficiency. A more immediate benefit accrues when larger aircraft replace smaller aircraft, and when airlines operate with higher load factors. Also, there is scope still to reduce the distances flown by commercial aircraft, particularly on international routes.

The International Civil Aviation Organization (ICAO), recognising the need for the airline industry to take action, has issued limits for emissions of smoke, unburned hydrocarbons, carbon dioxide and nitrogen oxides. Some countries have acted on these guidelines, but Sweden, Denmark and Norway have imposed a carbon tax, based on the amount of fuel consumed, on airlines flying within their territories. It is possible that other governments will adopt similar measures in the future. This gives airlines an added incentive to reduce fuel consumption, but it is possible that the imposition of a tax on fuel usage could result in an increase in nitrogen oxides, particularly in the upper atmosphere where they are more problematic.

Too little, however, is known about the extent of the problems and about the efficacy of some of the 'solutions'. ICAO's 15-member Committee on Aviation Environmental Protection was reported to have decided in its December 1995 meeting that it should recommend reductions in nitrogen oxide emissions by 16 per cent (*Airline Business,* March 1996). This would be achieved by requiring modifications in new engines from 2000 and for modifications to engines in current use by 2008. Meanwhile, there are numerous studies in progress to improve understanding of the levels of emissions and the effects of those emissions. The Mozaic project, for example, involves five Airbus A340s carrying out measurements during flights to study concentrations of ozone and water vapour. In another study a Swissair Boeing 747 is measuring nitrogen oxides in the troposphere.

In dealing with emissions from commercial aircraft, we have seen that there are initiatives that can be taken by the managers of local airports, but the design of aircraft and their operation in the upper atmosphere requires a response at a different level. Individual airlines, and for that matter individual nations, can achieve considerable success simply through better management of the air transport system. Higher load factors, improved navigation to achieve shortest path flights, and the use of the most efficient aircraft for the transport task can all make significant contributions to the reduction of emissions. Airframe manufacturers and engine manufacturers are also responding to these challenges, but multilateral agreement on appropriate environmental management of commercial aviation is difficult to reach. The international airline industry has been experiencing very difficult financial times and those close to the industry are reluctant to impose requirements that will add to costs. Debate within ICAO is taking place, but one industry analyst was recently led to make the statement that:

> Airlines can argue quite rightly that they have played a large role in reducing aircraft noise and emissions and contributing to a cleaner, greener earth. But their reluctance to take the next step could result in harsh measures being imposed by those with less knowledge of the industry than an organisation like ICAO.
>
> (*Airline Business,* March 1996:29)

OIL POLLUTION IN THE STRAITS OF MALACCA—THE IMPACT OF SHIPPING

The last section of this chapter deals with the management of the marine environment and focuses on the difficulty of management in the absence of universally enforceable regulatory regimes. In effect, of course, this means that policies and strategies formulated to meet equity goals have little chance of success.

In the early 1990s 1.4 billion tons of oil were moved globally by more than 3,000 tankers over an average distance of 4,700 nautical miles. The threat of spills has been of concern for some time but major tanker accidents such as the grounding of the *Torrey Canyon* in 1967 and the *Argo Merchant* in 1976, which broke up, releasing 23,000 tonnes of oil into the sea, focused global attention on the problem of how to preserve the marine environment. More recently, three shipping collisions in the Straits of Malacca—the *Royal Pacific* and *Tefu 51,* the *Nagasaki Spirit* and *Ocean Blessing* and the *Maersk Navigator* and *Sanko Harbour*—resulting in major oil spills and the loss of life focused the problem on a regional scale, and led to a call by the littoral states—Malaysia, Indonesia and Singapore—for more stringent marine pollution control measures.

All spills are serious, but special problems are created when they occur in an international and relatively narrow waterway such as the Straits of Malacca. These Straits separate the Indonesian island of Sumatra from Peninsula Malaysia and Singapore and form the main sea link between the Indian Ocean, South China Sea and the Pacific Ocean (Figure 13.2). It is a funnel-shaped waterway varying in width from 3 to 300 miles. The depth of the Strait is highly variable making it dangerous for navigation—in the Philip Channel alone there are 37 areas where the depth is less than 23 m—the minimum draught for a 250,000 dwt tanker. Further hazards include rocks, dangerous reefs and cross currents as well as wrecks in reported approximate positions; a seabed of hard rock and granite makes groundings very dangerous (Mochtar, 1994:12).

The Straits of Malacca are among the world's most heavily used waterways—up to four hundred vessels transit each day, 40 per cent of which are predominantly foreign-owned oil carriers. The dangers from pollution have increased with economic development in East Asia which requires increasing volumes of crude oil. Japan's oil consumption alone increased by 20 per cent annually after the late 1960s and its annual crude oil imports by the early 1980s had reached 600 million tons—the bulk of this passing through the Straits. Moreover, ships passing through the Straits are now larger and there are more of them; the number of accidents has increased with congestion, as has the severity of spills with increased vessel size. During the period 1977–1993 there were 71 shipping casualties in the Straits, 17 per cent involving tankers. This is a relatively small number of casualties, given the high traffic densities, but it only takes one or two major tanker accidents to cause very serious damage in an ecologically vulnerable area (Gold, 1994:5).

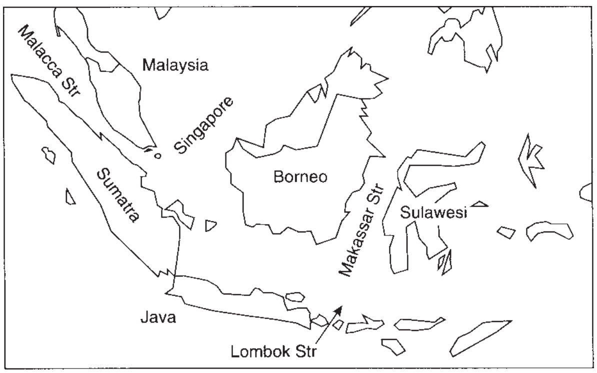


Figure 13.2 Straits of Malacca, Lombok and Makassar

The risk of pollution in the Straits is not new, but has always existed in a waterway whose physical characteristics make navigation difficult. Concern over this risk intensified over the last decade or so for a number of reasons. First, the increased use of the Strait, which is the major shipping link between the Middle East and East Asia, and the increased size of the vessels with the emergence of the supertanker. Second, there are problems associated with an ageing fleet— age of a vessel has been identified as a major factor in ship loss. As a ship ages a general deterioration in condition occurs, stresses build up which attack the structural soundness of the vessel—a particularly significant factor in large bulk carriers. Serious maritime casualties involving older bulk carriers have revealed the vulnerability of such ships to breaking up, rapid sinking when ruptured or to be engulfed in flames (Raja Malik and Bin Raja, 1993:6). The third reason is of an economic nature and concerns the growing practice of flagging out. The high cost of shipping has led to the transfer of ships from traditional flag states to Flag of Convenience (FOC) and second registries. FOC states offer tax and investment incentives but frequently have a less stringent approach to ship inspections and many FOC vessels fail to comply with agreed international standards.

A plethora of international conventions exists pertaining to laws of the sea and marine pollution, although many of them are either outdated or difficult to implement. The traditional formulation of the International Law of the Sea stipulated that ships had freedom to sail the oceans. The 1958 Convention on the Territorial Sea and Contiguous Zone formalised this tradition which meant that a vessel could proceed with 'innocent passage' so long as it did not threaten the peace, order and safety of coastal states (Mochtar, 1994). Although this tradition remains the basis of the law of the sea at present, it is questionable whether it is appropriate. The marine environment is now a valuable source of wealth in terms of food and of raw materials, not only as a transport lane, and it can be validly argued that no passage that leads to its destruction can be regarded as 'innocent'.

International regulatory regimes exist that deal specifically with the prevention of pollution and that clearly define the responsibilities of each party. Under the 1982 UN Convention on the Law of the Sea (UNCLOS), for example, 'vessels transiting international straits must comply with generally accepted international regulations, procedures and practices for safety at sea and for the prevention of marine pollution' (Gold, 1994:12). This rule was designed to protect coastal states without imposing unreasonable burdens on passing vessels, but the question of how compliance can be enforced is anything but straightforward. UNCLOS does provide enforcement powers to protect the marine environment but the procedure is complex and, in the context of transiting vessels, impracticable; a claim can be brought through diplomatic channels for a breach of treaty obligations, but its success depends entirely on the willingness of the flag state to enforce it (Gold, 1994:5). Coastal or boarder states do not, under UNCLOS, have the right to suspend passage.

The enforcement of international conventions in the Straits of Malacca brings with it a particular set of problems and difficulties. It entails the dispute over whether the Straits are, in fact, 'international'. Both Malaysia and Indonesia have argued that the Straits lie within their respective territorial waters and are, therefore, national. This has been rejected by users and the issue has never been resolved—the Straits remain formally 'international' as defined by UNCLOS (Gold, 1994:3).

Difficulties exist not only because conventions are not uniformly adhered to, or because agreement on definitions cannot be reached: the compensation regime as it stands is problematic and this can impede restoration. When spills do occur, access to immediate assistance is vital in minimising the environmental damage, but delays in receiving compensation payments are common for a number of reasons. First, a polluter must be identified. In the case of major disasters where ships break up and major spills occur, this is not a problem. But the majority of spills (84 per cent) are less then 7 tonnes and may not be readily detected (White, 1994:2). In the absence of an offender, against whom can a claim be lodged? Second, there are a number of categories for compensation and this, in itself, may be contentious. Among the more straightforward reasons for compensation claims are those for clean-up measures or property damage directly resulting from a spill, such as contamination of fishing gear, boats and pleasure crafts. It also includes economic loss associated with property damage, such as fishermen being prevented from carrying out their occupation because of damage to boats or oil on the surface of the sea. One of the most contentious areas in compensation claims, however, relates to environmental damage. The marine environment has a value to society beyond that which it confers on those who depend upon it for their livelihood. Controversy arises when compensation is sought for damage to natural resources which are neither commercially exploited nor used in an economic sense.

Solutions at a regional level may also be difficult to enforce, although instances of unilateral imposition of transit restrictions do exist. Following a collision between two tankers in the Strait of Messina in 1985, the Italian government issued a decree that merchant ships must navigate on the right side of the traffic separation line; pilotage is compulsory for merchant ships over 15,000 dwt, or for ships over 6,000 dwt if they carry oil or other substances that are harmful for the marine environment; and navigation in the Strait is prohibited for all oil-carrying tankers exceeding 50,000 dwt (Scovazzi, 1994:21).

The imposition of the 'Grandfather Clause' and the restriction on aged vessels are further policies aimed at the prevention of spills. The age of a ship has been identified as a major factor in ship losses and the imposition of the 'Grandfather Clause' has been considered an effective means of minimising spills in territorial waters elsewhere. The USA, for

BOX 13.2 MANAGEMENT STRATEGIES FOR THE STRAITS OF MALACCA

Implementing effective management strategies in the Straits of Malacca has not been without its problems. As it is an international waterway, passage through it cannot be prevented, nor can flag restrictions be imposed (Scovazzi, 1994). None of the attempts to implement any of numerous other strategies has been successful. The imposition of the Grandfather Clause, for example, while enhancing safety and reducing the risk of pollution would, if implemented, have serious repercussions on domestic transport, particularly on coastal trades. Both Indonesia and Malaysia are oil exporting countries and this restriction would also affect local vessels. And the legal question remains: can the Grandfather Clause be enforced legally by the littoral states? Where these strategies have been successfully implemented—in the USA and the European Community—the clause has been imposed in territorial waters and in ports, both instances where the coastal and port states can legally restrict passage.

Other strategies have been proposed but rejected, either because of user opposition or because agreement could not be reached among the littoral states. A toll was considered, for example, to be levied on all ships passing through the Straits. This would provide funds which could be used to improve the safety of navigation and to combat pollution. However, the proposal was rejected by users on the grounds that since the Straits are international waterways, how could a charge be legitimately levied, and by whom?

Proposals to restrict tanker size, as occurred in the Strait of Messina, have also been difficult to implement. While some Japanese users agreed to abide by a 200,000 dwt restriction, limitations based on size or dead weight were rejected by the majority of users for a number of reasons. First, it was considered discriminatory and unfair to westbound tankers, generally returning in ballast, or to tankers travelling to the east not fully loaded. Second, modern methods of design indicate that tonnage measurement may be less relevant to size and for limitation purposes an Underwater Clearance (UWC) would be more appropriate. As a result a UWC of 3.5 meters has been recommended and adopted as an alternative criterion for size limitation.

Re-routing strategies through the Straits of Lombok and Makassar have also been considered. These have economic implications, however, and on that basis have been opposed by users, as a detour through the Straits of Lombok and Makassar adds approximately 1,600 nautical miles and an additional two days to the voyage, at an added cost. In addition, Singapore opposed the re-routing strategy from the outset because it would by-pass the port of Singapore altogether. It can be argued that re-routing is not in any event a long-term solution, merely a displacement of the problem—a shifting of potential polluters from the Straits of Malacca to the Straits of Lombok and Makassar.

Compulsory pilotage through the narrower and more difficult passages has not been considered an appropriate solution and was also rejected on the basis that over 90 per cent of marine casualties were due to groundings caused by poor seamanship, rather than by the physical characteristics of the Straits (Mochtar, 1994:30).

The Traffic Separation Scheme (TSS), a routing scheme introduced in the early 1980s, has proved an effective means of controlling maritime casualties in the Straits. Raja Malik (1993) points out that since its introduction there have been no recorded shipping casualties within the Scheme. Unfortunately, the Scheme covers only the northern part of the Straits and accidents have continued to occur in those areas outside the Scheme. It is now proposed to extend the Scheme, but this will require concurrence not only among the littoral states but also among users. Undoubtedly, the intended changes will be scrutinised, especially by East Asian countries, to ensure that any impediment of passage is strictly in accordance with international law. The extension of the TSS will mean that ship masters will no longer be free to exercise the option of drawing their own courses. This places the additional onus on the littoral states of ensuring safe navigation. They will have the responsibility of ensuring that ships are able to position themselves with reasonable ease, that the route is safe from navigational hazards, unobstructed and with ample manoeuvring width. This calls for additional navigational aids, verification of chartered depths, hydrographic surveys, marking and removal of wrecks, and the establishment of monitoring and communication capability to issue navigational warnings. While this will undoubtedly make the Straits safer, it is costly and the question arises of who will pay for it. Again, safety and equity are mutually exclusive conditions.

