ENVIRONMENTAL HYDROLOGY

Water Science and Technology Library

VOLUME 15

Editor-in-Chief
V. P. Singh, *Louisiana State University, Baton Rouge, U.S.A.*

Editorial Advisory Board

M. Anderson, *Bristol, U.K.*
L. Bengtsson, *Lund, Sweden*
A. G. Bobba, *Burlington, Ontario, Canada*
S. Chandra, *Roorkee, U.P., India*
M. Fiorentino, *Potenza, Italy*
W. H. Hager, *Zürich, Switzerland*
N. Harmancioglu, *Izmir, Turkey*
A. R. Rao, *West Lafayette, Indiana, U.S.A.*
M. M. Sherif, *Giza, Egypt*
Shan Xu Wang, *Wuhan, Hubei, P.R. China*
D. Stephenson, *Johannesburg, South Africa*

The titles published in this series are listed at the end of this volume.

ENVIRONMENTAL HYDROLOGY

edited by

VIJAY P. SINGH
Louisiana State University,
Baton Rouge, U.S.A.



SPRINGER-SCIENCE+BUSINESS MEDIA, B.V.

Library of Congress Cataloging-in-Publication Data

Environmental hydrology / edited by Vijay P. Singh.
p. cm.
Includes index.
ISBN 978-90-481-4573-7 ISBN 978-94-017-1439-6 (eBook)
DOI 10.1007/978-94-017-1439-6
1. Hydrology. 2. Hydrology--Environmental aspects. I. Vijay, V. P. (Vijay P.)
GB665.E58 1995
551.48--dc20 95-11259
CIP

ISBN 978-90-481-4573-7

02-0897-100 ts

Printed on acid-free paper

All Rights Reserved
© 1995 Springer Science+Business Media Dordrecht
Originally published by Kluwer Academic Publishers in 1995
No part of the material protected by this copyright notice may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying, recording or by any information storage and retrieval system, without written permission from the copyright owner.

To the People of the United States of America

Table of contents

Preface xiii

1 What is environmental hydrology? 1
V.P. Singh

1.1 Environmental Continuum . . . 1
1.2 Integrated Environmental Management . . . 1
1.3 Water Continuum . . . 2
1.4 Integrated Water Management . . . 3
1.5 Classification of Hydrology . . . 7
1.5.1 Properties of Water . . . 7
1.5.2 Sources of Water . . . 7
1.5.3 Scientific Content . . . 7
1.5.4 Solution Technique . . . 7
1.5.5 Area of Emphasis . . . 8
1.5.6 Basin Size . . . 8
1.5.7 Basin-Type or Land Use . . . 8
1.6 Definition of Environmental Hydrology . . . 8
1.7 Scope of Environmental Hydrology . . . 9
1.8 Role of Environmental Hydrology . . . 10

2 Watershed acidification modelling 13
A.G. Bobba, D.S. Jeffries, W.G. Booty and V.P. Singh

2.1 Introduction . . . 13
2.1.1 Watershed Acidification Modelling . . . 15
2.1.2 Interaction between Lake and Terrestrial Systems . . . 16
2.1.3 Flow Pathways in Watersheds . . . 18
2.1.4 Bio-geochemical Processes: Chemical Species . . . 22
2.1.5 Basic Mechanism of Leaching - Mobile Anions . . . 23
2.1.6 Effect of Acid Rain on Leaching . . . 24
2.1.7 Soil Buffering Systems . . . 25
2.2 Watershed Acidification Models . . . 27
2.2.1 ILWAS Model . . . 27
2.2.2 Birkenes Model . . . 30
2.2.3 TMWAM Model . . . 33
2.2.4 Enhanced Trickle-Down Model . . . 40
2.2.5 RAINS Model . . . 42
2.2.6 MAGIC Model . . . 44
2.4 Stochastic Analysis to Predict Acid Shocks . . . 46
2.4.1 Number of Exceedances . . . 47
2.4.2 Time of Exceedance . . . 47
2.4.3 Time of Events . . . 48
2.4.4 Duration of Exceedances . . . 49
2.4.5 Magnitude of Exceedances . . . 49

2.5 Application of Models to the Turkey Lake . . . 49
2.5.1 Statistical Analysis . . . 51
2.5.2 Acidification Events . . . 56
2.6 Discussion . . . 60
2.7 Conclusions . . . 62

3 Climatic change **69**
M.A. Mimikou
3.1 Introduction . . . 69
3.2 A Brief Overview of the Literature . . . 70
3.3 Climatic Variability and Change . . . 72
3.3.1 Global Greenhouse Warming – Past, Present and Future Conditions 72
3.3.2 Methods for Evaluating Global and Regional Climatic Changes . . 75
3.3.3 Impacts of Climatic Changes . . . 76
3.4 Hydrological Effects of Climatic Change . . . 77
3.4.1 Hydrological System Response to Climatic Changes . . . 77
3.4.2 Methods for Evaluating Regional Hydrological Effects of Climatic Changes . . . 78
3.4.3 Sensitivity of Major Climatic and Hydrological Variables to Global Warming . . . 82
3.5 Climatic Change Impacts on Water Resource Planning and Management . 88
3.5.1 Basic Considerations . . . 88
3.5.2 Sensitivity of Some Critical Water Management Issues to Climatic Change . . . 90
3.6 Future Research Directions . . . 97

4 Understanding river hydrology **107**
B.L. Finlayson and T.A. McMahon
4.1 Introduction . . . 107
4.2 Data . . . 108
4.2.1 Types of Data and Their Sources . . . 108
4.2.2 Data Problems . . . 108
4.2.3 Synthetic Data . . . 109
4.3 Characteristics of the River Hydrograph . . . 110
4.3.1 The Generation of the River Hydrograph . . . 110
4.3.2 Hydrograph Separation . . . 112
4.3.3 Flow Duration Curves . . . 113
4.3.4 High Flows . . . 115
4.3.5 Low Flows . . . 118
4.4 Variability . . . 118
4.4.1 Seasonal Regime . . . 120
4.4.2 Hydrograph Variability . . . 120
4.4.3 Interannual Variability . . . 121
4.5 Catchments and Channels . . . 123
4.5.1 Fluvial Geometry . . . 123
4.5.2 Ungauged Streams . . . 124
4.6 Regulated Rivers . . . 127
4.7 Waterways and Watersheds . . . 129
4.7.1 Influent and Effluent Streams . . . 129
4.7.2 Environmental Aspects . . . 130
4.8 Concluding Statement . . . 131

5 Transport of reacting solutes in rivers and streams 137
R.L. Runkel and K.E. Bencala
5.1 Introduction . . . 137
5.2 Physical Transport Processes . . . 139
5.3 Sources of Solutes . . . 147
5.4 Geochemical Processes . . . 152
5.4.1 Rates of Reaction . . . 153
5.4.2 Precipitation/Dissolution . . . 155
5.4.3 Sorption/Desorption . . . 157
5.4.4 pH, the Master Variable . . . 159
5.5 Coupling of Hydrologic and Geochemical Processes . . . 160
5.6 Summary . . . 163

6 Water and contaminant transport in the vadose zone 165
P.J. Wierenga and M.L. Brusseau
6.1 Introduction . . . 165
6.2 Water Storage . . . 166
6.2.1 Water Content . . . 166
6.2.2 Energy Relationships . . . 166
6.3 Water Movement . . . 170
6.3.1 Steady Flow . . . 170
6.3.2 Hydraulic Conductivity . . . 170
6.3.3 Transient Flow . . . 171
6.3.4 Infiltration . . . 172
6.3.5 Redistribution . . . 175
6.4 Contaminant Transport . . . 175
6.4.1 Equations For Nonreactive Solutes . . . 175
6.4.2 Equations for Reactive Solutes . . . 177
6.4.3 Transport in Uniform Soil: Steady flow . . . 178
6.4.4 Transport in Uniform Soil: Transient flow . . . 180
6.5 Nonideal Contaminant Transport . . . 180
6.5.1 Heterogeneity . . . 180
6.5.2 Preferential Flow . . . 182
6.5.3 Nonlinear, Rate-Limited Sorption . . . 183
6.5.4 Facilitated Transport . . . 186
6.5.5 Vapor and Immiscible-Liquid Phase Transport . . . 186
6.5.6 Coupled Physical, Chemical, and Biological Processes . . . 187
6.6 Conclusion . . . 188

7 Transport of moisture and solutes in the unsaturated zone by preferential flow 193
F. Stagnitti, J.-Y. Parlange, T.S. Steenhuis, J. Boll, B. Pivetz and D.A. Barry
7.1 Introduction . . . 194
7.2 Preferential Moisture and Solute Transport . . . 195
7.3 Sampling for Preferential Flow in the Unsaturated Zone . . . 199
7.4 Characterisation of Soil Structure . . . 200
7.5 Modelling Preferential Transport . . . 203
7.6 Transport of Pesticides . . . 210
7.7 Influence of Macropores on Biological Degradation of Pesticides . . . 213
7.7.1 Biodegradation of p-Nitrophenol (PNP) . . . 213
7.7.2 Biodegradation of 2,4-D (2,4-dichlorophenoxyacetic acid) . . . 216
7.8 Summary and Conclusions . . . 218

8 Groundwater contamination modelling 225
A.G. Bobba and V.P. Singh
8.1 Introduction . . . 225
8.2 Classification of Groundwater Contamination Models . . . 227
8.2.1 Models . . . 227
8.2.2 Model Development . . . 227
8.3 Groundwater Modelling . . . 228
8.3.1 Principles and Concepts used in Groundwater Modelling . . . 229
8.3.2 Darcy's Law . . . 229
8.3.3 Hubbert's Force Potential . . . 231
8.3.4 Conservation of Mass . . . 232
8.4 Flow Models . . . 232
8.4.1 Basic Assumptions . . . 232
8.4.2 Two Dimensional Horizontal Flow . . . 233
8.4.3 Definition of boundary and initial conditions . . . 233
8.4.4 Initial conditions . . . 235
8.5 Contaminant Transport Models . . . 236
8.5.1 Advective Dispersion Phenomena . . . 236
8.5.2 Basic Assumptions . . . 236
8.5.3 Advective Dispersion Equation in Cartesian Coordinates . . . 238
8.5.4 Boundary and Initial conditions . . . 238
8.6 Hydrodynamic Dispersion . . . 239
8.6.1 Effects of Dispersion . . . 241
8.6.2 Quantification of Dispersion . . . 242
8.6.3 Determination of Coefficient of Molecular Diffusion . . . 243
8.7 Chemical and Biological Activity . . . 244
8.7.1 Chemical Processes . . . 244
8.7.2 Linear Adsorption . . . 246
8.7.3 Freundlich Isotherm . . . 246
8.7.4 Langmuir Isotherm . . . 246
8.7.5 Biological Processes . . . 247
8.8 Development of Contaminant Transport Models . . . 247
8.8.1 Analytical and Numerical Models . . . 248
8.8.2 Types of models . . . 248
8.8.3 Analytical Models . . . 250
8.8.4 Analytical Element Method Models . . . 251
8.8.5 Finite Difference Models . . . 252
8.8.6 Finite Element Models . . . 260
8.8.7 Application of Galerkin's method . . . 264
8.8.8 Application of Green's theorem . . . 265
8.8.9 Boundary Element Models . . . 271
8.8.10 Velocity Based Modelling . . . 272
8.9 Accuracy of Numerical Models . . . 272
8.9.1 Numerical Dispersion . . . 273
8.10 Available Numerical Models . . . 273
8.10.1 General Guidance for Groundwater Contamination Model Use . . . 274
8.10.2 Data Requirements . . . 276
8.10.3 Physical Framework . . . 276
8.11 Stresses on System . . . 277
8.11.1 Groundwater Flow . . . 277
8.11.2 Contaminant Transport (in addition to above) . . . 277
8.12 Physical Interpretation . . . 277

8.12.1 Effect of Velocity Ratios on Contaminant Transport Model 278
8.12.2 Effect of Adsorption on Contaminant Concentration 279
8.12.3 Effect of Velocity Ratios on Concentration 283
8.12.4 Effect of Dispersivity on Contaminant Transport 284
8.12.5 Contaminant Recovery Phase 287
8.13 Application of Groundwater flow and Contaminant Transport Models . . . 288
8.13.1 Prot Granby Radioactive Disposal Site, Ontario, Canada 288
8.13.2 Application of SUTRA model to Lambton county, Ontario, Canada 295
8.13.3 Numerical Results 306
8.14 Limitations and Source of Error in Groundwater Contamination Models . . 310
8.14.1 Model Misuse 313
8.15 Improvements in Modelling 314
8.15.1 Output Formats 314
8.15.2 Data Collection Methods 314
8.15.3 Further Development of Flow and Transport Models 314
8.16 Further Research in Chemical Kinetics 315
8.17 Summary 315

9 Modeling subsurface transport of microorganisms **321**
Y. Tan and W.J. Bond
9.1 Introduction 321
9.1.1 Development of Mechanistic Models of Transport 322
9.1.2 Approaches to the Pore-Scale Distribution of Bacteria 323
9.2 Processes Affecting Transport and Fate of Microorganisms 324
9.2.1 Movement of Microorganisms 324
9.2.2 Growth and Decay of Microorganisms 327
9.2.3 Attachment and Detachment Processes 331
9.2.4 Mathematical Description of Attachment and Detachment 334
9.3 Transport of Microorganisms in Porous Media 337
9.3.1 General Transport Equation for Microorganisms 337
9.3.2 Special Cases of the Microbial Transport Equation 339
9.3.3 Simultaneous Transport of Substrates with Bacteria 340
9.4 Comparison of Models with Experiments 342
9.4.1 Transport of Bacteria 342
9.4.2 Transport of Viruses 346
9.4.3 Summary of Experimental Studies 347
9.5 Summary and Conclusions 347

10 Assessment and control of loading uncertainty for managing eutrophication and toxic chemical fate in lakes **357**
T.C. Young and D.M. Dolan
10.1 Introduction 357
10.2 Loadings and Sources of Uncertainty 359
10.2.1 Tradeoffs 360
10.3 Assessing Loading Uncertainty 360
10.3.1 Demonstration 362
10.4 Managing Loading Uncertainty 363
10.4.1 Tributary Mouth Loadings 365
10.4.2 Point Sources 369
10.4.3 Non-Point Sources 370
10.4.4 Data Handling Improvements 373
10.5 Conclusions 373

11 Modeling watershed water quality 377
A.S. Donigian, J.C. Imhoff and R.B. Ambrose, Jr.
11.1 Introduction . . . 377
11.2 Evolution of Watershed-Scale Water Quality Modeling . . . 378
11.3 Why Model? . . . 379
11.3.1 Role of Modeling in Water Quality Analyses . . . 380
11.3.2 Environmental Analyses That Benefit from Modeling . . . 380
11.4 Introduction to Watershed Models and Modeling . . . 381
11.4.1 Modeling Fundamentals . . . 381
11.4.2 Relevant Model Features . . . 383
11.4.3 Watershed Water Quality Modeling Components . . . 387
11.5 Methods and Models for Nonpoint Source Water Quality Modeling . . . 387
11.5.1 Urban Runoff Modeling Methods . . . 387
11.5.2 Urban Runoff Models . . . 392
11.5.3 Non-Urban Runoff Modeling Methods . . . 395
11.5.4 Non-Urban Runoff Models . . . 399
11.5.5 Guidance on Nonpoint Source Model Selection and Use . . . 405
11.6 Methods and Models for Surface Receiving Water Quality Modeling . . . 407
11.6.1 Methods . . . 407
11.6.2 Guidance on Receiving Water Model Selection and Use . . . 415
11.7 The Future of Watershed Water Quality Modeling . . . 416
11.7.1 Algorithm/Computation Enhancements . . . 417
11.7.2 User Interaction . . . 419
11.8 Closure . . . 421

12 Eco-hydrological consequences of environmental degradation: Hydrology, ecology and environmental impacts 427
P.H. Templet and J. Sorensen
12.1 Introduction . . . 427
12.2 Environmental Change . . . 429
12.2.1 A Partial List of the Agents of Change . . . 429
12.3 Comparative Hydrology and Landforms . . . 429
12.4 Weighting Impacts . . . 431
12.5 Overview of The Planning Process . . . 437
12.5.1 Water Conservation . . . 439
12.6 Detailed Planning and Impact Assessment Process . . . 439
12.6.1 Description of the Proposed Action . . . 440
12.6.2 Screening . . . 441
12.6.3 Scoping . . . 446
12.6.4 Identification and Selection of Alternatives . . . 448
12.6.5 Presentation . . . 455
12.6.6 Mitigation of Adverse Impacts . . . 456
12.6.7 Review and Comment . . . 457
12.6.8 The Decision . . . 458
12.6.9 Auditing or Monitoring Impacts . . . 460
12.7 Conclusion . . . 460

Glossary of terms 463

Contributors 475

Subject Index 477

Preface

In environmental planning, management and restoration, hydrology plays a fundamental role. Collection of hydrologic concepts that apply to solution of environmental problems constitutes the subject matter of environmental hydrology. A book exclusively devoted to discussion of these concepts and their application to environmental issues is lacking. Because of the increasing significance of environmental management, an attempt has been made to cover some aspects of environmental hydrology in this book. It is hoped that publication of this book will encourage others to write comprehensive texts on this subject of enormous interest and growing significance.

The subject matter of this book is divided into twelve chapters. Introducing environmental hydrology in the first chapter, watershed acidification modeling is discussed in Chapter 2. The precipitation of acidic materials from the atmosphere to terrestrial and aquatic systems has received international attention, and is, therefore, an appropriate start for the subject of environmental hydrology. If the climate as we know it today changes in a significant way, its implications are so gigantic that it is even difficult to fathom its full impact on the human race. Because of considerable expenditure of resources and debate now being devoted to climate change, its discussion is given in Chapter 3. River hydrology constitutes the core of surface-water hydrology and much of chemicals introduced into our land and water resources find their way into surface waters through rivers. Chapter 4 is devoted to discussion of river hydrology.

Chemicals introduced into the environment migrate through rivers and streams, unsaturated zone, and/or groundwater formations. The next four chapters deal with the fate and migration of these chemicals through the water continuum. Specifically, Chapter 5 discusses the fate and migration of reacting solutes in rivers and streams, physical mechanisms influencing solute transport, and the spatial and temporal distribution of the sources contributing the solute mass. Chapters 6 and 7 focus on solute transport in subsurface unsaturated zone. Basic principles of water storage, and of flow and contaminant movement through unsaturated soil are emphasized in Chapter 6. Transport of water and solutes through the vadose zone under field conditions is presented in Chapter 7, with emphasis on the nature and impact of preferential flow. Chapter 8 discusses the role of groundwater contamination models in planning, management, and regulation of groundwater systems, with a focus on generic and site specific contamination. A case history from Canada is presented to illustrate the current modeling technology. Besides chemicals, microbial transport is receiving increasing attention these days, partly because of the role bacteria and viruses play in environmental pollution as well as in waste remediation. Subsurface transport of microorganisms is presented in Chapter 9, summarizing the progress made to date in this area of growing significance.

The last 3 chapters deal with water quality and its modeling. Of particular concern are loadings of nutrients and toxic chemicals. Chapter 10 reviews major sources of uncertainty associated with tributary loads, other nonpoint sources, and point sources, with particular emphasis on fluvial loadings. It also discusses alternative monitoring and sampling designs. Water quality modeling is presented in Chapter 11. It summarizes the historical development of watershed-scale water-quality modeling, presents the best of the current generation of modeling tools, and explains potential areas for future model development and research. The last Chapter 12 dwells upon eco-hydrological consequences of environmental degradation. It provides a guide to those interested in impact assessment or more detailed analyses.

The editor expresses his deep appreciation to the various contributors who generously contributed to the book despite their busy schedule.

V.P. SINGH
Baton Rouge, Louisiana

CHAPTER 1

What is environmental hydrology?

V.P. Singh

1.1 Environmental Continuum

Air, soil, and water constitute the environmental continuum, and are vital components for sustaining life on earth. These components are interactive and interactions amongst them are complex. Stated another way, a management change imposed on one component of the environmental continuum has effects that propagate to other components, and some of these effects are unknown and cannot be quantified. The interactive nature requires that the environment is managed and protected as a cohesive whole (or a system). This can be accomplished through an integrated approach to environmental management that clearly has to be interdisciplinary. Such an approach may entail management of individual components with proper accounting of their interactions and dependencies. Thus, the management of air, soil, and water must be undertaken in an integrated manner.

1.2 Integrated Environmental Management

Worldwide, environmental systems are rarely managed as a cohesive whole. That is partly because air, soil, and water resources have been and continue to be managed by independent organizations having little interaction amongst them. For example, at the federal level in the United States there is the Environmental Protection Agency (EPA) charged with the responsibility of protecting the environment, establishing guidelines for environmental protection, and overseeing environmental restoration works. Other federal departments such as the U.S. Department of Agriculture, U.S. Army Corps of Engineers, U.S. Bureau of Reclamation, U.S. Geological Survey, etc., are also concerned with safeguarding the environmental quality of land and water resources within their jurisdiction. Then there are the departments of natural resources or environmental quality at the state level. These departments have the responsibility of protecting the environment at the state and local level. In a similar vein, there is a plethora of organizations in the United

V. P. Singh (ed.), Environmental Hydrology, 1–12.
© 1995 *Kluwer Academic Publishers.*

States that have authorities to manage nation's water resources at the local, state, regional, and national level. The U.S. Bureau of Reclamation and the U.S. Army Corps of Engineers are two of the federal agencies with broad authority for water resources management. Then there are the departments of water resources or natural resources at the provincial (state) level having authority to manage water resources in their individual states. At the local (county or parish) level, there are the water boards, commissions, and the like that manage local water resources. The nation's land resources are managed in an analogous manner. The management of land and water resources at all these levels must be accomplished in conformity with environmental protection guidelines.

1.3 Water Continuum

Water exists in different forms – as water vapor in the atmosphere, as liquid water in oceans, streams, lakes, etc. (or hydrosphere) and lithosphere, and as snow and ice on the land surface. The flow of water observed in rivers evolves as a continuum comprising surface flow (overland and channel flow), interflow, and baseflow. These components occur concurrently but their relative magnitudes vary with time. For example, for a burst of rainfall, surface runoff predominates during the rising part of the streamflow hydrograph, interflow during the early part of its recession, and baseflow during the delayed part of its recession. The mechanisms governing these components are different but are influenced by dynamic interactions prevailing between them. The factors controlling flow generation can be summarized as climate, geology, topography, soil characteristics, vegetation, and land use. The relative significance of each factor varies in time and space.

The occurrence of surface runoff over the entire watershed is based on the assumption that rainfall intensity exceeds the infiltration capacity over the whole watershed. However, in many geographic regions surface runoff is rarely observed. Many watersheds yield well-defined hydrographs from storms whose rainfall intensity is less than the infiltration capacity of the soil-vegetation-land use (SVL) complex. This leads to the existence of the various sources of water generating a storm hydrograph: (1) Overland flow may occur over the entire watershed or a portion thereof. (2) Overland flow may occur on an area saturated by downslope groundwater flow as well as infiltrating water. (3) Subsurface flow may occur due to rapid infiltration and transmission through unsaturated and saturated zones. Saturated areas grow during a storm and shrink during interstorm periods. Therefore, the initial size of the contributing area depends upon the time history of rainfall during the prestorm period and the growth depends on the time history of rainfall during the storm.

Solute transport in surface and subsurface waters is affected by a complicated mix of physical, chemical, and microbiological processes. Because water is the principal carrier of solutes and sediments, a complete understanding of water continuum is critical to determining the fate and migration of these contaminants. For example, huge quantities of chemical fertilizers and herbicides are used in agriculture every year in the United States, Canada, and Europe. The release and migration of nutrients, pesticides, and other chemicals in runoff from agricultural

lands is both an economic loss and a threat to the quality of surface and ground waters (Ahuja, 1990). Significant amounts of these chemicals find their way in surface and subsurface waters and may cause critical contamination if these waters form part of the drinking water supply. The high water solubility of these chemicals affords them remarkable mobility in soils under certain conditions; contamination is increased in soils with low adsorption capacity or when groundwater levels are low. Irrigation management practices influence soil salinity or solute movement to groundwater. The gaseous reaction products are produced from solutes moving with water in unsaturated subsurface regime and their migration is affected by water (Angelakis et al., 1993). To protect, improve, or remediate water quality depends upon development and implementation of proper farming systems that incorporate major processes affecting agrichemical fate and migration (Burkart et al., 1990).

Storm runoff from urban areas is an important nonpoint pollution source. An analysis of different components of discharge in a river provides the key to understanding the processes controlling the transport and transformation of pollutants. This analysis aids in quantifying the influence of land-use change on river water quality and in developing strategies for improving the water quality of rivers. Effluents from nonpoint sources may, for instance, be limiting factors for the stability of aquatic ecosystems.

On the other hand, acid precipitation has had a significant effect on water quality in North American, European, and Scandinavian rivers, lakes, and reservoirs (Bobba and Lam, 1990). There is strong interaction between hydrological and chemical processes, and the watershed acidification processes are complex. Hydrology plays a dominating role in acidification modeling–prediction of long-term acidity of watersheds. Identification of flow paths and components of streamflow are the key to predicting short-term chemical changes in stream-water chemistry (Neal et al., 1992).

1.4 Integrated Water Management

The integrated water management should be accomplished within a spatial unit called watershed through the instrument of modeling. The integrated water management can be viewed as a three or more dimensional process centered around the need for water, the policy to meet the needs and the management to implement the policy. Water elements encompassing physical, chemical, and biological aspects of water quantity and quality may constitute the first dimension. Water uses, including agriculture, water supply, energy generation, industry, fish production, recreation, transportation, etc., may constitute the second dimension. Clearly, these water uses have to be accomplished following a well-defined management policy balancing the demand for water amongst different uses. The strategy to implement this policy may form the third dimension. The management must be dynamic and evolve with time, in response to changing needs and objectives. Thus, the fourth dimension may be formed by these needs and objectives.

The water elements, their interactions, and the effects on them of natural as well as external constraints, as shown in Figure 1.1, constitute the foundation upon

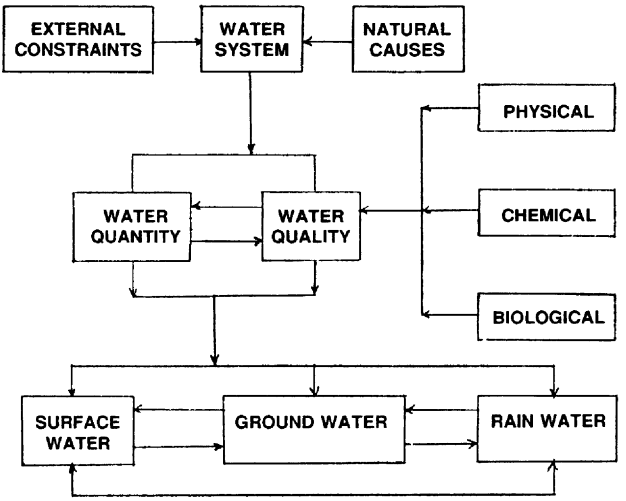


Fig. 1.1. Elements of a water system and their interactions.

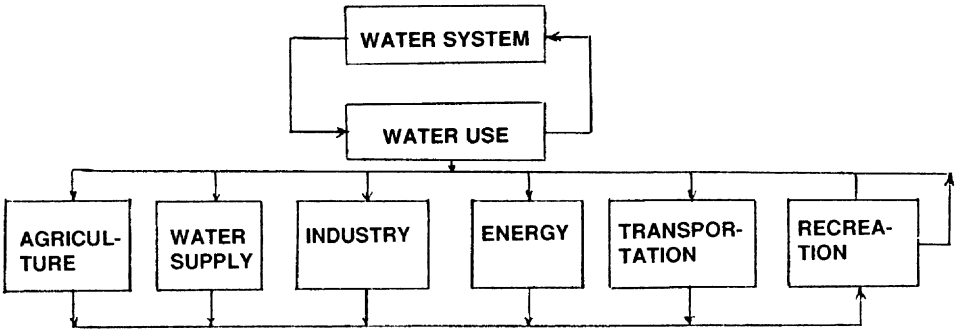


Fig. 1.2. Water use and its interaction with water system.

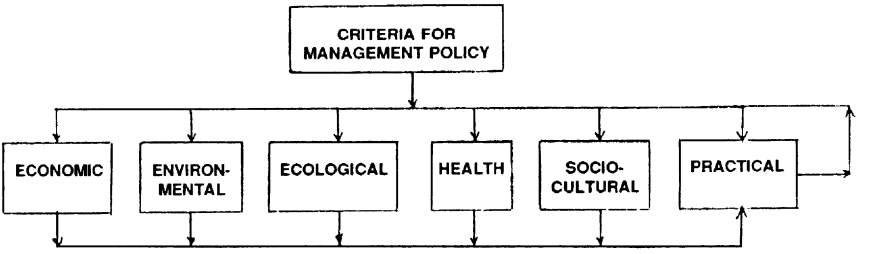


Fig. 1.3. Criteria for information of a management policy.

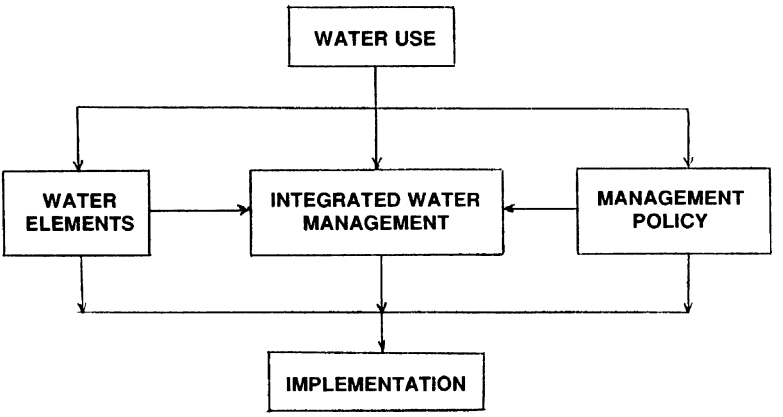


Fig. 1.4. Integrated water management.

which the edifice of integrated water management is to be built. External constraints such as economic, demographic, transportation and other forms of development directly influence one or the other water elements. Likewise, climatic vagaries, climatic change, and climatic extremes, and a host of natural hazards are some of the natural causes that greatly influence the water elements and have significant impact on the integrated water management.

The need for water is derived from a variety of activities in which it is used as shown in Figure 1.2. These activities are vital for human society–its existence and development. Because usable water is limited in its availability, it has an economic value. Furthermore, different activities require water of differing quality. For example, for domestic use, water of high quality is needed. Clearly, all uses of water cannot be supported to the fullest extent and a management policy has to be developed that can prioritize water use following established criteria. There may be conflicts and interactions amongst different water uses, and these, in turn, interact with water elements. The management policy has to incorporate all of these considerations.

The criteria, as shown in Figure 1.3, that form the foundation of the management policy must be developed following such considerations as cost effectiveness, economic efficiency, environmental impact, ecological and health considerations, socio-cultural aspects, to name but a few. The criteria must be practical, implementable, and must be acceptable to the society at large. Furthermore, they must be capable of responding to changing needs of the society. The resulting policy must satisfy such interactive social goals as equity, efficiency, environmental quality, etc.

Once a management policy is established, a strategy, including administrative infrastructure, has to be employed to undertake the integrated water management as shown in Figure 1.4. The components of integrated water management are interactive, and hence the administrative set up must be flexible and be responsive to changing goals. Thus, the integrated water management requires integration of the various components discussed earlier — physical, biological, chemical, social, eco-

nomic, ecological, health, and environmental. This can be accomplished through development and application of mathematical models. The physical, chemical, biological, environmental, and ecological components and their models must be embedded in the development of comprehensive hydrologic models.

A model for integrated water management helps develop implementable solutions to water resources and environmental problems by combining into an optimization scheme all the essential component models. The model incorporates or accumulates all of the interactive forces or influences. Hence, it aids the decision-making process and keeps the policy results within the intersection of the social goals of the management policy and the legal constraints.

Fundamental to integrated water management is hydrologic modeling. The hydrologic models abound in hydrologic literature (Singh, 1989), and the state of art of hydrologic modeling is reasonably advanced, especially when viewed in the context of practical application. However, these models place much greater emphasis on water quantity than on water quality, and have yet to become common planning or decision-making tools. To that end, two milestones will have to be achieved. First, these models will have to be transformed and packaged at the level of a common user who is not necessarily a hydrologist. For some models, this has already begun to occur. Second, these models will have to be integrated with social, environmental, ecological, economic and management models yielding information that is easily interpreted or understood by the user.

A majority of hydrologic models simulate watershed response either without consideration of water quality or inadequate consideration thereof. These models will have to be made more versatile and more comprehensive to include biological as well as chemical facets of watershed response. If these models are to be used for environmental or ecological management then this versatility becomes an essential attribute of the models. A decision maker wants to know as much about a watershed as he can, not just water-quantity information.

In the past few years, with revolution in information and communication technology, new vistas have unfolded that hydrologic models have not taken full advantage of. Examples include remote sensing and satellite technology (RSST), GIS, artificial intelligence, environmental isotopes, and risk and reliability analysis. The RSST is a powerful technology for data acquisition over extensive and even otherwise inaccessible areas. The GIS technology can be usefully employed for data management – data retrieval, manipulation, organization and the like – that may be entirely suitable for development of distributed models. Artificial intelligence allows one to make hydrologic models user-friendly and usable by nonhydrologists. Environmental isotope techniques (Payne, 1988) can be meaningfully employed in field studies of a wide range of hydrologic problems, including flow gaging, sediment transport, recharge estimation, groundwater dating, arid zone hydrology, waste-disposal hydrology, interaction of surface and ground waters, leakage between aquifers, salinization, to name but a few. Risk and reliability analysis allows for providing information on the accuracy of model results that the user wants to know about. It is this information that in many cases becomes crucial in selection of a model and determines the economic worth of data as well as of the model.

1.5 Classification of Hydrology

Hydrology can be classified depending upon the source of water, characteristics of water, the subject matter (or scientific content) to be included, the technique of solution to be employed, or the area of emphasis.

1.5.1 Properties of Water

Depending upon the properties of water to be emphasized in study, hydrology can be classified as (1) water-quantity hydrology, and (2) water-quality hydrology. Many people equate water-quality hydrology to environmental hydrology. Each branch includes consideration of water of all sources.

1.5.2 Sources of Water

It is customary to classify hydrology based on the source producing water. Accordingly, this classification consists of (1) atmospheric hydrology, (2) surface-water hydrology, (3) unsaturated-zone (or vadose-zone) hydrology, (4) saturated-zone (or groundwater) hydrology, (5) lake hydrology, and (6) snow and ice hydrology. Atmospheric hydrology is sometimes referred to as hydrometeorology. Surface-water hydrology is also called stream hydrology. Unsaturated-zone hydrology and the saturated-zone hydrology can be combined to form subsurface-water hydrology. Lake hydrology and the snow and ice hydrology can be merged with surface-water hydrology. Each branch of hydrology must encompass physical, biological, and chemical aspects of water.

1.5.3 Scientific Content

Depending upon the scientific direction to be emphasized, hydrology can be classified as (1) physical hydrology, (2) chemical hydrology, and (3) biological hydrology. Traditionally, physical hydrology has been the area of emphasis, largely because of its ubiquitous application to water resources engineering design. However chemical hydrology as well as biological hydrology are receiving increasing attention these days due to heightened awareness of environmental problems.

1.5.4 Solution Technique

The techniques of solving hydrologic problems can be distinguished as either mathematical or statistical. The mathematical techniques can be analytical, numerical, digital, or systemic. Similarly, statistical techniques can be empirical, probabilistic, or stochastic. Thus, based on solution techniques, hydrology can be classified as (1) mathematical hydrology, and (2) statistical hydrology. Mathematical hydrology can be further classified as (i) parametric hydrology, (ii) numerical hydrology, (iii) digital hydrology, and (iv) systems hydrology. Similarly, statistical hydrology may consist of (i) empirical hydrology, (ii) probabilistic hydrology, and (iii) stochastic hydrology.

1.5.5 Area of Emphasis

Depending upon the particular application of hydrology to be emphasized, hydrology can be classified as (1) agricultural hydrology, (2) atmospheric hydrology, (3) coastal hydrology, (4) desert hydrology, (5) forest hydrology, (6) geographical hydrology, (7) environmental hydrology, (8) ecosystems hydrology, (9) mountainous hydrology, and (10) wetlands hydrology. Much of the emphasis in this classification is on the surficial hydrology.

1.5.6 Basin Size

A basin, terrestrial or lithologic, constitutes the spatial unit within which many of the hydrologic processes occur and are dealt with. The evolution of these processes varies from one basin to the other and with the basin size. Therefore, based on the basin size, hydrology can be classified as (1) small-basin hydrology, (2) medium-basin hydrology, and (3) large-basin hydrology.

1.5.7 Basin-Type or Land Use

The soil, vegetation, and land use (SVL) complex plays a fundamental role in shaping the hydrologic behavior of a watershed. Therefore, hydrology can be distinguished, based on basin type or land use, as (1) agricultural hydrology, (2) rural hydrology, (3) desert hydrology, (4) mountainous hydrology, (5) wetland hydrology, (6) urban hydrology, (7) transportation hydrology, and (8) forest hydrology.

1.6 Definition of Environmental Hydrology

In a broad sense, environmental hydrology is all inclusive-embracing all branches of hydrology. However, in practice, environmental hydrology has a more restricted connotation. It is primarily restricted to water-quality hydrology. In this sense, it can be defined as the science dealing with the space-time variability of water quality and its evolution in the hydrosphere, in streams, in lakes, in soil, as well as in the lithosphere. Thus, it includes occurrence, distribution, and variability of water quality in surface water, vadose-zone, and ground water. Also included is the water quality in the atmosphere, e.g., acid precipitation. It is in this sense that environmental hydrology is perceived in this book. It should, however, be emphasized that quantity and quality of water are intertwined and should conjunctively be dealt with.

The emergence of environmental hydrology can be traced to some of the landmark problems that aroused public awareness of the environmental contamination caused by both civilian and military establishments, and of the seriousness of these problems both in terms of human health and economic costs. This awareness led to passage of a series of environmental laws since 1970 in the United States.

In 1970, the United States Congress passed the National Environment Policy Act that heralded the beginning of a new era in environmental protection. The Act was based on the notion that man and nature must exist in productive harmony. A sweeping federal and state effort to clean up the nation's waters was initiated in

1972 through the passage of the Federal Water Pollution Control Act Amendments. The Act set water quality control goals and established provisions for reducing and eliminating water pollution. Two sources of water pollution – point and nonpoint – were identified in the law. Point sources are end-of-the-pipe discharges, including those from municipal sewage, treatment plants, industries, and concentrated animal feeding operations. Nonpoint-source pollution is caused by runoff or melting snow that infiltrates into soil and enters waterways. Examples include urban stormwater, agricultural runoff, construction runoff, acid mine drainage, and forestry runoff. Nonpoint pollution also arises from septic tanks, poor landfills, and underground waste areas which seeps through the soil, picking up pollutants and transporting them into waterways and groundwater. Through proper management of land and water resources, these pollution sources can only be treated, but cannot be collected and treated.

Another major piece of legislation in environmental protection is the Toxic Substances Control Act of 1976. This Act requires cleaning up of environmental residues after they have been produced as well as prevention of harmful residues in the first place. The Environmental Protection Agency (EPA) is, under this Act, required to ensure the safe manufacture, use, and distribution of potentially dangerous chemicals.

Two other important laws are the Resource Conservation and Recovery Act (RCRA) passed in 1976, and the Comprehensive Environmental Response Compensation and Liability Act (CERCLA), also known as "Superfund", passed in 1980. Simultaneously, problems relating to nuclear wastes began to receive increasing attention. As a result, in 1982 the Nuclear Waste Policy Act was passed in the United States that provided the regulatory framework and the timetable for developing the first repository for nuclear wastes. In fact, by 1975, most countries had initiated programs for environmentally acceptable storage of nuclear wastes (Schwartz, 1988).

Although serous problems at the Rocky Mountain Arsenal (Konikow, 1976), and the Idaho National Engineering Laboratory (Robertson et al., 1974) had received attention, it was the problem of contamination from hazardous wastes at Love Canal in Niagara Falls, New York, that provided the real impetus to what we perceive today as environmental awareness and forever changed the face of hydrology and water science. Indeed, Love Canal focused the world's attention on environmental contamination resulting from indiscriminate development. It prompted action to begin cleanup of contaminated sites here in the U.S. and served as a catalyst for similar actions in other countries.

1.7 Scope of Environmental Hydrology

Cleaning up our nation's streams, rivers, and lakes has been of growing concern. Clean water is not only essential to human health and general well being but also to economic welfare. Clean water is essential for drinking water supplies, agricultural use, industrial use, commercial and residential development, recreational use, and for a healthier environment.

The subject matter of environmental hydrology can be defined to include chemical as well as biological quality of surface, unsaturated, and saturated waters. Because of significant pollution occurring in the atmosphere due to activities on the land, acid precipitation is a legitimate component of environmental hydrology. Human activities are believed to be also influencing climate and accumulated evidence points to a likely climate change that may have far reaching hydrological implications. Thus, a discussion of climate change also constitutes a part of environmental hydrology. A study of eco-hydrological consequences of the damage done and being done to the environment also lies within the realm of environmental hydrology. This essentially charts out the scope of this book.

To further grasp its scope, it may be worthwhile to survey a sample of environmental problems whose solution depends on critical input of environmental hydrology: (l) Design of waste-disposal facilities such as treatment plants, lagoons, landfills etc.; (2) quantification of environmental pollution; (3) design of pollution abatement schemes; (4) hazardous waste management; (5) aquacultural production; (6) construction of wetlands; (7) wetlands restoration; (8) lake restoration and management; (9) health-risk assessment; (10) health-hazard mitigation; (11) groundwater clean up; (12) strategies for controlling encroachment of saltwater intrusion; (13) environmental planning, management and restoration; (14) dynamic interaction between components of environmental continuum; (15) integrated approach to physical, chemical, and biological aspects of ecosystems, etc.

1.8 Role of Environmental Hydrology

Water quantity and water quality processes constitute an integral part of the natural hydrologic environment. These two processes are in continuous dynamic interaction so that proper assessment, development, and management of water resources require a full understanding of these processes. Although water quantity has historically been the primary factor controlling the use of water resources, water quality is critical in terms of pollution control and environmental management — a very important issue in a world increasingly aware of environmental concerns — and is increasingly becoming critical in determining the amount of available water that can be used to meet a specific water demand.

The general trend in water quality management to date has been to gather and use information on water quality variables for purposes of planning, design, and operation of water resources and wastewater treatment systems. However, growing concern for environmental quality has given rise to a new trend in respect of the impact of water quality variables on human health and life conditions. Thus, there is the need for a better understanding of how water quality processes evolve in space and time under natural and man-made conditions.

The virtual dependence of modern society on chemical products has been fraught with enormous price-environmental degradation, health hazard, climate change, and pollution of precious natural resources. Indiscriminate generation of huge quantities of hazardous substances and their disposal without cognizance of ensuing repercussions have been nothing short of man-made calamities. This spans both civilian and military establishments. A 1992 Department of Defense (DOD)

report to the U.S. Congress lists 10,924 active hazardous waste sites at more than 1,800 domestic military installations, with sites in each of the 50 states. Of the total installations, about 100 are included in the Superfund National Priorities List (Sidel, 1993). Military bases generate large quantities of a wide range of toxic substances, including fuels, pesticides, solvents, polychlorinated biphenyls, and phenols. The DOD estimated that domestic military bases produced about 80,000 tons of hazardous waste for off-site treatment and disposal in 1991. Some authorities view the figure for hazardous waste releases into air, water, and on-site as large as 500,000 tons. Can we clean up in the near future the mess we have created over the past four decades? Have we learnt the lessons we need to learn and not to repeat the mistakes of the past? Do we have the technology to restore the natural environment? Questions such as these are being frequently asked these days, and this is a healthy sign of public awareness. However, as regards environmental clean up, we have not even scratched the surface. The EPA has estimated, using current DOD, Department of Energy (DOE), and General Accounting Office data, that the total cost of bringing U.S. military facilities into compliance with environmental laws and of repairing the damage caused by them could easily surpass US$400 billion. Of course, the real cost of all environmental damage cannot be expressed in dollars, because some health and environmental costs are intangible and some losses are permanent. Although economic costs of this mess and its clean up are being debated, its costs in terms of human health, plant life, or the entire ecosystem have yet to receive the attention they deserve. Major water quality issues that need collective public and government attention include: (1) Contamination of drinking water supplies, (2) risk to human health and cost of clean up, (3) prevention and mitigation of environmental hazards, (5) impact of toxic and other contaminants on the aquatic environment, (4) impact of land use on water quality, (6) chemical contamination of fish and aquatic life, (7) impact of airborne pollutants on water quality, (8) public perception of water quality, (9) conflicts over water quality use, (10) water quality monitoring, (11) pollution of waters used for recreation, (12) impact of water resources projects on the aquatic ecosystems, etc. Environmental hydrology will a critical role in addressing all of these issues.

References

Ahuja, L. R., 1990. Modeling soluble chemical transfer to runoff with rainfall impact as a diffusion process. Soil Science Society of America Journal, Vol. 54, pp. 312-321.

Angelakis, A. N., Kadir, T. N., and Rolston, D. E., 1993. Analytical solutions for equations describing coupled transport of two solutes and a gaseous product in soil. Water Resources Research, Vol. 29, No. 4, pp. 945-956.

Bobba, A. G. and Lam, D. C. L., 1990. Hydrological modeling of acidified Canadian watersheds. Ecological Modelling, Vol. 50, pp. 55-32.

Burkart, M. R., Onstad, C. A. and Bubenzer, G. D., 1990. Research on agrichemicals in water resources. Transactions, American Geophysical Union, Vol. 17, No. 29, pp. 980-981, and 988.

Konikow, L. F., 1976. Modelng chloride movement in the alluvial aquifer at the Rocky Mountain Arsenal, Colorado. U.S. Geology Survey Water-Supply Paper 2044, 43 pp.

Neal, C., Robson, A., Reynolds, B., and Jenkins, A., 1992. Prediction of future short-term stream chemistry - a modelling approach. Journal of Hydrology, Vol. 130, pp. 87-103.

Payne, B. R., 1988. The status of isotope hydrology today. Journal of Hydrology, Vol. 100, pp. 207-237.

Robertson, J. B., Schoen, R., and Barraclough, J. T., 1974. The influence of liquid waste disposal on the geochemistry of water at the National Reactor Testing Station, Idaho, 1952-1970. U.S. Geological Survey, Open-File Report IDO-22053, 231 pp.

Schwartz, F. W., 1988. Contaminant hydrogeology - dollars and sense. Journal of Hydrology, Vol. 100, pp. 453-470.

Sidel, V. W., 1993. Cleaning up: risk and risk reduction at military sites. U.S. Water News, August, p. 7.

Singh, V.P., 1989. Hydrologic Systems II. Watershed Modeling. Prentice-Hall, Englewood Cliffs, New Jersey.

CHAPTER 2

Watershed acidification modelling

A. Ghosh Bobba, Dean S. Jeffries,
William G. Booty and Vijay P. Singh

Abstract. A wide variety of watershed acidification models are currently being used to address water quality issues from national to local field scale due to acid precipitation. The complexity of the watershed process and the need to provide information quickly has increased the reliance on these models. To properly utilize a watershed acidification model, the user must clearly understand the goals and objectives of the modeling project, and the capabilities and limitations of the watershed acidification model. An overview of process descriptions, assumptions, constraints and other considerations which enter into the development of deterministic watershed acidification models is given in this chapter. Various approaches to watershed acidification models from very simple lumped models to complex models describing flow and biogechemical reactions in watershed systems are described. Four watershed acidification models (TMWAM, ETD, ILWAS and RAINS) are reviewed and a comparison of model performance is presented for Turkey Lakes watershed, Ontario, Canada.

2.1 Introduction

The deposition of acidic materials from the atmosphere to terrestrial and aquatic ecosystems popularly called as 'acid rain' is a phenomenon that has received international attention. Over the past three decades, this environmental concern has been the impetus for an exponential increase in scientific and political activity.

Most atmospheric acidity is due to anthropogenic emission of S and N compounds, primarily in the forms of SO_2 and NO_x. These chemicals are then oxidized in the atmosphere by heterogeneous oxidation of dissolved gases in liquid aerosols and homogeneous oxidation in the gas phase both involving a large number of intermediate substances. The essence of the acidic deposition issue is that these acidic substances return in wet or dry form to lakes, croplands, forests, and other natural ecosystems. Wet deposition little influenced by anthropogenic acids has a

V. P. Singh (ed.), Environmental Hydrology, 13–68.
© 1995 *Kluwer Academic Publishers.*

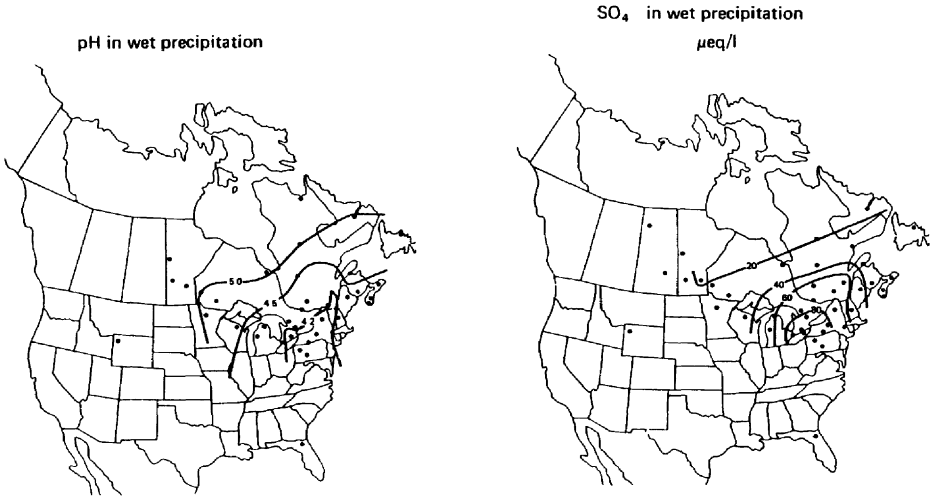


Fig. 2.1. Weighted mean pH and SO_4 concentrations in wet precipitation at different stations.

theoretical "normal" pH of approximately 5.6. Recent studies suggest "normal" precipitation pH actually falls within a range from slightly below 5.0 to greater than 6.0 in natural environments. Extensive monitoring in recent years finds average wet deposition pH values range between 4.0 and 4.5 in the eastern United States, eastern Canada (Figure 2.1), Europe, and elsewhere. Concern over the effect of this acidic deposition on natural ecosystems has fueled the tremendous interest in this environmental topic.

Most of the acidity in acid precipitation is due to the strong mineral acids H_2SO_4 and HNO_3 (Lindberg et al., 1981) to account for 70% of the acidity in precipitation at the Walker Branch Watershed in eastern Tennessee. A large proportion of the mineral acid component of precipitation is due to S compounds, primarily as a result of SO_2 emissions. Glass et al. (1980) found that 60 to 70% of the acid input was attributable to H_2SO_4 while only 30 to 40% was associated with HNO_3 for sites in New York and New England.

The major sources of S are the burning of coal and petroleum products and ore smelting. Electric utilities currently contribute roughly two-thirds, and industry another one quarter of total anthropogenic SO_2 emissions. Industry has reduced these emission loads from 11 to 7 million tons in the last 20 years. Most NO_x emissions are associated with transportation (40%); however, electric utilities and industry together supply an additional 55% of the total. The mix of S and N in precipitation is not static. In the 1980's, SO_2 emissions have declined; however, increases in NO_x emissions of equal or greater magnitude have occurred (Hileman, 1981). Of the total S in the atmosphere of the northeastern United States, 90% is thought to be contributed by anthropogenic sources.

Dry deposition is the dry phase removal of gases and particulates from the atmosphere. It has been shown that the dry deposition component of the total atmospheric input of acidity can be equal to, or greater than, wet deposition input

in some cases. Evidence suggests that the dry deposition contribution to total acidity increases relative to the wet deposition component as one moves closer to the emission source.

The mechanisms of wet and dry deposition can be categorized into five general processes. For wet deposition, S and N are removed from the atmosphere as (a) rainout, or (b) washout. Rainout involves transfer of materials to cloud droplets before they make their descent as a raindrop, while washout is the transfer of materials to a falling raindrop. Rainout usually involves condensation of the cloud droplet around aerosol nuclei containing $(NH_4)_2SO_4$, H_2SO_4, or gaseous SO_2 and NO_x compounds. Dry deposition of acidic materials from the atmosphere occurs as a result of (a) surface absorption of gases, (b) particle sedimentation, and (c) impaction of aerosols on vegetation, soil water or snow.

The acid-base status of water moving through a lake-watershed system is affected by many biogeochemical and hydrologic processes acting separately and in combination. Biogeochemical reaction rates depend on numerous environmental factors, some of which vary with time (e.g. temperature, moisture, concentrations of chemical species in solution) and others that are relatively time invariant (e.g. soil structure, watershed topography, lake morphometry, plant community type). Some processes (e.g. sulphate reduction) always produce alkalinity, others (e.g. nitrification) always produce acidity, and still others (e, g. cation exchange) can either produce or consume alkalinity depending on environmental factors. Rate constants for the time dependent process also vary from practically instantaneous (cation exchange) to centuries (weathering of resistant minerals). There are too many relevant processes to discuss each one within this chapter. Detailed descriptions can be found in standard texts on soil science, hydrology, geology, limnology, micrometeorology, forest ecology, aqueous chemistry and geochemistry.

2.1.1 Watershed Acidification Modelling

Acid rain was one of the most important and carefully evaluated environmental issues of the 1980's. One of the mechanisms of evaluation was the development and application of watershed acidification models. The usual purpose of modelling is to predict events in time and space. There are generally three categories of mathematical models used in hydrology: stochastic models, lumped models and distributed models. Lumped models are frequently labeled conceptual models when they attempt to represent the ecosystem structure being modelled. The response of a model to inputs is determined by a set of parameters, typically, the equation coefficients. The parameters may be obtained independently by appropriate measurement or they may be obtained through fitting, i.e. by adjusting them with the help of past data such that prediction comes as close as possible to observation. In many cases model paramaterization is a mixture of the two. A model with fitted parameters is in general site specific, i.e. can only be used for prediction in time.

Models have been developed to simulate the effects of precipitation chemistry on surface water quality. Water acidification is usually defined in terms of a loss of alkalinity or Acid Neutralizing Capacity (ANC). The first being measured directly by titration while the second is calculated indirectly as the difference between base

cations and acidic anions. In soil, a decrease in base saturation can be used. The complete modelling of the hydrogeochemical response of watersheds requires coupling of a hydrological model with a geochemical model. The hydrological model quantitatively describes the flow paths of the moisture starting with precipitation, followed by interception by the vegetation and canopy, the percolation through the soil layers, lateral flow through the layers, and the flow into streams, bogs and lakes. The geochemical model describes the chemical reactions that occur which produce or consume ANC.

Relating the emission of sulphur and nitrogen oxides to the acidification of aquatic ecosystems requires a watershed modelling approach. Not all acidic deposition enters the rivers and lakes directly; most undergoes chemical changes in the soil layers before reaching the streams. To quantify the impact of acidic deposition requires an understanding of the watershed processes included and their expression in the model. Model results need to be verified with data collected from a number of catchments to ensure that it is spatially transferable. Once assured of the models validity, it may be used to establish an accurate environmental assessment of the acidic deposition impact.

A number of process-oriented mathematical models have been developed that simulate the flow of acidic input through terrestrial systems and the resulting chemical response of surface waters. Examples of these mechanistic models include: the Turkey-Mersey Watershed Acidification Model (TMWAM) (Lam et al., 1988; Bobba and Lam, 1990), the ILWAS model (Chen et al., 1983), the Birkenes model (Christophersen et al., 1982), the ETD model (Schnoor et al., 1984; Nikolaidas et al., 1988, 1989; Nikolaidas, 1987), the MAGIC model (Cosby et al. 1986) and the RAINS model (Kamari, 1985).

2.1.2 Interaction between Lake and Terrestrial Systems

The biogeochemical processes that produce and consume alkalinity and the environmental factors that determine the rates of these processes are distributed nonuniformly throughout the watershed system. They vary both laterally and vertically within the terrestrial and aquatic components of the system. The hydrologic characteristics of the ecosystem (flow paths, flow rates, component storage capacities) determine the relative contribution of different input components (direct precipitation, surface runoff, interflow and groundwater) to the acid base status of the lake.

To illustrate the differing ways in which these processes may be linked, consider a hypothetical set of three lake–watershed systems. The first system (Figure 2.2a) contains a small, high elevation drainage basin lake surrounded by a steep sloped watershed. The entire basin is contained by impermeable bedrock. The only source of water and other substances to the system is atmospheric deposition within the watershed's boundaries as defined by its surface topography. Water can leave the system by two routes, evapotranspiration and the lake's surface outlet. The mean retention time of water in this type of lake tends to be short unless the lake surface area comprises a large proportion of the overall basin drainage area.

The second system (Figure 2.2b) contains a seepage lake with no surface inflows or outflows. Such systems typically exist in relatively flat terrain having

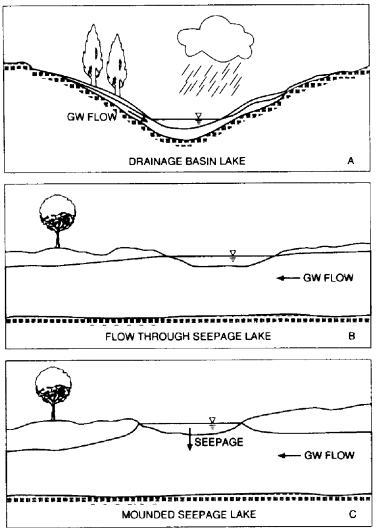


Fig. 2.2. Different types of lake-watershed systems.

very deep, highly permeable soils. The boundary of the groundwater system that serves as a source of water to the lake can be very different than the boundary of the watershed as defined by the surface topography (Figure 2.2b). Because of the flat terrain and deep, highly permeable soils there is no surface flow or interflow into the lake as there is in the drainage system. The sources of water for the lake are precipitation falling on its surface and groundwater inflow. Water leaves the lake either by evapotranspiration or groundwater outflow. Unlike the drainage system where groundwater must go through the lake to leave the system, in seepage systems, groundwater can bypass the lake and not interact with it.

The third system (Figure 2.2c) is a special type of seepage system, called a mounded system. It differs from the seepage system above, in that it sits on top of a high spot in the groundwater table, so that groundwater inflow to the lake is practically zero. An equivalent hydrological situation would be a perched seepage lake that was separated from the groundwater table by a low permeability soil/sediment layer.

For the drainage system (Figure 2.2a), the largest potential sources of alkalinity are found in the soil, i.e. cation exchange, mineral weathering, and sulphate adsorption. The relative strength of these sources will depend on the chemical and physical nature of the soil. For example, higher base saturation increases the alkalinity generation by cation exchange, weathering of carbonate minerals more rapidly generates alkalinity than weathering of silicates, and older soils tend to have higher sulphate adsorption capabilities. Because of the mean residence time of water in the lake, sulphate reduction in the lake sediments is usually of lesser significance as an alkalinity source. In general, the deeper the flow path that the water follows in the soil, the greater will be the alkalinity generated. Thus in this type of system, the relative distribution of water flow through different depths of soil will be a major determinant of the acid-base status of the lake. Since deeper

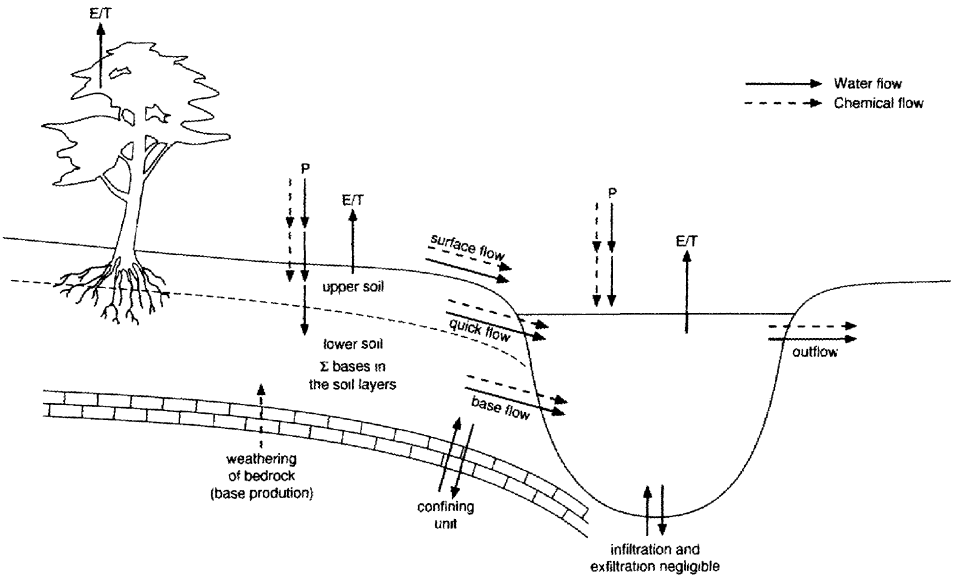


Fig. 2.3. Schematic illustration of the general flow and chemical pathways from watershed to lake.

soil favours greater relative flow along deeper flow paths, depth of soil can be an important indicator of the overall acid neutralizing capacity of the system.