Clearly, effective solutions require regional consensus. But it is not a regional or national problem only. The management and protection of the marine environment is of global concern and therefore requires internationally implementable solutions.

example, has effectively banned tankers exceeding 15 years of age from operating in its waters, and the European Parliament passed a resolution urging authorities in the Union to disbar oil tankers more than 15 years old from using member state ports. Fully laden oil tankers have been discouraged from using the Strait of Bonifacio altogether (Scovazzi, 1994:22).

CONCLUDING COMMENTS

Transport in the modern era has provided the technological means to facilitate the movement of passengers and goods. While effective transport systems have become indispensable, they have frequently been developed in isolation of environmental protection to the extent that transport has been cited as a major contributor to the ills of twentieth-century society.

Central to the issue is the growing concern about environmental degradation in the form of air, noise and marine pollution, traffic congestion and global warming.

Although governments are developing policies to combat some of these problems, in many instances it is a matter of 'too little too late'. In addition, many issues transcend national boundaries and so require an international, rather than a national, approach. Effective environmental management in the future depends on adequate resources and an integrated global approach if attempts to reverse some of the ills are to succeed. In an increasingly global environment international strategies will require equity across nations; where this responsibility is not shared, outcomes are bound to remain sub-optimal.

REFERENCES

Airline Business, March 1996.

Apogee Research Inc. (1994) *Cost and Effectiveness of Transportation Control Measures (TCM's): A Review and Analysis of the Literature,* Bethesda, Md., USA: Apogee Research Inc.

Australasian Transport News, 1994.

Austroads (1994) *Cost of Personal Travel,* Sydney, Australia: Austroads.

Bureau of Transport and Communication Economics (BTCE) (1995) *Greenhouse Gas Emissions from Australian Transport,* Report 88, Canberra, Australia: Australian Government Publishing Service.

Cambridge Systematics Inc. (1994) *The Effects of Land Use and Travel Demand Management Strategies on Commuting Behaviour,* Washington DC, USA: Federal Highway Administration, US Department of Transportation.

Chapman, P. (1992) 'The Adelaide O'Bahn: how good in practice?', *Papers of the Australasian Transport Research Forum,* Part 1: 83–100.

Davis, S.C. (1995) *Transportation Energy Data Book: Edition 15,* Tennessee, USA: Oak Ridge National Laboratory.

Dobes, L. (1995) *Greenhouse Gas Emissions in Australian Transport in 1900 and 2000,* Bureau of Transport and Communication Economics Occasional Paper 110, Canberra, Australia: Department of Transport.

Downs, A. (1962) 'The law of peak-hour expressway congestion', *Traffic Quarterly* 16: 393–409.

Downs, A. (1992) *Stuck in Traffic: Coping with Peak-Hour Traffic Congestion,* Washington DC, USA: The Brookings Institution.

Federal Airports Corporation (1993) *Monthly Air Quality Monitoring Report,* Federal Airports Corporation, Sydney, Australia, December.

Gold, E. (1994) 'Transit services in international straits: towards shared responsibilities?', paper presented at the International Conference on the Strait of Malacca: Meeting the Challenges of the 21st Century, Kuala Lumpur, Malaysia, 14–15 June.

Gomez-Ibanez, J. (1995) 'Pitfalls in estimating whether transport users pay their way', paper presented at the conference on Measuring the Full Social Costs and Benefits of Transportation, sponsored by the Bureau of Transportation Statistics, and held at the Beckman Centre, University of California at Irvine, USA, 6–8 July.

Goodwin, P.B. (1995) *The Case For and Against Urban Road Pricing,* London, UK: House of Commons Transport Committee, Third Report, Volume II, Minutes of Evidence, HMSO, pp. 49–64.

Goodwin, P.B., Hallett, S., Kenny, F. and Stokes, G. (1991) *Transport: The New Realism,* Report to Rees Jeffrey Road Fund, Transport Studies Unit, University of Oxford, Oxford, UK.

Howarth, D. and Moss, J. (1995) 'A planning strategy for Sydney Airport', in *Airport Engineering Innovation, Best Practice and the Environment,* Airports '95 Conference, Institution of Engineers Australia, Sydney, 9–11 October, pp. 65–69.

Kinhill (1991a) *Proposed Third Runway Sydney (Kingsford Smith) Airport. Draft Environmental Impact Statement,* Report prepared for the Federal Airports Corporation, Sydney, Australia.

Kinhill (1991b) *Proposed Third Runway Sydney (Kingsford Smith) Airport. Supplement to the Draft Environmental Impact Statement,* Volume 1, Report prepared for the Federal Airports Corporation, Sydney, Australia.

McCubbin, D.R. and Deluchi, M.A. (1995) 'Health effects of motor vehicle air pollution', paper presented at the conference on Measuring the Full Social Costs and Benefits of Transportation, sponsored by the Bureau of Transportation Statistics, and held at the Beckman Centre, University of California at Irvine, USA, 6–8 July.

McDonnell Douglas (1995) *World Economic and Traffic Outlook,* California, USA: Economic Research Department, Douglas Aircraft Company.

Marchetti, C. (1992) *Anthropological Invariants in Travel Behaviour,* Laxenburg, Austria: International Institute of Applied Systems Analysis.

Mitchell McCotter & Associates Pty Ltd and Peter W. Stephenson & Associates Pty Ltd (1994) *Sydney*

(Kingsford Smith) Airport. Draft Air Quality Management Plan, Report prepared for the Federal Airports Corporation, Sydney, Australia.

Mochtar Kusuma-Atmadja (1994) 'Overcoming navigational hazards and marine pollution in the Straits of Malacca and Singapore', paper presented at the International Conference on the Strait of Malacca: Meeting the Challenges of the 21st Century, Kuala Lumpur, Malaysia, 14–15 June.

Neutze, M. (1995) 'Funding urban infrastructure through private developers', *Urban Policy and Research* 13, 1: 20–28.

OECD (1991) *The State of the Environment,* Paris, France: OECD.

Raja Malik Saripulazan and Bin Raja Kamaruzaman (1993) 'Safety of the Strait and other outstanding issues', paper presented at the National Conference on the Strait of Malacca, Kuala Lumpur, Malaysia, 11 November.

Roy, J., Brotchie, J. and Marquez, L. (1995) 'A study on urban residential density vs. transport energy consumption', paper prepared for the EASTS conference, Manila, The Philippines, September.

Scovazzi, T. (1994) 'Management regimes and responsibility for international straits', paper presented at the International Conference on the Strait of Malacca: Meeting the Challenges of the 21st Century, Kuala Lumpur, Malaysia, 14–15 June.

Small, K.A. (1991) 'Transportation and the environment', in R.Thord (ed.) *The future of Transportation and Communication,* Borlange, Sweden: Swedish National Road Administration, pp. 217–230.

Small, K. and Song, S. (1992) 'Wasteful commuting: a reso*lution', Journal of Political Economy* 100, 4: 888–989.

Thomas, W. (1996) 'Stratospheric jet traffic: a sky-high problem', *Ecodecision,* Winter: 61–63.

Wachs, M., Taylor, B.D., Levine, N. and Ong, P. (1993) 'The changing commute: a case-study of the jobs-housing relationship over time', *Urban Studies* 30, 10: 1711–1729.

Walsh, M. (1993) 'Highway vehicle activity trends and their implications for global warming: the United States in an international context', in D.L.Greene and D.J.Santini (eds) *Transportation and Climate Change,* Washington DC, USA: American Council for an Energy-Efficient Economy, pp. 1–50.

White, I.C. (1994) 'Liability and compensation for oil pollution from tankers', paper presented at the International Conference on the Strait of Malacca: Meeting the Challenges of the 21st Century, Kuala Lumpur, Malaysia, 14–15 June.

SUGGESTED READING

Dittmar, H. (1995) 'A broader context for transportation planning: not just an end in itself', *Journal of the American Planning Association* 61, 1: 7–13.

Lave, C.A. (1990) 'Things won't get a lot worse: the future of U.S. traffic congestion', paper presented at the 1990 Annual Meeting of the US Transportation Research Board, Washington DC.

McLoughlin, B. (1991) 'Urban consolidation and urban dispersal: a question of density', *Urban Policy and Research* 9, 3: 148–156.

Mensah, T. (1994) 'Flag, port and coastal state jurisdiction: a case for a joint approach', paper presented at the International Conference on the Strait of Malacca: Meeting the Challenges of the 21st Century, Kuala Lumpur, Malaysia, 14–15 June.

Newman, P. (1995) 'Support for the construction and maintenance of freeways has decreased', *World Transport Policy and Practice* 1, 1: 12–19.

Raper, D.W. (1990) *Airport and Air Quality. A Review Paper,* Manchester, UK: Manchester Airport plc.

Small, K.A. and Kazimi, C. (1995) 'On the costs of air pollution from motor vehicles', *Journal of Transport Economics and Policy* XXIX, 1: 7–32.

Wachs, M. (1994) 'Will congestion pricing ever be adopted?', *Access* 4: 15–19.

SELF-ASSESSMENT QUESTIONS

1 Improving urban public transport through fare decreases and service level increases will have a significant positive influence on reductions in carbon dioxide:
   (a) Definitely.
   (b) Impact will be minimal no matter how big a change.
   (c) Depends on the dominating influence on the automobile.
   (d) In general, it is not a preferred policy in isolation.

2 Evidence to date suggests that improvements in transport technology will have a greater impact on cleaning up the environment than will strategies to modify travel behaviour
   (a) Definitely.
   (b) Not true at all.
   (c) Depends on the strength of government policy.
   (d) True in the absence of efficient pricing of transport use.

3 Congestion pricing must be introduced if there is to be any noticeable impact of reducing greenhouse gas emissions and improving air quality:
(a) Definitely.
(b) Not true at all.
(c) Only effective if accompanied by physical planning codes.
(d) Will only be acceptable and hence successful if the revenue is allocated equitably.

4 The typical major airport can effect improvements to local air quality only through attention to the layout of its runways and taxiways, the auxiliary power it provides to aircraft while they are on the ground, and better refuelling and fuel storage practices.
(a). Definitely.
(b) Not true at all.
(c) Emissions from ancillary vehicles required to operate the airport can be converted to alternative fuels to reduce remaining emission problems.
(d) Ancillary vehicles can be converted to alternative fuels, but additional measures can be taken to reduce emissions from vehicles required for the ground access of passengers and cargo.

5 Improved engine technology has the potential to reduce emissions of nitrogen oxides from aircraft.
(a) Definitely.
(b) Not true at all.
(c) Only possible if fuel efficiency is sacrificed.

6 Imposition of a carbon tax on airlines flying within a country's airspace will lead to reductions in fuel use and hence reduce the problems of harmful emissions.
(a) Definitely.
(b) Not true at all.
(c) Only effective if airlines switch from jet aircraft to turboprop aircraft.
(d) Only effective if airlines do not move more of their flight operations into the upper atmosphere.

7 The means of resolving maritime pollution problems globally is by the application of UN Conventions.
(a) Definitely.
(b) Sometimes.
(c) Never.

8 Rerouting strategies must be introduced if the present pollution problems are to be eliminated
(a) Definitely.
(b) Not at all.

14

SUSTAINABLE TOURISM

Atanaska Nikolova and Luc Hens

SUMMARY

Tourism is directly dependent on the quality of the natural and cultural environment. Unfortunately, no form of tourism exists that does not cause environmental stress. Tourism can be a threat to the environment; the challenge is to find a way towards sustainable tourism development that harmonises' economic benefits with the protection of natural diversity and the cultural identity of the destination areas.

The objectives of this chapter are to examine critically the multidimensional impact of tourism; to introduce the definition, goals and principles of sustainable tourism development; and to focus on the development policy and impact control measures appropriate for achieving environmental, socio-cultural and economic sustainable forms of tourism development.

The gap between the theoretical approach to sustainable tourism development and the character of current tourism development is stressed, and the new challenge of research and planning in tourism is argued and illustrated.

ACADEMIC OBJECTIVES

This chapter is intended for students interested in the interface between tourism, environmental protection and sustainable development. After reading this chapter, they should have a better understanding of the real costs and benefits of current tourism developments; be able to identify sustainable forms of tourism among the great variety of existing and proposed types of tourism; and be familiar with environmentally friendly policies and management instruments for putting sustainable tourism into action.

BACKGROUND

Tourism has become one of the world's leading economic sectors. The disadvantages of non-sustainable tourism, such as ecological disruption, pressure on resources, air, water and soil pollution, visitor pressure, disturbance and damage to the ways of life and social structure, all lead to a growing public awareness and the need to establish a new type of tourism development that harmonises socio-economic benefits with the protection of the environment.

During the past ten years alternative forms of tourism have substantially been promoted. There is no precise definition of alternative tourism; often it is simply defined as those forms of tourism that are not mass tourism. This may entail a wide variety of activities, for example, green, soft, responsible tourism and eco-tourism.

Work by Krippendorf (1975) and Jungk (1980) has been particularly important in arousing the sensibility of the public to the ecological damage caused by tourism. They developed the concept of 'hard' and 'soft' tourism (Table 14.1), and popularised the importance of an adequate regional tourism development strategy.

Eco-tourism is the alternative type of tourism that has been given most attention. A frequently used definition is as follows:

> tourism to relatively undisturbed natural areas with the specific objective of studying, admiring, and enjoying the scenery, flora and fauna, as well as any existing cultural manifestations.
>
> (Boo, 1991)

Table 14.1 Strategies of tourism development: selected elements

Hard tourism	*Soft tourism*
Development without planning	First planning, then development
Each community plans for itself	Planning for larger areas
Widespread and scattered construction	Concentrated construction to save land
Construction for identified need	Determining limits of final extension
Tourism in the hands of non-local promoters	Local population participates and makes decisions
Develop all facilities to their maximum capacity	Develop all facilities for average capacity

Source: Becker, 1995

More recently the eco-tourism debate has focused on the degree of human responsibility towards the environment. The passive view is that eco-tourism is tourism that minimises disturbance of the environment; the active view argues that it should positively contribute to sustainable management of the environment.

At present the concept of sustainable tourism is still relatively undeveloped. But it is important to understand that sustainable tourism means far more than encouraging low impact tourist activities, such as walking or cycling. It is an approach that encompasses all elements of the industry, from global hotel corporations to tour operators to the very small accommodation establishment. Sustainable tourism involves non-consumptive use of natural resources, protects the culture and well-being of local communities and promotes equity within and between generations.

The aim of this chapter is to analyse critically the multidimensional (environmental, economic, socio-cultural) impact of tourism; to introduce the definition, goals and principles of sustainable tourism development; and to discuss possible ways of achieving environmental, socio-cultural and economic sustainable forms of tourism development.

IMPACTS OF TOURISM DEVELOPMENT

Tourism

> Tourism includes a heterogeneous set of activities and services related to the temporary moving of individuals outside their usual place of residence, for purposes of entertainment, rest, culture, health care, and generally for reasons other than income-earning activities.
>
> (WTO, 1994)

Travel and tourism is the world's largest industry, transporting more than 528 million people internationally and generating US$ 322,000 million in receipts in 1994 (WTO, 1995). It is a major economic force, generating in 1995 an estimated US$ 3.4 trillion in gross output, creating employment for 211.7 million people, producing 10.9 per cent of world gross domestic product (GDP), investing US$ 639.9 billion in new facilities and equipment, and contributing more than US$ 637 billion to global tax revenue.

According to the projections of the World Tourism Organisation (WTO) (1992) a strong and sustained growth in tourism is expected. There are a number of reasons to predict such a sustained expansion:

- Tourism has become part of the accepted lifestyle and aspirations of people.
- Rising disposable incomes.
- Increased levels of urbanisation, leading to de mands for escape to tourist destinations.
- Increased mobility.
- Higher level of education and knowledge.
- Increased leisure time, particularly related to in creased life expectancy and earlier retirements.
- The global media network stimulates the desire to travel.

Complementary aspects of tourism include the fact that the market is dominated by very few big transnational companies (TNCs). Their dominance is most pronounced in developing countries. In OECD countries the market is segmented by the presence of many small and medium-sized enterprises. There is an outspoken relationship between both positive and negative impacts of tourism on the environment and on the evolution of the market.

The overall growth in tourism is one cause of the increased pressure on the environment.

Relationship between tourism and environment

In its broader definition, environment encompasses all of people's natural and cultural surroundings. The natural environment is what exists in nature—climate, weather, air, the land and its soil and topography, geology, water features, flora, fauna and ecological systems. The built environment consists of man-made physical features, principally buildings of any type and other structures, infrastructure development, archaeological and historical sites.

The close relationship between tourism and environment is recognised internationally. Three aspects of the tourism-environment relationship are fundamental:

- Many features of the physical environment are an attraction for tourists.
- Tourist facilities and infrastructure constitute one aspect of the built environment.
- Tourism development and tourist use of an area generate environmental impacts.

Impacts of tourism on the environment are often more intensive because tourism tends to develop preferentially in fragile and vulnerable environments, such as on small islands, in coastal, marine, mountainous and alpine areas and at archaeological and historic sites, as these places offer important resources or attractions for tourists.

The type and extent of environmental impacts as well as of socio-economic impacts also relate closely to the type and intensity of the tourism development undertaken.

Environmental impacts

Tourism can generate either positive or negative environmental impacts, depending on how its development is planned and managed.

Positive impacts

The analysis of existing studies shows that tourism has great potential to maintain and improve the environment in various ways.

Conservation of natural areas: Tourism can help to justify and pay for the conservation of important natural areas and the development of national or regional parks and reserves. Preservation of native wildlife for tourist viewing has proven its economic success, especially with the present significant greening of the market.

Conservation of archaeological and historic sites and architectural character: Tourism provides incentives and helps the conservation of archaeological and historic sites that might otherwise deteriorate or disappear. In South and South-East Asia, much of the archaeological and historic preservation taking place can be economically justified (in these lower income countries) because it provides attractions for tourists.

Improvement of environmental quality: Tourism can provide incentives to 'clean up' the environment through control of air, water and noise pollution, littering and other environmental problems, and to improve environmental aesthetics through sea- and landscaping programmes, building design, sign controls and improved buildings maintenance.

Enhancement of the environment: Development of well-designed tourist facilities may enhance the rural and urban landscape.

Improvement of infrastructure: The local infrastructure of airports, roads, water, sewage and solid waste disposal systems and telecommunication can be improved through the development of tourism.

Increasing environmental awareness: Observation of the tourist's interest in nature and realisation of the importance of conservation to the economic success of tourism can encourage local awareness and the interest of the tourist in the environment.

Negative impacts

Various negative or undesirable impacts can be generated by tourism. It is unlikely that all the listed impacts would take place in the same area because types of impact often depend on the kind of tourism development and the specific environmental characteristics of the tourism area. The scale of tourism development in relation to the carrying capacity of the environment influences the extent of the impacts.

Water consumption: Tourism is recognised as an important resource consuming industry. Tourists may consume disproportionate quantities of local resources. For example, the average tourist in Barbados consumes eight times the amount of water the average resident does (USOTA, 1993).

Water pollution: If no sewage disposal system has been installed for hotels, resorts and other tourist facilities, there may be pollution of groundwater or discharge into a nearby river, lake or coastal area. It is common for many coastal towns of the province of Alicante (Spain) to double their population during the summer months. Often raw, untreated wastewater pours into the sea as a result of accidental spills, overloading or blocked sewer lines (Zoffmann *et al.*, 1989).

Waste disposal problems: Littering of debris and packaging materials is a common problem in tourism areas. Improper disposal of solid waste from hotels, restaurants and resorts can generate significant environmental, health and aesthetic problems (for an example see Box 14.1).

Air pollution: Tourism is generally considered a 'clean industry', but air pollution from tourism development can result from excessive use of vehicles used by and for tourists, especially in those areas that are accessible only or mainly by road. Exhaust from private cars and coaches damages trees and wildlife. In Switzerland 70 per cent of domestic and foreign tourists arrive at their destination by car. As a result, significant increases of nitrogen oxides, hydrocarbons, carbon monoxide, lead, etc. are recorded (BUWAL, 1992). Long-distance travel linked with tourism in exotic countries has been favoured by transnational companies (TNCs) and causes environmental overload.

Energy: Energy consumption in many tourism facilities is often inefficient and possibilities to use sustainable energy sources (such as solar or water energy for example) are often overlooked.

Noise pollution: Noise generated by a concentration of tourists, tourist road and off-road vehicles, such as dune buggies and snowmobiles, aeroplanes and motor boats may reach uncomfortable and irritating levels for nearby residents and other tourists.

Visual pollution: Visual pollution may result from several sources:

- Poorly designed hotels and other tourist facilities incompatible with the local architectural style or poorly integrated into the natural environment.
- Use of unsuitable building materials on external surfaces.
- Badly planned layout of tourist facilities.
- Inadequate and inappropriate sea- or landscaping.
- Use of large and ugly advertising signs.
- Overhead utility (electric and telephone) lines and poles.
- Poor maintenance of buildings and landscape.

Dramatic and in some cases chaotic tourism development is considered to be one of the most dangerous and environment deteriorating human activities. In the island of Crete (Greece) alone, there are more than 5,000 illegally built units that contribute to coastal disfigurement (Peterson and McCarthy, 1989).

Ecological disruption: Several ecological problems can result from uncontrolled tourism development and use. Recreational activities can affect physical and chemical properties of the soil, changes in compaction, organic matter, moisture and erosion. Vegetation's deterioration includes changes in species and habitat and mechanical damage. Impacts on wildlife include disturbance, alteration of habitats and hunting.

Ecological damage results when the carrying capacity of the area has been exceeded. Some intensively used sites, such as Buccaneers Cove and the tuff cone on Isla Bartolome in Ecuador, show substantial signs of walking-track erosion (Kenchington, 1989). In Kenya, the central circuit of the Ambaselli National Park has been turned into a semi-desert by visitors' vehicles (Tourism Concern and WWF, 1992). Tourism development has already caused substantial damage to inshore reefs near Hurghada, Egypt from infilling, sedimentation and overfishing for marine curios (Hawkins and Roberts, 1994).

Socio-economic and cultural impacts

Tourism can have both positive and negative economic and socio-cultural impacts. This largely depends on the type and intensity of the tourism developed and on the characteristics of the host society. Whether impacts are considered positive or negative depends, in part, on criteria such as income earned, but is also subject to the perception of the host

BOX 14.1 CASE STUDY 1: TOURISM INDUSTRY PRESSURE ON THE SAND BEACH QUALITY OF SLANTCHEV BRIAG RESORT, BULGARIA

The Bulgarian Black Sea coast is well known for its beautiful weather, attractive landscape and spectacular sea and beaches. Slantchev Briag Resort is one of the biggest and oldest Bulgarian Black Sea resorts. It is characterised by leisure-oriented mass tourism. Today the resort is less popular because of the economic, social and environmental problems it suffers. As a result, the need for overall assessments of the current situation and future development of the alternative, modern, environmentally oriented tourism is recognised both by regional and local authorities, and by the public.

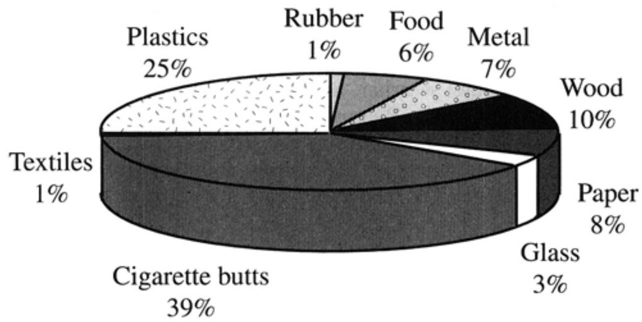
An extensive investigation was set up during the holiday season of 1995 into mass tourism pressures on the biotic and abiotic aspects of the sand beach ecosystem. Results were obtained according to the following agenda:

- An inventory of the legislation; existing administrative structures; environmental information, awareness and education of the staff and tourists in the area.
- Set up and implement a graphic method to determine precisely the recreational capacity and tourist density in the different parts of the beach.
- An inventory of the existing places of entertainment situated along the beach and the related adverse pressure on the sand beach ecosystem.
- Set up an experimental method to determine the quality, composition, spatial and time distribution of the litter remaining on the sand.
- An inventory of the organisation, management and control of the cleaning and maintenance of the beach.
- Analysis of the microbiological ecological and sanitary-hygienic quality of the sand.
- Analysis of the state of the dune formations, and of threatened and rare species in the region.

The following conclusions were reached:

1. Mass tourism development has caused significant deterioration and alteration of the natural sand beach ecosystem of Slantchev Briag Resort.
2. The main factors responsible for the adverse anthropogenic influence on the sand beach quality are:
   - extensive tourist density;
   - improper construction and irregular everyday activity of the entertainment places on the beach;
   - refuse contamination of the sand.
3. Tourist density distribution on the beach is extremely irregular and does not correspond to the officially established thresholds.
4. Overcrowding of the beach results in resource degradation, increasing pollution and decreasing of the visitor's vacation quality.
5. A great part of the beach area (3,855 m^2) has been deteriorated by the construction activity. There are 22 places of entertainment along the beach, of which nine are all-brick buildings.
6. Places of entertainment on the beach are considered both as significant sources of income and as suitable and necessary sites for tourists. At the same time they have many negative effects:
   - Uncontrolled building has caused alteration of landscape and deterioration of the unprotected sand dunes.
   - Pollution with building refuse has caused a change in the physical and chemical characteristics of the sand.
   - Irregular functioning of the entertainment places is a threat to human health and has caused pollution of groundwater and sand.
   - Daily activities of the entertainment places are found to be one of the main sources of solid waste on the beach.
   - Inadequate maintenance of the entertainment places causes negative aesthetic impacts.
7. Refuse contamination of the beach depends on the tourist density, tourist culture and behaviour and on the quality of cleaning and maintenance of the sand beach.
8. The prevalent types of beach litter collected are cigarette butts (39 per cent), plastics (25 per cent), paper (8 per cent), etc. (see Figure).
9. Solid waste remaining on the beach could adversely affect human health, quality of life, quality of the sand beach ecosystem, as well as the development and economic efficiency of tourism.
10. The main anthropogenic factor influencing the microbiological ecological quality of the sand is the general hygiene of the beach. As a result of the greater refuse contamination of the south part of the beach, relatively high

→



Composition (in per cent) of the solid waste collected at the beach of Slantchev Briag, Bulgaria

values of the measured parameters – organic carbon content and number of the psychrophillic saprophytic micro-organisms (aerobic and anaerobic) – were detected.

11 The microbiological sanitary-hygienic quality of the sand depends both on the tourist density and general hygiene of the beach. Regular and thorough sand cleaning at the central part of the beach made for the high sanitary quality of the sand, characterised by the coli-indices number in the surface sands, despite the highest tourist density detected.

12 The present state of the dunes is characterised by decreasing mobility, gradual stabilisation, altered structure of vegetation, etc. The outlook for the preservation of these unique natural formations is not optimistic. The main factors in the complete deterioration of the sand dunes in the northern part of the beach and the significant alteration of the landscape in the southern beach region are uncontrolled constructional activity, large-scale afforestation, movement of machines, horses and people, refuse contamination and occasional grazing.

13 Governmental protection policy has led to preservation of part of the sand dunes and to favourable conditions for threatened and rare species in the region.

Source: Nikolova, 1995

community. Most often, different community groups have varying reactions to tourism developments. Moreover in evaluating socio-economic impacts it is always important to ask the question: positive or negative for whom?

Positive impacts

Economic benefits: Direct economic benefits include employment, income and international tourism-foreign exchange. This might lead to improved living standards of the local community and overall national and regional economic development. An important indirect economic benefit of tourism is that it serves as a catalyst for the development or expansion of other economic sectors.

Conservation of cultural heritage: Tourism can be a major stimulus for the conservation of important elements of the cultural heritage of an area. These elements include:

- conservation of archaeological and historic sites and interesting architectural styles;
- conservation and revitalisation of traditional arts, handicrafts, dance, music, drama, customs and ceremonies, dress and certain aspects of traditional lifestyle;
- direct or indirect financial contribution to the maintenance of museums, theatres, cultural facilities and activities.

Renewal of cultural pride: Residents' sense of pride in their culture can be reinforced or even renewed when they observe tourists appreciating that culture. In multicultural countries regional tourism can help maintain the cultural identity of minority groups and indigenous people.

Cross-cultural exchange: Tourism can promote cross-cultural exchange between tourists and residents. In this way people learn more about one another's cultures, which can foster greater and better mutual understanding and respect.

Negative impacts

Tourism may also generate negative impacts or reduce the efficiency of the positive ones.

Loss of potential economic benefits: Loss of potential economic benefits of the local area can occur, and local resentment sometimes stems from seeing that many tourist facilities are owned and managed by outsiders. The World Bank estimates that 55 per cent of gross tourism revenues to developing countries actually leaks back to the developed ones (Boo, 1991). Potential foreign exchange earnings are reduced when imported goods and services are used. For example, it is estimated that at least 80 cents in every American dollar brought into the Bahamas are spent to pay for imported foodstuffs (Tourism Concern and WWF, 1992).

The costs caused by tourists in a given area offset some of the economic benefits. For example, the National Trust in England is spending over US$ 150,000 a year to repair footpaths on its Lake District estate (English Tourist Board, 1991).