In contrast, differences in the soil depth are not important to the seepage systems of Figure 2.2, since there is practically no direct flow of water to the lake from the upper soil layers. The acid-base status of the lake depends strongly on the relative contribution of water input from precipitation versus groundwater and this in turn depends on the topography of the groundwater system. Hence from a hydrologic perspective, the mounded system would appear to be the most sensitive to change in the chemical composition of precipitation. However, since the mean water residence times of seepage lakes is typically rather long, resulting sediment reduction of sulphate is a potentially significant source of alkalinity. Therefore the sensitivity of seepage lakes cannot be judged solely on a hydrologic basis.

2.1.3 Flow Pathways in Watersheds

The pathways that water follows (Figure 2.3) as it moves through watersheds influence the discharge response of streams, the residence time of water in lakes, and the rate of recharge of the groundwater reservoir. The nature of the flow paths is particularly important in regulating the rapidity with which water enters streams. Initial interest in the movement of water along flow paths developed from recognition that these processes affect the volume and timing of stream flow and consequently the generation of floods during wet periods and the maintenance of low flows during dry periods. Until relatively recently, research on flow paths and runoff processes concentrated on the physics of these phenomena relative to modelling and prediction of floods and low flow conditions.

Hydrologic flow paths also influence what material the water contacts in its passage through the watershed and the contact time. It is becoming increasingly

apparent that flow paths play a major role in determining water quality. Waters following shallow pathways (Figure 2.3) may contact highly weathered mineral materials for short periods of time, providing little opportunity for geochemical weathering reactions, but allowing potentially significant interactions with organic matter, soil exchange sites, and dry deposition materials. Waters following deeper pathways may contact unweathered bedrock, will have a much longer residence time, and will undergo more extensive geochemical reaction. The chemical and isotopic changes that occur in water as it moves along various flow paths can be powerful tools in deciphering the mechanisms controlling water movement in watersheds.

Part of the precipitation (rain or snow) that falls on a watershed is lost by evapotranspiration and the remainder eventually leaves as runoff or deep groundwater recharge. Runoff does not move uniformly through a watershed but follows preferred pathways dictated by the local topography, vegetation, soil characteristics, and geology. Rainwater and meltwater can follow a number of different surface and subsurface pathways enroute to a stream channel, lake, or the groundwater reservoir. Precipitation falling directly on the stream channel or lake surface does not intercept any of the terrestrial components of the watershed and is called channel precipitation. It generally makes up no more than a small percentage of the flow during storms. However, if the precipitation event is strongly acidic, measurable changes in the chemistry of the surface waters of lakes and streams can occur.

Rain falling on other parts of the watershed may run off along the surface or infiltrate and travel along subsurface pathways. Horton (1933) introduced the concept of overland flow to explain the rapid generation of stream flow by storm events. He postulated that during most rainfall events, the infiltration capacity of the soil is exceeded, and excess rainfall then flows over the watershed surface to a stream channel or lake. This direct surface path might explain the rapid increase in stream discharge observed in many watersheds following a rainfall event. However, subsequent field investigations have shown that overland flow as originally postulated by Horton is not commonly observed over large areas, except in watersheds with extensive exposures of bare rock. Overland flow in undisturbed watersheds occurs only under special conditions and usually from very small areas along stream channels where the soil has been saturated. A variation of overland flow in which infiltrated water moving along subsurface flow paths reemerges at the surface has been described as 'return flow' by Musgrave and Holton (1964).

The possibility that only part of a watershed may contribute runoff during storms was suggested by Moldenhauer et al. (1960). Betson (1964) recognized that certain limited parts of a watershed sometimes contribute overland flow to streams, which led to the development of the concept of partial-area contributions. According to this concept, in certain limited parts of watersheds, the infiltration capacity of the soil is exceeded and overland flow occurs. The areas where this is observed are usually confined to the valley bottoms along stream channels and such riparian areas expand and contract in response to levels of shallow groundwater in the watershed.

The evolution of the variable contributing area concept of storm flow generation, widely accepted as the correct conceptual model for forested watersheds in humid temperate climates, was reviewed by Hibbert and Troendle (1987). Hewlett

and Hibbert (1967) measured lateral subsurface flows to streams and suggested that during rainfall events ephemeral streams expanded headward by collecting shallow surface runoff and overland flow along their channels. Ragan (1968) investigated variable area contributions in a small forested watershed near Essex Centre, Vermont. He found that no overland flow occurred on the upland portions of the hillslopes during storms and that only a small portion of the watershed contributed flow to the storm hydrograph. The contributing area varied with storm duration and intensity and with antecedent moisture conditions. No interflow through the soil above the groundwater table was observed, and the bulk of the flow was derived from rapid response of the groundwater table adjacent to the stream. Dunne and Black (1970) examined the role of variable-area contributions to runoff in a small, steeply sloping watershed used as pastureland. They observed that significant amounts of storm runoff were produced only on small areas of hillside close to the stream where the water table reached the surface. These areas varied seasonally and during individual storms. Fluctuations in the areas were related to topography, soils, antecedent moisture, and rainfall intensity. Many others have investigated runoff production mechanisms and generally confirmed the variable source mechanisms controlling factors.

Much more significant is the diffuse, generally downward movement of water through soils, suggesting that subsurface pathways must be important for water flow through watersheds. This idea is not new; however, many hydrologists have believed that water cannot travel quickly enough by subsurface flowpaths to produce runoff peaks in river hydrographs. Schumacher (1864) noted that the movement of water through soils is controlled by large pores in which the water is not held by capillary forces, implying rapid subsurface flow. Lawes et al. (1882) recognized two types of water draining from soils: (a) rainwater travelling rapidly through open channels in the soil with little change in composition and (b) water discharged more slowly from the pores of a saturated soil. Atkinson (1978) has applied the terms "pipeflow" and "matrixflow" to waters following these pathways. Lowdermilk (1934) recognized shallow seepage as a source of water feeding streams. Hursh (1936) termed waters that moved along subsurface pathways more rapidly than normal groundwater flow "subsurface stormflow". Hursh and Brater (1941) suggested that subsurface flow contributions to streamflow could account for a major portion of the hydrograph peak.

Displacement of groundwater into streams by rainfall was suggested by Riesbol (1942) and Hursh (1944). Horton and Hawkins (1965) showed experimentally that percolation of rainwater through the soil to the water table proceeds throughout most of the flow path by downward displacement of water previously retained by soils at field capacity. Hewlett (1961) demonstrated that water draining from the soil contributed to stream baseflow. Anderson and Burt (1978) and Anderson and Kneale (1980) measured the dynamic behaviour of variable-source areas using an automatic tensiometer system which provided contributions data on soil-water potential in hillslope hollows. Rapid infiltration of rain causes convergent downslope movement of soil water into the hollows, enlarging the saturated wedge at the base of the slope and increasing the throughflow discharge to the stream. Freeze (1972, 1974) examined the roles of subsurface sources contributing to streamflow and concluded that the importance of the subsurface response of watersheds has

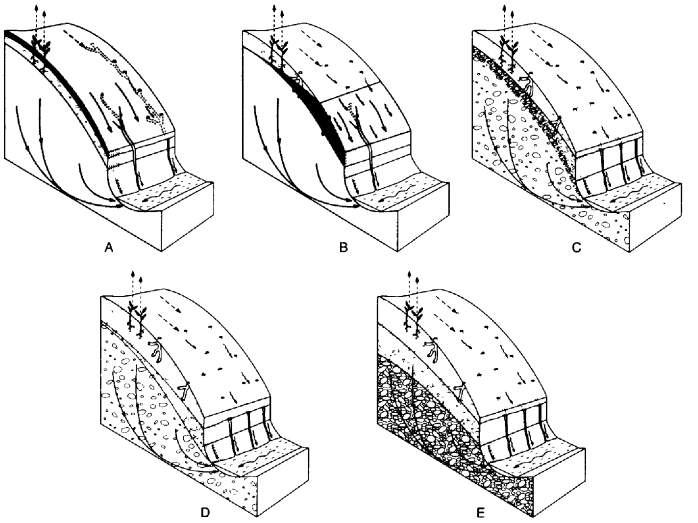


Fig. 2.4. Schematic view of some watershed runoff processes occurring during rainfall or snowmelt in humid areas. Modified from Atkinson (1978). (A) Hortonian overland flow, (B) saturated overland flow; (C) subsurface storm flow; (D) matrix throughflow; and (E) groundwater flow.

been greatly underrated. In an isotopic study of two hydrogeologically diverse watersheds, Sklash and Farvolden (1982) demonstrated that "groundwater plays a much more significant role in the runoff process than contemporary theories and conventional studies suggest".

In summary: (a) The existence of subsurface flow has been proven by numerous hydrologists working in the field (Hewlett, 1961; Hewelett and Hibbert, 1967; Nutter and Hewlett, 1971; Ragan, 1968; Weyman, 1970; Knapp, 1974; Dunne and Black, 1970). (b) Subsurface flow can be the main contributer to peak discharge (Whipkey, 1965; Rawls and Asmussen, 1973; Arnett, 1974; Harr, 1977; Piligrim et al., 1978; Mosley, 1979, 1982), or is at least an important contributer (McCaig, 1982). (c) The area contributing to peak discharge is only a part of the watershed area. The contributing area is saturated (discharge area) or of low permeability. In the former case, saturation overland flow is occurring (Hewelett and Hibbert, 1967), and in the latter, Hortonian overland flow. (d) The contributing area can increase and shrink depending upon saturation conditions (Dunne and Black, 1970) or infiltration capacity (Beston, 1964). (e) The intensity of rainfall has minor influence on the character of the peak runoff event (Hewlett et al., 1977; Taylor and Pearce, 1982. (f) Saturation overland flow occurs most readily at the bottom of slopes, in hallows and concave slope profiles and poorly developed soils (Kirby and Chorley, 1967; Anderson and Burt, 1978; Huff et al. 1982). In humid weather conditions five different storm runoff processes can be distinguished, their occurrences depending upon soil type and geometry of the hillslope (Figure 2.4).

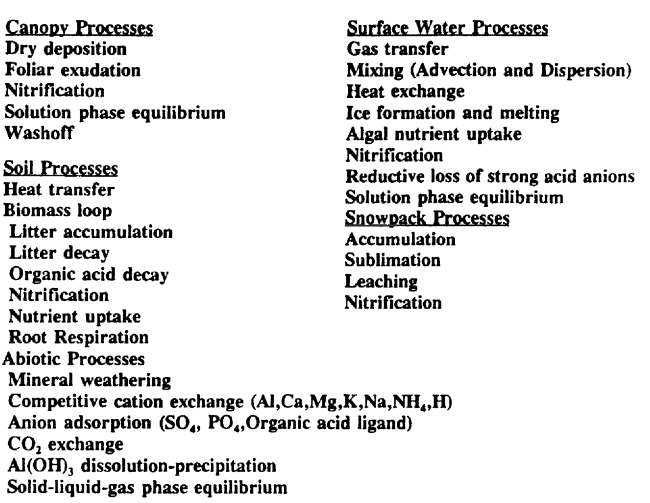


Fig. 2.5. Biological, chemical and physical processes operating in a typical temperate watershed.

2.1.4 Bio-geochemical Processes: Chemical Species

Biogeochemical processes identified in Figure 2.5 influence the overall acid-base status of the solution chemistry and its effect on terrestrial and aquatic biota. In order to understand the role of any single process, a mechanistic analytic framework is needed that explicitly includes the individual chemical species involved. For instance, an analysis of the role of nitrification requires explicit inclusion of ammonium and nitrate. The concentrations of the species above the line in Figure 2.6 can be treated as independent variables, while those below the line can be calculated.

Hydrogen ion is not treated as an independent variable because its mass within solution is not a conservative property. It is easier to calculate hydrogen ion concentration from the concentrations of the constituents above the line than to try to measure its mass balance.

Silicic acid (H_4SiO_4), $pK_A = 9.5$, does not generally affect the acid-base status of natural water systems, however, it provides information on the magnitude of weathering, since it is a product of many important primary mineral weathering reactions.

The detailed Al speciation included in Figure 2.6 arises from concern about potential toxicity, the large variation of toxicity among species, and the high concentrations that can occur under acidic conditions.

Because of the overwhelming difficulty in specifying the chemical nature of a complex mixture of organic acids, the organic acid chemistry is usually modelled based on the general titration and complexation properties of the overall mixture. Figure 2.6 identifies two such model acids, one trivalent and the other monovalent.

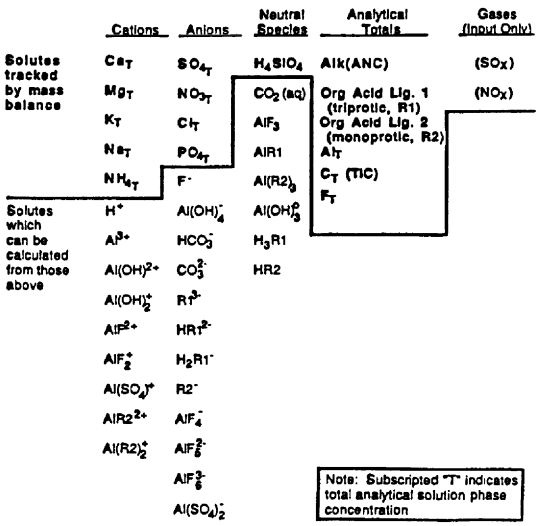


Fig. 2.6. Solution-phase chemical constituents in watershed.

2.1.5 Basic Mechanism of Leaching - Mobile Anions

A watershed is a very complex system, and a large number of processes influence acidification. Some of the important processes are:

- Canopy attenuation of precipitation composition;
- CO_2-pressure in soil;
- Uptake and release of cations by vegetation;
- Cation exchange in the soils;
- Weathering;
- Oxidation and reduction reactions (sulphur and nitrogen compounds) in both terrestrial and aquatic environments;
- Watershed storage and release of incoming S and N compounds; and
- Release of organic acids and other weak acids from organic soils.

Watershed acidification models are based on the "mobile anion concept" (Seip, 1980) with sulphate as a key ion. Figure 2.7 shows schematically the inputs to and outputs from a watershed. In acid streams, the bicarbonate concentration is very low. In the watersheds we have considered that Na^+ and Cl^- are found in nearly equivalent amounts both in precipitation and runoff, and they may as a first approximation be ignored. The importance of sulphate for the transport of cations, including H^+ and Al^{3+}, then becomes evident. To explain acidification it

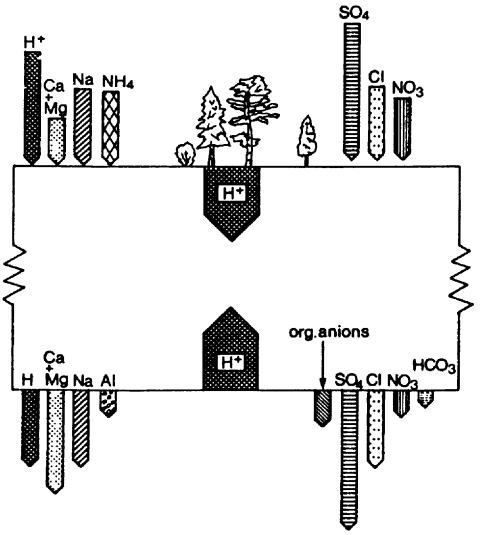


Fig. 2.7. Schematic illustration of the input and output of ions for a catchment. The broad arrows inside the box represent production and consumption of H^+ in the watershed.

is important to consider the possibilities for leaching as well as the size of the large reservoir and the production and consumption of H^+ indicated by the two broad arrows inside the "catchment" in Figure 2.7.

An important potential effect of acidic deposition is the enhancement of cation leaching. The basic process of leaching can best be understood by applying the principles of chemical equilibrium to the soil solution. Since soil is largely dominated by negative exchange sites associated with the surfaces of organic materials and clay minerals, adsorption reactions tend to restrict the mobility of cations. In order for leaching to occur, cations displaced from the exchange sites must be accompanied by equivalent anions in solution to maintain electroneutrality. Therefore the rate of nutrient leaching is determined by the presence and mobility of the anion carriers (Cronan, 1980; Johnson and Cole, 1977; McColl and Cole, 1968; Nye, 1981).

In temperate and tropical regions, the leaching process in undisturbed forest soils of pH values greater than 5 is dominated by the bicarbonate ion, which largely originates from intensive soil respiration. (Johnson et al., 1977). In arctic and subalpine forest ecosystems, where soil microbial activity is strongly inhibited by low temperatures, organic acids may constitute the major leaching agents (Cronan, 1980; Johnson et al., 1977; Ugolini et al., 1977).

2.1.6 Effect of Acid Rain on Leaching

Acid precipitation introduces potential cation carriers (SO_4 and NO_3) to the soil system in addition to the proton load. Alterations in the soil microbiological community may cause decreased nutrient availability due to changes in organic matter decomposition and carbon cycling. Chemical alterations brought about by

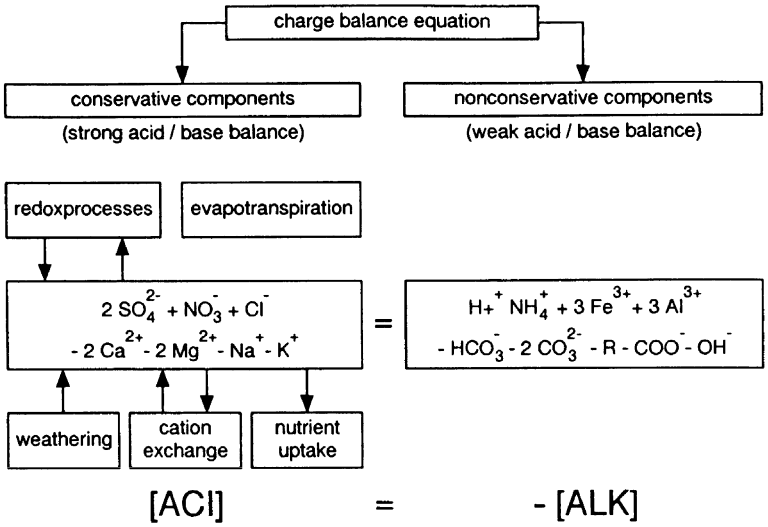


Fig. 2.8. The acid-base concept of fresh waters.

acidic deposition may cause some trace metals in the soil solution to reach levels that negatively affect trees, either directly or indirectly. These substances include Cd, Pb, Zn, Cu, Ni, V, Co, Mn, with the greatest concern being Al.

The sulphate anion can be removed (demobilized) through soil adsorption processes typically involving Fe and Al oxides and, to a far lesser extent, through biological uptake. In highly weathered soil enriched in Al and Fe sesquioxdes, the mobility of sulphate may be strongly restricted. Only immobilization through uptake by plants or microorganisms can effectively restrict this form of anion loss. Biological uptake of nitrate will have little influence on the overall leaching process in soils with inherently high nitrogen content.

2.1.7 Soil Buffering Systems

Depending on the pH range, differing buffering systems predominate in the soil (Table 2.1), although they are operative more or less simultaneously depending on the prevailing conditions. It is important to understand how these systems are affected by a changing deposition. In particular, it is essential to consider the problems related to reversibility of soil processes and acidification.

Buffer systems starting from a perpective of water and soil acidification cation exchange and the role of carbon dioxide and weak acids will be especially discussed. Aluminium dynamics beside the aspect of buffering properties, aluminium and its speciation is of considerable interest to the biological effects in soils and water (Figure 2.8). Weathering is the long term effect of acid deposition on soils and water is always associated with the weathering capacity of the soil. Sulphur and nitrogen dynamics are of special importance in interpreting the effects on acidification of changes in sulphur and nitrogen deposition. Sulphur adsorption/desorption processes and their reversibility are essential. Can a soil be saturated by nitrogen?

TABLE 2.1
Classification of buffering systems

pH-range	Buffer system
6.2 – 8.6	Calcium buffer
5.0 – 6.2	Silica buffer
4.2 – 5.0	Exchange buffer
3.0 – 4.2	Aluminium buffer
< 3.0	Iron buffer

How is nitrification and denitrification regulated in soils? These are key questions to answer in order to predict long-term acidification.

pH is defined as the negative common logarithm of the activity of hydrogen ions:

$$pH = -\log(H^+). \tag{2.1}$$

Charge balance equation:

$$\text{Cations} = \text{Anions}. \tag{2.2}$$

Cations:

$$[Na^+]+[K^+]+[H^+]+[NH_4^+]+2[Ca^{2+}]+2[Mg^{2+}]+ 2[Fe^{2+}]+ +3[Fe^{3+}]+3[Al^{3+}]+ \cdots = \tag{2.2a}$$

Anions:

$$[HCO_3^-]+[NO_3^-]+[HS^-]+[Cl^-]+[R\text{–}COO^-]+[H_3SiO_4^-]+ +[OH^-]+2[SO_4^{2-}]+2[CO_3^{2-}]+ \cdots , \tag{2.2b}$$

$$\text{Aci (acidity)} = [NO_3^-]+[Cl^-]+2[SO_4^{2-}]+ \cdots -[Na^+]-[K^+]-2[Ca^{2+}]-2[Mg^{2+}], \tag{2.3}$$

$$\text{Alk} = [HCO_3^-]+[R\text{–}COO^-]+[OH^-]+2[CO_3^{2-}]+[NH_4^+]-[H^+], \tag{2.4}$$

$$\text{Aci} = -\text{Alk}, \tag{2.5}$$

$$pH = f(\text{Aci}). \tag{2.6}$$

After Stumm et al. (1983), the acidification of aquatic systems should be understood as a large-scale natural titration reaction.

$$\text{Aci} = f(H^+). \tag{2.7}$$

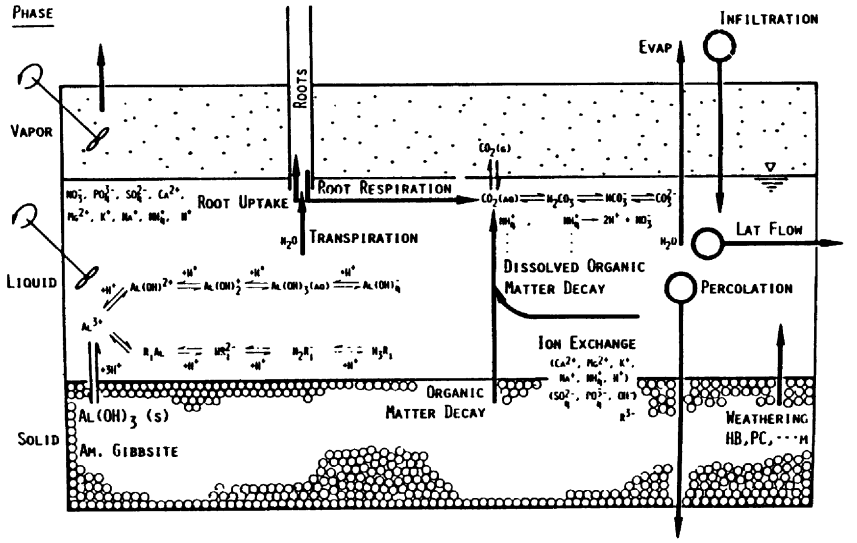


Fig. 2.9. Idealized soil layer showing major processes considered in the ILWAS model (from Gherini et al., 1985).

Then Aci and Alk are equal, i.e.

$$\text{Aci} + \text{Alk} = 0. \tag{2.8}$$

Self-purification capacity of terrestrial water resources against acidification is determined by their acid buffer capability. Buffering of an aqueous solution should be understood as its capability to counteract pH value changes after the addition of a strong acid or base. The following heterogeneous buffer reactions should be taken into account.

* Carbonate solution;
* Alumino -silicate weathering;
* Ion exchange;
* Aluminium mobilization;
* Ferrous mobilization etc.

The acid-base metabolism is decisively influenced by redox process as shown in Figure 2.8.

2.2 Watershed Acidification Models

2.2.1 ILWAS Model

The Integrated Lake Watershed Acidification model (ILWAS, Figure 2.9) was developed to simulate biogeochemical processes within a watershed and ascertain the effect of acidic deposition on surface water chemistry (Chen et al., 1983; Gherini et al., 1985). The model divides the ecosystem into a series of connected

TABLE 2.2
Hydrological and hydrogeochemical model parameters for ILWAS model

Hydrological Parameters	Geochemical Parameters
Land slope	Soil cation exchange capacity
Soil Permeability	Soil cation selectivity coefficients
Soil depth	Soil anion adsorption coefficients
Soil field capacity	Mineral weathering rate coefficients
Soil saturated moisture capacity	Soil nitrification rate coefficient
Seasonal evapotranspiration	Initial soil and surface water solute concentrations

homogeneous compartments that include distinct soil layers. It is one of the most complex models presently in use.

The model consists of two basic modules: hydrological and chemical. The hydrological module can be operated independently from the chemical module. The hydrologic processes simulated include canopy interception, throughfall, snow accumulation and melting, evapotranspiration, sublimation, unsaturated groundwater flow, saturated groundwater flow, stream and lake hydraulics. The chemical processes simulated include dry deposition, foliar exudation, snow leaching, nitrification, oxidation, decomposition, nutrient uptake, root respiration, mineral weathering, cation exchange, anion adsorption, Al dissolution-precipitation, and chemical equilibration in water.

The hydrologic state variables are interception storage on the canopy, snowpack depth on the lake and land surfaces, the soil horizon water contents, the water volume in the stream segments, the ice thickness on the lake and the water volume in the lake layers. The chemical state variables are the masses of each chemical species in the different phases (solid, liquid, and gas) of each model compartment. The chemical variables simulated include: SO_2, NO_x, H^+, NH_4^+, Al (total monomeric species), Ca^{2+}, Mg^{2+}, K^+, Na^+, SO_4^{2-}, NO_3^-, Cl^-, F^-, phosphate species, ANC, TIC, H_4SiO_4, and two organic acid analogues. Another state variable which influences the concentrations of dissolved species is temperature.

ILWAS is basically a lumped parameter model that attempts to minimize the heterogeneity of watershed parameters by dividing the watershed into horizontally 'homogeneous' sub-basins, stream and lake segments. Each sub-basin is segmented vertically into compartments representing canopy, snowpack and soil layers. The lake is also segmented vertically for the simulation of heat and chemical fluxes. The model time step is variable with a maximum time step of one day. However, much of the input data such as precipitation chemistry is supplied in monthly averaged form. Many of the model process coefficients are also supplied as monthly values. Model outputs can be daily, weekly, or annual as specified by the user.

The characteristics of the soil/overburden control hydrological pathways and rates of flow of water as it passes through the watershed. These flowpaths and rates determine what exchange, sorption, weathering, or biological reactions are most important. This is achieved by the ILWAS model by coupling its relatively

comprehensive hydrologic module to the chemical module in such a way that the differing chemical reactions receive different weights during model execution. The principle of conservation of mass is used to calculate water balances for the snowpack, soil layers, stream segments, and lake layers. Rate equations are used to relate water transport to driving forces in processes such as groundwater flow and evapotranspiration.

The ILWAS model permits the user to specify up to 5 primary minerals in the soil/overburden and bedrock of a catchment. Weathering is calculated as a function of the mass of mineral present and the H^+ ion concentration. The weathering function also uses a weathering rate coefficient that must be input by the user. Stoichiometric equations for the weathering reactions are used to account for H^+ consumption and cation liberation.

For phosphate and sulphate, the model uses a linear isotherm to represent anion adsorption. The removal of the strong acid anion from solution results in an increase in the solution phase alkalinity. The rate of the nitrification reaction, which results in a net decrease in alkalinity, is modeled using a Michaelis-Menten rate expression.

Each soil is assumed to have a specific cation exchange capacity. The exchange sites are occupied by the six major cations, Ca^{2+}, Mg^{2+}, K^+, NH_4^+, Na^+ and H^+. These are assumed to be in equilibrium with the cations in the soil solution. The exchange equations are all written with Ca^{2+} as the reference cation. Divalent–monovalent cation exchange is represented by the use of Gapon equations and the divalent-divalent by Kerr equations.

Aluminum dissolution and complexation in the model is simulated by both equilibrium and rate limited processes and the model user must specify the process that applies in each soil horizon. Aluminum can be in equilibrium with solid phase gibbsite, with hydroxide, fluoride, sulphate, and organic acid.

Terrestrial basin nutrient uptake is also simulated in the model. It uses a monthly varied coefficient to relate annual productivity with monthly nutrient demand. Depending on the specified chemical composition of the biomass, nutrient uptake may not remove the same number of equivalents of base cations as strong acid anions. The imbalance results in a change in the soil solution alkalinity.

The ILWAS chemical module contains all of the equations and variables necessary to provide a quantification of the "salt effect". The salt effect describes the phenomenon whereby H^+ and Al^{3+} concentrations in soil solution increase in response to an increase in total mineral acid anion concentration. The alkalinity depression may be accompanied by little reduction in soil solution pH due to cation exchange, but may result in a major reduction in surface water pH.

Soil water generally degasses CO_2 when it emerges, and the pH rises. In the ILWAS model, pCO_2 dynamics are simulated through root respiration and organic matter decay functions. A mass balance equation is used to calculate the sum of the aqueous and gas phase CO_2 in each layer and Henry's law is used to determine the partitioning between the gas and aqueous phases.

In most surface waters, over 95% of the alkalinity is present as bicarbonate. However, in dilute surface water systems that are susceptible to acidification, organic acid ligands and aqueous aluminum species become significant components of the buffering system. The alkalinity equation includes these other species in the

ILWAS model. In-lake processes that generate or consume alkalinity such as algal productivity and decay are also included in the model.

The ILWAS model represents the watershed as a series of Continuously Stirred Tank Reactors, (CSTR). Eulerian numerical integration techniques are used for the solution of the simultaneous equations. Large tri-diagonal matrices are generated based on the mass balances of ionic species in the layered systems. These matrices are then solved in the model by subroutine TRDIAG. A number of numerical solution techniques are used in the model to solve the chemical process equations, including the Newton-Raphson technique, the delta-squared technique, the bisection method, and the regula-falsi method.

The model inputs are of two types: time variant meteorological, air quality and precipitation data that are used to drive the model, and time invariant data that describe the hydrological and chemical properties of the watershed (Table 2.2). Input file RHUNIT contains the lake-watershed physical data. The lake-watershed chemical data and rate coefficients are read into the model from file RQUNIT. The largest input data file is RMUNIT which contains the daily meteorological data. Input file RSUNIT contains the monthly (or daily) precipitation and air quality data.

The model output includes 1) error messages generated by an internal subroutine (CHECK), 2) input echo (printout of input data in formatted tables), 3) a hydrologic summary, 4) a water quality summary, 5) flux data, 6) cycle charts, and 7) plot files. The hydrologic summaries can be user- specified for a daily, monthly, or yearly basis. The summaries provide values for the state variables for various points in the watershed. The water quality summaries can also be user-specified for the same time intervals as the hydrologic data. These summaries provide the masses of chemical constituents stored in each model compartment. The flux data describe the ion fluxes associated with each individual process and are always output as the cumulative amount in meq for the period of simulation. The cycle charts can be user specified for the individual chemical species. They show the storage, change in storage, reaction, and fluxes between the various model compartments. The plot files are user specified and can be of volumetric flow rates and solution chemical concentrations at various points in the watershed on a daily, weekly, or annual basis.

2.2.2 Birkenes Model

The Birkenes model (Figure 2.10) was originally developed for examining the effects of acidic deposition in Norway (Christophersen et al., 1982; Wright, 1984). Because of their importance in acidification chemistry and the need for simplicity, only the cations H^+, Mg^{++}, and Al^{n+} and the anion SO_4^{2-} were included in the model. The Birkenes model is divided into two layers or reservoirs (Table 2.3). The hydrological model consists of two compartments, with the upper (A) representing a mostly organic soil horizon providing quick flow, and the lower (B) deeper soil layers providing base flow. A snow reservoir is included when needed. Water may be rooted from the upper compartment directly to the stream or to the lower compartment. The present model includes the possibility of "piston flow", which

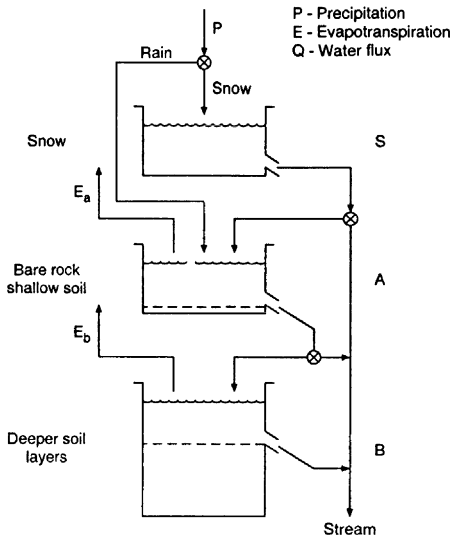


Fig. 2.10. The structure of Birkenes model (from Christophersen et al., 1984)

allows water from A to flush out an equivalent amount of water from B, thereby altering the chemical quality but not the quantity entering the stream.

The mobile anion concept is a key feature of the model. The relative concentrations of cations in the water in the upper reservoir are determined by exchange reactions. In the upper reservoir, wet and dry deposition, adsorption/desorption and mineralization of S were considered, while within the lower reservoir only SO_4^{2-} adsorption/desorption was included. It was assumed that: a) all S entering the watershed as wet deposition is SO_4^{2-}; b) dry deposition includes SO_2 gas and SO_4^{2-} aerosols; and c) S exists only as dissolved SO_4^{2-} in stream water. Dry deposition was estimated using mean concentrations in the air multiplied by deposition velocity. Christophersen et al., (1982) noted that "processes such as biological activity and mineralization of sulphur compounds thus cannot be considered quantitatively but are included indirectly in the model".

The model assumes that SO_4^{2-} of atmospheric origin accumulating in the upper zone remains potentially water soluble. The aqueous concentration of SO_4^{2-} is only a function of the total amount of SO_4^{2-} in the solid phase as indicated by the following relationship, which may implicitly include organic S transformations as well as adsorption-desorption reactions:

$$C_a = K_a \cdot F_a, \tag{2.9}$$

where K_a is a linear, equilibrium SO_4^{2-} adsorption-immobilization term in $m^2\,L^{-1}$; C_a ($mg\,L^{-1}$) is solution SO_4^{2-} concentration; and F_a is the solid phase SO_4^{2-} concentration ($mg\,m^{-2}$). This formulation is the same as that used in the ILWAS model.

The main equations determining the concentrations in discharge from the upper reservoir are:

TABLE 2.3
Summary of processes included in Birkenes model (Christophersen et al., 1982)

Parameter	Upper Reservoir	Lower Reservoir
H_2O	Precipitation, evapotranspiration, infiltration and discharge to stream	Infiltration, evapotranspiration, discharge to stream
SO_4	Wet and dry deposition, adsorption/desorption, mineralization	Adsorption/desorption
Ca + Mg	Ion exchange	Release by weathering and adsorption/desorption
H^+	Ion exchange and equilibrium with gibbsite	Consumption by weathering, adsorption/desorption, equilibrium with gibbsite
Al^{3+}	Equilibrium with gibbsite	Equilibrium with gibbsite, adsorption/desorption

The charge balance:

$$[H^+] + [Na^+] + [Ca^{2+}] + [Mg^{2+}] + [Al^{3+}] = [SO_4^{2-}] + [Cl^-] + [HCO_3^-]. \quad (2.10)$$

Equations determining cation concentrations:

$$K_{Al} = [Al^{3+}]/[H^+]^3, \quad (2.11)$$

$$K_{Ca} = [Ca^{2+}]^3/[Al^{3+}]^2, \quad (2.12)$$

$$K_{Mg} = [Mg^{2+}]/[Ca^{2+}]. \quad (2.13)$$

$[Na^+]$ is assumed not to be generated by chemical processes in this reservoir, i.e. it is conservative in the upper reservoir or at least of minor importance compared to Ca^{2+} and Mg^{2+}.

Equations determining anion concentrations:

$$[HCO_3^-] = KP_{CO_2}/[H^+]. \quad (2.14)$$

P_{CO_2} is the CO_2 pressure which is assigned a fixed value of 10 times atmospheric P_{CO_2}.

$$[SO_4^{2-}] = F_A/(K_A + W_A). \quad (2.15)$$

F_A is the amount of water soluble sulphate both in the solutions and in solid phases, W_A is the water level and K_A is a constant. Mineralization adds sulphate to F_A during dry periods and when there is a thick snow pack.

$$S_{\text{miner A}} = K_{\text{min A}}(A_{\min} - 0.2W_A)/A_{\min}. \quad (2.16)$$

W_A is the water level in Reservoir A and $K_{\text{min A}}$ is a constant mineralization rate. If $W_A > A_{\min}$, $S_{\text{miner A}}$ is set equal to zero. When there is a snowpack with a

water equivalent of > 5 cm, mineralization is assumed to occur at double the rate of that in Equation (2.16). During wet periods, sulphate may be removed from the solution:

$$S_{\text{red A}} = (T_s - 5.0)K_{\text{red A}}/10.0 \quad (\text{for } 5 < T_s < 15°\text{C}), \tag{2.17}$$

where T_s is the soil temperature and $K_{\text{red A}}$ the maximum reduction rate. At $T_s > 15°\text{C}$, $S_{\text{red A}} = K_{\text{red A}}$, and at $T_s < 50°\text{C}$, $S_{\text{red A}} = 0$. Like Na^+, Cl^- is assumed to be conservative.

For the lower reservoir, Equations (2.10) and (2.11) are assumed valid, but the constant K_{Al} in Equation (2.11) has different values in the two reservoirs. H^+ ions are consumed by weathering, producing Ca^{2+}, Mg^{2+} and Na^+ ions. The amount of Ca^{2+} , Mg^{2+} and Na^+ obtained is equivalent to the amount of H^+ ion removed in the weathering process, which is calculated by the following equation:

$$H^+_{\text{wthr}} = K_{wB}([H^+]_B - [H^+]_{EB}), \tag{2.18}$$

where H^+_{wthr} is the amount of H^+ weathering, K_{wB} is a rate constant, W_B is amount of water, $[H^+]_B$ the $[H^+]_{EB}$ in Reservoir B. Equation (2.14) is used for $[HCO_3^-]$, but with a CO_2 pressure varying with temperature. Cl^- is assumed to be conservative. The SO_4^{2-} concentration is controlled by two pools of sulphur, one soluble and one fixed. Transfer between these two pools is dependent upon the amount of water in the reservoir and the difference between the current and equilibrium SO_4^{2-} concentrations. A small amount of reduction is included and, unlike in Reservoir A, the amount is solely dependent upon the water level in the reservoir. However, its contribution is small. Otherwise, the concentration is calculated using an equation similar to Equation (2.15) with a different value for F_A.

For the stream, the Equations (2.10) and (2.14) are assumed to apply. The CO_2 pressure is set at four times the partial pressure of CO_2 in the atmosphere. Equation (2.11) is only in force if the product $[Al^{3+}]\,[H^+]^{-3}$ exceeds a critical value, in which case $Al(OH)_3$ is assumed to precipitate. The constant was set at a high value and no removal of Al^{3+} occurred. The concentrations of Al^{3+}, as well as other cations and anions, are obtained by mixing the discharges from two reservoirs.

2.2.3 TMWAM Model

The Turkey-Mersey Watershed Acidification Model (TMWAM) has four major submodels that deal with hydrology, sulphate, cations and DOC respectively. (Lam et al., 1988, 1989; Bobba and Lam 1988, 1989, 1990). The hydrology submodel drives the other models. and the authors emphasize the importance of having an accurate hydrological model before attempting to model the flux of chemical constituents (Figure 2.11).

The hydrological model (Figure 2.12) consists of mass balance calculations for snow accumulation and melting, soil moisture budgeting, runoff generation, and hydrological routing. Structurally, the model represents the watershed as a set of

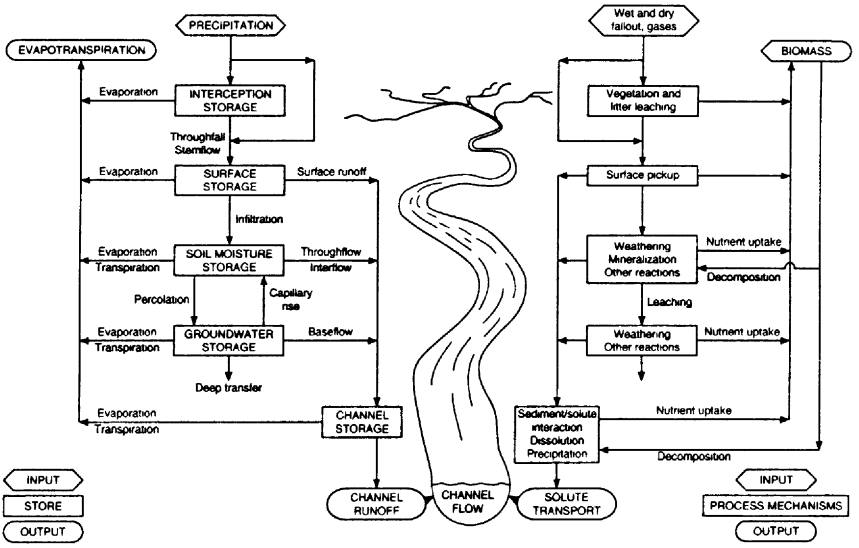


Fig. 2.11. Schematic representation of the precipitation/runoff processes operating in a basin and associated transported mechanisms governing solute concentration in streamflow.

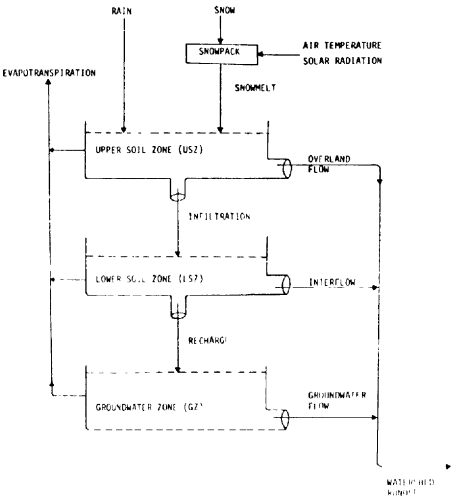


Fig. 2.12. Conceptualized TMWAM hydrological model (from Bobba and Lam, 1990).

subbasins, each with six compartments: a) atmospheric inputs, b) the snowpack, c) the upper soil reservoir (surface litter zone); d) a lower soil reservoir (unsaturated zone, including the root zones of most of the perennial vegetation); e) a groundwater reservoir; and g) a stream.

The model uses three basic rules of water movement. The soils through which water flows are divided into three compartments, called the upper, lower and groundwater reservoirs. Water entering the upper reservoir normally infiltrates

downward to the lower reservoirs, unless the later reservoirs are full. When this occurs, runoff originates from the upper reservoir. Such runoff is called surface runoff. Surface runoff water tends to have relatively little opportunity for contact with soils with high acid neutralizing capacities, and thus its quality tends to be similar to that of precipitation. Runoff from the lower reservoirs is called interflow and base flow (or groundwater flow). The base flow is the only component of flow during dry periods, since the upper reservoirs become dry. The groundwater flow tends to be older water (i.e. the water has been in contact with soils for longer periods of time) and is thus typically higher in base cation concentrations. Given these substantial differences between the quality of groundwater flow and subsurface flow, it is evident how important it is that the model accurately estimate the relative contribution of each to runoff.

The hydrology model accounts for precipitation, evapotranspiration, infiltration to the lower reservoirs, and discharge of water from subsurface reservoirs to the stream.

The aerial infiltration (I) is proportional to the volume remaining in the upper soil zone:

$$I = \mathrm{NS}\left(1 - \frac{\mathrm{USZM}}{\mathrm{USZC}}\right), \tag{2.19}$$

where NS, USZM, USZC are net supply, upper soil zone moisture content and upper soil zone capacity, respectively.

The outflow (O) from any of the three zones is linearly proportional to the reservoir storage (S), i.e

$$O = \alpha S. \tag{2.20}$$

Evapotranspiration (e_t) is proportional to the potential evaporation rate (e_p) required by heat balance considerations and to the storage, i.e.

$$e_t = Be_pS, \tag{2.21}$$

where B is linear reservoir constant, namely B_u, B_l and B_g for evapotranspiration rate at USZ, LSZ, and GWZ, respectively. The continuity equations for the three zones are:

$$\frac{\partial(\mathrm{USZM})}{\partial \mathrm{t}} = \mathrm{NS}\left(1 - \frac{\mathrm{USZM}}{\mathrm{USZC}}\right) - (\alpha_0 + \alpha_p + B_u e_p)\,\mathrm{USZM}, \tag{2.22}$$

$$\frac{\partial(\mathrm{LSZM})}{\partial \mathrm{t}} = \alpha_p \mathrm{USZM} - (\alpha_i + \alpha_{dp} + B_l e_p)\,\mathrm{LSZM}, \tag{2.23}$$

$$\frac{\partial(\mathrm{GZM})}{\partial \mathrm{t}} = \alpha_{dp} \mathrm{LSZM} - (\alpha_g + B_g e_p)\,\mathrm{GZM}, \tag{2.24}$$

where α is a linear reservoir constant, namely α_o and α_p for overland flow and percolation in the USZM; α_i and α_{dp} for interflow and deep percolation in the LSZM, and α_g groundwater flow in the GZM, respectively. The above system

equations are first order ordinary differential equations and are solved by analytical methods (Bobba and Lam, 1984).

The atmospheric deposition of air pollutants occurs in both the dry form and the wet form. It is assumed that for each time step, the contaminants from both dry and wet deposition are accumulated and added to the snowpack contaminant concentration. Snowmelt is a hydrological event that may vary greatly from year to year. Marmorek et al. (1989) reviewed models of snowmelt and concluded that simple models of snowpack accumulation and ablation have been relatively successful in predicting the main physical characteristic of snowmelt at a number of locations. There is a substantial body of literature on the physics of snowmelt, and it is quite clear that a forest canopy modifies the melting process. The delay of snowmelt under a forest canopy compared to clear cut locations could substantially alter hydrological pathways. Within a single watershed, or within a region, this type of spatial variability could produce important differences in hydrological response and, hence, in surface water composition.

In large river basins that develop large, seasonal snowpacks, snowmelt induced runoff usually produces higher peaks than those of rain events. On the other hand, in small watersheds with short water retention times, intense rainstorms can produce higher runoff if the snow pack is "ripe". If the snowpack is dry, all the rain may be retained by the snowpack. In many watersheds, the annual discharge maximum occurs when rain falls during the snowmelt period or just after the snowpack has disappeared, leaving a wet watershed. Semkin and Jeffries (1988) found that in the Turkey Lakes Watershed, Ontario, spring rainfall can have a deleterious effect on surface waters because soils are already saturated with meltwater. If the rainfall is intense, the high permeability of any remaining snowpack, combined with a high water table, will permit a rapid influx of rainwater to streams and lakes. If rain events are common throughout the period when a snowpack is present, such as in southwest Pennsylvania, snowpack melting may be frequent and intermittent, and the meltwater discharge is less dramatic than in locations with colder winters where a lengthy snowpack accumulation occurs (Sharpe et al., 1984).

In the case of sulphate ions, a simple mass balance equation for the sulphate concentration in the snow pack is:

$$\frac{\partial(VC)}{\partial t} = \beta(t) - mC_mA, \tag{2.25}$$

$$\beta(t) = R_dC_dA + PC_wA, \tag{2.26}$$

where V is the volume of snowpack (m^3) as computed from the hydrological model, C is the concentration of sulphate in the snowpack (g/m^3), β is dry deposition rate (m/d), C_d dry deposition concentration (g/m^3), A is watershed area (m^2), P is precipitation rate (m/d), C_w is precipitation concentration (g/m^3), m is the melting rate (m/d), and C_m is concentration of the melt water (g/m^3). Thus, the melting rate m in Equation (2.25) is an important model parameter. In general, potential snowmelt,

$$m = 0, \qquad T < 0\,^{\circ}\mathrm{C}, \tag{2.27}$$

$$m = a\mathrm{DD}, \qquad T > 0\,^{\circ}\mathrm{C}, \tag{2.28}$$

where m is daily snowmelt rate (m^3/d); a is a proportionally constant for snowmelt per degree-day ($m^3\ °C^{-1}\ d^{-1}$); T is air temperature (°C), estimated as the average of the daily maximum and minimum temperature; and DD degree-days per day (°C d/d), computed as the integral of air temperature with time over those portions of the day when the temperature is above the freezing point.

Since the fluctuation of air temperature during the diurnal cycle is not always known, a triangular distribution can be assumed. The resulting expression for degree-day-is:

$$\mathrm{DD} = 0 \text{ for } T_{\max} \leq 0, \tag{2.29}$$

$$\mathrm{DD} = \frac{1}{2}\frac{T_{\max}^2}{T_{\max} - T_{\min}} \text{ for } T_{\min} < 0 < T, \tag{2.30}$$

$$\mathrm{DD} = T \ 0 \leq T_{\min}, \tag{2.31}$$

where $T_{\max}$ is maximum air temperature (°C), and $T_{\min}$ minimum air temperature (°C). The proportionality constant a is a model parameter to be determined during model calibration. Thus, snowmelt at any point in time is highly dependent upon the condition of the snowpack as well as the heat energy flux. The snowpack condition is generally identified in terms of its "cold content" and water holding capacity or liquid-water deficiency. Basically, energy is required to raise the snowpack temperature to 0 °C and to melt enough snow to saturate the water-holding capacity of the snowpack before liquid water can reach the soil surface. Once these conditions are met, any additional heat input to the snowpack causes snowmelt runoff. Although these processes of a snowpack are generally difficult to quantify, Equations (2.27) to (2.31) represent a simple approach to define the melting rate.

In the hydrogeochemical model (Figure 2.13), the Birkenes model has been modified for Canadian conditions. The 3 major modifications are: addition of one more soil layer in between the top and bottom layers in the original model to make it compatible with the hydrological model, elimination of the Al component in the Birkens model, since for the pH range 5.5–7.00, Al kinetics are not important.

Kinetic equations are essentially the same equilibrium constants for other variables. The equations have been discussed in detail in Birkenes Model. For the upper soil reservoir, the mass balance of SO_4 is computed from wet and dry deposition, snowmelt and oxidation of S as sources, and reduction and adsorption of sulphate as sinks. The computed water flow and moisture content in the soil are considered to allow this mobile anion to enter, to be retained, or leave the soil layer. Sulphate concentrations are computed for each of the soil layers, although the sources and the sinks of sulphate in these layers involve only adsorption (desorption) and reduction. The concentrations of H, Ca^{2+}+Mg^{2+} and HCO_3^- for the upper soil layer are then computed from the charge balance equation. The chemical kinetics for soil layers B and C are similar, except that layer C often is a thicker and has slower hydrological constants. Sea salt correction has been applied to the major ions so that $[Na^+]$ is mainly governed by cation exchange processes and $[Cl^-]$ is not included in the model. The third major modification is coupling the organic acid model with the inorganic acid model.

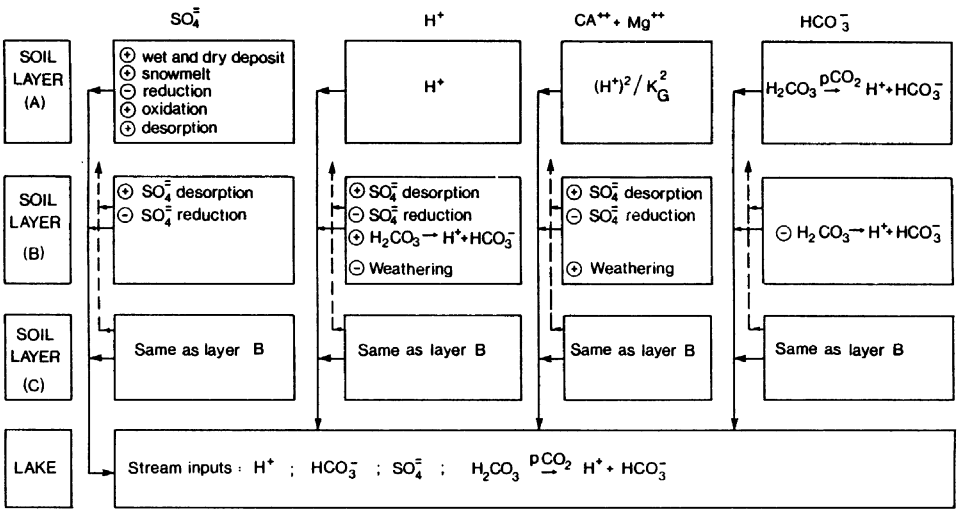


Fig. 2.13. Schematic diagram of the interfacing of hydrological and hydrogeochemical models (from Lam et al. 1988).

In addition to inorganic ions, organic anions [A] are also considered in the modified TMWAM model. All these ions are subjected to hydrological transport among the soil layers and their flow paths are determined by the evapotranspiration and hydrological flows between layers and to the streams. The concentrations of a particular ion in a given layer therefore depends on the amount of the ion entering and leaving that layer and on the soil moisture content. In the case of sulphate there are processes such as sorption/desorption and oxidation/reduction, which act as internal sources or sinks that can increase or decrease the amount of the ion within the moisture content volume V of the layer, i.e.

$$\frac{\partial V[\mathrm{SO_4}]}{\partial t} = \text{sorption} - \text{desorption} + \text{oxidation} - \text{reduction} + \\ + \text{deposition} + \text{snow}. \tag{2.32}$$

At each soil layer and in the stream, charge balance is required.

$$[\mathrm{H^+}] + [\mathrm{Ca^{2+}}] + [\mathrm{Mg^{2+}}] + [\mathrm{Na^+}] + [\mathrm{Al^{3+}}] = [A^{3-}] + [\mathrm{SO_4^{2-}}] + [\mathrm{HCO_3^-}]. \tag{2.33}$$

The organic anion $[A]$ equation commonly used is the one developed by Oliver et al. (1983),

$$[A'] = \frac{R[\mathrm{DOC}]K'}{(K' + [\mathrm{H^+}])}, \tag{2.34}$$

where R is the charge density and is estimated by Oliver et al. (1983) to be 10 μeq $[A']$ $\mathrm{mg^{-1}C}$ of DOC. The constant K' is a function of $[\mathrm{H^+}]$. Here $[A']$

is an aggregated organic anion determined by laboratory methods which does not distinguish free organic anions from those bound by metals such as Fe and Mn. It is with these considerations that we can use the following approach. First, we use the charge balance approach for defining $[A]$ as required by the mass balance equation which excludes metal-organic complexes. Secondly, we consider a triprotic organic acid with mono-, di-, and tri-valent anions: H_2A^-, HA^{2-} and A^{3-}. The three dissociation constants are represented by

$$H_3A \overset{K_1}{\Longleftrightarrow} H_2A^- + H^+ , \tag{2.35}$$

$$H_2A^- \overset{K_2}{\Longleftrightarrow} H_2A^{2-} + H^+ , \tag{2.36}$$

$$H_3A^{2-} \overset{K_3}{\Longleftrightarrow} A^{3-} + H^+ , \tag{2.37}$$

$$[\text{organic anions}] = [A] = [H_2A^-] + [HA^{2-}] + [A^{3-}]. \tag{2.38}$$

It follows that

$$\frac{K_1}{[H^+]}\left(1 + \frac{K_2}{[H^+]} + \frac{K_2K_3}{[H^+]^2}\right) = \frac{[A]}{H_3A} . \tag{2.39}$$

If $[A] = [H_3A] = R[\text{DOC}]$, then

$$[A] = R[\text{DOC}]\frac{aK_1}{aK_1 + [H^+]} , \tag{2.40}$$

where $a = 1 + K_2/[H^+] + K_2K_3/[H^+]^2$. Thus, although we can use a triprotic organic acid, the constants can be combined into one equation which completes the set of equations needed for solving the unknown variables. By varying the values of the equilibrium constants K_1, K_2 and K_3, the composition of $[A]$ can be changed, for a given pH. Empirically, as in the case of Oliver et al. (1983), we relate $[A]$ concentration to the DOC concentration through the charge density R.

The following initial values are required before calibrating the model: water amounts in each reservoir; water soluble sulphate on the solid phase in upper reservoirs and sulphate concentration in the lower reservoir.

The following input values are required to run the model: the daily precipitation volume; the daily precipitation SO_4 and air SO_2 concentration; the daily maximum and minimum air temperature, and the daily solar radiation.

The following parameters are required for the hydrological model: maximum storage for upper reservoir; maximum and minimum storage for lower reservoirs; routing, infiltration and evapotranspiration parameters; and reservoir drainage parameters; and for the hydrogeochemical model: dry deposition velocities for SO_4 and SO_2; SO_4 mineralization rate (Upper reservoir); cation production rate during dry periods (upper reservoir); SO_4 mineralization rate (upper reservoir); cation production rate during dry periods (upper reservoir); SO_4 adsorption/desorption equilibrium constant (upper reservoir); SO_4 adsorption/desorption equilibrium constant and half time to equilibrium (lower reservoirs); Gibbsite solubility product; lime potential for ion exchange; and weathering rate half time (lower reservoir).

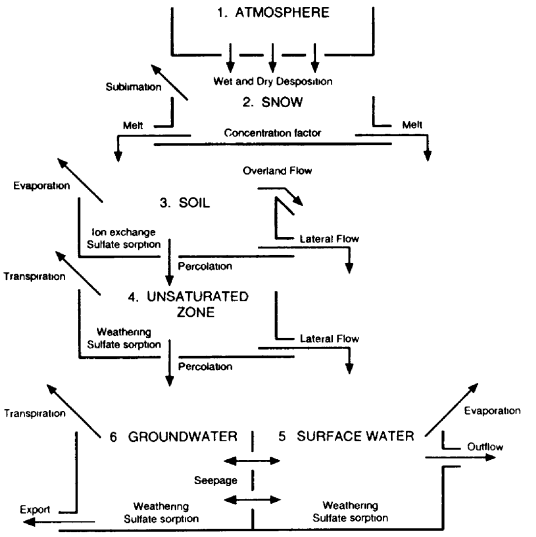


Fig. 2.14. The structure of hydrologic balance in the Trickle Down Model (from Schnoor et al., 1984)

2.2.4 Enhanced Trickle-Down Model

The Enhanced Trickle-Down model (ETD, Figure 2.14) is a compartmentalized model treating each of 9 compartments as a CSTR. This model consists of three submodels. The hydrologic submodel component simulates snowmelt, interflow, overland flow, groundwater flow, frozen ground driven processes, seepage, and evapotranspiration. The alkalinity submodel is a set of mass balance equations. Alkalinity is considered to be conservative. Sources of alkalinity included in the model are: ion exchange, chemical weathering, sulphate adsorption and sulphate reduction. The sulphate submodel, like the alkalinity model, is a set of mass balance equations.

The model inputs are time-series of precipitation, and alkalinity, sulphate and chloride deposition. The system of 20 differential equations of the four submodels is solved by using a variable time step 4th order Predictor-Corrector numerical scheme.

The model was developed to perform acid precipitation assessments on a number of seepage lakes. The objective of the modelling effort was to provide an "in-between" approach; that is, the model is only detailed enough to solve for alkalinity concentrations in surface water, soil water and groundwater. Alkalinity (acid neutralizing capacity, ANC) was selected as the master variable because it represents the aggregate of all acid/base reactions that occur in the watershed. Thus, any reaction that occurs in the watershed (e.g., sulphate reduction, biomass synthesis, ammonia oxidation, or chemical weathering) must be formulated in terms of its effects on alkalinity via stoichiometry. The model is developed to calculate the alkalinity, pH, and aluminum concentrations of a lake. By adjusting

the acidity of precipitation or the amount of dry deposition, the effects of various loading scenarios can be determined.

The model is based on an alkalinity mass balance. The key processes considered are chemical weathering, ion exchange, and hydrologic flow paths, and chemical weathering is the key. Based on the observed water chemistry, it can be shown that the predominant watershed reaction that modifies precipitation chemistry is the release of cations (Ca^{2+}, Mg^{2+}, Na^{+}, K^{+}) coupled with base production and the possible formation of bicarbonate alkalinity. Weathering and ion exchange phenomenon can be important processes in sediment-water as well as soil-water interactions, so the kinetics discussed here apply also to the sediment-water interface as well as soil-water interactions.

The model uses a kinetic approach which makes sense since primary minerals are at chemical equilibrium with respect to the aquatic and terrestrial environment at these locations with the possible exception of gibbsite. Even ion exchange is not likely to be at chemical equilibrium during precipitation events, and especially considering the possibility of macropore flow.