Economic and employment distortions: Economic distortions can take place geographically if tourism is concentrated in only one or a few areas of a country and region, disconnected from corresponding development in other places. This situation may engender resentment among residents in undeveloped areas. On the other hand, inflation of local prices of land and certain goods and services may take place, in which case tourist demand can create financial hardship for residents. Employment generation is often an important argument in convincing local authorities of the soundness of new tourism development. In most instances employment is overestimated and original figures are not realised. Moreover in more remote areas local citizens generally obtain menial, low paid and often seasonal jobs.

Overcrowding and loss of amenities for residents: If there is overcrowding of amenity features, shopping, community facilities and congestion of the transportation system by tourists, residents cannot conveniently use them and will become irritated and resentful.

Cultural impacts: In extreme cases, there may be a loss of cultural character, self-respect and overall social identity through submergence of the local society by the alien cultural patterns of seemingly more affluent and successful tourists. Misunderstanding and conflict can arise between residents and tourists because of differences in languages, customs, religious values and behaviour patterns. Tourism development has played a major role in the destruction of ancient Hawaiian sacred sites. In Egypt tourists have been the target of terrorist attacks in what appears to be an extreme example of cultural clash, that between western culture and a militant form of Islam (IFTO, 1993).

Social problems: Problems of drugs, alcoholism, crime and prostitution may be exacerbated by tourism.

The analysis shows that tourism can severely impact the environmental and socio-economic context of a society. Whether and to what extent this will happen depends greatly upon the early recognition of these phenomena and the management response to them. The aim of this approach is to provide integrated policies and impact control measures that do not generate negative environmental and socio-economic impacts, but reinforce the positive ones. Sustainable tourism is such an approach and is fast gaining attention.

SUSTAINABLE TOURISM DEVELOPMENT: RECOMMENDATIONS FOR GOOD PRACTICE

Sustainable tourism

The concept of sustainable development has received considerable attention in recent years. It has grown out of concerns that the needs and ways of life of an ever-growing world population are outstripping the planet's capacity to support us. Sustainability is defined as 'the rearrangement of technological, scientific, environmental, economic and social resources in such a way that the resulting heterogeneous system can be maintained in a state of temporal and spatial equilibrium' (WCED, 1987). The strategy for sustainable living developed in 1991 by the World Conservation Union (IUCN), the World Wide Fund for Nature (WWF) and the United Nations Environmental Programme (UNEP) recognises that 'humanity must live within the carrying capacity of the Earth .... We must adopt life styles and development paths that respect and work within nature's limits' (IUCN *et al*., 1991).

Sustainable tourism is a very real application of the idea of combining conservation principles with development in the area of tourism. Its definition will undoubtedly continue to evolve over the next decade. In the Brundtland report on Our Common Future (WCED, 1987), a number of notions were advanced that contribute to a definition.

> Sustainable tourism development can be thought of as meeting the needs of present tourists and host regions while protecting and enhancing opportunity for the future. It is envisaged as leading to management of all resources in such a way that we can fulfil economic, social, and aesthetic needs while maintaining cultural integrity, essential ecological processes, biological diversity and life support systems.
> (WTTC *et al.,* 1996)

The goals of sustainable tourism are:

1 to develop greater awareness and understanding of the significant contributions that tourism can make to the environment and economy;
2 to promote equity in development;
3 to improve the quality of life of the host community;
4 to provide a high quality of experience for the visitors;
5 To maintain the quality of the environment (Inskeep, 1991).

There is no universally accepted strategy to achieve sustainable tourism development. Each situation is different and requires different approaches and solutions. However, if tourism is to be truly beneficial to the natural and social environment and sustainable in the long term, resources should not be overconsumed; natural, social and cultural diversity should be protected; tourism development should be integrated into a national and local strategic planning framework; local people should be involved in tourism planning and implementation; and continuous research and monitoring must be put into practice (Tourism Concern and WWF, 1992). These principles of sustainability must be implemented by the whole tourism industry in close consultation with the public at large. Authorities have a crucial role in this process.

Recently, all parties involved in the tourism industry—national tourist authorities, tourist trade organisations and communities—have responded to public awareness of the harmful environmental and social impacts of current tourism development. Regulatory and self-regulatory measures have been introduced to reduce the tourism industry pressure on the host communities and the environment. This paper gives some examples of good practice, discusses their effectiveness and pays attention to the need of partnerships between tourism industry, governments, local communities, academics and the public at large to achieve the advocated goals and principles of sustainable tourism.

Sustainable tourism policy

The concept of sustainable tourism is increasingly supported by the international political organisations and institutions such as the European Union (EU). A number of EU Directorates are involved with tourism and environment. Directorate General (DG) XI develops environmental policy. Its fifth environmental action programme 'Towards Sustainability' promotes sustainable development (Commission of European Communities, 1992a). Special attention is paid to tourism as one of the five sectors of key importance. The Tourism Division within DG XXIII is responsible for tourism projects and policies. Tourism policy during the 1980s was slow to recognise environmental aspects, and as a result there is no EU tourism legislation relating specifically to the environment. But it is encouraging that in 1992 the first comprehensive tourism policy and a three-year action plan for development of an environment friendly tourism was adopted by the EU (Commission of European Communities, 1992b).

In response to the need for tourism that is not harmful to the natural and cultural environment, hosts such as Canada, Australia, Hawaii, Ireland, the UK, etc. have developed their own national strategy towards sustainable tourism.

The 1992 United Nations Conference on Environment and Development (UNCED) was the most important international political forum to address the problems and discuss ways of achieving sustainable development. Agenda 21, adapted to the travel and tourism industry, has been developed by the World Travel and Tourism Council, the World Tourism Organisation and the Earth Council. It identifies three core tools to achieve the sustainability objectives:

- Introduction of new, or strengthening of existing, regulation to ensure the protection of human health and environment.
- Use of free market mechanisms, by which the price of goods and services should increasingly reflect the environmental costs of resource inputs, manufacture, use, recycling and disposal subject to the country specific conditions.
- Industry-led voluntary programmes, which aim to ensure responsible and ethical management of products and processes from the point of view of health and safety environmental aspects.

It is clear from the principles advocated that the Travel and Tourism Agenda 21 is highly tourism industry driven. There is, for example, no guarantee that industry-led voluntary programmes will contribute in a significant way to realising environmental and socio-economic targets. On the other hand, it is encouraging that industry has its sustainable tourism agenda ready.

In response to Agenda 21, national, regional and local authorities have also developed strategies and policies of sustainable tourism. Often, however, these plans face enormous inertia in implementation, so that they often resemble window dressing rather than efficient policy.

Consumers, the tourists themselves, are poorly organised when it comes to sustainable tourism. They remain the most passive party even though the demand for more sustainable forms of tourism is increasing in many places. Eco-tourism is growing at annual rates of 25–30 per cent world-wide. In spite of this, the majority of tourists in many countries are non-green tourists.

Sustainability is a long-term aim, but the foundations for a sustainable future need to be laid immediately. It is essential to adopt and develop close partnerships between all parties involved in the tourism industry at international, national and regional levels in order to implement sustainable tourism policies and action plans.

Self-regulatory measures

There are increasing signs that many businesses are beginning to recognise that environmental policies can be beneficial for business, and more importantly that they may be essential for long-term economic survival. A number of organisations within the travel, tourism and leisure industries have adopted environmental codes of conduct on a voluntary basis. Such codes seek both to improve environmental performance and to demonstrate that self-regulation is a viable, cost-effective complement to government intervention. The actual implementation of environmental management is often perceived as the most critical and difficult step. The numerous schemes offering information and advice, especially the availability of guidelines on 'best practices' and 'self-help' green audit kits, do provide important starting points. Table 14.2 lists examples of the contribution of the tourism industry in achieving sustainable tourism.

Marketing and eco-labelling

Marketing plays an important role in tourism. The marketing strategy for sustainable tourism provides full and honest information about the tourism product on offer, and involves identification, appraisal and constant review of supply of both natural and cultural resources as well as of demand side factors. The main goals of the marketing process are to make tourists aware of their potential impacts on the local physical and socio-cultural environment, to educate visitors in advance of their arrival, and to promote tourism appropriate to the capacities of the destination areas, in terms of scale, number and types of tourism.

Eco-labelling of tourism products attempts to extend marketing strategies that can effectively embrace aspects of environmental policy and establish the 'environmental seals of quality'. The principles of such schemes are based on identifying a range of environmental criteria or best practices that can be applied to tourism products and services.

Developing an environmental purchasing policy

The initiation of a positive environmental purchasing policy brings a range of direct and indirect benefits to tourism business as well as to the local economy. The basic principles of a good environmental purchasing policy are:

Table 14.2 Possible practices of sustainable tourism and responsibility as identified by the tourism industry

Responsibility	*Cost-cutting*	*Value added*	*Long-term investment*	*Legislation*
Tour operators	Recycle brochures Rely less on luxury resorts	Sympathy booklets Clean beaches etc. Pressure suppliers Diversity product	Staff training Research Lobby governments to provide infrastructure	Lobby governments to ensure satisfactory laws
Travel agents	Recycle brochures	Provide special advice	Staff training	
Hotels	Recycle waste Reduce fuel use	Information packs Sympathy booklets Pressure suppliers	Staff training	
Carriers	Reduce fuel use	Sympathy booklets	Research Staff training	
Tourism associations	Urge recycling and reduced consumption	Information packs Sympathy booklets	Research Staff training Lobby governments to provide infrastructure	Lobby governments to ensure satisfactory laws
Government tourism offices	Waste management	Education of tourists	Provide adequate infrastructure Research Staff training	Controls on tourists and development Land-use zoning Tourist tax

Source: Tourism Concern, 1996

- avoid products made with, or containing, envi ronmentally damaging materials;
- buy in bulk only what you really require;
- avoid goods that are over-packaged;
- buy recycled and recyclable products;
- purchase good quality, repairable products;
- Buy locally produced goods (Williams and Shaw, 1996).

The ultimate goal of the environmental purchasing policies is not only to improve the business 'green-rating', but also to modify the supplier's views and activities.

Energy management
Tourism organisations, especially hotels and resorts, use substantial amounts of heat and power. A starting point in identifying energy savings is to review where the best opportunities exist. The alleviation in this context is the 'green audit' kit, which helps to evaluate the energy performance of the organisation. This information can be the basis for an action plan to increase the energy performance.

Water management
The high level of water consumption within the tourism sector, together with high and rising costs of obtaining water, make the need for water conservation critical. In purely business terms the efficient use of water can make significant environmental and cost savings. However, there are important long-term reasons for better water management within the tourism industry. The most important step that has to be undertaken by individual businesses is an audit of current uses within the enterprise. An effective water management plan should be developed and implemented. Typically in hotels, the greatest potential to make savings of water are in guest-rooms, laundries, kitchens, swimming pools and golf courses.

Waste management re-using, reducing and recycling
Waste disposal is a world-wide problem, but it is especially critical to the tourism industry. The presence of rubbish tips and polluted land and water areas has a direct economic impact on the tourism industry. At the same time tourism must take much of the responsibility for the generation of large amounts of solid waste materials, including cans, bottles and jars, kitchen waste, old fixtures and equipment and potentially hazardous waste. Business waste

management strategies based on sustainability make economic and environmental sense within tourism. The objective of such management practice is to reduce to a minimum the amount of solid waste generated. There are three important waste management steps:

- To list all the waste items a given business disposes of; methods of disposal and their costs; yearly amounts involved; types of hazardous waste that have to be separately treated.
- To identify for each item if it is possible to reduce amount used; re-use all or some of the items; recycle them; dispose of residual waste in the safest way possible.
- Prepare an action plan that introduces a 'no waste' campaign; prioritise actions that are relatively easy to implement.

Transport

Transport is a critical part of tourism. Tourism projects involving traffic management and transportation have suggested a wide array of solutions, although particular stress has been placed on the outline active promotion of public transport. However, it is increasingly clear that solutions must involve more integrated approaches that help to promote existing public transport as well as developing new transport provisions to meet visitors' needs.

Training

The importance of staff training is widely recognised. Good quality staff is critical for the long-term success of the tourism industry. Staff tasks within the tourism industry are extremely varied, and this needs to be reflected in training programmes. However, training that integrates environmental practices into all levels of tourism management is currently rare to non-existent. Development of new staff training programmes, which increase awareness of local nature and culture and introduce the principles of sustainable tourism and nature conservation, is needed to educate the new generation of managers.

Tourist information and education

Tourists' behaviour and their attitude to the place they are visiting is critically important for achieving the goals of sustainable tourism development. If tourist behaviour is characterised by respect, care and tolerance, then all aspects of tourism can benefit. The need for environmentally oriented visitor management is increasingly recognised by the representatives of the tourism industry. Managers should be able to use explanation, persuasion, information and education. Instruments to reach these goals are easy to apply, non-regulatory and have long-term benefits. The idea is not to stuff users with knowledge, but to raise their awareness of the environment.

The ECOMOST programme (Box 14.2) is a good example of the voluntary initiative taken by the tourism industry in close partnership with academics to achieve sustainable tourism.

The self-regulatory measures adopted by the tourism industry are important management tools for reducing the environmental damage of tourism. However, they cannot prevent unethical behaviour by individual companies that choose to act in a destructive way. The need to adopt a wider regulatory framework based on partnerships between tourism business, governments and local communities is now increasingly recognised. It is the responsibility of national and regional authorities to develop and adopt an efficient tourism planning procedure and a system for evaluating the environmental performance of tourism companies in helping to prevent or limit environmental and social damage.

Planning for sustainable tourism development

Tourism has the potential to bring economic prosperity and environmental improvement to the destinations in which it operates. On the other hand, poorly planned and managed tourism can harm the resources on which it is based. Environmental and cultural degradation can be avoided by the adoption of environmentally sound planning procedures.

Figure 14.2 provides an overview of the planning process. The basic steps are: study preparation; determination of the objectives; survey; analysis and synthesis; policy and plan formulation; recommendations; implementation and monitoring. Effective and comprehensive national/regional planning process should involve the following principles and requirements:

1 National/regional tourism planning authorities and trade organisations should work in close consultation with local communities and the

BOX 14.2 CASE STUDY 2: THE ECOMOST PROJECT

ECOMOST is an acronym for European Community Models of Sustainable Tourism. The project was developed by the International Federation of Tour Operators (IFTO). The starting point was Mallorca, the largest of Spain's Balearic Islands. It has one of the highest concentrations of tourists in the world. Mallorca is dependent on tourism – half of its income is derived from the holiday business – but the industry faces difficulties. The project was set up to study its decline. It had to assess whether the problem was temporary, and could be reversed, or whether there was some inherent weakness that would result in serious stagnation.