Neutralization processes that occur within the lake are presently calculated in the Trickle-Down model as a zero-order input of alkalinity at the sediment-water interface. The rate constant for neutralization is determined from model calibration with observed or field data. Aluminum concentrations are derived from alkalinity and pH output variables of the model and are calculated assuming saturation of microcrystalline gibbsite.

Ion exchange is another key process that is included in the Trickle-Down model. Cation exchange reactions can occur between hydrogen ions and base cations producing the same net effect as mineral dissolution but with more rapid kinetics. Using a calcium base exchange site ($-R$) as an example:

$$\mathrm{Ca}{-}R + 2\mathrm{H}^{+} \rightarrow \mathrm{H}_2{-}R + \mathrm{Ca}^{2+}. \tag{2.41}$$

The kinetics of this reaction are assumed to be directly proportional to the sum of exchangeable bases in the upper soil horizons.

$$\frac{\mathrm{d}(VA)}{\mathrm{d}t} = kMV_e(r_0 - r), \tag{2.42}$$

where:

A = alkalinity concentration, eq/m^3;
V = water column within the compartment, m^3;
t = time, days;
k = ion exchange rate constant, day^{-1};
M = bulk density of soil, kg/m^3;
V_e = total environmental volume of the compartment, m^3;
r_0 = initial sum of exchange bases, eq/kg;
r = exchangeable base depletion, eq/kg.

The rate constant, k, can be a function of hydrogen ion or neutral salt concentration to the fractional power, temperature, mean soil particle size, or flow rate

through the soil, but only hydrogen ion dependence has been tested so far. The exchangeable base depletion term, r, allows estimation of delayed response systems. As r increases during the course of a long-term simulation, the base exchange rate decreases until such time as all the exchangeable bases are depleted ($r_0 - r = 0$). Mineral dissolution (chemical weathering) would continue, and it could serve to replenish the depleted exchangeable bases. This is a mechanism that is easily facilitated in the Trickle-Down model, and plans exist for implementation in the model code. To assess delayed response systems, it will be necessary to know the size and quantity of weatherable minerals (e.g., feldspars, and amphiboles) in the soil horizons.

Hydrologic processes are another key element of the model. The model is similar to the ILWAS model which was reported by Chen et al., (1983). It includes snowpack hydrology, overland flow or stream flow, infiltration, interflow, and groundwater exchange between the lake and the groundwater aquifer.

The model can be resolved into both its temporal and spacial increments. For the temporal scale, the model requires daily precipitation inputs and can be used to simulate many decades. The compartments are divided into horizontal slices of variable thickness: a snowpack compartment, approximately a half meter upper soil horizon, a variable unsaturated zone, and a saturated zone. Storage of water in each compartment occurs during and following precipitation events, so the water volume of each compartment is variable over time for a given watershed.

The key assumptions of the model are as follows: a) chemical weathering and ion exchange are the main processes that neutralize acidic deposition and control water chemistry; b) the rates of these reactions are dependent on the fractional order hydrogen ion activity in solution, the mineralogy of the soils, base exchange, base exchange capacity, and factors determining the hydrologic flow paths (soil permeability, hydraulic conductivity, and depth to bedrock); c) sulphate retention in the watersheds to which the model has been applied is relatively small and can be neglected; and d) free monomeric aluminum concentrations can be estimated from microcrystalline gibbsite solubility in pore waters.

The model inputs are thermodynamic constants for carbonate equilibrium and aluminium internal to the model can be obtained from the literature. Vertical permeability and weathering reaction rates are initially estimated from literature values but are checked or modified by model calibration. All other input variables must be measured in the field or obtained from observed data.

The model can be used to determine seasonal and long-term risk of acidification. The output consists of all input data, followed by a tabular output of flows, alkalinity concentrations, and pH values at each print interval for each compartment (snow, soil, unsaturated zone, saturated zone and lake). An annual budget of the sources and sinks of alkalinity is printed.

2.2.5 RAINS Model

The model constructed is being used as a component of the IIASA, Regional Acidification Information and Simulation model (RAINS). The model consists of four modules (Figure 2.15). The model computes the lake acidity distributions for

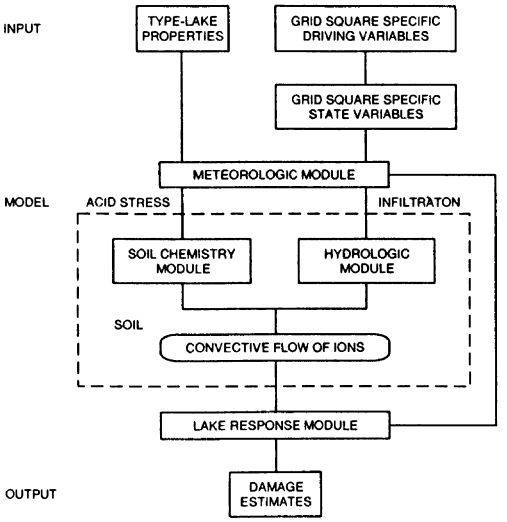


Fig. 2.15. The structure of the RAINS model (from Kamari)

the lakes within a given lake region as a function of watershed characteristics and local deposition. Details of the model are presented by Kamari (1985).

Each watershed is divided conceptually into four spatial sectors: snowpack, upper soil layer, lower soil layer, and lake. Different modules of the model compute the hydrology and the flux of ions contributing to the acidity and alkalinity of the lake water. The meteorologic module regulates the input flows of water and deposition to the soil and directly to the lake. The hydrological and soil chemistry modules together determine the flow of ions leached from the terrestrial watershed to the lake. New equilibrium concentrations in the lake water are then computed in the lake response module.

The meteorological module transforms monthly sulphur deposition into acid stress to various sectors of the watershed. The monthly precipitation is broken down into rain and snow according to local mean monthly temperature. Snowpack accumulates and melts at a temperature-dependent rate. Other equations in this module account for storage of wet and dry deposition in snowpack, release with meltwater, and direct hydrogen ion deposition to soil and lake.

The hydrological module routes precipitation into quickflow (surface runoff), base flow (groundwater flow) and flow between soil layers. These flow components are computed based on rates of precipitation and evapotranspiration and geomorphological characteristics such as soil depth, surface slope, hydraulic conductivity and volumetric water content of soil.

The soil chemistry module uses the same analytical approach as the IIASA forest soil acidity submodel to estimate hydrogen ion concentration in the different soil layers. However, acid stress in this module is input on a monthly, rather than annual basis. This monthly input is based on deposition, snowmelt rate, relative amount of rainfall versus snowfall, and other considerations derived from

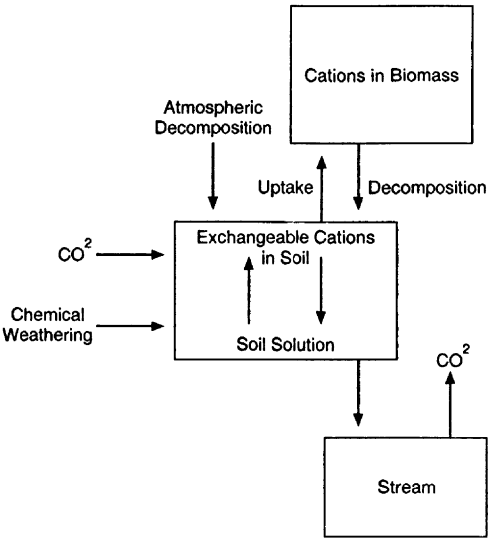


Fig. 2.16. Schematic view of the MAGIC model (from Cosby et al., 1986).

the meteorological module. The loads of ions which contribute to acidity and alkalinity of the lake are then computed from a mass balance equation.

The lake module calculates the pH and alkalinity of the lake based on the ion load. The loads are assumed to be mixed within a mixing layer which depends on climate and season. Finally, the change in lake acidity is calculated according to equilibrium reactions of inorganic carbon species.

2.2.6 MAGIC Model

The MAGIC model (Model of Acidification of Groundwater In Catchment) shown schematically in Figure 2.16 is a relatively simple, yet process-oriented model for estimating catchment response to acidic deposition. The model is mainly intended for studies of long term changes in water quality, and no attempts are made to model events. The model has roots in the Reuss (1980), Reuss-Johnson (1985, 1986) conceptual system. The model is based on mathematical equations for the following processes (Cosby et al., 1985a,b): a) Anion retention by soils (e.g. sulphate adsorption by soil), b) Cation exchange by watershed soils, c) Alkalinity generation by carbonic acid dissociation, d) Dissolution of aluminium minerals and mobilization of aluminium, and e) Weathering of minerals as a source base of cations (Ca, Mg, Na, K).

The general model contains three compartments (Figure 2.16) that are assumed to be internally homogeneous. This assumption implies that vertical stratification in the soils in the watershed is unimportant, or equivalently, that all water reaching the streams contacts and has its chemical quality determined by a single layer within the soil column.

The most serious effects of acidic deposition on catchment surface water quality are thought to be decreased pH and alkalinity and increased base cation and

aluminum concentrations. In keeping with the aggregated approach to modelling whole catchments, a relatively small number of important soil processes, processes that could be treated by reference to average soil properties, could produce these responses. MAGIC model assumes that atmospheric deposition, mineral weathering and exchange processes in the soil and soil water are responsible for the observed surface water chemistry in the watershed. Alkalinity is generated in the subsurface system by the formation of bicarbonate from dissolved CO_2 and water.

$$CO_2 + H_2O = H^+ + HCO_3^-. \tag{2.43}$$

Bicarbonate ion concentrations in soil water are calculated using the familiar relationships between the partial pressure of CO_2 and hydrogen ion activity in the water

$$[HCO_3^-] = K_c\{P_{CO_2}/[H^+]\}, \tag{2.44}$$

where the combined constant K_c is known for a given temperature (Stumm and Morgan, 1970). The free hydrogen ion produced by this mechanism reacts with an aluminium mineral in the soil. The model assumes the following equilibrium relationship between Al and H^+.

$$3H^+ + Al(OH)_3\ (s) = Al^{3+} + 3H_2O. \tag{2.45}$$

Generally, the cation exchange sites on the soil matrix have higher affinity for the trivalent aluminium cation than for di- or monovalent base cations. An exchange of cations between the dissolved and adsorbed phase results:

$$Al^{3+} + 3BCX(s) = AlX_3(s) + 3BC^+, \tag{2.46}$$

$$2Al^{3+} + 3BCX(s) = 2AlX_3(s) + 3BC^{2+}, \tag{2.47}$$

where, X is an adsorbed phase, BC^+ and BC^{2+} are base cations.

The net result of these reactions is the production of alkalinity. As CO_2 partial pressure or the availability of base cations on the soil exchange sites increases, the equilibrium reactions proceed further to the right-hand side in each case resulting in higher alkalinity. When the solution is removed from contact with the soil matrix and is exposed to the atmosphere (e.g. entering the stream channel), the CO_2 partial pressure declines and the pH increases. Because the solution is no longer in contact with the soil, cation exchange reactions no longer occur and the alkalinity and base cation concentrations remain unchanged.

The process of acidification is controlled in part by the rate at which the exchangeable base cations in the soil are depleted. This in turn is affected by the rate of re-supply through weathering of base cations from the soil. Leaching of base cations is affected mainly by the concentration of strong acid anions and base cations in the solution moving through the soil. As anions increase in concentration, there must be an equivalent increase in cation concentration to maintain a charge balance.

The model calculates the concentrations of three strong acid anions in both soil and stream water (SO_4^{2-}, Cl^-, NO_3^-). Sulphate has an adsorbed phase in soil and the relationship between adsorbed sulphate and the concentration of dissolved sulphate in soil water is assumed to follow a Langmuir isotherm. If ions derived from atmospheric deposition are accompanied by H^+, as is the case for acid deposition, the excess H^+ will initially displace base cations from the soil exchange sites. As the base saturation declines, aluminum and H^+ become increasingly important in maintaining the charge balance. The water discharged to the stream becomes more acidic as the acidic deposition persists. The model assumes that Al^{3+} and four base cations are involved in cation exchange between soil and soil solution. The parameters describing the cation exchange process in the model are the selectivity coefficients, (one coefficient for each base cation, Ca^{2+}, Mg^{2+}, Na^+, K^+) and the cation exchange capacity (CEC). Further details of the equations and the model structure have been given by Cosby et al. (1985).

2.4 Stochastic Analysis to Predict Acid Shocks

Acid shocks are very important acidification events. The stochastic model is developed to predict shocks. The stochastic model is based on the developments in the theory on extreme values (Bobba et al., 1990; Bobba and Ventresca, 1990) and represents an attempt to develop a more general description and prediction of acid shocks. The model converts pH to H^+ concentration to predict the probability distributions for higher acidity flow events. The stochastic processes $X(t)$ is defined as the maximum term among a random number of variables in an interval of time $[0, t]$. Since the number of peak hydrogen ion concentration discharges in $[0, t]$ exceeding a certain reference level H_0 and the magnitudes of these peaks are random variables, the foregoing model seems to conform well to the acid shock phenomenon in acidified watersheds (Figure 2.17).

In the present analysis of high hydrogen ion concentrations, a reference level H_0 is selected on the basis of toxicity to the aquatic environment, and only those simulated daily flows having hydrogen ion concentration above this level are considered. Within a time period t, for $0 < t < -365$ days, there are $m(t)$ higher hydrogen ion (low pH) concentration flow events, where the beginning of the ith event (at time T_i) is marked by the upward crossing of acidograph above H_0. The duration of the event, D_i, is the time between the upcrossing and the next downcrossing of H_0. Within each higher hydrogen ion concentration flow event, the local maximum above the base level H_0 denotes the peak magnitude of the acidograph. The probability distributions of the random variables, m, H, D and T are considered separately. Two assumptions are considered in this analysis: (a) higher hydrogen ion concentration flow events are independent of each other; (b) the events are generated by common hydrological processes and this avoids the problem of compounding events from different populations.

A partial series is defined as the number of peak hydrogen ion concentration events, H_{p1}, H_{p2}, ..., H_{pm} that exceed a given reference level H_0, and which occur at corresponding times, T_1, T_2, ..., T_m. The duration of each event D_1, D_2, ..., D_m, is the time between the upcrossing and successive downcrossing

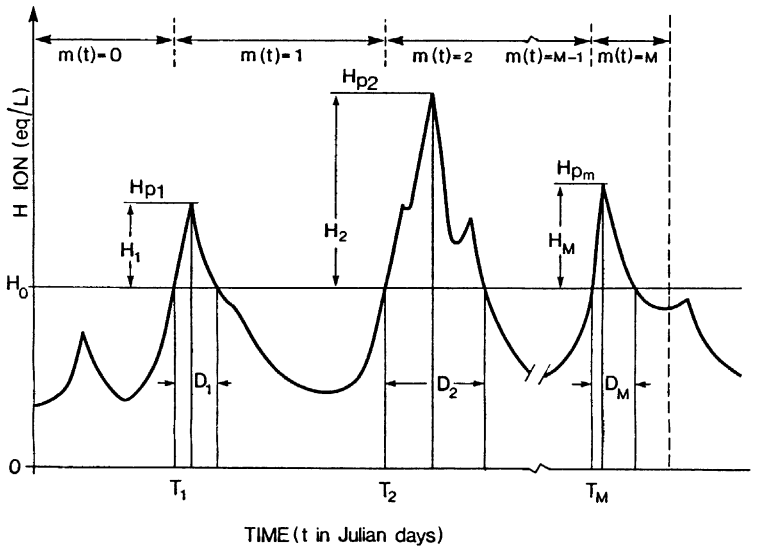


Fig. 2.17. Schematic representation and definition of high $[H^+]$ flow variables.

of H_0. The local maximum of each event above H_0 is denoted by H_1, H_2, ..., H_m. The stochastic prediction models of these variables using various simple probability distributions requires estimates of a total of four parameters. Todorovic and others have provided much of the theoretical groundwork for partial series analysis (Todorovic and Zelenhasic, 1970; North, 1980; Waylen, 1981; Bobba et al., 1990).

2.4.1 Number of Exceedances

Cramer and Leadbetter (1967) showed that the number of crossings of a base level (H_0) approaches a Poisson process as H_0 rises. The density function of a Poisson distribution is:

$$P(M) = [\exp(-A) \cdot A^M]/M! \text{ for } M = 0, 1, 2, \ldots , \tag{2.48}$$

where M is the number of crossings in a year and A is a parameter greater than zero. By applying the method of moments, A is estimated as the mean number of crossings per year:

$$A = \frac{\text{TNC}}{\text{N}} , \tag{2.49}$$

where TNC is the Total Number of Crossings (TNC) in N years of historical record.

2.4.2 Time of Exceedance

High H^+ concentration peaks are generated by precipitation and snow melt and tend to cluster at certain time periods in each year. It is expected therefore that for

$0 < t \leq 365$ days, the average number of higher flows, $\lambda(t)$, will vary within a year. Equation (2.48) may be generalized by a nonhomogeneous Poisson process (Ross, 1976).

$$P[m(t)] = \{\exp[-\lambda(t)] \cdot \lambda(t)^{m(t)}\}/m(t)! \text{ for } m(t) = 0, 1, 2, \ldots, \quad (2.50)$$

where $P[m(t)]$ is the probability of observing m events up to time, t. $\lambda(t)$ is the expected number of crossings up to t, estimated as:

$$\lambda(t) = \frac{\text{TNE}(t)}{N}, \quad (2.51)$$

where TNE(t) is the Total Number of Events (TNE) from time 0 up to time t, observed in N years of historic record.

The empirical expression describing the temporal variation of $\lambda(t)$ is difficult. An alternative approach is to examine the probability distribution $G(T)$ of the time at which higher flow events occur each year, where:

$$G(T \leq t) = \frac{\lambda(t)}{A}, \quad (2.52)$$

because from Equations (2.49) and (2.51), the empirical distribution of $G(T)$ can be obtained as $G(T \leq t) = \text{TNE}(t)/N$. This is equivalent to the observed relative cumulative frequency of events up to day t.

In this case, the derivation of $G(T)$ greatly simplifies the description of the expected number of higher flow events. $\lambda(t)$ can be obtained as:

$$\lambda(t) = G(T \leq t)A. \quad (2.53)$$

In this case, Equation (2.53) rather than Equation (2.51) will be used to estimate $\lambda(t)$.

2.4.3 Time of Events

In the case of higher flow events which are generated by a single dominant process, the probability distribution $G(T)$ can be approximated by a Gaussian distribution, given by:

$$G(T) = \frac{1}{x\sqrt{2\pi}} \int_{-\infty}^{\text{TNC}} \exp[-(s-u)^2/2x^2]\, ds, \quad (2.54)$$

where u is the mean of date exceedance, and x^2 is the variance of date exceedance.

In the investigation of spring snowmelt events, the lower bound of the integral is taken to be 0, without loss of information. Generally, snow melting processes occur at the same time of each year but with year to year variations due to changing hydrometeorlogical conditions. Statistically, this translates to a fixed variance about a long term mean date of hydrogen ion concentration.

2.4.4 Duration of Exceedances

The Duration (D) of Poisson distributed exceedance constitutes the length of time between up and successive down crossings and therefore should follow an exponential distribution (F) as given by Cramer and Leadbetter (1967):

$$F(D \geq d) = \exp(-d/O), \tag{2.55}$$

where O is a parameter greater than zero. Using the method of moments, O can be estimated by the mean duration of events.

2.4.5 Magnitude of Exceedances

The probability of a given exceedance magnitude (H_i) for a fixed truncation level (H_0) is given by the exponential distribution (M), it is a special case of the gamma distribution (Zelenhasic, 1970; Taesombat and Yevjevich, 1978):

$$M(H_i \geq H_0) = \exp(-H_i/h), \tag{2.56}$$

where h is a parameter greater than zero. Using the method of moments, h can be calculated by the mean of peak hydrogen ion concentration flow.

$$h = \sum_{i=1}^{\text{TNC}} (H_{pi} - Q_0)/\text{TNC}. \tag{2.57}$$

Proper selection of H_0 is key. The critical level of H_0 should be set at a sufficiently high value to meet the assumption that the individual high H_0 events are independent (Cramer and Leadbetter, 1967).

2.5 Application of Models to the Turkey Lake

The Turkey Lakes Watershed Study (Figure 2.18) was initiated in 1980 to define the impact of atmospheric deposition of acidifying substances and other pollutants on an undeveloped aquatic and terrestrial terrain in central Ontario, Canada. The Turkey Lakes Watershed (TLW) contains a chain of 4 lakes (Figure 2.18) that exhibits a gradient in chemical composition within a single defined system. The watershed is undisturbed, has high relief, and climatically, has a fairly wet environment. Elevation at the lowermost stream gauging station is 340 m AMSL while the highest point is atop Batchwana Mountain (630 m) on the northern edge of the basin-giving an overall relief of 290 m. An average annual precipitation at low elevations was 1212 mm from 1981–1984 and may be up to 15% more at higher elevations. In comparison, Sault Ste. Marie, Ontario, which lies south of the highlands containing the watershed, receives only 935 mm yr^{-1}.

The watershed is almost entirely underlain by Precambrian silicate greenstone with only small outcrops of more felsic igneous rock occurring north of Batchwana Lake and near the main inflow to Little Turkey Lake (Figure 2.18). Regional fault systems trending in approximately NW-SE and SW-NE directions have exerted

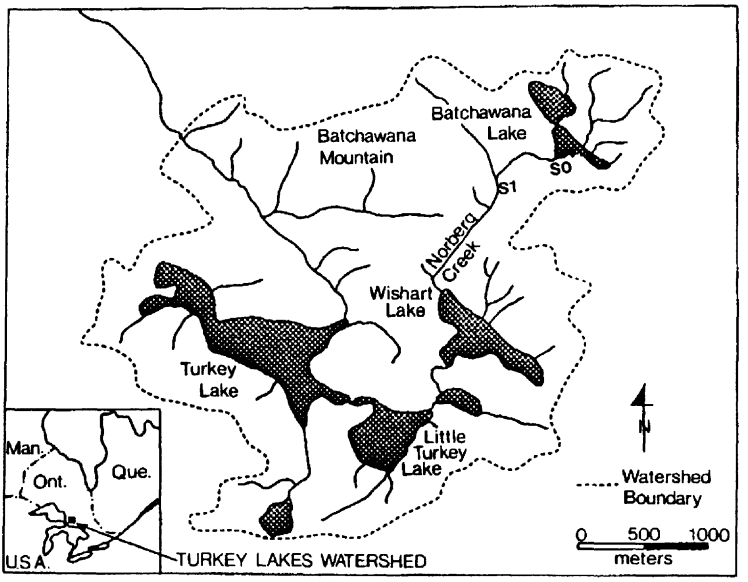


Fig. 2.18. Location map of the Turkey Lakes Watershed showing catchment boundary.

some control over water drainage giving the rather angular drainage pattern observed. A two component glacial till overlies the bedrock. Till thickness varies from < 1 m at high elevation locations to 1–2 m at lower elevations, and with the occasional occurrence of extremely deep till sequences (up to 70 m) when valleys in the bedrock have been entirely filled. The mineralogy of the till is more felsic than the underlying bedrock showing that the material was likely primarily derived from the large granitic intrusions that occur just north of the basin. Finally, the tills contain a small but measurable amount of $CaCO_3$ (0–2%) that increase with depth, and is higher (on average) at lower elevation locations. The most frequently encountered soil types on upland sites within the basin are Ferro-Humic and Humic-Ferric Podzols.

The headwater lake in the TLW is Batchawana Lake which is separated into two distinct basins (North and South). The outflow stream draining Batchawana Lake South is called Norberg Creek and traverses a rapid change in elevation (497 to 388 m, AMSL) prior to Wishart Lake — the shallowest lake in the basin. Little Turkey Lake and finally Turkey Lake follow Wishart Lake. The outflow from the watershed enters the Batchawana River below the perimeter of the basin, and thence passes on to Lake Superior. The relatively high precipitation causes the lakes to flush fairly rapidly.

The ILWAS, RAINS, ETD, and TMWAM models were applied to the Batchawana sub-basin of the TLW in order to evaluate and compare the accuracy of the models in predicting several common hydrological and chemical parameters. The Batchawana sub-basin is the headwater sub-basin of the TLW and is approximately 2.06 km^2 in surface area.

The ILWAS, ETD and TMWAM models operate on a daily time step but the RAINS model operates on a monthly time step. Since not all of the models are

capable of simulating all of the major cations and anions, only simulation results for pH, alkalinity, and SO_4 (which are common to all the models) are compared. All of the models were applied for the period January 1, 1981 to December 31, 1984, for data measured at station S1 (Figure 2.18). The models were initially calibrated for the period January 1, 1981–December 31, 1982 and then run for the final 3 years without further calibration. All models were run using common physical and chemical input constants, parameter values, and meteorological input data. Every effort was made to calibrate the different models to the best degree of accuracy, and where available, the model developer's procedure for model calibration was followed. The models were all calibrated by trial and error rather than by an automatic numerical optimization procedure. The model parameters were adjusted to minimize the errors between the observed and model calculated outputs, based on the principle of least squares. Additional statistics are used to evaluate model performance.

2.5.1 Statistical Analysis

Several statistical tests were used to assess the performance of the model. These tests compare the computed and observed flows and ion concentrations, and include: (1) Root Mean Square Error (RMSE), (2) Coefficient of Determination (CD), (3) Coefficient of Efficiency (CE) and (4) Coefficient of Residual Mass (CRM). These methods were developed by Aitken (1973) and have been successfully applied in hydrological modelling (Aitken, 1973; Bobba and Lam, 1990; and Loague et al., 1991).

Root Mean Square Error (RMSE): The root mean square error of the data is given by

$$\text{RMSE} = \left[\sum_{i=1}^{n} \frac{(C_i - O_i)^2}{n}\right]^{0.5} \frac{100}{O_m}, \tag{2.58}$$

where C_i are the computed values; O_i are the observed values, n is the number of sample data, and O_m is the mean of the observed data.

Coefficient of Determination (CD): The coefficient of determination is, in effect, the square of the correlation coefficient. It was defined by Aitken (1973) in the following manner, so as to emphasise the relationship between the coefficient of determination and the coefficient of efficiency:

$$\text{CD} = \frac{\sum_{i=1}^{n}(C_i - O_m)^2}{\sum_{i=1}^{n}(C_i - O_i)^2}. \tag{2.59}$$

The coefficient of determination is always less than one where a large value indicates good correlation and a low value indicates poor correlation. The coefficient measures the degree of association between estimated and observed results or the amount of scatter about the line of best fit but it does not indicate the existence of any systematic errors.

Coefficient of Efficiency (CE): The coefficient of efficiency was defined by Aitken (1973) is an analogous manner to the coefficient of determination:

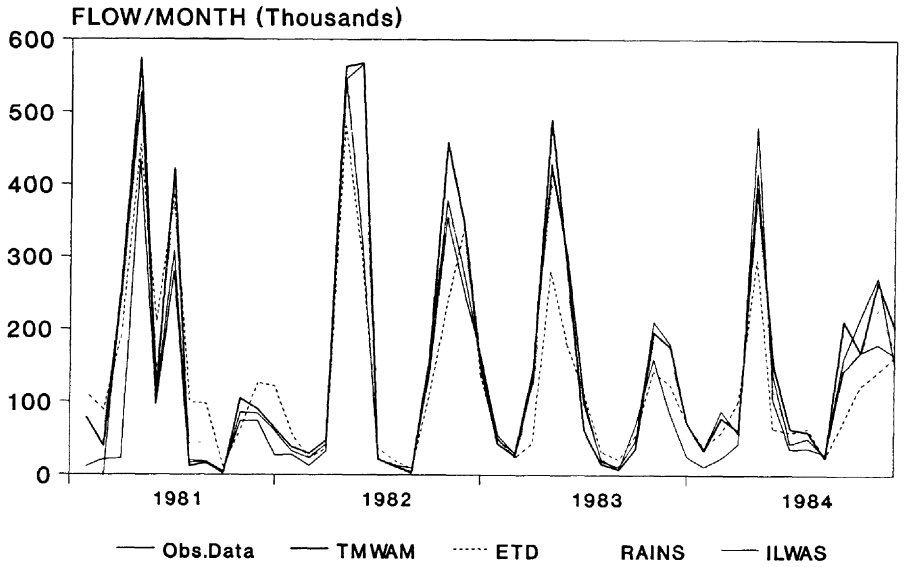


Fig. 2.19. Comparison of model simulated versus observed monthly runoff data for Batchawana Lake Watershed (after Booty et al., 1992).

$$\mathrm{CE} = \frac{[\sum_{i=1}^{n}(O_i - O_m)^2 - \sum_{i=1}^{n}(C_i - O_m)^2]}{\sum_{i=1}^{n}(O_i - O_m)^2} . \tag{2.60}$$

Again the value of CE is always less than one. If the results are highly correlated but biased (i.e., they do not plot randomly about the 45° line on a graph paper of the observed versus estimated results) then the value of CE will be less the value of CD.

Coefficient of Residual Mass (CRM):

$$\mathrm{CRM} = \frac{[\sum_{i=1}^{n} O_i - \sum_{i=1}^{n} C_i]}{\sum_{i=1}^{n} O_i} . \tag{2.61}$$

Linear and rank correlation coefficients and regression slopes based on a statistical comparison of computed and observed values are also used to indicate model performance. In the rank correlation, the correlation coefficient (r) and the regression slope (s) are derived from the ordered ranks of both the measured and computed data and are good indicators of episodic coincidences. A correlation coefficient and slope close to 1.0 are indicative of good model performance (Bobba and Lam, 1988). The coefficient of efficiency (CE) (Bobba et al., 1988) is a measure of the improvement on the variance not explainable by the mean value, but explainable by the particular model which produces the computed values (q_c's). The coefficient quantifies the relative improvement of the predictions over a benchmark model (i.e., the mean observed value). Any positive value of E can be interpreted as improvement over the mean-as-model, the closer to +1 the better.

The model simulated versus observed monthly watershed outflow data are presented in Figure 2.19 and the associated model performance statistical data are

presented in Table 2.4. The TMWAM and RAINS models show very similar simulated flows and have almost identical statistical performance values. The ILWAS model, which has the most sophisticated hydrologic submodel, performed equally well as the TMWAM and RAINS models in predicting monthly outflows from the watershed, except with respect to its coefficient of efficiency. The hydrologic submodel used in the ETD model is intermediate in complexity between the ILWAS and TMWAM and RAINS models. In this particular application, ETD shows the largest deviation from the measured values for both the monthly and cumulative flows, as reflected by all four model performance statistics (Table 2.4). With respect to the linear correlation coefficient, the RAINS and TMWAM models performed the best with values of $r = 0.987$ and 0.986, respectively, followed closely by the ILWAS model with $r = 0.967$, and the ETD model was the least precise with $r = 0.896$. The rank correlation coefficient indicates how well the models simulate the timing and magnitude of the peaks of flows as compared to the linear correlation coefficient. In this case, the TMWAM, ILWAS, and RAINS models perform equally well, with $R = 0.988$, 0.985, and 0.984, respectively. Considering the coefficient of efficiency, the RAINS and TMWAM models perform the best relative to the mean-as-model, with $E = 0.970$ and 0.961 respectively, followed closely by the ILWAS model with $R = 0.915$, and finally the ETD model with $R = 0.762$.