To achieve this, researchers devised the 'Model of Sustainable Tourism' with the intention that it could subsequently be used as an instrument to perform similar analyses of tourism in other destinations.

The leading guideline states that sustainable tourism development adheres to three main principles:

- Ecological sustainability to protect the ecology and biological diversity – tourism development should always be within the carrying capacity of the ecosystem.
- Social and cultural sustainability to preserve a society's identity – an important tenet of sustainability is that decision-making should involve all parts of the community.
- Economic sustainability to ensure economic efficiency and the management of resources so that they can support future generations.

According to the project developers, there are three main requirements for maintaining a successful tourist destination over the long term:

- The population should remain prosperous and keep its cultural identity.
- The place should remain attractive to tourists.
- Nothing must be done that damages the ecology.

Those three aims are the basis for a sustainable tourism, but for them to be achieved a fourth element is required:

- An effective political framework.

This means laws that guard the principles of sustainability; an efficient and integrated planning procedure to allow everyone involved in providing tourist facilities, and the local population, to participate in formulating a tourism policy.

The ECOMOST model works by breaking down those aims into their components (referred to as 'targets' in the study) and then subdividing those components into 'indicators'. The indicators are used to check the performance of the components. The critical value for each of the indicators is determined. The model is designed as a checklist to detect any danger to sustainability.

The checklist focuses on the following main topics:

- Population – including targets such as preservation of the population's prosperity; preservation of economic efficiency; preservation of cultural identity.
- Tourism – with the main targets of maintaining satisfaction for guest and tour operators, and maintenance and modernisation of accommodation.
- Ecology – assessing environmental consciousness, carrying capacity and preservation of the landscape.
- Policies – aimed at assessing the existence or quality of effective tourism and ecologically oriented legislation; regional planning; adequate participation by industry and public in the planning process.

The following factors led to the development of the Action Plan for tourism development in Mallorca: research findings concerning economic, socio-cultural and environmental impacts of the current tourism development; quality of the existing management, law and policies; and the particular experience of the Municipality of Calvia.

Each indicator is compared to the critical value or requirements, and assessed in terms of how near a place is to attaining sustainability. The priorities of the Action Plan are determined by how much needs to be done to the components to bring them as close as possible to their optimum state. Degrees of urgency for action are classified as need for immediate action; action needed as soon as possible; action required eventually; no obvious need for action at the moment. It is determined which body is mainly responsible for any measures to be taken, and which others have a part to play. Responsibility is shared between local communities, regional authorities, national government, national and international tour operators, the European Community and tourists.

→

The ECOMOST project is not the universal blueprint and solution for every type of problem that different tourism destinations face, but it does provide useful guidelines from which other regions in the world might benefit. If adopted, the principles of ECOMOST would provide a tourism that ensures enjoyment for those paying for it and profit for its suppliers, while safeguarding the future of host places. *Source:* IFTO, 1993

general public to determine the type, scale and objectives of tourism development.

2 Goals and objectives should be developed on the basis of the general goals of sustainable tourism development, taking into consideration specific features of the proposed development and the area.
3 The survey activities should be carefully organised and effectively conducted to supply objective and correct information on environmental characteristics and quality; socio-cultural patterns and trends; existing tourism legislation and regulations; and existing tourist facilities, services and infrastructure.
4 Policy and plan formulation should be developed on the basis of the quantitative and qualitative analysis and synthesis of the survey information. The concept of carrying capacity serves as an anchorpoint for the preparation and evaluation of alternative development policies and structure plans. Carrying capacity is the ability of an ecosystem to maintain itself and to support specific human uses or activities indefinitely without suffering adverse effects. Three types of carrying capacity should be assessed:

- Environmental—the degree to which an ecosystem, habitat or landscape can accommodate the various impacts of tourism and its associated infrastructure without damage being caused or without losing its 'sense of place'.
- Cultural and social—the level beyond which tourism developments and visitor numbers adversely affect local communities and their ways of life.
- Psychological—the level beyond which the essential qualities that people seek in the area would be damaged by tourism developments (FNNPE, 1993).

Once carrying capacity has been assessed, the next step is to set standards of environmental quality which should be maintained by sustainable tourism. Management methods should be developed to maintain these standards and a monitoring programme established to check their effectiveness.

Unfortunately, there is no simple way to measure carrying capacity and identify thresholds. A much-used concept in practical tourism and recreational management is 'limits of acceptable change' (LAC) (Herath, 1996). This human-oriented way of thinking concerns one or other of the actors in the tourism system (local community, tourists or travel trade) and clarifies how much change (in nature, economy, social relations and culture) they can accept. The LAC procedure is based on the development of a series of steps which lead to the development of a number of desired conditions. It also identifies the actions necessary to maintain these conditions.

Another method of making carrying capacity operational is based on the concept of the environmental utilisation space (EUS). This is a measure of the relative pressure of a community on the available environmental functions. It represents the quantity of natural resources and services that ecosystems can provide without reducing their productive capacity or generating irreversible changes in their essential parts. This concept links all available environmental functions to the production and consumption patterns in a particular country according to different economic sectors, including the tourism sector (Goeteyn, 1996).

It can be calculated that the maximum acidification that ecosystems in Europe are able to stand on a long-term basis is 1,400 to 2,400 acid equivalents. The current deposition rate, however, totals 5,000 to 6,000 acid equivalents, of which 10 to 15 per cent can be attributed to the tourism sector, mainly via heating, electricity use and transport. Thus to reach a sustainable level a reduction from 50 to 70 per cent is needed.

The concept of carrying capacity is a dominant way of thinking on the issue internationally. It is essentially based on the assumption that there are limits to what the environment can assimilate without major irreversible (negative) impacts. Moreover, the assumption is that science can clarify what and how large these limits and thresholds are. In practical tourism management it has been very difficult to identify thresholds, and to react promptly enough to prevent long-lasting problems.

- Policy and plan formulations should be based on the identification and evaluation of all the economic, natural and socio-cultural costs and benefits of the proposed tourism development.

Environmental impact assessment (EIA) is an instrument used to aid and improve the decision-making process. Its main goal is prevention of damage and degradation as a result of human activity. EIA provides the decision-makers with an account of the implication of a proposal before any decision is made. According to EC Directive 85/337/EEC the following types of tourism development projects are recommended for EIA: ski-lifts, cable cars, road constructions, harbours, marinas, airports, holiday villages and hotel complexes.

A project-level EIA typically evaluates the impacts of large-scale tourism projects. Recently, the need for evaluation of the effects of many independent (even small) tourism projects has been realised. Project assessments have definite limitations. Interactions and cumulative effects of subsequent developments in the same area and/or over time in, for example, the same riverbed are as a rule not taken into account in a project-by-project assessment. Also the tourism or energy policy of a new government can have more impact on the environment than many tourism projects. To deal with the impacts on these project decision-making levels, strategic impact assessment (SIA) has been established. SIA provides an opportunity for evaluation of plans, programmes or policies at national/regional level. It ensures consideration of cumulative impacts in policy-level alternatives and programme-wide mitigation measures.

Although subject to limitations, such as inadequacy of data, subjectivism in prediction of impacts and their magnitude, EIA and SIA can be powerful tools in achieving high environmental design standards, and keeping development within carrying capacity of the area. For a more complete discussion on EIA and SIA the reader is referred to Chapter 10 in Volume I of this series of textbooks (Devuyst, 1998).

- During and after implementation, tourism development should be monitored to ensure that it is accomplishing its objectives, following the development schedule and not generating any economic, environmental or socio-cultural problems. If problems are detected, remedial measures must be taken to bring the development back on track.

Indicators are often used as a technique in monitoring. An environmental indicator is a simple, objective and reliable measure reflecting a particular environmental condition. The establishment of indicators is generally based upon models. The pressure-state-response model is in widespread use in this area (UNEP and DPCSD, 1995). This model ensures that the different major aspects of sustainability (social, economic, institutional, environmental) are converted into different types of indicators:

- Pressure indicators reflect the social and economic elements putting stress on the environment (and culture).
- State indicators describe the actual conditions of the environment.
- Response indicators reflect how either the environment itself or policy try to correct the induced changes.

Indicator establishment for sustainability is very much at the centre of contemporary policy research. There is actually no set of universally agreed indicators for sustainable tourism that act both as good parameters for monitoring and as early warning signs for management corrections.

Maintenance of sustainable tourism

Environmental care systems (ECS) are important instruments in ensuring that an organisation—including organisations in the tourism sector such as hotels, resorts, airports, cities—is performing in terms of sustainability. They essentially involve the setting of targets, critical examination of the operations,

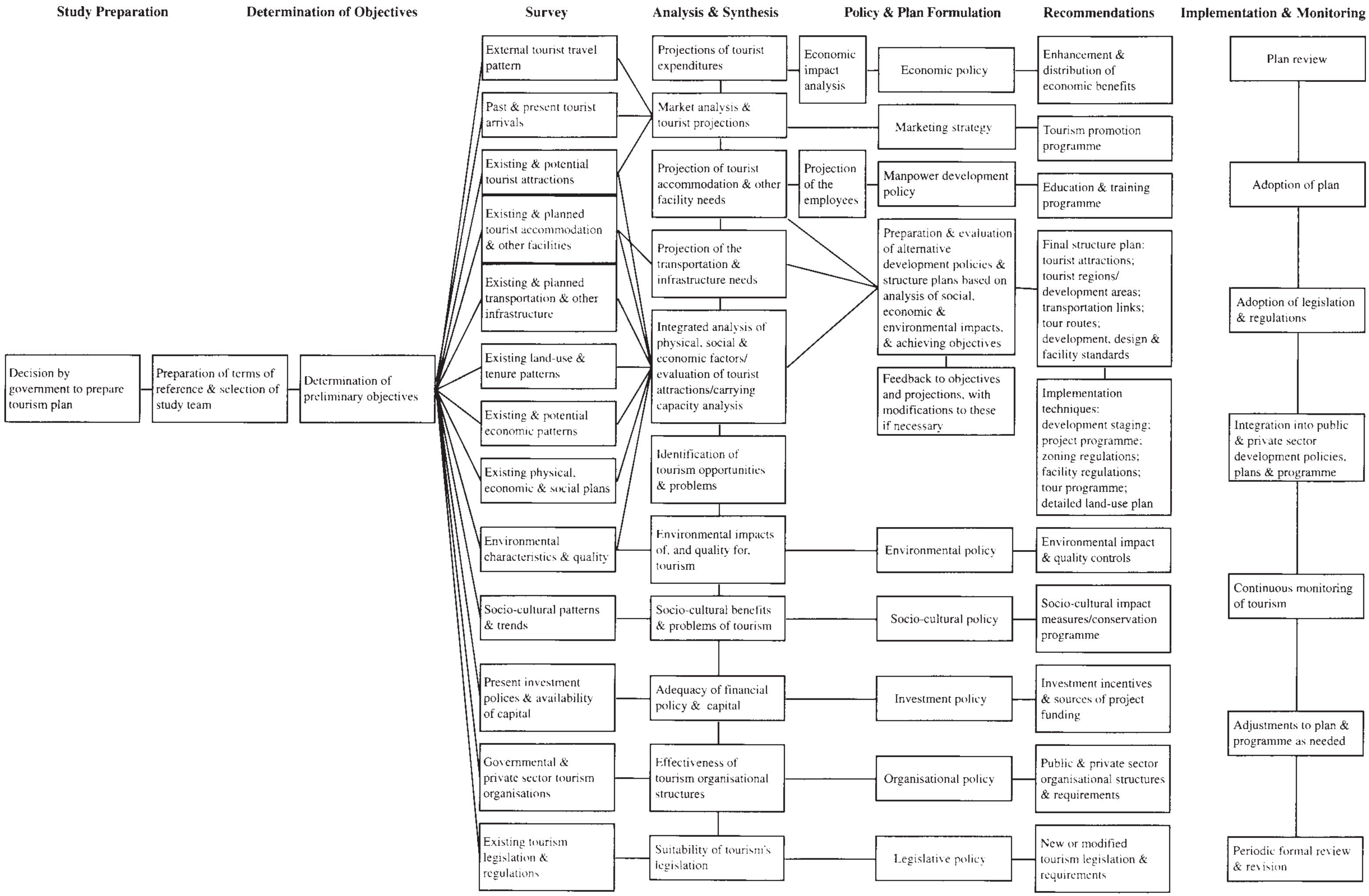


Figure 14.1 Process for preparing a comprehensive tourism development plan

Source: Inskeep, 1991

identification of areas for improvement and the implementation and evaluation of higher-performing environmental tools.

A central element in an environmental care system is the audit, which is often defined as:

> A management tool comprising a systematic, documented, periodic and objective evaluation of how well environmental organisation, management and equipment are performing with the aim of helping to safeguard the environment by:
> 1 facilitating management control of environ mental practices;
> 2 assessing compliance with company policies, which would include meeting regularity requirements.

In practice, environmental care systems handle and measure the performance of rational energy and water use policies, limit air pollution, minimise waste, watch the quality of environmental contaminants in food, option for environmental friendly materials in daily use and a high quality green environment without useless pollution from pesticides and/or fertilisers. Usually, environmental care systems not only cause less environmental pressure but also save costs and can therefore pay for themselves.

The results of ECS should be communicated in yearly environmental reports.

A more detailed discussion on environmental care systems is provided in Chapter 13 in Volume I of this series of textbooks (De Weerdt, 1998).

CONCLUSIONS

'Sustainable tourism' was originally discussed and developed, conceptually and theoretically, on national and international levels. It is therefore not surprising that there are deficiencies in its realisation on the local level. There is still a wide gap between the theoretical concept, goals and principles of sustainable tourism and the character of present tourism development. Specific sustainable development action plans should be developed and implemented for each destination. A close partnership is required between the tourism authorities, tourism trade organisations, local communities and the public at large in order to achieve qualitative economic development and conservation of natural and cultural resources.

It is equally clear that the tourism industry has most probably invested more than any other target group in stating what sustainable tourism means for them. While this is a very useful contribution to the debate, it is not sufficient on its own. The other target groups should be involved on a basis of equity.

In many countries, moreover, authorities are highly discrete when it comes to the sustainable management of the resources they own and control. It is of the utmost importance for nature reserves and for sites of cultural and/or archaeological importance that management plans are established that encompass the whole environmental context of these areas.

Obviously, the disharmony between current resource-consuming and disruptive tourism development, on the one hand, and the theoretical concept of 'sustainable tourism', on the other, is caused by the lack of effective mechanisms for realising integrated policy and management at national, regional and local levels. The challenge will be to forge partnerships between the different interest groups in tourism, and achieve harmony between tourism, environment and host population.

REFERENCES

Becker, Ch. (1995) 'Tourism and Environment: A Changing Framework of Values', in A.Montanari and A.Williams (eds) *European Tourism. Regions, Spaces and Restructuring,* Chichester, UK: John Wiley and Sons Ltd, pp. 207–219.

Boo, E. (1991) *Ecotourism: The Potentials and Pitfalls,* Vol. 2, 'Country Case Studies', Washington DC, USA: World Wildlife Fund.

BUWAL (1992) *The Swiss Report. UNCED 1992,* Berne, Switzerland: Swiss Federal Office of Environment, Forest and the Landscape (BUWAL).

Commission of European Communities (1992a) *A European Community Programme of Policy and Action in Relation to Environment and Sustainable Development,* Vol. 2, 'Towards Sustainability', COM (92) 23 final, Luxembourg: EU Official Publication Office.

Commission of European Communities (1992b) 'Council Decision on a Community Action Plan to Assist Tourism', *Official Journal* 92/421/EEC, Luxembourg: EU Official Publication Office.

Devuyst, D. (1998) 'Environmental Impact Assessment', in B.Nath, L.Hens, P.Compton and D.Devuyst (eds) *Environmental Management in Practice,* Vol. 1, London, UK: Routledge.

De Weerdt, H. (1998) 'Environmental Auditing and Environmental Management Systems', in B.Nath, L.Hens, P.Compton and D.Devuyst (eds) *Environmental Management in Practice,* Vol. 1, London, UK: Routledge.

English Tourist Board (1991) *Tourism and the Environment: Maintaining the Balance,* London, UK: English Tourist Board.

Federation of Nature and National Parks of Europe (FNNPE) (1993) *Loving Them to Death? Sustainable Tourism in Europe's Nature and National Parks,* Paris, France: Grafenau.

Goeteyn, L. (1996) 'Measuring Sustainable Development at the National and International Level', in B.Nath, L.Hens and D.Devuyst (eds) *Sustainable Development,* Brussels, Belgium: VUB University Press, pp. 161–180.

Hawkins, J.P. and Roberts, C.M. (1994) 'The Growth of Coastal Tourism in Red Sea: Present and Future Effects on Coral Reefs', *Ambio* 23, 8: 503–508.

Herath, G. (1996) 'Ecotourism and Nature Conservation: The Use of the Safe Minimum Standards', paper presented at the 40th Annual Conference of the Australian Agricultural and Resource Economy Society, 13–15 February 1996, University of Melbourne, Australia.

Inskeep, E. (1991) *Tourism Planning. An Integrate and Sustainable Development Approach,* New York, USA: Van Nostrand Reinhold.

International Federation of Tour Operators (IFTO) (1993) *Planning for Sustainable Tourism. The ECOMOST Project,* Geneva, Switzerland: IFTO.

IUCN, UNEP and WWF (1991) *Unsere Verantwortung für die Erde—Strategic für ein Leben in Einklang mil Natur und Umwelt (Our Responsibility for the Earth: Strategy for a Life in Harmony with Nature and the Environment),* Gland, Switzerland: IUCN, UNEP, WWF (in German).

Jungk, R. (1980) 'Wieviel Touristen pro Hektar Strand?', *Geo* 10: 156 (in German).

Kenchington, R.A. (1989) 'Tourism in the Galapagos Islands: The Dilemma of Conservation', *Environmental Conservation* 19, 1: 227–232.

Krippendorf, J. (1975) *Die Landschaftsfresser, Tourismus und Erholungslandschaft—Verderben oder Segen?,* Berne, Switzerland (in German).

Nikolova, A. (1995) 'Tourism Industry Pressure on the Sand Beach Quality of Slantchev Briag Resort', Master's Thesis, University of Chemical Technology and Metallurgy, Sofia, Bulgaria.

Partsch, K. (1991) *Rapport sur la situation des Alpes,* Sontholen, Germany: Alpen- und Europabüro (in French).

Peterson, I. and McCarthy, L. (1989) 'Greece Tackles Overbuilding in Tourist Area', *Environmental Conservation* 15, 2: 166–168.

Tourism Concern (1996) *Sustainable Tourism. Moving from Theory to Practice,* Godalming, UK: World Wide Fund for Nature.

Tourism Concern and World Wide Fund for Nature (WWF) (1992) *Beyond the Green Horizon: Principles for Sustainable Tourism,* Godalming, UK: World Wide Fund for Nature.

UNEP and DPCSD (1995) The Role of Indicators in Decision Making', in *Discussion Papers on Indicators of Sustainable Development for Decision Making,* Ghent, Belgium, 9–11 January, Brussels, Belgium: Bureau du Plan.

US Office of Technology Assessment (USOTA) of the Congress of the United States (1993) 'Science and Technology Issues in Coastal Ecotourism', in *Tourism Management* 14, 4: 307–316.

Williams, A. and Shaw, G. (1996) *Tourism, Leisure, Nature Protection, and Agri-Tourism: Principles, Partnerships and Practice,* Exeter, UK: Tourism Research Group, University of Exeter.

World Commission on Environment and Development (WCED) (1987) *Our Common Future,* Oxford, UK: Oxford University Press.

World Conservation Union (IUCN), World Wide Fund for Nature (WWF) and United Nations Environmental Programme (UNEP) (1991) *Caring for the Earth—A Strategy for Sustainable Living',* Gland, Switzerland.

World Tourism Organisation (WTO) (1992) *Tourism Trends and the Year 2000 and Beyond,* Sevilla, Spain.

World Tourism Organisation (WTO) (1994) *Sustainable Tourism Development: A Guide for Local Planners,* Madrid, Spain.

World Tourism Organisation (WTO) (1995) *Yearbook of Tourism Statistics,* Madrid, Spain.

World Travel and Tourism Council (WTTC), World Tourism Organisation (WTO) and Earth Council (1996) *Agenda 21 for the Travel and Tourism Industry. Towards Environmentally Sustainable Development,* Madrid, Spain.

Zoffman, C., Rodriguez-Verta, F., Perez-Fillol, M., Ruiz-Bevia, F., Torreblanca, M. and Colom, F. (1989) 'Microbial and Nutrient Pollution along the Coasts of Alicante, Spain', *Marine Pollution Bulletin* 20, 22: 74–81.

SUGGESTED READING

Briguglio, L., Archer, B., Jafari, J. and Wall, G. (eds) (1996) *Sustainable Tourism in Islands and Small States. Issues and Policies,* Island Studies Series, London, UK: Pinter.

English Tourist Board (1991) *Tourism and the Environment: Maintaining the Balance,* London, UK English Tourist Board.

Federation of Nature and National Parks of Europe (FNNPE) (1993) *Loving Them to Death? Sustainable Tourism in Europe's Nature and National Parks,* Paris, France: Grafenau.

Green Globe (1994) *Waste Minimisation* (available for the following sectors: Hotels, Office Facilities, Restaurants and Catering Facilities, Tour Operators and Travel Agents, Transport, Visitor Attractions) London, UK
Green Globe (1994) *Water Management* (available for the following sectors: Hotels, Office Facilities, Restaurants and Catering Facilities, Tour Operators and Travel Agents, Transport, Visitor Attractions) London, UK.
World Tourism Organisation (WTO) (1994) *Sustainable Tourism Development: A Guide for Local Planners,* Madrid, Spain.

SELF-ASSESSMENT QUESTIONS

Make a survey on a tourist destination area with which you are familiar and write an essay covering one of the following problems:

1 Environmental impacts of the tourism development.
2 Economic and socio-cultural impacts of the current tourism development.

Guidelines for this survey

1 The general geographic characteristics of the area should be described to provide the background for overall understanding, and the basis for analysis of particular relevant aspects of the environment.
2 The location of the area should be specified and mapped (in case 1).
3 Several characteristics of the natural environment should be surveyed (in case 1):
- Climatic patterns.
- Topography.
- Wildlife and vegetation.
- Geology.
- Ecological systems.

4 Economic and socio-cultural factors should be surveyed (in case 2), including:
- population characteristics, such as number, distribution, mitigation patterns, employment patterns, age-sex profiles, education level, etc.;
- cultural patterns, such as social structure, value systems, customs, lifestyles and attitudes to tourism development;
- economic patterns, such as the major component of the economy, share in gross national product, income levels and distribution, type and value of exports and imports and other economic factors that relate to analysis of the interface between tourism and economic development of the area.

5 The character of the current tourism development should be identified, including:
- existing and planned tourist accommodations and other facilities;
- existing and planned transportation and other infrastructure;
- existing and potential tourist attractions.

6 The character of the institutional elements should be analysed, including:
- present development policy and plans;
- existing tourism legislation and regulations;
- present investment policy;
- existing overall system of governmental, regional and local authorities, private tourism companies and NGOs;
- existing tourism legislation and regulations;
- existing environmental tourism education and staff training programmes.

7 Detected positive and negative environmental impacts of the current tourism development should be described and analysed (in case 1).
8 Detected positive and negative economic and socio-cultural impacts of the current tourism development should be described and analysed (in case 2).
9 Finally, analysing the survey information and taking into consideration the principles of sustainable tourism development, the respondent's own view on the appropriate policy and impact control measures should be included.

GLOSSARY

Accuracy: the degree of agreement between a measured value and the true value, usually expressed as a positive or negative percentage of the full scale (see below).
Acid soil: soil condition in which the release of hydrogen ions exceeds that of hydroxyl ions, and the soil pH is less than pH 7.0—very acid soils have pH less than pH 5.0.
Acidification: the process by which soils become acid, pH values being reduced to less than pH 5.0, with an increase of hydrogen over hydroxyl ions.
Aggregation: the process in which soil particles coalesce and adhere to form soil aggregates, which are the 'building bricks' of soil structure.
Agricultural industry: the people and processes involved in farming.
Agro-ecosystem: an ecosystem in which management is applied to enable crop and/or animal products to be harvested.
Air quality standard: legal limit on the level of air pollutants in the ambient air, generally aimed at protecting human health.
Aleutian low: area of the North Pacific subject to extended periods of low barometric pressure now thought to influence salmon, squid and zooplankton.
Alkaline soil: soil condition in which the release of hydroxyl ions exceeds that of hydrogen ions, and the soil pH is greater than pH 7.0 (with an upper limit of pH 10.0) — soil base saturation of 100 per cent produces a pH of 7.0 or higher.
Alternative tourism: type of tourism characterised by fewer tourists, changes in tourism types (more extensive, farming, etc.), education of tourists, host and managers and/or less consumerism. It is alternative as compared to the most widespread types of tourism.
Amplitude: the maximum value of a sinusoidal quantity.
Anoxia: the condition where oxygen does not reach body tissues or is not utilised by the tissues; shortage of oxygen.
Aquifer: a layer of permeable rock or unconsolidated material that is capable of storing water and is underlain by impermeable material.
Available soil water: water in soil that can be readily absorbed by plant roots and which is held loosely on soil surfaces under a pressure of 0.3 to about 15 bars.
Available water capacity: AWC is the available water stored by a soil, measured as a weight percentage of the soil. It is placed between field capacity and wilting point.
Ban: prohibition (e.g. to use, import and sell, etc.).
Benthic species: fish and other marine life that dwell on or near the ocean floor.
Benthos: living at the bottom of a sea or a lake.
Best available technology (BAT): techniques that have been shown through actual use to be superior to comparable technologies that are currently considered practical.
Best available technology not entailing excessive cost (BATNEEC): adds a cost factor to the idea of best available technology.
Best practice environmental control technology: see best available technology.
Biocenosis: an aggregate of living creatures (micro-organisms, plants, animals) dwelling in one and the same habitat.
Biodiversity: an ecological concept composed of five key components: species diversity (number of species), genetic diversity (variety of genes within a given population), habitat diversity (variety of biotic and abiotic elements), successional diversity (spectre of plant communities between early and late ecological succession), landscape diversity (landscape as a mosaic of forests in different succession stages).
Biological control: pest and disease control techniques that exploit antagonistic relationships between organisms.
Biological diversity or biodiversity: diversity of species of plants and animals in an ecosystem.
Biomass: weight of living organisms in a particular place.
Biosphere: a part of the planet, which includes the aggregate of living creatures and in which permanent life is possible.
Biota: living organisms.
Biotron: photographic equipment developed to make good imaging of underground biota.

Broad-leaved species: also known as hardwood, nonconiferous, or angiosperm species. This is a botanical classification, based on structure of seed and wood, and not all broad-leaved species have broad leaves. Trees are either broad-leaved or coniferous.

Buffering capacity: the ability of a soil, because of the buffers like clay, humus and carbonates it contains, to prevent rapid changes in pH when acids are added to the soil.

Built environment: man-made physical features, including buildings, infrastructure, archaeological and historical sites.

Bulk capacity: mass per unit volume of soil, sampled as a clod or core, after being dried to a constant weight at 105°C (to exclude moisture).

Calibration: the comparison of an instrument's measurement capability with a standard, in order to report, or to eliminate by adjustment, any deviation in the accuracy of the instrument.

Capacity expansion rate: added energy facility expansion, in discrete or continuous terms, over a fixed time horizon.

Capillarity: the process (otherwise called capillary action) by which soil moisture may move as skins or films of moisture through the fine or capillary pores of the soil, under surface tension force between the moisture and soil particles. Capillary moisture moves from wetter areas of the soil (under low surface tension force) into drier areas (under high surface tension force).

Carbon monoxide: a colourless, odourless gas slightly lighter than air. As an asphyxiant it can cause stress, headaches, fatigue, respiratory problems and ultimately death.

Carbon quota: the ratio of fossil fuel to total energy use measures in appropriate energy efficiency units.

Carnivore: meat-eating animal.

Carrying capacity: level of biomass and/or yield that can be sustained in the long term in relation to the productivity of a particular ecosystem or region with regard to the diversity, structure and trophodynamics of populations and communities within that system.

Charge: price demanded for service or goods.

Clay: either mineral particles less than 2 μm in diameter, *or* a class of soil texture, *or* silicate clay minerals.

Cleaner production: a continuous and integrated preventive approach aimed at lowering environmental risks and impacts of processes and products. It is based on reducing the quantity and toxicity of emissions and wastes at source and before they leave the production process, and on minimising environmental impacts over the entire product life cycle, from materials extraction to final disposal.

Clogging: obstructing, filling up to hinder free passage, action or function.

Closed-cycle economic system: economic organisation and structure based on cyclic systems, from production to reproduction, where what is made mainly originates from what has already been made, or utilised and wasted. In this circular context, each operator's output becomes the next operator's input, as in biological systems.

Cluster analysis: various methods of grouping variables according to the magnitudes and interrelationships among their correlation coefficients.

Coagulation: clotting of body fluids or blood.

Cohort: all fish born in the same year.