The models show a significant range in their abilities to simulate the monthly and seasonal changes in water chemistry. The predicted versus observed stream pH, alkalinity, and sulphate values are presented in Figures 2.20, 2.21 and 2.22, respectively, and the associated statistical data on model performance are shown in Table 2.4. Visually, it is apparent from Figures 2.21 and 2.22 that the RAINS model tends to overpredict the fluctuations of observed pH and alkalinity. The ETD model tends to underpredict the fluctuations of observed pH but does the best job of simulating the observed alkalinities. The TMWAM model appears to do the best job of simulating the observed pH data and is second to the ETD model in simulating the measured alkalinity data.

The TMWAM predictions for SO_4 have the highest correlation with the observed data ($r = 0.682$, $R = 0.680$, $s = 0.819$). It is also the only model that has a positive coefficient of efficiency ($E = 0.423$). In this case, the other three model's predictions of sulphate concentration cannot be considered as improvements over the mean-as-model. TMWAM predictions most closely simulate both the magnitude and the timing of the observed SO_4 stream concentration peaks, as confirmed by the rank correlation coefficient of 0.680 (Table 2.4). The ILWAS model also performs well in predicting the timing of the peaks in stream SO_4 concentration, with $R = 0.610$; however, the magnitudes of the peaks, especially in the spring of 1983 and 1984, are significantly different than the observed values. As was the case for pH, the ETD model predictions for SO_4 show the smallest fluctuations of magnitude over the period of study. However, the timing of the peaks do not coincide very well with those of the observed data, as seen by the low rank correlation coefficient ($R = 0.233$). The RAINS model underpredicts the stream sulphate concentrations over the entire period of study. It also fails to predict the timing of the peaks, as verified by the very low rank correlation coefficient shown in Table 2.4 ($R = 0.153$).

TABLE 2.4
Statistical evluation of model performance for Turkey lake watershed (from Booty et al. 1992)

Model	Linear correlation coefficient (r)	Rank correlation coefficient (R)	Slope (s)	Coefficient of Efficiency (CE)
		Runoff		
ILWAS	0.967	0.985	1.100	0.915
TMWAM	0.986	0.988	0.912	0.961
ETD	0.896	0.892	1.181	0.762
RAINS	0.987	0.984	1.065	0.970
		pH		
ILWAS	0.293	0.613	−6.772	−2.965
TMWAM	0.438	0.713	0.648	0.264
ETD	0.549	0.477	1.616	0.240
RAINS	0.549	0.477	1.616	0.24
		Alk		
ILWAS	0.442	0.511	0.441	0.275
TMWAM	0.598	0.533	0.851	0.340
ETD	0.781	0.687	0.959	0.608
RAINS	0.356	0.407	0.281	−1.048
		SO_4		
ILWAS	0.488	0.610	0.334	−4.314
TMWAM	0.682	0.680	0.819	0.423
ETD	0.274	0.233	0.444	−0.904
RAINS	0.126	0.153	−0.172	−6.479

The MAGIC model has been applied to the headwater sub-basin of the Turkey Lakes Watershed in order to evaluate the long term trends due to constant and varying acid loadings. The constant load is fixed at the average level (i.e. 58 ueq/L) as observed during 1982–1986. The varying load follows the reduction plans for Canada (scenario 2) scheduled for completion by 1994, and for the U.S. (scenarios 3 and 4) by 1995 and 2000, respectively. Figures 2.23 and 2.24 show the sub-basin outlet stream water chemistry results due to different sulphate loads in the watershed. The constant sulphate deposition results in a depression in pH of approximately 1.2 units (Figure 2.23) and a depression in alkalinity of 185 ueq/L by the year 2124 (Figure 2.24). When the sulphate input loading is reduced according to the reduction scenarios (input sulphate concentration varying from 58

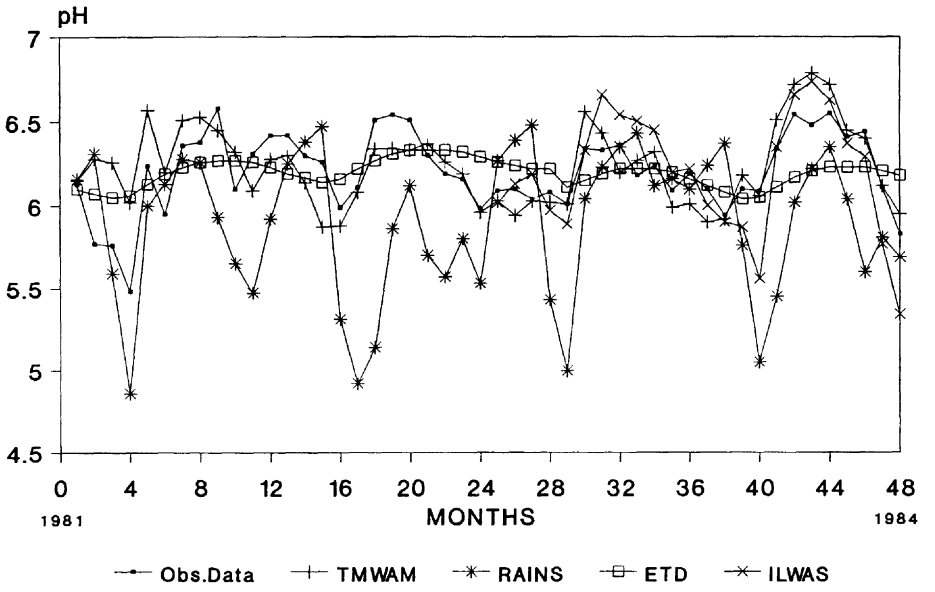


Fig. 2.20. Comparison of model simulated versus observed monthly averaged outflow pH values for Batchwana Lake Watershed at station S1 (from Booty et al., 1992).

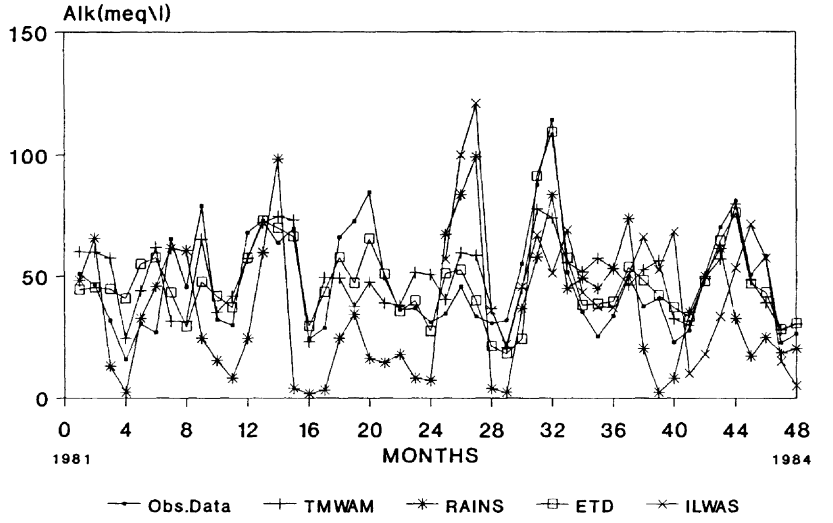


Fig. 2.21. Comparison of model simulated versus observed monthly averaged outflow alkalinities concentrations for Batchawana Lake Watershed at station S1 (from Booty et al. 1992).

to 33 ueq/L), the model simulates an almost immediate response (Figure 2.24). By the end of the simulation (i.e. 2124) the pH has only been depressed by 0.4 pH units and the alkalinity by 97 ueq/L, showing the effects of the reduction. In both cases, it requires approximately 100 years to reach equilibrium.

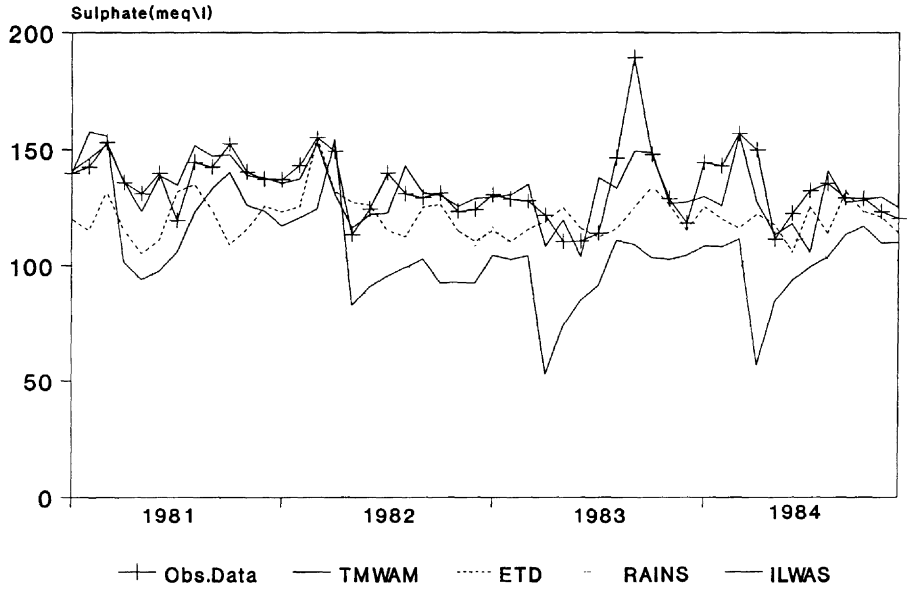


Fig. 2.22. Comparison of model simulated versus observed monthly averaged outflowsulphate concentrations for Batchawana Lake Watershed at station S1 (from Booty et al., 1992).

Not all of the parameters used in the MAGIC model have been directly measured in the field, and some of the parameters may exhibit a significant degree of heterogeneity within a watershed. For example, the MAGIC parameter EMX (maximum sulphate adsorption capacity) for the soils in the Turkey Lakes watershed has not been thoroughly measured. Foster (1985) examined 13 B horizons in five soil profiles within watershed 31 of the Turkey Lakes Watershed and found that the greatest adsorption was 40 ppm SO_4 , and three samples adsorbed no SO_4. In Figure 2.25 the results of a sensitivity analysis where the value of EMX was varied from 25 to 100 ueq/g (for a constant load of 58 ueq/L) are shown for the outlet stream pH for Batchawana Lake sub-basin. In Figure 2.26 the results are shown for the same sub-basin for the reduced input loading scenarios mentioned above. In both figures it can be seen that the predicted stream pH is sensitive to the value of EMX chosen for the watershed soil, for a significant period of time. For both constant and variable input loadings, by the end of the simulations, the differences in predicted pH of the outlet stream due to the selection of different values of EMX are becoming insignificant. However, from these sensitivity analyses it can be seen that a significant amount of uncertainty in MAGIC model predictions is possible, and this must be taken into consideration when the model is applied to large areas over which significant variations in model parameters, such as soil chemical properties, exist.

2.5.2 Acidification Events

All the models simulate significant pH depressions (0.5 units) during the snowmelt and storm events. The occurrence and duration of these episodes are related to

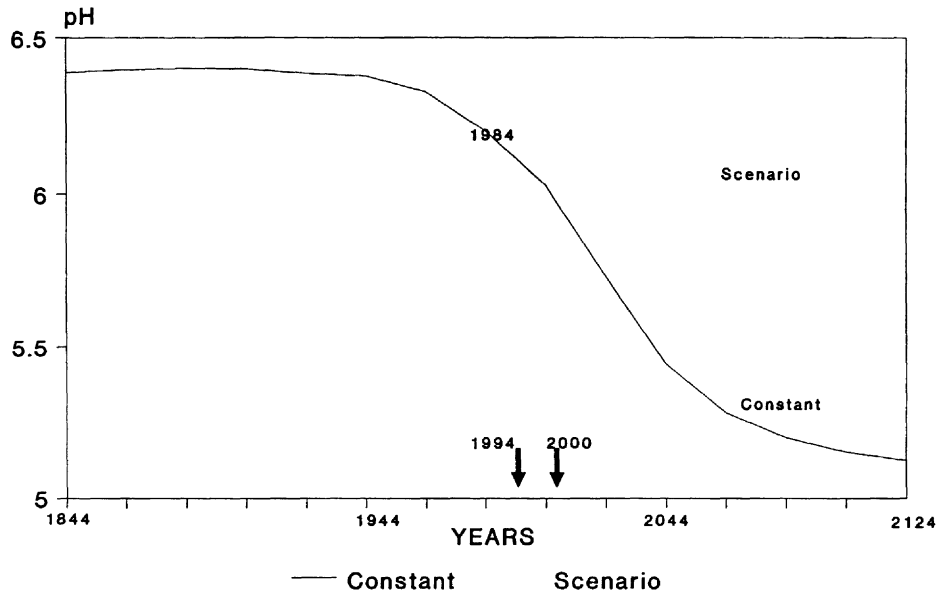


Fig. 2.23. MAGIC model predictions of Batchawana Lakes sub-basin outlet chemistry pH for different sulphate loadings.

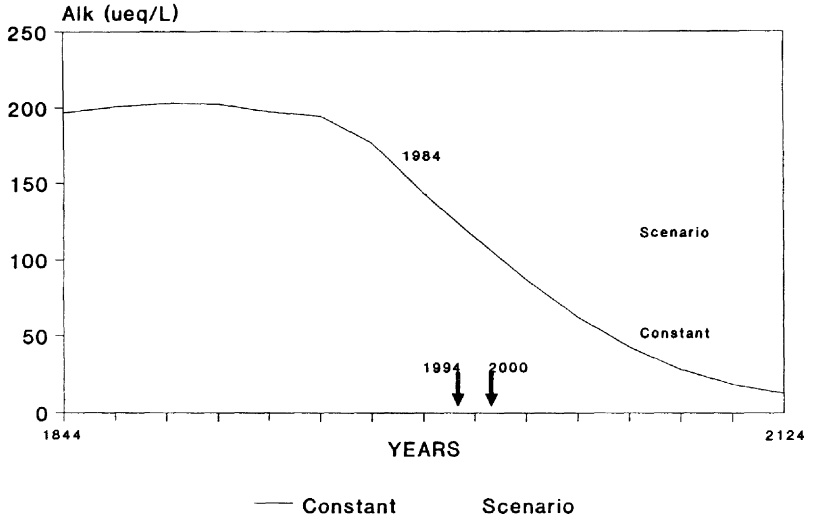


Fig. 2.24. MAGIC model predictions of Batchawana Lakes sub-basin outlet chemistry alkalinity for different sulphate loadings.

weather events. Statistically, the magnitude, duration, frequency and timing of high hydrogen ion concentration can be related to flow events using the method by Bobba et al. (1990).

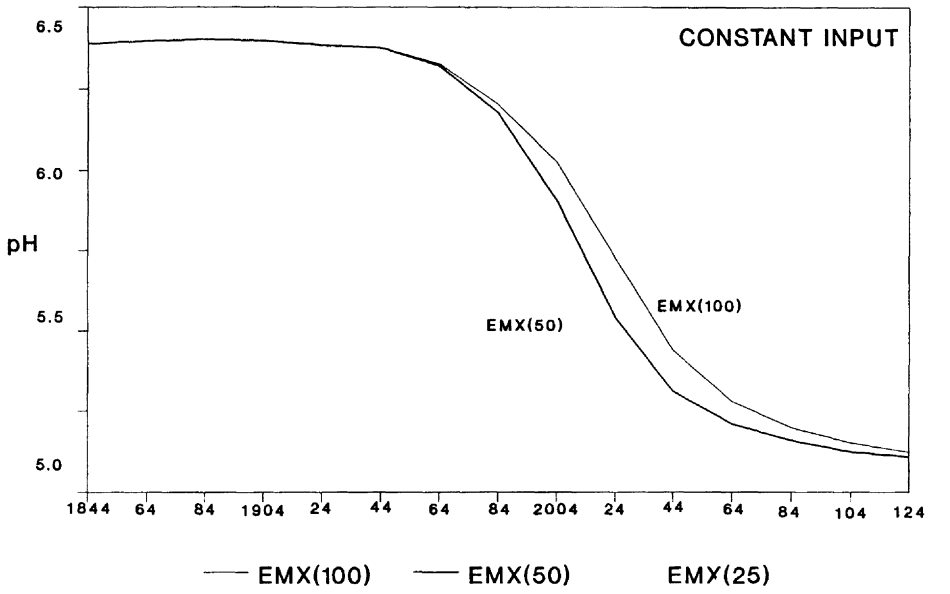


Fig. 2.25. Sensitivity of MAGIC model predicted stream pH (Batchawana Lakes sub-basin outlet) to variations in model parameter EMX (maximum sulphate adsorption capacity) for constant sulphate loadings.

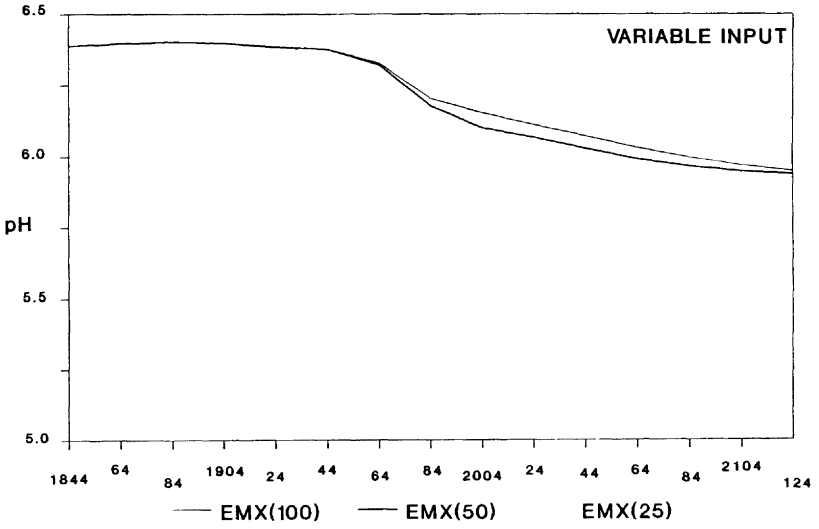


Fig. 2.26. Sensitivity of MAGIC model predicted stream pH (Batchawana Lakes sub-basin outlet) to variations in model parameter EMX (maximum sulphate adsorption capacity) for variable sulphate loadings.

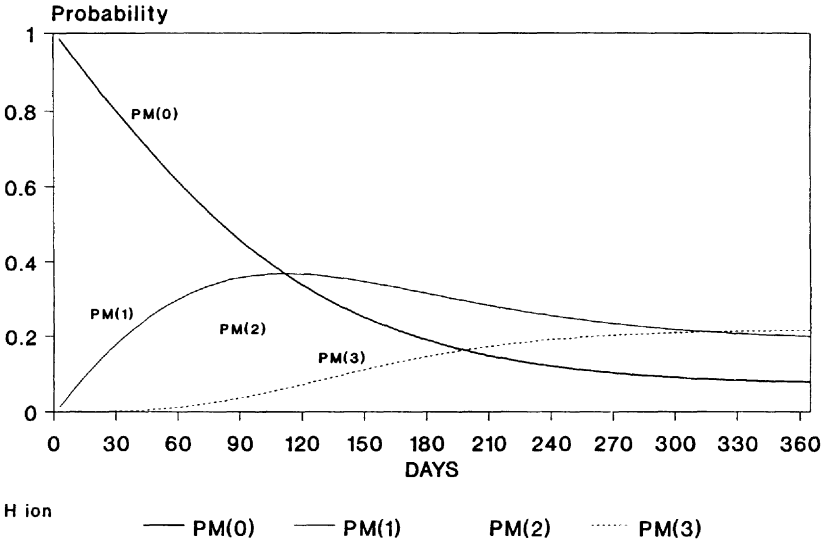


Fig. 2.27. The probabilities of the number and timing of acidification events above the selected truncation level (pH = 6.0, $[H^+] = 1.0 \times 10^{-6}$) for the Turkey Lakes Watershed (from Bobba et al., 1990).

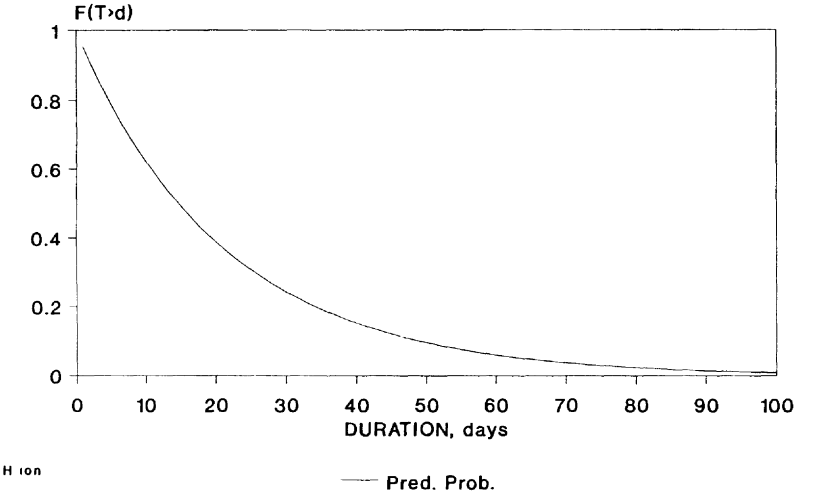


Fig. 2.28. The probabilities of the duration of high hydrogen ion concentrations above the selected truncation level ($\langle pH=6.0 \rangle$, $[H^+] = 1.0 \times 10^{-6}$) for the Turkey Lakes Watershed (from Bobba et al., 1990).

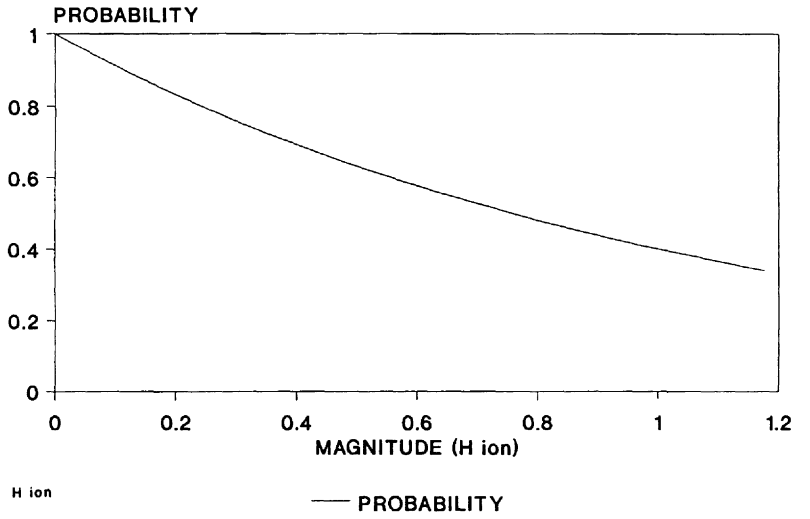


Fig. 2.29. The probabilities of the magnitude of high hydrogen ion concentration above the selected truncation level (pH = 6.0, $[H^+] = 1.0 \times 10^{-6}$) for the Turkey Lakes Watershed (from Bobba et al., 1990).

Figure 2.27 shows the probability of acid shock events (measured as H ion concentration, for pH < 6) in the Batchawana Lake outflow. For the period of observation, there was a maximum of 3 acidification events per year, as indicated by a maximum PE (Probability of Events) of 3 on Figure 2.27. The probability of zero acidification events per year starts at 100% and rapidly decreases to approximately 20% over the first 180 days. The probability of 2 events can be seen to reach a maximum of approximately 30% after 225 days, after which the probability decreases only slightly.

In Figure 2.28 the probability of event duration is presented. It can be seen, for example, that 50% of such events last less than 15 days in the Turkey Lakes Watershed. Figure 2.29 shows the probability of the magnitude being greater than the selected H^+ concentration (1×10^{-6} eq/L) for the Turkey Lakes Watershed. For example, there is a 50% probability of the H^+ concentration of the event being 1.75 times greater than 1 eq/L. These episodic statistics can vary from location to location, but two case studies (Bobba et al., 1990) show a strong relationship between pH depression and flow events that could be important when considering biological damages.

2.6 Discussion

Watershed acidification models that assess long term responses are very difficult to verify. Validation of these types of models requires long time series records to determine whether the model estimates match the observed watershed responses. All but one of the four models which have been applied to the Turkey Lakes Watershed headwater sub-basin have been found to reasonably predict monthly or seasonal fluctuations in the water acidity.

To date, the acidification models have been applied only to single watersheds or watershed chains in Canada. In many cases, the chosen catchments have been highly sensitive to acidification and strongly acidified. A regionalized model structure is needed to allow examination of chemical responses on an extensive geographical scale.

As it has been shown in numerous studies (Johnson et al., 1969; Jeffries and Semkin, 1982; Bobba and Lam, 1990; Rustad et al., 1986), watershed stream chemistry is strongly correlated with stream hydrology. It might be assumed that if the hydrological submodel of the ETD model was improved so that its performance was more similar to those shown in Table 2.4 for the other three models, its ability to predict stream chemistry parameters such as pH, alkalinity, and sulphate may also be improved.

The results of these comparisons of model performance for Batchawana Lake watershed outflow are presented only as an example of how well the models perform for this one site, at one point, and based on monthly averaged data. If we were to apply the models to another period of observations, and /or to different watershed parameters, the relative performance of the models may be expected to be different than those presented in this study. However, the main point is to show that basic watershed parameters such as basin outflow, outflow pH, alkalinity, and sulphate are simulated to different degrees of accuracy by the four different models considered in this study. A great deal more effort is still required in order to fully test and validate the models, and more importantly, to help to understand which key processes or differences in model structure are responsible for the different levels of performance shown by the models. Many models are calibrated and verified over similar short periods of observation and then are used to make future predictions for longer periods of time (i.e. 10, 20, 50 years). It is clearly advantageous to evaluate the relative uncertainties in model predictions using multiple model predictions for measured data before making long term forecasts of model responses to changes in such inputs as acid loadings.

For long term prediction of acidity the time scale used is typically one year. This means that inputs are expressed as yearly averages. With this requirements there is apparently not much data available for parameter fitting. Parameters must therefore be estimated independently which requires much field work. The prediction results do not look as glorious as those from the models with fitted parameters but prediction can certainly be extended in space and can be supplemented by simulation (e.g. effects of liming, of clear cutting of drainage and so on). It is often said that mathematical models aid understanding of processes in a watershed. The model is then equivalent to a working hypothesis. Fitted parameters can then be interpreted physically so as to support or disqualify the hypothesis. There is, however, one difficulty here. Fitted parameters are always correlated and often fitting can be considered as good within a fairly wide range of combinations of parameters. Orthogonal sets of parameters can be fitted but the physical interpretations of these is very obscure indeed. Independently determined parameters are, of course, a better check on model structures. However, the pathways of water in a basin are well known in principal and since they are fundamental there is not much choice of relevant model structure for simulating chemical constituents.

A final word about distributed models. The fundamental unit here is a groundwater flow that is based on the hydraulic properties of the ground and the topography of the groundwater surface. The water balance of the unsaturated zone and the divergence of the groundwater flow determine changes in groundwater levels. and, consequently, also the geographic delineation of groundwater recharge and discharge areas. All parameters can be obtained independently. Flow paths will appear as results of the modelling and, hence, all the prerequisites for reducing the fate of chemical substances provided the mineralogy and chemistry of the soil and rocks are known. For research purposes such a model seems to be the only rational alternative.

2.7 Conclusions

The TMWAM, ETD, and ILWAS models are all lumped parameter, deterministic models that were designed to operate on a daily time step or less in order to be able to examine watershed processes in real time. The RAINS and MAGIC models were developed to examine catchment responses to acidic precipitation over long intervals (monthly and annual, respectively). A summary of the watershed acidification models is given in Table 2.5.

TMWAM consists of four major submodels that deal with hydrology, sulphate, cations and DOC. It is a modified version of the Birkenes model, using one more (3) soil compartment and a separate sub-model to account for A^- (Lam et al., 1989). The ETD model determines the concentrations of Cl^-, SO_4^{2-}, alkalinity and pH in an upper soil zone, an unsaturated zone, bedrock aquifer, and lake compartments. The ILWAS model is the most complex of all the watershed acidification models. The hydrologic module routes water from tree top to lake outlet. Both equilibrium and rate limited expressions are used to describe mass balances for 18 chemical constituents.

The application of the ILWAS model to the Batchawana Lake subbasin of the Turkey Lakes Watershed required a substantial amount of effort, especially in the calibration phase. The model was found to be capable of simulating the hydrologic flows with reasonable accuracy. However, this was only possible by using values for hydrologic parameters such as hydraulic conductivity which were substantially different from the field/laboratory measured values. The model was found to show a bias towards the generation of shallow interflows over deeper infiltration and the generation of groundwater flows. The simulated snowpack volumes were found to match the observed values within the range of error of measurement.

The model could not be calibrated to accurately simulate lake temperature profiles and ice formation on the lake. Lake temperatures were only reasonably well simulated during the fall and early spring when the lake was well mixed. During the soil chemistry calibration process, it was discovered that the model predicts an expected increase in soil % base saturation but an unexpected decrease in soil solution alkalinity and Ca^{2+} concentration when the % of $CaCO_3$ in the lower soil horizon is increased. The model was only able to predict the lake outlet alkalinity with reasonable accuracy. Model predicted NO_3^- concentrations are consistently 3–5 times greater than the measured values.

TABLE 2.5
Comparison of different types of watershed acidification models

Model	Time step	Hydrology	Inorganic S	Organic S	Reference
ILWAS	Simulation Daily time step	Linked to hydrology model	Sulphate adsorption fitted to linear isotherm; weathering based on stoichiometric release from minerals	Only mineralization from organic matter and plant uptake based on decomposition rates and plant uptake nutrient needs	Chen et al. 1983
Birkenes	Simulation Daily time step	Linked to hydrology model (Two layer model)	Soluble sulphate, combined effect of both organic and inorganic S in solid phase; linear adsorption assumed	Same as inorganic S	Wright 1981; Christophersen et al., 1982
TMWAM	Simulation Daily time step	Linked to hydrology model (Three layer model)	Same as above but coupled with organic acids	Same as inorganic S	Lam et al., 1988; Bobba and Lam, 1990
Trickle-Down	Simulation Daily time step	Linked to hydrology model	Sulafte adsorption treated as linear isotherm	Not considered	Schnoor et al., 1984
RAINS	Simulation Monthly time step	Linked to hydrology model	Not considered	Cosby et al., 1984	
MAGIC	Simulation Monthly and Yearly time step	Linked to hydrology model	Chemical equilibria reactions used to describe relationships among dissolved and adsorbed ions	Not considered	Cosby et al., 1985

RAINS (Regional Acidification Information and Simulation model System) computes the lake acidity distributions for the lakes within a given region as a function of watershed characteristics and local deposition.