Command and control: a directive strategy of environmental policy characterised by authorities who provide regulations. To ensure that the target groups of environmental policy apply the regulations a strict control and monitoring policy is necessary. Examples of command and control measures in environmental policy are, for example, standards, acts for clean-up and liability determination and pollution permits.

Compaction: a process by which soil particles are packed closer together, reducing pore space and increasing built density values.

Composting: a biological decomposition process that turns wet, organic, degradable waste into a soil conditioner which is used to improve soil quality.

Concentrate and contain: concerns a procedure whereby waste is buried in such a way that any non-controlled contact with the environment is believed to be excluded for long periods, so that radioactive or hazardous compounds cannot be taken up in biological cycles.

Conifer or coniferous species: softwood or gymnosperm species, distinct from broad-leaved (hardwood) species; mostly evergreen.

Cradle to the grave: an approach entailing the environmental aspects over the whole lifetime of a product, from its crude resource and energy demand over the different stages of production and consumption to its final treatment as waste.

Delay and decay: involves a procedure based upon temporary storage of highly radioactive waste in a suitable cooled vessel, until a sizeable fraction of radioactivity has been decayed and radiation and heat generation have declined to more tractable levels.

Dematerialising: the increasing incidence of the immaterial components over the material ones in the value of goods. These are conceived and realised as service goods, where R&D, design and technical assistance are assuming an increasing importance in all economic activities.

Denitrification: the biological reduction of nitrate to ammonia, resulting in a loss of nitrogen from the soil into the atmosphere.

Deposit-refund system: in buying a good the price is raised by a mark-up that will be refunded (partly) when the good is delivered at a particular place (and in a particular shape).

Deposition: fall out; the process that removes pollution from the air into the water or the soil.

Detrivore: organism that feeds on detritus or organic matter, mostly litter and woody debris.

Dilute and disperse: low radioactive or hazardous substances are discharged into the atmosphere or into water. It is assumed that the absorptive power of these compartments is able to retain, dilute, decompose or neutralise the waste.

Dispersion: the process by which the structure or aggregation of the soil is broken down, so that each particle is separated and dispersed from the others.

Driftnetting: capturing fish using a net suspended horizontally in the water which catches fish by the gills.

Durability: time span during which a product performs its function, starting from the delivery to a final consumer and ending when the product is discarded as a waste. It is generally referred to as multiple-use good.

Dustfall can: instrument to measure dust deposition.

Dwelled environment: a part of the environment where the interaction of an organism with its nearest medium (organic or mineral) occurs.

Eco-efficiency: low level of environmental burdens per output unit. It is reached by the delivery of competitively priced goods and services that satisfy human needs and bring quality of life, while progressively reducing ecological impacts and resource intensity throughout the life cycle, to a level at least in line with the Earth's estimated carrying capacity.

Eco-tourism: tourism to relatively undisturbed natural areas with the specific objective of studying, admiring and enjoying the scenery, flora and fauna, as well as any existing cultural manifestations.

Ecosystem: complex of living organisms and their environment.

Edaphic environment: from the Greek *edaphos* meaning ground or soil, the term describes the soil environment.

El Niño: infusions of warm surface waters originating in the mid-Pacific into areas off Ecuador and Peru occurring irregularly every 3–10 years for 6–18 months, during which nutrients are prevented from rising to the surface from the colder layers of water below causing disruption of the food chain and significant reductions in marine populations.

Emergency standard: (legal) limit designed to protect against acute incidents.

End-of-pipe technology: refers to pollution management at the moment pollution has been generated. Contrasts to a prevention approach in environmental management and policy. Waste incinerators, air pollution filters and wastewater treatment plants are examples of end-of-pipe technology.

Environment friendly: poorly defined term generally referring to a situation causing less or minimal stress for the environment.

Environmental auditing: a systematic documented periodic evaluation of how well environmental management systems and equipment are performing with the aim of helping to safeguard the environment by facilitating management control of environmental practices; assessing compliance with company policies, which would include meeting regulatory requirements.

Environmental care system: see environmental management system.

Environmental competitiveness: a company should not only be able to compete at the level of prices and quality of products. Other elements such as safety, occupational health and environmental care are also in the focus of management. The elements that provide the company with an environmental record and allow it to compete at the level of environmental quality belong to the environmental competitiveness sphere.

Environmental impact assessment (ELA): an instrument used to aid and improve the decision-making process. The objective is to determine the potential environmental, social and health effects of a proposed development project. It attempts to assess these effects in a form that permits a logical and rational decision to be made. Attempts can be made to reduce or mitigate any potential adverse impacts.

Environmental management system: the organisational structure, responsibilities, practices, procedures, process and resources for determining and implementing environmental policy.

Environmental stress: economic, social or population factors changing the quality of air, water, soil or ecosystems.

Environmental utilisation space: the quantity of natural resources and services which the ecosystems can provide without reducing their productive capacity or generating irreversible changes in their essential parts.

Exclusive economic zone: a zone of 200 natural miles off the shore of coastal states as defined in the Law of the Sea, allowing a state to exploit this zone exclusively.

Fail safe reactor: a new generation of nuclear reactors (on the drawing board) that are inherently safe because of preventive diagnostic steps.

Fall time (90 per cent): the time interval between initial response and 90 per cent response after a decrease in concentration.

Field capacity: the water that is retained by the soil after excess has drained away freely (by gravity); measured as a percentage of over-dry soil.

Fine earth: the fine fraction (less than 2 mm diameter) of the soil, which includes the sand, silt and clay fractions and is used for soil analytical tests.

Fisher-Pry model: a model of technology substitution in which the new emerging technology replaces the existing one in a dynamics governed by the logistic curve.

Food chain: the series of interactions that occur among organisms in efforts to obtain food and energy. Because of energy loss in each transfer, most food chains involve only four or five links from beginning transfer of energy from primary producers through a series of consumers to the ultimate large consumer.

Food web: interconnected food chains that may consist of grazing food chains involving direct consumption or detritus food chains which involve decomposition of dead matter.

Forest: (a) noun (ecological sense of term): a plant community predominantly of trees and other woody vegetation, growing more or less closely together; (b) verb: to cover with trees; the term includes both afforest and reforest.

Forestry: science, business and art of creating, conserving and managing forests and forest lands for the continuing use of their resources, material or other.

Frequency: the number of times per second at which the sound pressure disturbance oscillates between positive and negative values. This oscillation rate is called frequency. It is the reciprocal of the fundamental period.

Full scale: the maximum measuring limit for a given operational range.

Gasification: the conversion of a solid or liquid combustible at high temperature into a gaseous fuel by means of a reactive gas (O_2, CO_2, H_2O, etc.).

General circulation model: a dynamic physical model of atmosphere gas interaction characterised by structural equations subject to large-scale modelling and simulation.

Genetic diverstiy: heritable variation within and between species.

Global commons: goods and values belonging to the world community.

Global warming potential: for some substances that contribute to the enhancement of the greenhouse effect parameters have been developed in the form of Global Warming Potential (GWP). These parameters can be used to express the potential contribution of these substances to the greenhouse effect in a single effect score. The GWP is a relative parameter which uses CO_2 as a reference. In this way atmospheric emissions (in kg) can be converted to CO_2 emissions (in kg) resulting in an equivalent greenhouse effect.

Green tourism: tourism characterised by minimal impacts on environment and landscape.

Guideline: an estimation of the highest allowable concentration which still guarantees a reasonable health or environmental quality condition. A guideline has scientific but no legal value.

Habitat: a place that is home to an organism or group of organisms, either plant or animal.

Hazardous waste: contains chemicals or compounds in excess of a given concentration as reported in *ad hoc* inclusive lists.

Herbivore: organism feeding on plants.

Homotherm: having the same temperature.

Humus: decomposed organic matter, usually on the soil surface, that is dark brown and amorphous, having lost most traces of plant and animal material from which it came.

Hydrologic cycle: the path followed by water in nature. Atmospheric water is precipitated as rain, snow, sleet, hail and dew. Some of the water falling on the surface will run off directly to streams, rivers, lakes and oceans. The remainder will percolate into the ground. A portion of this will go into the water table and be stored. Some will be taken up by vegetation and eventually transpired back into the atmosphere. Some of the groundwater storage will be withdrawn for use and be discharged to receiving waters. Waters on the surface of the ground and on the surfaces of water bodies will evaporate, thus completing the cycle.

Incineration: the controlled burning of solid, liquid or gaseous combustible waste to produce gas and residues containing as little combustible material as possible.

Industrial metabolism: in analogy to biological sciences, the integrated set of physical processes that convert raw materials and energy, plus labour, into finished products and wastes in a (more or less) steady-state condition. Industrial metabolism may refer to the whole economic system or to a single enterprise, the latter being the analogue of a living organism in biology. This analogy between firms and organisms can be carried further, resulting in the notion of 'industrial ecology'. Just as an ecosystem is a balanced, interdependent quasi-stable community of organisms living together, so its industrial analogue may be described as a balanced, quasi-stable collection of interdependent firms belonging to the same economy. The interactions between organisms in an ecosystem range from predation and/or parasitism to various forms of co-operation and synergy. Much the same can be said for firms in an economy.

Infiltration: the downward entry of water into the pore space of the soil. The capacity of a soil to absorb water, depending on vegetation cover, moisture content and volume of pore space, is the infiltration capacity.

Integrated farming: the practical expression of the concept of sustainable agriculture. Integrated farming seeks to reconcile objectives for production and environmental protection.

Integrated fertiliser management: use of the optimum combination of mineral fertilisers, organic manures and biological fixation of nitrogen together with appropriate techniques of soil management for the maintenance of soil fertility.

Integrated pest management: the use in optimum combination of a range of techniques for control of pests in agriculture. The techniques include selective use of pesticides, biological control, disease resistance, interference with reproduction of the pest and environmental manipulations such as systems of mixed planting and crop rotation.

Integrated rural development: an approach to development of rural areas which brings together all aspects of rural enterprise and society.

Interference: an undesired positive or negative output caused by a substance other than the one being measured.

Irrigation: the process of artificially increasing the water available to plants. Water may be introduced by spraying on to the soil surface, or to root systems through channels and ditches.

Kiss and ride: used to describe being driven to a rail or bus station and continuing the journey by rail or bus.

Lag time: the time interval between a step change in the input concentration at the instrument inlet and a reading of 90 per cent of the ultimately recorded concentration.

Large marine ecosystem (LME): an ecologically distinct area of ocean at least 200,000 square kilometres in area treated as a single management zone in ecosystem management approaches.

Leaching: the removal of soil compounds in solution and in physical suspension by weak acid solution (leachate) from the A into the B horizon (and beyond).

Limits of acceptable change: clarify how much change (in nature, pollution, economy, social relations, culture) the acceptors in the tourism system can accept without suffering from adverse effects.

Linearity: the maximum deviation of the measured value from the corresponding value on a straight line curve drawn between upper and lower calibration points.

Longlining: capturing fish using a line carrying multiple baited hooks.

Loudness: a means for rating a noise based on the judgement of a person or a listening jury.

Lower detectable limit: the smallest amount of input concentration that can be detected as the concentration approaches zero, usually twice the noise level at zero concentration.

Material balance: a comparison between input and output flows of materials and products allowing the characterisation of yield and efficiency, as well as the nature and amount of individual directed or fugitive waste streams.

Maximum sustainable yield (MSY): the level of catching producing the highest yield that can be continued without depleting the stock.

Mesofauna: small invertebrates.

Methaemoglobinaemia: a condition in which infants, usually less than two months old, fed on formulae prepared with water having high nitrate concentration, become cyanotic. Nitrate is reduced in the intestine to nitrite, which combines with haemoglobin, rendering it incapable of carrying oxygen.

Mycorrhizae: symbiotic fungi living inside or outside the root system of most plants.

Natural environment: refers to what exists in nature: climate, weather, soil, water, air, fauna, flora and ecosystems.

NIMBY syndrome: 'not in my backyard' syndrome, refers to what is believed to be a general attitude of consumers unwilling to accept the environmental consequences of pollution and/or waste treatment.

NIMT syndrome: 'not in my term' syndrome refers to a widespread attitude among politicians to delay sensitive decisions so that it becomes unnecessary to handle them during their term of responsibility.

Nitrification: the oxidation of ammonia to nitrite, and nitrite to nitrate, by micro-organisms, so increasing the store of nitrogen in the soil.

Nitrogen oxides (NO_x): include nitric oxide (NO) and nitrogen dioxide (NO_2). Nitric oxide is a colourless, odourless gas. Nitrogen dioxide is a reddish brown toxic gas, which adds to acidification and photochemical pollution. It can cause throat and eye irritation and exacerbate asthma and increase susceptibility to infections.

Noise: (a) any disagreeable or undesired sound; (b) sound, generally of a random nature, the spectrum of which does not exhibit clearly defined frequency components. Note: by extension of the above definitions, noise may consist of electrical oscillations of an undesired or random nature. If ambiguity exists as to the nature of the noise, a term such as acoustic noise or electrical noise should be used.

Noise dose: the noise dose is a measure of the total A-weighted sound energy received by an employee, and is expressed as a percentage of the allowed daily noise dose. It therefore depends not only on the level of the noise but also on the length of time that the employee is exposed to it.

Noisiness: subjective rating scale from basic acoustic measurements.

Octave: the interval between two frequencies which have a frequency ratio of two.

Open-cycle economic systems: economic organisation and structure where the physical inputs needed for production are mainly taken from nature, thus contributing to resource depletion, and waste materials are released into the environment contributing to its pollution.

Optimum yield (OY): a yield, generally lower than the MSY, from the level of catching that maximises the excess of revenue over costs.

Organic farming: systems of agriculture that conform to the standards of the International Federation of Organic Agriculture Movements and which avoid the use of synthetic agro-chemicals especially through exploiting biological control, crop rotations and recycling.

Ozone: a bluish gas about 1.6 times as heavy as air. It is indicative of photochemical pollution. It influences the respiratory capacity. It is highly reactive and capable of attacking surfaces and rubber materials. It is toxic to trees, crops and vegetation.

Parasitoid: in general an insect that deposits its eggs inside a living prey, usually another invertebrate.

Park and ride: a term used to describe access to a rail or bus station when the car is parked at the station.

Pheromone: a chemical secreted to the outside by an animal and which causes a behavioural response in others of the same species, for example, insect pheromones that attract mates.

Phon: the numerical value of the sound pressure level of a 1000 Hz tone having the same loudness as the considered sound.

Photic zone: the near-surface layer of water, in coastal areas typically 10 to 30 metres deep, through which sufficient light penetrates to permit growth of phytoplankton.

Physical environment: synonym for natural environment.

Phytoplankton: tiny floating aquatic plants that form the first link in the marine food chain.

Poaching: poaching or puddling is the physical process in which soil structure is destroyed by the impact pressure from trampling by animals or wheeled machinery, while the soil is saturated.

Polder: a low-lying flat area reclaimed from the sea and protected by embankments from invasion by the sea. Polder drainage is controlled by pumping into the sea.

Policulture: cropping in the same field or plot several crops at the same time.

Pollination: operation performed by pollinators or by other factors such as wind, rain, etc.

Pollinator: an animal, mostly insects, capable of disseminating pollens among flowers.

'Pollutee suffers and pays' principle: collateral of the 'polluter pays' principle. Refers to the fact that the polluter only pays if discovered and prosecuted. In a number of other cases the bill is presented to the pollutee who not only pays but also suffers from the consequences of pollution.

'Polluter pays' principle: suggests that the polluter should bear the cost of preventing and controlling pollution. The intent is to force polluters to internalise all the environmental costs of their activities so that these are fully reflected in the costs of the goods and services they provide.

Pollution prevention pays: refers mainly to in-plant industry waste management practices, directed towards identifying sources of avoidable losses and contaminations and eliminating them if possible by simple means or changes to procedures.

Polycyclic aromatic hydrocarbons: produced by the incomplete combustion of fuels. They entail a complex range of chemicals, some of which are carcinogenic.

Precision: the level of agreement between repeated measurements of the same value, expressed as the average deviation of the data from the mean.

Predator: organism feeding on other organisms.

Principal component analysis: a branch of factor analysis aimed at successfully analysing the greatest, second greatest and successively smaller sources of variation.

Productivity: the mass of living matter synthesised by plants, animals or an aggregate of species comprising a biocenosis, per unit of area and per unit of time.

Proximity principle: calls for waste elimination at a reasonably short distance from its point of generation.

Pyrolysis: decomposition by heating in the absence of oxygen.

Quality assessment: procedures that monitor the quality control mechanism and evaluate the quality of the data produced.

Quality control: the mechanism used to reduce and maintain random and systematic errors within tolerable limits.

Re-industrialisation: the transformation process in manufacturing industries—started in the 1980s—having very different, and even diametrically opposed features from those of the industrial revolution: flexibility against rigidity; high and increasing product diversification in place of simplification; companies 'as project' rather than 'as structure'; increasing importance of 'non-material' inputs and of intellectual employment; further intensification of capital requirements in relation to labour.

Re-producing: a set of activities that lead to closing the circle of material convection employing potentially circular production processes, which re-use products and materials to the utmost and minimise the use of non-renewable sources.

Reclamation: the process or processes of change that create fertile and productive agricultural land from a previously unproductive state, by using techniques such as drainage, irrigation, deep ploughing, lining and fertilisation.

Recruitment: addition of individuals of a particular age ('recruits') to the fish stock, the age most commonly defined as the transition from larvae to juvenile or the age of initial vulnerability to capture by gear commonly used in fishing the stock.

Recyclability: aptitude of a material to be recovered and re-used, at the end of a given product's lifespan, for obtaining new products, either of the same type (closed-loop recycling) or of different—usually lower-grade—type (open-loop recycling).

Refugium: place in which living organisms have moved in, usually forced by human activities (e.g. agriculture, urbanisation), but which is still hospitable enough to allow reproduction.

Regenerative capacity: the ability of natural systems to recover from environmental pressure.

Relict vegetation: the set of plants belonging to the native forest in a modified rural landscape.

Remote sensing: technique of observation at large distances, often referring to satellite images of the Earth's surface.

Resource steady state: a condition where the rate of depletion of exhaustible resources should be no greater than the rate of increase of man-made substitutes. As far as renewable resources are concerned, the steady state is achieved if resources are exploited at a rate no greater than their natural replacement rate.

Responsible tourism: in the sense of responsible towards the environment, the term covers at least two possible views: (a) passive: tourism that minimises disturbance of the environment; (b) active: tourism that contributes to sustainable management of the environment.

Rise time (90 per cent): the time interval between initial response and 90 per cent response, after a step increase in inlet concentration.

Rizotron: photographic equipment developed to take images of underground root systems.

Road dust: highly carcinogenic chemical; also referred to as fugitive dust.

Road pricing: method of charging for the use of roads based on an allocative efficiency criterion such as prices should reflect the marginal cost of use including externalities such as congestion.

Rotation: crop sequence in the same field.

Salinisation: the process of accumulating soluble (mainly sodium) salts in surface horizons, by an upward capillary movement from saline groundwater.

Saturation: the maximum amount of water that can be contained in rock or soil when all the pore space is filled with water. It is measured as a percentage of the dry weight of rock or soil.

Set-aside policy: the policy of taking arable land out of production to reduce crop surpluses, often accompanied by compensation payments to farmers.

Social engineering: engineering approach incorporating social and ethical values.

Soil conservation: the management of soil to sustain future yield, by protection from soil erosion and deterioration through chemical pollution.

Soil erosion: the removal of material from the land surface, by running water, moving ice, wind and by gravity.

Soil horizon: relatively uniform and horizontal soil layers that are formed by processes of change (weathering and organic decay) and relocation of materials within the soil.

Soil management: a range of practices and operations on soil that encourage the production of agricultural crops and have regard for sustained yield in the future.

Soil structure: the aggregation of particles into larger units, called the peds of soil structure. It may be compared with scaffolding (inside the soil).

Soil texture: the texture or feel of the soil fine earth depends on the relative composition of sand, silt and clay fractions present.

Solid waste: a generic name for different waste streams that might be handled or eliminated in a way that is common for real solids in industrial waste disposal centres. Therefore it also includes non-solids, such as liquors, sludges, slurries, pastes and solid residues from gas or wastewater treatment.

Sound: a disturbance that propagates through an elastic material at a speed characteristic of that medium, (a) the sensation of hearing excited by an acoustic oscillation; (b) acoustic oscillation of such a character as to be capable of exciting the sensation of hearing; (c) an oscillation in pressure, stress, particle velocity, etc. in a medium with internal forces.

Sound pressure: the incremental variation in pressure above and below atmospheric pressure, which, in turn, is normally about 1.013×10^5 N/m^2 in metric units at sea level. Instead of N/m^2, the unit pascal is often used.

Span drift: the change in instrument output over a stated period, usually 24 hours; of unadjusted continuous operation when input concentration is a stated up-scale value, usually expressed as a percentage of the full scale.

Spectrometry: analytical chemical technique based upon the principle that each substance emits a particular spectrum of radiation if irradiated at a defined frequency.

Steady state: a stationary equilibrium situation in which interacting variables control each other in a balanced manner.

Stewardship (environmental): the responsibility of looking after the environment.

Strategic environmental assessment (SEA): the application of EIA at the level of policies, plans, programmes or other human activities that are more than an individual project.

Subsidy: amount of money paid by a public authority to a polluter for every unit less of pollution that is emitted below a fixed bench-mark level. The reverse of a tax or a charge.

Sulphur dioxide (SO_2): colourless acidic gas with a choking taste. It adds to acidification, can impair the respiratory system and corrodes stonework and other materials.

Superfund legislation: enables blame for environmental damage to be apportioned to a selected one of many past mine owners and for that company to be charged an estimated cost for work which the government contracts-in to clean up and rehabilitate the damaged site.

Suspended particulate matter: covers a range of fine solids or liquids dispersed in the atmosphere. Especially fine particles are associated with mortality incidence and increased heart and lung disease.

Sustainability: the long-term non-impairment of the capacity of the natural system to regenerate raw material inputs and to absorb waste outputs of the human economy.

Sustainable agriculture: systems of farming that maintain the resources upon which production depends.

Sustainable management: supports multiple uses (including biodiversity preservation, timber harvesting, extraction of bush meat and other non-wood products, soil and water conservation, tourism recreation and enjoyment of natural amenities) based on an ecosystem concept that allows utilisation of forests without undermining their use by present and future generations. Different systems of management would be required for each category of forest depending on the intended output.

Sustainable tourism: tourism matching environmental, social and economic aspects world-wide and transgenerational. It involves non-consumptive use of natural resources, protects the culture and well-being of local communities and promotes equity into and between generations.

Technology transfer: transfer of capital goods, engineering services and equipment designs: the physical items of the investment, complemented by training in the skills and know-how for operating the plant and equipment.

Threshold: value that should not be exceeded to guarantee a defined (e.g. environmental) quality level.

Tourism: includes a heterogenous set of activities and services related to the temporary moving of individuals outside their usual place of residence for purposes of entertainment, rest, culture, health care, and generally for reasons other than income-earning activities.

Toxic waste: see hazardous waste.

Transformation: the process of changing from one state into another state.

Trawl netting: capturing fish using a dragged net open at one end.

Uncertainties: (a) benign: those gaps in knowledge that make two good strategies better; (b) malign: those gaps in knowledge that make it impossible to determine whether the outcome of a particular policy will be tolerably good or bad.

Underdrainage: the artificial installation of various types of pipe drains under the top soil of poorly draining land (otherwise field drainage).

Volatile organic compounds: released in vehicle exhaust gases as burned fuels and emitted by the evaporation of solvents and motor fuels. They play an important role in photochemical pollution. Some of them as benzene and 1,3-butadiene are carcinogens that cause leukaemia.

Wait and see approach: attitude in which one waits for further (scientific) results before taking action.

Waste: anything that is discarded, will be discarded or has to be discarded on the basis of prevailing legal codes.

Watershed: area draining ultimately to a particular watercourse or body of water.

Willingness to pay: economic concept that measures the amount in dollar terms that an individual is willing to pay to secure a service. It is closely associated with the concept of consumer surplus which is a measure of the difference between what a person has to pay and is willing to pay.

Wilting point: the point in the drying-out of soil at which the soil surface holds moisture with greater surface tension than plant roots can exert to attract moisture. Plants wilt die for lack of moisture intake.

Zero drift: the change in instrument output over a stated period, usually 24 hours, of unadjusted continuous operation when the input concentration is zero, usually expressed as a percentage of the full scale.

Zooplankton: tiny marine animals that feed on phytoplankton and are in turn fed upon by larger marine creatures.

ANSWERS TO SELF-ASSESSMENT QUESTIONS

CHAPTER 1

1 The European Soil Charter of the EU (1972 and 1982) includes evaluation of soil quality for current and possible uses, state of soil degradation, as well as assessment of the soil's resilience to chemical pollution. Soil protection policy must be international because of distances involved, and frequent monitoring of soil is necessary.
2 Soil management (here) is for optimal and sustained agricultural production, and allows for measures for soil improvement by drainage and controlled fertiliser application, compatible with comments in answer to question 1. Soil protection of very sensitive soils may severely limit agricultural use.
3 (a) Soil reclamation of the Dutch polders involves the selection of marine off-shore areas with ad equate clay content, the building of sea dykes and pumping stations, the leaching of saline salts and liming with gypsum,
(b) Soil reclamation of Danish heathlands involves stabilisation of coastal sand dunes, tree plant ing of shelter belts and other soil surface protection measures, liming and fertiliser application. Sprinkler irrigation is required.
4 (a) Wind erosion is most likely in extensive, exposed and level plains in Central and Northern Europe and where the climate has a high frequency of wind.
(b) Water erosion is most common in wet climates or in semi-arid areas with flash floods. Steep slopes and valley terrain increases the risk. Both types of erosion are more likely with reduction of a protective plant cover.
5 Water used by plants is capillary water, which can be maintained as skins of moisture by soil drainage or by irrigation, in the zone above the water table.
6 The best type of soil structure for agriculture is a well-bound stable structure, usually with crumb-shaped or blocky peds and 35–55 per cent pore space within the soil.
7 Organic matter is important in soil management because it helps bind soil structure, holds and exchanges nutrient ions, increases water-holding capacity, is a store of certain plant nutrients (N, P, S) and gives physical protection to the soil surface.

CHAPTER 2

1 The wrong answer is (d).
2 Correct answer: (1+(b)); (2+(d)); (3+(a)); (4+(c)).
3 The best answer is (d).
4 The best answers are (b), then (c).
5 The wrong answer is (b).

CHAPTER 3

1 (c) and (e), 2 (b), 3 (c), 4 (b), 5 (c).

CHAPTER 4

1 (a), 2 (b), 3 (a), 4 (c), 5 (c).

CHAPTER 5

1 The decibel is most commonly used to describe sound level; it is logarithmically related to sound pressure level. The decibel alone is not useful for describing loudness because perceived loudness depends on the frequency of the sound. Consequently weighted dB scales have been developed.

2 The factors which need to be considered include the sound level, the source (point, line or area), the emission patterns (temporal, spatial, mobility) and generation factors such as the design of the machinery and the factors which affect transmission such as distance. Other factors can be wind direction, turbulence, temperature inversion, absorption by grassland, corn, low shrubs or trees.

3 The safe limit is 85 dB(A) and is equivalent to a dose of 1. Calculation of the noise dose is:

	90 dB(A)	*93 dB(A)*	*96 dB(A)*	*99 dB(A)*	*Total dose Σ t_i/T_i*
Actual exposure time (t_i)	16	4	10	10	
Maximum permitted exposure (T_i)	40	20	10	5	
t_i/T_i	0.4	0.2	1	2	3.6

3.6 is an unacceptable noise dose.

4 There is no exact answer to this question but you should include the following points: economic, health and welfare effects of the surrounding community; a safety programme for workers; the methods that are being considered to minimise sound such as the orientation of runways, rules against take-off and landing of aircraft during the night and on weekends; and the type of aircraft permitted to use the airport. You might also consider the noise impacts of increased motor traffic due to the expansion of the airport

CHAPTER 6

1 (d), 2 (b), 3 (c), 4 (a), 5 (b), 6 (c), 7 (a), 8 (b).

CHAPTER 7

1 (a), 2 (a), (c) and (d), 3 (a), (b) and (c), 4 (c), (d) and (e), 5 (b) and (c).

CHAPTER 8

1 (b), 2 (c), 3 (a), 4 (c), 5 (b), 6 (b).

CHAPTER 9

1 (b), 2 (c), 3 (c), 4 (b), 5 (a).

CHAPTER 10

1 (c), 2 (a), 3 (b), 4 (c), 5 (c), 6 (b).

CHAPTER 11

1 (a), 2 (b), 3 (a) and (b), 4 (c), 5 (a) and (b) 6 (a), (b) and (c).

CHAPTER 12

1 (b), 2 (b), 3 (c), 4 (b), 5 (a), 6 (c), 7 (a), 8 (b).

CHAPTER 13

1 (c), 2 (c), 3 (a), 4 (a), 5 (a), 6 (d), 7 (b), 8 (b).

INDEX

Note: Italic page numbers refer to charts and tables.