MAGIC (Model of Acidification of Groundwater In Catchments) is a relatively simple model which assumes that all water reaching the streams contacts and has its chemical quality determined by a single layer within the soil column. Due to its small input data requirements, it is the easiest model to apply to new catchments.

The stochastic model has been successfully applied to predict acid shock event probability distribution of frequency, duration and magnitude above toxicity level for aquatic environment.

Acknowledgments

The authors are grateful to David C.L. Lam for his major contribution in watershed acidification modelling. The authors are also very grateful to H.M. Seip, N. Christophersen, S.A. Gherini, R.K. Munson, M.M. Lang, J.L. Schnoor, N.P. Nickolaidas and J. Kamari for providing the various model codes and for useful discussions. The encouragement and support from J.H. Carey and W. Strachan are also appreciated.We are greatful to R. Semkin, R.A. Bourbonniere, and M. E. Thompson for their interaction at various stages of the watershed acidification modelling development. The authors acknowledge the publishers to permit to publish the figures in this chapter. The senior author (A. Ghosh Bobba) dedicated this chapter to Lord Venkateswara, Tirupati, India.

References

Anderson, M.G.and Burt, T.P., 1978. The role of topography in controlling throughflow generation. Earth Surf. Processes, 3: 331-344.

Anderson, M.G., and P. Kneale 1980. The influence of low-angled topography on hillslope soil-water convergence and stream discharge. J. Hydrol. 57:65-80.

Arp, P.A., and S. Ramnarine 1983. Verifying the effects of acid precipitation on soil leachates - a comparison between published records and model predictions. Ecol. Modelling, 19, 119-138.

Ashkar, F. and Rousselle, J. 1981. Design discharge as a random variable: A risk study. Water Resources Research. 17:577-591.

Atkinson, T.C., 1978. Techniques for measuring subsurface flow on hillslopes. In: M.J. Kirkby (Editor), Hillslope Hydrology, Wiley, New York, N.Y., pp. 73-117.

Betson, R.P., 1964. What is watershed runoff? J. Geophys. Res., 69: 1541-1551.

Bobba, A.G. and D.C.L. Lam 1984. Application of linearly distributed surface runoff model for watershed acidification problems. In Proc. Canadian Hydrology Symp., U. of Laval, Quebec, No. 15-1984, Vol. 1, 381-398.

Bobba, A.G. 1988. Acid deposition- Soil acidification. Literature review. NWRI Contribution # 88-109. National Water Research Institute, Burlington, Ontario, Canada L7R 4A6

Bobba, A.G. and D.C.L. Lam 1988. Application of a hydrological model to the Turkey Lakes watershed. Can. J. Fish. Aquat. Sci. Vol. 45, 81-87.

Bobba, A.G. and D.C.L. Lam 1989. Application of a hydrological model to acidified watersheds: A study on Mersey River and Moosepit Brook, Nova Scotia. Water, Air and Soil Pollution

Bobba, A.G. and Lam, D.C.L. 1990. Hydrological modelling of acidified Canadian watersheds.Ecological Modelling 50, 5-32.

Bobba, A.G. and Ventresca, U. 1990. User's manual to predict the probability of peak flows from watersheds. NWRI Unpublished Report.

Bobba, A.G., D.C.L. Lam, W.G. Booty, D.S. Jeffries, and R.L. Thomas, 1990. Stochastic analysis of acid shocks generated by mixed hydrological processes. Water, Air and Soil Pollution, 53: 239-250.

Booty, W.G. 1983. Watershed acidification model and the soil acid neutralization capacity concept, Ph.D. dissertation, 194 pp., McMaster Univ., Hamilton Ont. Canada.

Booty, W.G., A.G. Bobba., D.C.L. Lam., and D.S. Jeffries, 1992. Application of four watershed acidification models to Batchawana watershed, Canada. Env. Pollut. 77: 243-252.

Booty, W.G., and J.R. Kramer, 1984. Sensitivity analysis of a watershed acidification model. Philos. Trans. R. Soc. London, B305, 441-449.

Brownlee, K.A., 1965, Statistical theory and Methodology in Science and Engineering, Wiley, New York, N.Y. 590P.

Charette, J.Y., H.G. Jones, W. Sochanska, and J.M. Gauthier, 1984. Changes in lakewater quality during spring runoff in a northern Boreal forest subjected to acidic precipitation. Proceedings of the Canadian Hydrology Symposium, No. 15-1984, pp. 201-220.

Chen, C.W., Gherini, S.A., Hodson, R.J.M., & Dean, J.D., 1983. The ILWAS study: Model principles and applications procedures. EPRI EA-3221, Project 1109-5, Vol. 1, Electric Power Research Inst., Palo Alto, CA.

Christophersen, N., Seip, H.M., Wright, R.F., 1982. A model for streamwater chemistry at Birkenes, Norway. Water Res. Research. 18, 977-996.

Clair, T.A., F. Barlocher, P. Brassard, and J.R. Kramer, 1989. Chemical and microbial diagenesis of humic matter in freshwaters. Water, Air, and Soil Poll.

Clair, T.A., and J.R. Kramer, 1989. Interactions between mineral and organic acids in surface freshwater. Environ. Sci. and Tech. (submitted).

Cosby, B.J., Hornberger, G.M., Galloway, J.N. & Wright, R.F., 1985. Modelling the effects of acid deposition: Assessment of a lumped parameter model of soil water and stream chemistry. Water Res. Research. 21, 51-63.

Cosby, B.J., Hornberger, G.M., Galloway, J.N., and Wright, F.R., 1985. Time scales of catchment acidification: A quantitative model for estimating freshwater acidification. Environ. Sci. Technol. 19, 1144-1149.

Cosby, B.J., G.M. Hornberger, J.N. Galloway, and R.F. Wright, 1985. Modelling the effects of acid deposition: Assessment of a lumped parameter model of soil water and stream chemistry. Water Res. Research. Vol. 21, 51-63.

Cosby, B.J., R.F.Wright, G.M. Hornberger, and J.N. Galloway, 1986. Modelling the effects of acid deposition: Estimating long-term water quality responses in a small forested catchment. Water. Res. Research. Vol. 22, 1591-1601.

Cramer, H. and Leadbetter, M.R., 1967. Stationary and related stochastic processes, Wiley, New York, N.Y. 348p.

Cronan, C.S., 1980. Controls on leaching from coniferous forest floor microcosms. Plant soil 56: 301-322.

Davis, J.E. and Goldstein, R.A., 1988. Simulated response of an acidic Adirondack lake-watershed to various liming mitigation strategies. Water Res. Research. 24, 525-532.

Dunne, T. and Black, R.D., 1970. An experimental investigation of runoff production in permeable soils. Water Resour. Res., 6; 478-490.

Dunne, T. and Black, R.D., 1970. Partial area contributions to storm ruoff in a small New England watershed. Water Resour. Res., 6: 478-490.

Evans, L.D., Hendery, G.R., Stensland, G.J., Johnson, D.W., and Francis, A.J. 1981. Acidic precipitation: Considerations for an air quality standard. Water, Air, Soil Pollution. 16:469-509.

Foster, N.W. , 1985. Acid precipitation and soil solution chemistry within a maple-birch forest in Canada. Forest Ecology and Management, 12, 215-231.

Freedman, B. and T.A. Clair , 1987. Ion mass balances from four acidic brownwater streams in Nova Scotia. Can. J. Fish. Aquat. Sci., 44, 538-548.

Freeze, R.A. , 1972. Role of subsurface flow in generating surface runoff. 1. Base flow contributions to channel flow. Water Resour. Res. 8: 609-623.

Freeze, R.A., 1974 Streamflow generation. Rev. Geophys. Space. Phys. 12:627-647.

Gherini, S.A., L. Mok, R.J.M. Hudson, G.F. Davis, C.W. Chen, and R.A. Goldstein, 1985. The ILWAS model: Formulation and application. Water Air and Soil Pollution, Vol. 26, 425-460.

Glass, N.R., Glass, G.E., and Rennie, P.J. 1980. Effects of acid precipitation in North America. Environ. Intern. 4, 443-452.

Harr, R.D., 1977. Water flux in a soil and subsoil on a steep forested slope. J. Hydrol., 33: 37-58.

Hewlett, J.D., 1961. Soil moisture as a source of baseflow from steep mountain watersheds. U.S. Dep. Agric. For. Serv., 132: 11pp.

Hewlett, J.D. and Hibbert, A.R., 1967. Factors affecting the response of small watersheds to precipitation in humid areas. In: W.E. Sopper and H.W. Lull (Editors), Forest Hydrology. Pergamon, New York, N.Y., pp. 275-290.

Hewlett, J.D., Fortson, J.C. and Cunningham, G.B., 1977. The effect of rainfall intensity on storm flow and peak discharge from forest land. Water Resour. Res., 13: 259-266.

Hibbert, A.R., and C.A.Troendle, 1987. Streamflow generation by variable source area, pp. 111-127. In: W.T. Swank and D.A. Crossley, Jr., eds. Forest Hydrology and Ecology at Coweeta. Springer-Verlag. New York.

Hilmen, B. 1981. Acid precipitation. Environ. Sci. Technol. 15: 1119-1124.

Howell, G.D. 1989. Seasonal patterns of mineral and organic acidification in two streams in southwestern Nova Scotia. Water, Air, and Soil Poll. 14: 165-175

Horton, R.E. 1933. The role of infiltration in the hydrological cycle. Trans. Am. Geophys. Union, 14: 446-460.

Horton, J.H., and R.H. Hawkins, 1965. Flow path of rain from the soil surface to the water table, Soil Science, 100, 377-383.

Huff, D.D., O'Neill, R.V., Emanuuel, W.R., Elwood, J.W. and Newbold, J.D., 1982. Flow variability and hillslope hydrology. Earth Surf. Processes, 7: 91-94.

Hursh, C.R. 1936. Storm-water and adsorption. In: Discussion of list of items with definitions. Report of the committee on Adsorption and Transpiration. Am. Geophys. Union, Trans. 17:301-302.

Hursh, C.R. 1944. Report of the subcommittee on subsurface flow. Am. Geophys. Union, Trans. Part v, 743-746.

Hursh, C.R., and E.F. Brater, 1941. Separating storm hydrographs into surface- and subsurface-flow. Trans. Am. Geophys. Union, 22: 863-871.

Jeffries, D.S., and Semkin, R.G., 1982. Basin description and information pertinent to mass balance studies of the Turkey Lakes watershed. Turkey Lakes Watershed Rep. TLW-82-01, Can. Centre Inland Waters, Burlington, Ont. Canada.

Jeffries, D.S. and Semkin, R.G., 1983. Changes in snowpack, stream, and lake chemistry during snowmelt in the Turkey Lakes Watershed. VDI-Berichte 500:377-386.

Jeffries, D.S., Semkin, R.G., Neureuther, R. & Seymour, M., 1988. Ion mass budgets for lakes in the Turkey Lakes watershed, June 1981-May, 1983. Can. J. Fish. Aquat. Sci., 45, 47-58.

Jeffries, D.S. 1991: Effects of acidic deposition on lake chemistry in southeastern Canada. NWRI contribution 91-69.

Jenne, E.A., Eary, L.E., Vail, L.W., Girvin, D.C., Liebetrau, A.M., Hibler, L.F., Miley, T.B., & Monsour, M.J., 1989. An evaluation and analysis of three dynamic watershed acidification codes (MAGIC, ETD, and ILWAS). U.S. Environmental Protection Agency, Washington, DC. PNL-6687. UC-11.

Johnson, N.M., Likens, G.E., Bormann, F.H., Fisher, D.W. and Pierce, R.S., 1969. A working model for the variation of stream water chemistry at the Hubbard Brook Experimental Forest, New Hampshire. Water Res. Research. 5, 1353-1363.

Johnson, D.W., and D.W. Cole, 1977. Sulfate mobility in an outwash soil in western Washington. Water, Air, and Soil Pollut. 7:489-495.

Johnson, D.W., Cole, D.W., Cessel, S.P., Singer, M.J., and Minden, R.V., 1977. Carbonic acid leaching in a tropical, temperate, subalpine, and northern forest soil. Arctic and Alpine Research 9:329-343.

Kamari, J., 1985. A model for analyzing lake water acidification on a regional scale. Report#CP-85-48. IIASA, 2361 Laxenberg, Austria.

Kerekes, J.J., S.T. Beauchamp, R. Tordon, C. Tremblay and T.L. Pollock , 1986. Organic versus anthropogenic acidity in tributaries of the Kejimkujek Watersheds in western Nova Scotia. Water, Air, and Soil Poll. 31, 165-173.
Kirkby, M.J. and Chorley, R.J., 1967. Throughflow, overland flow and erosion. Bull. Int. Assoc. Sci. Hydrol., 12: 5-21.
Knapp, B.J., 1974. Hillslope throughflow observation and the problem of modeling. In. K.J. Gregory and D.E. Walling (Editors), Fluvial Processes in Instrumented Watersheds. Inst. Br. geogr., London, pp. 23-31.
Lam, D.C.L., A.G. Bobba, D.S. Jeffries, and D. Craig 1988. Modelling stream chemistry for the Turkey Lakes Watershed: Comparison with 1981-84 data. Can. J. Fish. Aquat. Sci. Vol. 45, 72-80.
Lam, D.C.L., Bobba, A.G., Bourbonniere, R.A., Howell, G.D. and Thompson, M.E. 1989. Modelling organic and inorganic acidity in two Nova Scotia rivers. Water, Air, and Soil Pollut., 46, 277-287.
Lawes, J.B., J.H. Gilbert, and R. Warington., 1882. On the amount and composition of rain and drainage water collected at Rothamsted. Williams, Clowes and Sons, Ltd., London.
Lazerte, B.D. and P.J. Dillon, 1984. Relative importance of anthropogenic versus natural sources of acidity in lakes and streams in central Ontario. Can. J. Fish. Aquat. Sci., 41, 1664- 1667.
Lindberg, S.E., Shriner, D.S., and Hoffman, Jr., W.A. 1981. The interaction of wet and dry deposition with forest canopy. In initial draft of the proceedings for the effects of acid precipitation on ecological systems: Great Lakes Region. Institute of Eater Research, Michigan State University, East Lansing, Michigan, U.S.A.
Lowdermilk, W.C., 1934. The role of vegetation in erosion control and water conservation. J. Forestry, 32:534.
McCaig, M., 1982. Contributions to storm quickflow in a small headwater catchment: The role of natural pipes and soil macropores. Earth Surf. Processes, 8:239-252.
McColl, J.G. and D.W. Cole, 1968. A mechanism of cation transport in a forest soil. Northwest Science. 42: 134-140.
Marmorek, D.R., D.P. Bernard, C.H.R. Wedels, G. Sutherland, J.A. Malanchuk, and W.E. Fallon. 1989. A protocol for determining lake acidification pathways, Water, Air, Soil Pollution, 44:235-257.
Moldenhaur, W.C., W.C. Burrows, and D. Swartzendruber. 1960. Influence of rainstorm characteristics on infiltration measurements. Trans. Intern. Cong. Soil Sci., 7th, Madison, Wi., pp. 426-432.
Musgrave, G.W. and H.N. Holton. 1964. Chapter 12. In: V. Chow, ed. Handbook of Appled Hydrology. McGraw-Hill, New York.
Mosley, M.P., 1979. Streamflow generation in a forested watershed, New Zealand. Water Resour. Res., 15: 795-806.
Mosley, M.P., 1982. Subsurface flow velocities through selected forest soils, South Island, New Zealand. J. Hydrol., 55:65-92.
Nash, J.E., and Sutcliffe, J.V. 1970. Riverflow forecasting through conceptual models. 1. A discussion of principles. J. Hydrol. 10, 282-.
Nikolaidas, N.P., Rajaram, H., Schnoor, J.L., and Geogakakos, K.P. 1988. A generalized soft water acidification model. Water Resour. Res. 24, 1983-96.
Nikolaidas, N.P, Schnoor, J.L., and Georgakakos K.P. 1989. Modelling of long-term lake alkalinity responses to acid deposition. Journal of Water Pollution Control Federation, 61, 2, 188-199.
Nikolaidis, N.P., 1987. Modelling the direct versus delayed response of surface waters to acid deposition in northeastern United States. Unpub. Ph.D. Thesis, University of Iowa, iowa City, Iowa.
North, M., 1980. Time dependent stochastic model of floods. ASCE Journal of Hydraulics Division. 106:649-665.
Nutter, W.L. and Hewlett, J.D., 1971. Stormflow production from permeable upland basins. In: E.J. Monke (Editor), Biological Effects in the Hydrological Cycle. Purdue Univ., West Lafayette, Ind., pp. 248-258.
Nye, P.H., 1981. Changes of pH across the phizosphere induced by roots. Plant and Soil 61: 7-26.

Oliver, B.G., E.M. Thurman, and R.L. Malcom. 1983. The contribution of humic substances to the acidity of colored natural waters. Geochim. Cosmochim. Acta 47:2031-2035.
Piligrim, D.H., Huff, D.D. and Steele, T.D., 1978. A field evaluation of subsurface and surface runoff. II. Runoff processes. J. Hydrol., 38: 319-341.
Ragan, R., 1968. An experimental investigation of partial area contributions. Publ. Int. Assoc. Sci.Hydrol., 76:241-251.
Rawls, W.J. and Asmussen, L.E. 1973. Subsurface flow in Georgia Coastal Plain. J. Irrig. Drain. Div., Am. Soc. Civ. Eng., 99: 375-385.
Reuss, J.O. 1980. Simulation of soil nutrient losses resulting from rainfall acidity. Ecol. Modelling, Vol. 11, 15-38.
Reuss, J.O. and D.W.Johnson., 1986. Acid deposition and acidification of soils and waters. Ecol. Series, Vol. 59. Springer-Verlag, New York.
Reuss, J.O. and D.W.Johnson. 1985. Effect of soil processes on the acidification by acid deposition. J. Environ. Qual. 14:26-31.
Reuss, J.O. and D.W. Johnson., 1985. Effect of soil processes on the acidification of water by acid deposition. J. Environ. Qual. Vol. 14, 26-31.
Riesbol, H.S. 1942. Some aspects of subsurface water in hydrologic research on agricultural watersheds, pp. 85-108. In: F.T. Mavis, ed. Proc. Hydrology Conf., Tech. Bull. 27, Pennsylvania State Univ., College of Engineering, University Park.
Rustad, S., N. Christophersen, H.M. Seip, and P.J. Dillon, 1986. Model for streamwater chemistry of a tributary to Harp Lake, Ontario. Can. J. Fish. Aquat. Sci., 43, 625-633.
Schnoor, J.L., Palmer, W.D. Jr., & Glass, G.E. 1984. Modelling
impacts of acid precipitation for northeastern Minnesota. In Modelling of Total Acid Precipitation Impact, ed. J.L. Schnoor, Butterworth, Boston, Mass., 155-173.
Schnoor, J.L., W.D. Palmer, Jr., and G.E. Glass. 1984. Modelling impacts of acid precipitation for northeastern Minnesota, edited by J.L. Schnoor, pp. 155-173, In: Modelling of Total Acid Precipitation Impact, Butterworth, Boston, Mass., .
Schumacher, W., 1864. Die Physik des Bodens, Berlin.
Semkin, R.G., and D.S. Jeffries, 1988. Chemistry of atmospheric deposition, the snowpack, and snowmelt in the Turkey Lakes Watershed. Can. J. Fish. Aquatic. Sci. 45:38-46.
Sklash, M.G., and R.N.Farvolden, 1979. The role of groundwater in storm runoff. J. Hydrol.43: 45-65.
Stein, J., H.G. Jones, J. Roberge, and W. Sochanska, 1986. The prediction of both runoff quality and quantity by the use of an integrated snowmelt model, In: Modelling Snowmelt-Induced Processes, IAHS Publ. No. 155, 347-358.
Stumm, W., J.J. Morgan, and J.L. Schnoor., 1983. Saurer Regen, eine Folge der Storung hydrogeochemischer Kreislaufe, Naturwissenschaften 70:216-223.
Stumm, W., and J.J. Morgan ., 1970. Aquatic Chemistry: An introduction Emphasizing Chemical Equilibria in Natural Waters. Wiley-Interscience Publication, John Wiley & Sons, New York. 780 pp.
Taylor, C.H. and Pearce, A.J., 1982. Storm runoff processes and subcatchment characteristics in a New Zealand hill country catchment. Earth Surf. Processes. 7: 439-447.
Toddorvic, P., 1978. Stochastic models of floods. Water Resources Research, 14:345-356.
Todorovic, P., and E. Zelenhasic, 1970. A stochastic model for flood analysis. Water Resource Research. 6: 1641-1648.
Ugolini, F.C., R. Minden, H. Dawson, and J.Zachara. 1977. An example of soil processes in the Abies amabilis zone of Central Cascades, Washington. Soil Sci. 124:291-302.
Waylen, P.R., 1981. Analysis of highflows in the Fraser River Catchment, British Columbia, Unpub. Ph.D. Thesis, McMaster University.
Weyman, D.R., 1970. Throughflow on hillslopes and its relation to the stream hydrograph. Bull. Int. Assoc. Hydrol. sci., 15:25-33.
Whipkey, R.Z., 1965. Subsurface flow from forested slopes. Bull. Int. Assoc. Sci. Hydrol., 10:74-85.
Zelenhasic, E., 1970, Theoretical Probability Distribution for Flood Peaks, Colorado State University, Hydrology Paper 42, 35p.

CHAPTER 3

Climatic change

M.A. Mimikou

Abstract. A global climatic change is likely to occur caused by the increasing atmospheric concentrations of carbon dioxide and other trace gases. These "greenhouse gases" are expected to change the radiative balance of the atmosphere, causing increases in temperature and changes in other climatic variables. This so-called "greenhouse effect" will have serious implications for the hydrologic balance and water resources. Current hydrological research suggests that plausible climatic changes will strongly influence various forms of water resources (surface runoff, soil moisture, groundwater), change sea-level, and affect water quality. These climatic change impacts entail important implications for future water resources planning and management associated with serious environmental and socio-economic problems.

3.1 Introduction

It is widely known that the atmospheric concentration of radiatively active "greenhouse" gases, principally carbon dioxide (CO_2), nitrous oxide (N_2O), methane (CH_4), and chlorofluoro-carbons (CFCs), has increased steadily over the past century. There is considerable consensus among the international scientific community that we are facing global warming due to this growing concentration of greenhouse gases in the atmosphere. This effect is known colloquially as "the greenhouse effect". It has been estimated that the average global temperature has increased 0.5–0.7°C since the turn of the century, whereas current estimates from models indicate that a doubling of atmospheric CO_2 and other gases may increase the annual temperature by 3.0 ± 1.5°C during the next 50–100 years. Such predictions are, of course, far from being certain and can, at best, be only considered as possible scenarios. Although the properties of the radiative greenhouse gases are rather well-understood, and despite the improvements in our understanding of atmospheric dynamics and large-scale climatic processes, the climatic change due to greenhouse gases is still not entirely understood, and a clear consensus has

V. P. Singh (ed.), Environmental Hydrology, 69–106.
© 1995 *Kluwer Academic Publishers.*

yet to emerge as to the overall direction that global climate is evolving toward. Intensive research efforts are spent in investigating the total climate and earth system as well as the global change. These efforts are effectively co-ordinated within international programs, such as the World Climate Research Program (WCRP) and the International Geosphere-Biosphere Program (IGBP). The WCRP has been initiated in 1979 as a joint undertaking of the International Council of Scientific Unions (ICSU) and the World Meteorological Organization (WMO), while the IGBP has been established by the ICSU in 1986.

One of the most important consequences of climatic change will be alterations in major climatic variables such as temperature, precipitation and evapotranspiration, subsequent changes in hydrologic cycles, and effects on the quantity and quality of water resources. At our present state of knowledge, even if it was assumed that the global warming will take place on the scale and over the period predicted, the mechanism to translate climatic changes to specific changes in temperature and precipitation over a watershed for the planning, management and operation of specific water projects, would still be poorly understood. The regional climatic change effects on hydrology and water resources have gained widespread research interest since the early '80s. Hydrological research suggests that plausible climatic changes will strongly influence various forms of water resources, such as mean and extreme surface runoff, soil moisture and groundwater, change sea-level, and affect water quality. These climatic change impacts entail important implications for future water resources planning and management associated with serious environmental and socio-economic dislocations.

3.2 A Brief Overview of the Literature

The earliest literature focuses mainly on the anthropogenic climatic changes, i.e. on how various human activities can influence local and global climate. For example, how might local activities and technical works, such as land use modifications and large reservoirs, alter local climate conditions? How might a growing number of environmental problems, such as the injection of ozone-reducing gases into the stratosphere, large scale deforestation in tropical zones, and the increase of radiatively active trace gases in the atmosphere, turn to change global climate? (SMIC, 1971; U.S. National Academy of Sciences, 1977). Growing atmospheric concentration of carbon dioxide - a direct by-product of the combustion of fossil fuels and of net deforestation - has gained early interest (Revelle and Suess, 1957), and has been recognized as having the largest single anthropogenic effect on climate. Simulation models of general atmospheric circulation have shown an increasing trend of the global annual temperature as a direct result of the increasing atmospheric concentration of CO_2 and of other radiative gases (U.S. National Academy of Sciences, 1977, 1979, 1983; Clark, 1982). Much of the related research work has been devoted on the consequences of a "doubling" of the atmospheric concentration of CO_2, a situation which, according to current "guess-estimates", can range sometimes between 2030 and 2090.

It is noticeable that very little of the early work on climatic change considered possible greenhouse warming impacts on hydrology and water resources. This is

possibly due to the uncertainty about the way that future climatic changes would affect precipitation patterns. Early attempts to evaluate the interactions among climate, hydrology, and water supply, were made in the study "Climate, Climatic Change, and Water Supply" by the National Research Council Panel on Water and Climate (1977), where the problem of the future water shortages due to climatic change was discussed. It was concluded that, despite problems caused by the insufficiency of the climatological forecast ability to aid regional water-resource planning or hydrologic design, certain useful measures could be implemented that would help to minimize the undesirable effects of future water shortages. Global scale studies with the simulation models of general atmospheric circulation have shown that a doubling of atmospheric CO_2 will lead to a global average increase in the intensity of the hydrologic cycle, and, more specifically, in the rates of precipitation and evapotranspiration (Clark, 1982; U.S. National Academy of Sciences, 1983; Manabe and Wetherald, 1985). In 1985, the Scientific Committee on Problems of the Environment (SCOPE) published a report on "Climate Impact Assessment" (Kates et al., 1985), in which climatic change impacts on water resources started to be considered in a systematic and integrated manner.

Since the '80s, the regional hydrological effects of greenhouse warming have begun to be evaluated extensively through detailed case studies in different places of the world. Key issues that gained widespread research interest are: the evaluation of regional changes in major climatic variables, the development of appropriate simulation models in basins of interest to assess climate change impacts on various forms of water resources as well as water management works, and the testing - validation of the different assessment methods. The methodologies used in these studies cover a wide spectrum of models and techniques for the assessment of the climatic change impacts on hydrology and water resource systems. In 1985, the World Meteorological Organization published a review of such methods including specific recommendations for testing and validating the different assessment methods (Klemes, 1985). This was followed by a comprehensive report on the sensitivity of water resource systems to climate change and variability (WMO, 1987). Later, and during the past few years, several international associations and organizations convened panels of experts, organized workshops, symposia and conferences, and published books including a variety of important articles on the subject (Solomon et al., 1987; Pearman, 1988; Commission of the European Communities, 1989; Conference on Climate and Water-The Publications of the Academy of Finland, 1989; WMO, 1989; AAAS, 1990; Houghton et al., 1990; Mac Cracken et al., 1990; Jager and Ferguson, 1991; Mintzer, 1992).

The very unusual weather patterns and hydrological conditions that have characterized much of the post-1987 period, created severe drought and flooding problems in many countries of the world. These problems stimulated considerable scientific and public interest and served to highlight an emergency consensus regarding the possible environmental, economic and societal impacts of climatic change. As climate change research progresses, it becomes increasingly clear that attempts should be made towards an integrated approach to study the entire scope of the problem from basic physical, climatological and hydrological issues to questions of environmental, political and economic nature.

3.3 Climatic Variability and Change

3.3.1 Global Greenhouse Warming – Past, Present and Future Conditions

Climate is a dynamic regime subject to natural variations on all time scales, from years to millennia, and to possible changes by human activities as well. Climate variability is defined as the variation from year to year and generally comes as a result of natural conditions, whereas climate change is defined as a significant change in climatic variables, from one period to another period, and can be caused both by natural conditions and by human activities. (Refsgaard et al., 1989). The largest single anthropogenic effect on climate that gained widespread interest during the past decade, is the greenhouse effect that is caused by the growing concentrations of the radiatively active gases, principally carbon dioxide, and other trace gases such as nitrous oxide, methane and chlorofluorocarbons. These increasing concentrations of greenhouse gases reduce the amount of the earth's radiation that escapes to space with consequent "greenhouse" warming of the lower atmosphere and the earth's surface. The record of the atmospheric concentrations of CO_2 (in ppm) since 1740 is shown schematically in Figure 3.1 (adapted from Siegenthaler and Oeschger, 1987; WMO/ICSU, 1990). In this figure, one can easily notice the significant acceleration of the rate of CO_2 concentrations increase during the past quarter century. It has been estimated (Siegenthaler and Oeschger, 1987) that the emissions of CO_2 have increased by 25% since the beginning of industrialization. This increase of atmospheric concentration of CO_2 results mainly from the combustion of fossil fuels and from deforestation causing an imbalance between the absorption and release of CO_2 by vegetation. On the other hand, the record of global-mean surface air temperature change during the past century, shown in Figure 3.2 (adapted from Hansen and Lebedeff, 1987; WMO/ICSU, 1990), indicates a considerable warming of the earth's a surface during the last hundred years. It has been estimated that the average global temperature has increased by 0.5–0.7°C since the turn of the century.

Moreover, the record in Figure 3.2 indicates ongoing temperature fluctuations of comparable amplitude, which do not constitute a long-term trend and must be ascribed to the natural climatic variability. Thus, the problem of detecting long-term temperature trends induced by human activities cannot be separated from that of understanding and predicting natural climatic variations. The warming realized so far is substantially less than could be inferred from the change in the atmospheric radiative energy budget due to the excess concentration of greenhouse gases over the past decades. However, it is still within the range of transient natural variations and cannot be separated from these with our present knowledge. Nevertheless, it is likely that the limits of natural variability will be exceeded within the next decades (WMO/ICSU, 1990).

The estimation of the future emissions and global budget of greenhouse gases originated from either human activities or natural exchanges between the atmosphere, the ocean or the land, as well as the role of clouds on the earth radiation budget, are only some of the unsolved problems causing considerable uncertainty in the prediction of future climatic changes. Despite these uncertainties, and re-

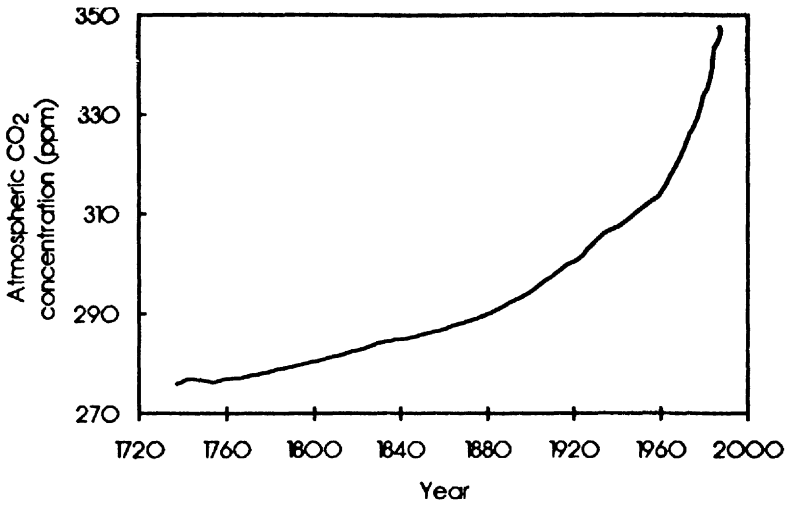


Fig. 3.1. Record of the atmospheric concentration of carbon dioxide (adapted from Siegenthaler and Oeschger, 1987; WMO/ICSU, 1990).

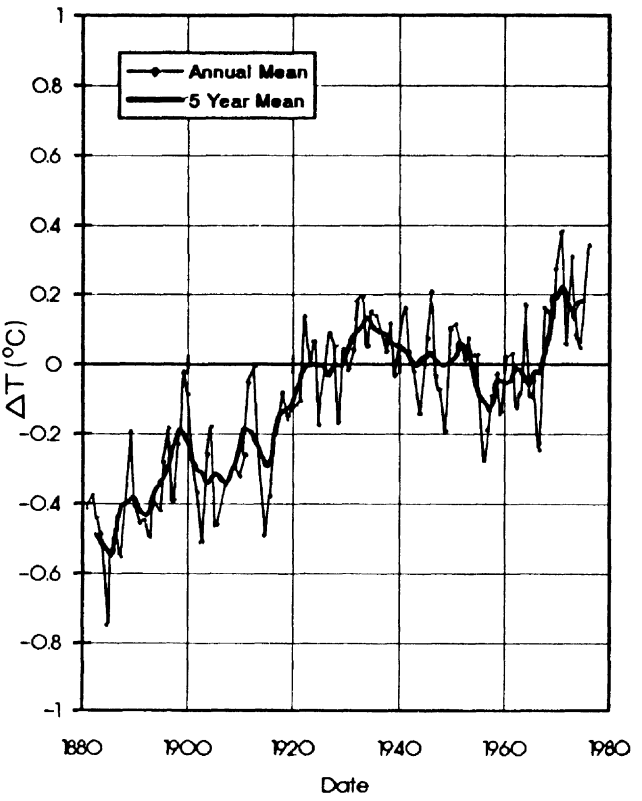


Fig. 3.2. Record of global-mean air temperature change at the surface of the earth (adapted from Hansen and Lebedeff, 1987; WMO/ICSU, 1990)

gardless of the ongoing efforts to control emissions of greenhouse gases, it is expected that, under any circumstances, considerable increase in the atmospheric concentration of greenhouse gases will occur in the next century. The climate response to this greenhouse forcing is expected to be much less than could be inferred directly from the imbalance of the atmospheric radiative energy budget, and delayed by decades because of the effect of oceanic thermal inertia (WMO/ICSU, 1990). Nevertheless, the climatic impacts of greenhouse gases appear to be inevitable, and the global greenhouse warming, to which we are already committed, will be realized. The WMO/UNEP Intergovernmental Panel of Climate Change (IPCC) adopted working assumptions consisting of three possible scenarios for the rate of increase of the atmospheric CO_2 concentrations: The high rate scenario is equivalent to doubling the early 20th century CO_2 concentrations by the year 2030, the medium rate scenario by the year 2060, and the low rate scenario by the year 2090 (WMO/ICSU, 1990). Present "guess-estimates" from state-of-the-art climate models of the atmosphere based on the long-term equilibrium climate change assumption, e.g. when the atmosphere has reached equilibrium with the ocean, indicate that this CO_2 doubling will lead to a global average equilibrium warming estimated to be 3.0°C with a probable error of ± 1.5 degrees (U.S. National Academy of Sciences, 1979; 1983). The global-mean temperature rise, due to CO_2 alone, is expected to be accelerated because of the growing concentration of other greenhouse gases (CH_4, N_2O, CFCs), and is likely to occur even before the doubling of atmospheric CO_2 alone (Berger, 1989). The global warming rate, consistent with an equivalent doubling of CO_2 concentration by the year 2050 and an estimated 4.0°C equilibrium temperature increase at that time, is estimated to be 0.3–0.4°C per decade (Hansen et al., 1988). This rate exceeds the fastest observed temperature trend during any 10-year period since 1880 (Figure 3.2), and is believed to cause a significant deviation from the prevailing climatic conditions. Such predictions are, of course, far from being certain, and can, at best, be considered as possible scenarios. Despite the significant improvements in our understanding of atmospheric dynamics and large-scale climatic processes, the greenhouse warming is still not entirely understood, and a clear consensus has yet to emerge as to the overall direction that global climate is evolving toward.

The World Climate Research Program has recently instituted three major climate change-research projects: The Tropical Ocean and Global Atmosphere (TOGA), the Global Energy and Water Cycle Experiment (GEWEX), and the World Ocean Circulation Experiment (WOCE). These research projects will advance our knowledge of basic phenomena, such as the interaction between tropical ocean (the Southern Oscillation and the associated El Nino phenomenon) and global atmosphere, the hydrologic cycle in the climate system, and the "delayed" transient response climatic changes due to the thermal inertia of the oceans. They will also provide the observations needed to produce this kind of global description, which will be essential to distinguish the warming trend due to the steady increase of greenhouse effect in the atmosphere from ongoing natural variations. In addition, the WCRP co-operates with the International Geosphere-Biosphere Program of ICSU to establish the scientific basis for quantitative assessments of changes in the earth's geo-biochemical cycles that control the CO_2 atmospheric concen-

trations. These programs constitute the international framework for a scientific understanding of global change.

3.3.2 Methods for Evaluating Global and Regional Climatic Changes

The climatic changes, due to the steady increase of the greenhouse effect in the atmosphere, are currently evaluated by using a variety of mathematical models with different levels of complexity, from simple models to complex, three-dimensional general circulation models. This variety of models reflects the wide range of simplifications and assumptions used to approximate the enormous complexities of the climate system, as well as the wide range of researchers' interests to study a broad spectrum of climatic changes. There are models dealing with the radiative balance of the earth/atmosphere system, which are simple in their structure, such as the **Energy-Balance Models** (EBMs), (North et al., 1981), and the **Radiative-Convective Models** (RCMs) (Manabe and Wetherald, 1967; Reck, 1974; Hansen et al., 1978). These models have a limited geographical and temporal domain, and exclude hydrologic processes. Improvements over these two types of models in the areas of atmospheric dynamics are offered by the **Statistical-Dynamical Models** (SDMs) (Mac Cracken et al., 1981) that are of intermediate complexity between the simple models and the most advanced climate models, which are the **General Circulation Models** (GCMs). The GCMs are detailed, time-dependent, three-dimensional numerical simulations that run on powerful computers. They include detailed representations of atmospheric motions and heat exchanges, and the important land-ocean-ice interactions as well. In addition, GCMs have been undergoing refinement for many years. Comprehensive reviews of the history and the improvements of GCMs are provided by many researchers, e.g. Dickinson (1982); Manabe (1982); Hansen et al. (1983); Meehl (1984); Washington and Meehl (1984); Mearns et al. (1990). GCMs are extensively used to evaluate CO_2 induced changes in various climatic and hydrological variables, such as temperature, cloud and ice cover, radiative fluxes, the water vapor distributions, evapotranspiration, precipitation, runoff and soil moisture. They are believed to be more accurate than other models, at least on a global scale, because of their higher resolution, as well as to be capable of providing better information on regional scales. Nevertheless, the complexity of GCMs and their demand for significant computer power that will increase when finer resolution is needed for reliable regional climate change assessments, are presently the main factors limiting their usefulness. GCMs exhibit other weaknesses as well, such as inadequate and limited representations of cloudiness, surface hydrology, soil moisture and turbulent heat transfer, and oversimplification in ocean modeling (North et al., 1981).

Despite their acknowledged limitations, GCMs are considered to be the most important method for obtaining information on possible future climatic changes that are likely to be encountered as atmospheric concentration of greenhouse-gases increases. Because of deficiencies and limitations in current GCMs, their climatic change predictions should only be considered as possible scenarios. A climate scenario is intended to be an internally consistent picture of future climatic conditions. In addition to the GCM-derived scenarios, climate scenarios may be constructed by a number of other means including: the use of *past instrumental* data,

where the differences between globally warm and cold (relative to the long-term mean) periods are used as indicators of future changes, the use of *paleoclimatic analogs* (usually the analog of the early Holocene warm period), methods of constructing *synthesis* scenarios combining features from a number of different methods, and most commonly *arbitrary (hypothetical)* change scenarios spanning ranges of plausible variations in climatic variables, e.g. a change in temperature of +0%, +2%, +4%°C combined with 0%, ±10%, ±20% changes in precipitation.

3.3.3 Impacts of Climatic Changes

One of the most important consequences of climatic change that will be felt both globally and regionally will be alterations in major climatic variables, such as temperature, precipitation and evapotranspiration, subsequent effects in hydrologic cycles, and changes in the quantity and quality of water resources, sea-level, plants and natural ecosystems. These climate change impacts entail important implications for future water resource planning and management associated with serious environmental and socio-economic problems.

The surface greenhouse warming, estimated from several state-of-the-art climate models in the literature (Hansen et al., 1984; Washingthon and Meehl, 1984; Manabe and Wetherald, 1987; Wilson and Mitchell, 1987), exhibits significant geographical and seasonal variation. The geographical pattern of temperature change indicates larger warming at higher latitudes than in the tropics, and lesser warming of the ocean compared to land. Because of the physical relationship between temperature and the saturation pressure of water vapor, warmer climate entails a wetter atmosphere, and consequently, a general increase in precipitation varying with geographical location and season. This means that annual, and especially winter, runoff may be increasing at higher latitudes. On the other hand, a warmer climate results in increased evapotranspiration, which affects directly runoff and soil moisture conditions, especially in summer, and deteriorates seriously soil aridity and groundwater depletion, particularly in arid mid-continental temperate regions. However, the predicted regional changes in temperature, and even more so in precipitation and evapotranspiration, exhibit very uneven distribution and differ widely from one model to another. It is generally accepted that regional scale changes in climate have not yet been modeled with confidence, therefore current models cannot provide reliable predictions of regional climatic changes and impacts.

The climate change impacts on hydrology and water resources gained widespread interest since the early '80s. Hydrological research, progressing mainly through case studies in different places of the world, strongly suggests that plausible climatic changes will have dramatic influence on hydrologic cycles and water resources.

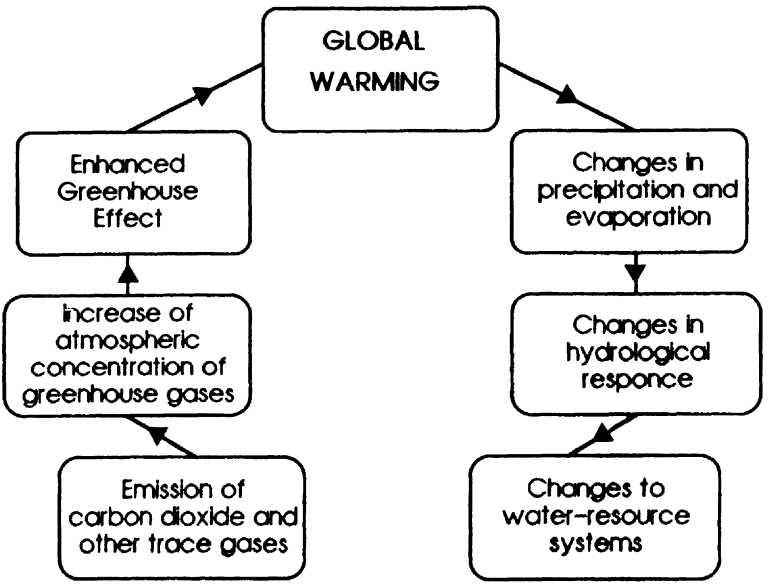


Fig. 3.3. The hydrologic cycle in the greenhouse effect.

3.4 Hydrological Effects of Climatic Change

3.4.1 Hydrological System Response to Climatic Changes

A very simple view of the hydrologic cycle in the greenhouse effect is illustrated in Figure 3.3. Of course, the full role played by the hydrologic cycle in the greenhouse effect cannot be inferred from this figure that highlights only consequences and oversimplifies the processes and feedbacks involved. The hydrologic cycle is instrumental in controlling and mediating greenhouse processes through its agency in other primary cycles that make up the global environmental system, such as momentum and energy transfer, carbon, etc. Pathways representing the influence of human activities through changes to the exploitation of water resources and catchment management on growing emission and atmospheric concentration of CO_2, would effectively complete the circle in Figure 3.3.

The hydrological response to the global greenhouse warming will be alterations in major climatic variables, such as precipitation and evapotranspiration, and consequent changes in hydrological variables, such as surface runoff, soil moisture, groundwater and sea-level. These climatic change impacts on various forms of water resources entail serious implications for future water resource planning and management, such as water supply, agricultural water use, reservoir reliability, energy production, land drainage, flood protection, etc., and will also affect water quality and treatment. As previously described, the expected global hydrologic changes include an overall increase in the intensity of the hydrologic cycle due to higher average temperatures and increased evapotranspiration, which, in turn, will mean higher global-average precipitation rates. But such general predictions are of little use for hydrologic studies which usually look at small-scale hydrologic

processes or at watersheds that range in size from tens of hectares to hundreds of square kilometers. The essential problem here is to evaluate the regional hydrological effects of such global climatic changes since hydrologic patterns are often characterized by small-scale physical phenomena as a result of which the scale of the surface grids of global climate models is large enough to provide the appropriate level of details to evaluate regional impacts. Therefore, other different methods must be used to study in more detail the climatic change effects on small-scale hydrologic phenomena.

3.4.2 Methods for Evaluating Regional Hydrological Effects of Climatic Changes

An assessment of the impact of climatic changes on hydrology and water resources involves the following steps (Klemes, 1985; Gleick, 1989):

* A quantitative estimate of changes in the long term indices of the major climatic variables, such as temperature, precipitation and evapotranspiration, by using quantitative climatic scenarios.

* A simulation of the hydrologic cycle for a basin of interest, using the scenarios developed in the previous step, as well as estimation of the climatically induced variations in various hydrological variables of interest.

* An assessment of the implications of the previously identified hydrologic variations for the performance and the reliability of water resource systems, such as aqueducts, dams, reservoirs, groundwater aquifers, etc.

The first two steps satisfy the needs of hydrologists, whereas the third one is necessary for water-resource planners, which are more interested in the effects of climatic changes on water supply, irrigation, reservoir reliability, energy production, and other water management issues.

General Circulation Models provide the best information on the overall response of the atmosphere to growing concentrations of carbon dioxide and other trace gases. At least in theory, GCM-derived estimates of variations in hydrological variables can be used directly to assess changes in water quantity and quality, and consequent implications for future water resource planning and management. Nevertheless, GCM-generated hydrological variables are extremely simplified and, at best, they can be considered to represent some hypothetical large-scale areal averages with little relevance to specific drainage basins (U.S. Environmental Protection Agency, 1984; Manabe and Wetherald, 1986; Schlesinger and Mitchell, 1985; Mitchell and Warrilow, 1987). For example, major hydrological variables, such as runoff and soil moisture, have limited realism since they are computed as secondary variables and calculated from simple difference equations set up to manipulate the first-order climatological variables, such as temperature, humidity and energy transfers (Washington and Meehl, 1983; 1984). Another major limitation of GCMs concerns their coarse resolution that is inappropriate for regional hydrological studies. By this it is meant that GCM resolutions usually range from 4.0° to 7.5° latitude by 5.0° to 10.0° longitude and cover grid areas of hundreds of

thousands of square kilometers. This is a scale several orders of magnitude larger than the scale of hydrologic interest that is usually limited to tens or hundreds of square kilometers. Uncertainty is also caused from the basic assumption of GCMs concerning equilibrium response to a doubled or quadrupled CO_2 concentration, and is not yet known whether it will correspond to regional patterns of changes in climate (Webb and Wigley, 1985). Efforts have been made to improve hydrologic parameterizations in GCMs, and modify GCM-outputs by obtaining sub-grid scale information (Bach, 1988). Also, efforts have been made to develop hydrologic models, so-called "macroscale models", specifically designed to use outputs from GCMs (Becker and Nemec, 1987). However, until such improvements occur, those interested in the regional hydrological effects of climatic changes must look for other methods of evaluation (Gleick, 1989). Nevertheless, GCM-predicted changes in temperature and precipitation are considered to constitute plausible climatic change scenarios, and there is a tendency of using them as inputs to other, external to GCMs, more detailed regional hydrologic models.

The use of regional hydrologic models to evaluate changes in hydrological variables of interest, such as runoff, soil moisture and groundwater, offers important advantages over the direct use of hydrologic outputs from GCMs. Firstly, there is a wide variety of models with moderate operating requirements. Secondly, the effects of a variety of climatic change scenarios constructed by either GCMs or other methods can be evaluated. And finally, methods combining both regional characteristics and outputs from GCMs can be developed, which can take advantage of the continuing improvements of global climate models (Gleick, 1989).

Several types of regional hydrologic models have already been developed to evaluate changes in important hydrological variables. A variety of classification schemes have been used to distinguish among these models (Gleick, 1986a; 1989), the most important distinction being made on the basis of the different techniques used to evaluate the regional hydrologic characteristics and extend the existing hydrologic records. The *empirical* and *statistical models* use empirical relationships and statistical properties of hydrologic processes (Stockton and Boggess, 1979; Revelle and Waggoner, 1983; Cohen, 1987; Palutikof, 1987). The *deterministic* or *conceptual models* make use of physically-based mathematical descriptions of hydrologic processes (Nemec and Schaake, 1982; Cohen, 1986; Gleick, 1986a,b; 1987a,b; Flaschka et al., 1987; Fitzerald and Walsh, 1987; Bultot et al., 1988b; Schaake, 1990; Mimikou et al., 1991a). It can be noted that *stochastic models* have limited use in evaluating hydrological effects of climatic changes, since they assume statistical equivalence between future and past. They are merely used in evaluating probabilities of dry and wet years in future sequences based on past hydrologic data, and therefore are useful in estimating climatically induced impacts on the reliability of various water management works (Mimikou et al., 1991b).

Considerable attention has been given to the physically-based conceptual models, which, although laborious to calibrate especially in the absence of sufficient field data, permit direct and integrated simulation of hydrologic processes. Many models have been developed with varying ability to represent small and large-scale basin characteristics with different time resolutions. A systematic presentation of the most significant factors to be considered when selecting a model to assess the regional hydrological effects of climatic changes has been reported by Gleick

(1986b; 1989). More specifically, he suggests the following set of criteria: (1) The inherent accuracy of the model, (2) the degree to which model accuracy depends upon the existing climatic conditions for which the model was calibrated, (3) the availability of input data, including comparative historical climatic data, (4) the accuracy of input data, (5) the model flexibility, ease of use, and adaptability, and, lastly, (6) the compatibility with general circulation models.

The *accuracy* of a model in reproducing existing hydrologic conditions is a very important factor for its performance in evaluating climatic change impacts. Hydrological research indicates that the efficiency of a model is significantly affected by specific characteristics of the watershed and the quality of the input data. An extensive intercomparison analysis of conceptual models was undertaken and published by the World Meteorological Organization (WMO, 1975), where it was shown that, even though all models performed well for humid basins, only the explicit moisture accounting models, such as water balance models, were superior in semi-arid and arid basins. In addition, it was observed that simple models performed better than complex models in basins with poor quality of input data. Both basin aridity and quality of input data are well recognized in the literature as important factors affecting the efficiency of hydrologic models (James and Burges, 1982; Linsley, 1982; Mimikou et al., 1992).

The degree of *dependence of a model on its initial calibration* is very significant for those circumstances where a climatic change affects the initially calibrated model parameters. A model should be validated in order to guarantee its accurate performance for evaluating conditions following a climatic change that are different from those for which the model has been initially calibrated. Strictly speaking, this type of validation is not possible until the climatic changes actually occur (Gleick, 1989). However, there are several validation tests in the literature (WMO, 1975; Klemes, 1985) that can assign at least some measure of credibility on the assumption that the initial model calibration is still valid. Among these tests is the differential split sample test, which implies that two periods with different historical climatic conditions must be chosen, such as a period of high average precipitation, and a period of low average precipitation. The model, intended for example to simulate a transition to a warmer and wetter scenario, should be calibrated on the dry, low average precipitation data set, and then validated for the wet period with high average precipitation. A comprehensive analysis of such operational validation testing of hydrologic simulation models is provided by Klemes (1985).

The *availability of input data* must satisfy both calibration and validation requirements of a model and it must be kept in mind that the value of a model for evaluating the effects of climatic changes is questionable, unless it can be validated using observed historical data.

The *quality* and *accuracy of input data* are also important for both calibration and validation of a model. Poor quality of input data leads to serious implications for achieving a good fit during calibration, and results in biases in model application.

Model flexibility and *adaptability* are important features of the structure of a model, permitting its direct applicability to many diverse watersheds exhibiting different patterns of flows. Simplicity and ease of use of a model improve its adaptability, and, as stated earlier, are not necessarily against model accuracy.

The need for *compatibility with global large-scale climate models*, especially with GCMs, arises from our need for the most detailed information on future climatic changes, provided by GCMs, which will improve our ability to understand the regional impacts of global change. By linking GCMs and regional hydrologic models, methods can be developed that combine the capabilities of both since detailed GCM-derived climatic change scenarios are incorporated into watershed models to evaluate regional hydrological effects of climatic changes (Gleick, 1987b; Bultot et al., 1988; U.S. Environmental Protection Agency, 1988; Lettenmaier and Gan, 1990). These methods of linking GCMs and regional hydrologic models will improve their performance, taking advantage of the continuing improvements in the resolution and hydrology of GCMs.

The most widely used models in the category of physically-based conceptual models are the water balance models that were first developed in the late 1940s and 1950s (Thornthwaite, 1948; Thornthwaite and Mather, 1955; 1957). Since then, they have been used in a variety of hydrological problems, and a great deal of experience has been gained on how they can be applied most efficiently (Sokolov and Chapman, 1974; Miller, 1977; Mather, 1978; U.S. Army Corps of Engineers, 1980; Gleick, 1986a,b; 1987a,b; Vandewiele et al., 1992). The capability of water balance models to incorporate soil moisture, snowmelting and evapotranspiration routines, in addition to their flexibility in using different time scales, render them an invaluable tool for the evaluation of regional hydrological effects of climatic changes.

A wide variety of studies appear in the literature during the past few years for the assessment of regional hydrological effects of climatic changes. The latter are evaluated by using plausible climate change scenarios constructed by either GCMs or by other methods, such as hypothetical climate scenarios, paleo- and recent- climate analogs, and so on. An overview of some of these hydrologic impact studies using different climatic change scenarios is given in Table 3.1. The studies in Table 3.1 are referenced with the name(s) of author(s), and the year of publication. Moreover, Table 3.1 includes information on the category and type of model, and the climate change scenario used in the studies. The scenarios used are mainly either outputs from GCMs or hypothetical; in the former case, the name of the researcher who produced the specific GCM-output run is given, while in the latter case, the accepted values of changes in temperature (ΔT), and precipitation (ΔP) are indicated. The results of such regional hydrologic studies should be interpreted as a sensitivity analysis to alternative climates rather than as predictions, mainly because of the uncertainty of the input climatic data (Lettenmaier and Gan, 1990). Hydrological research is progressing mainly through such case studies in different places of the world. So far, results suggest a strong regional influence of greenhouse warming on major climatic and hydrological variables, and indicate a considerable consensus on how these changes will manifest themselves. In the following sections some regional climatic change effects on major climatic and hydrological variables are presented.

TABLE 3.1
Overview of some hydrologic studies using different climate scenarios

Study	Category/type of model	Climate scenarios		
		GCM-Output (*)	Hypothetical ΔT (°C)	ΔP(%)
Stockton and Boggess (1979)	Empirical/Regression	——	±2	±10
Nemec and Schaake (1982)	Conceptual/SSMAM	——	+1,+3	±10,25
Revelle and Waggoner (1983)	Empirical/Regression	——	+2,+4	±10
Gleick (1986b)	Conceptual/Water Balance	——	+2, +4	±0,10,20
Palutikof (1987)	Empirical/Regression	(Recent climate analog	- instrumental	record)
Cohen (1987)	Empirical/Regression	GISS/GFDL	——	——
Flashka et al. (1987)	Conceptual/Water Balance	——	±2	±10, 25
Sanderson and Wong (1987)	Water Budget/Morton Model	GISS	——	——
Gleick (1987b)	Conceptual/Water Balance	GISS/GFDL/NCAR	——	——
Bultot et al. (1988b)	Conceptual/IRMB	GISS	——	——
Arnell and Reynard (1989)	Empirical/Regression	—	+3 to +5	+20,10~25(**)
Schaake (1990)	Conceptual/Water Balance	—	+2 to +5	±0, 10, 20
Cooley (1990)	Conceptual/NWSRFS	—	±2, +4	+10
Lettenmaier and Gan (1990)	Conceptual/NWSRFS	GISS/GFDL/OSU	——	——
Vehvilainen and Lohvansuu (1991)	Conceptual/HBV	GISS	——	——
Mimikou et al. (1991a)	Conceptual/Water Balance	——	+1, +2, +4	±0, 10, 20
Nash and Gleick (1991)	Conceptual/NWSRFS	GISS/GFDL/UKMO	+2, +4	±0, 10, 20

(*) GCM-outputs have been produced by different researchers, such as GISS, NCAR, OSU and UKMO
(**) Depending on the seasonal distribution of precipitation

Abbreviations in the table are defined as follows:
GISS: Coddard Institute for Science Studies, New York
GFDL: Geophysical Fluid Dynamics Laboratory, Princeton, New Jersey
NCAR: National Center for Atmospheric Research, Boulder, Colorado
OSU: Oregon State University, Corvallis
UKMO: United Kingdom Meteorological Office
NWSRFS: National Weather Service River Forecasting System
SSMAM: Sacramento Soil Moisture Accounting Model
IRMB: Daily-step hydrologic model developed by the Royal Meteorological Institute of Belgium

3.4.3 Sensitivity of Major Climatic and Hydrological Variables to Global Warming

3.4.3.1 Precipitation and evapotranspiration

The effects of global warming on precipitation and evapotranspiration appear to be more complex compared to the effects on air temperature. Because of the physical relationship between temperature and the saturation pressure of water vapor, global warming entails a general increase in precipitation. On the other hand, a warmer climate results in increased evaporation as well. Despite the broad consensus as to the general direction of changes in global precipitation and evaporation, there are considerable doubts concerning GCM-predicted displays of regional distributions for both climatic variables. These displays are believed to be far from realistic, and are not considered adequate to specify the regional climatic conditions for

many applications (Budyko, 1989; Schneider, 1990). Thus, several case studies have been conducted to investigate the changing regional patterns of precipitation and evapotranspiration under global warming by using watershed specific hydrologic models. For example, Bultot et al. (1988a,b) undertook an experiment conducted in Belgium and investigated the modifications of climatic variables in this geographical zone, i.e. at an intermediate latitude, under the doubling of CO_2 climatic conditions predicted by a general circulation model (Manabe and Wetherald, 1975), being fed into a regional daily hydrological model (the *IRMB model*). Among other interesting results, they found that: under the $2 \times CO_2$ climate conditions used in modeling, the potential and effective evapotranspiration will be strengthened throughout the year (annual increase by about 9% and 7%, respectively), the precipitation cells with snow cover will be shorter whereas winter rainfall and runoff will be strengthened, thus increasing the frequency of floods, in contrast to the summer conditions, which will become drier (Bultot and Gellens, 1989). Increased annual evaporation values ranging from 14% to 19%, were also estimated for a number of lakes and reservoirs in Poland under the hypothesis of CO_2 doubling (Jurak, 1989).

The effects of CO_2-induced climatic changes on the amount of precipitation falling as rain and as snow, and subsequent snow accumulation and melting, are shown to play a key role for the evaluation of changes in other hydrological variables, such as evapotranspiration, runoff, and soil moisture (Yeh et al., 1983; Wigley and Jones, 1985; Rango and Martinec, 1987; Gleick, 1987b; Lettenmaier and Gan, 1990; Cooley, 1990; Vehvilainen and Lohvansuu, 1991). For example, Lettenmaier and Gan (1990) in their study on hydrologic sensitivities of four mountainous catchments of the Sacramento-San Joaquim River basins, California, to three GCM-derived CO_2 doubling scenarios, found that an increase in rain-on-snow events accompanied by a major seasonal shift and substantial decreases in snow accumulation and melt would occur. This would result in increased winter runoff and soil moisture storage, whereas spring snowmelt, summer runoff and summer soil moisture would be decreased. The maximum actual monthly evapotranspiration patterns were shifted earlier in the year, but the net annual change of total evapotranspiration was little. Significant reductions of snow packs and snow water equivalents under warming conditions, resulting in variable, but significant, changes in stream flows, have been evaluated in other case studies as well, as, for example, in south-eastern Montana, USA, by Cooley (1990), and in Finland by Vehvilainen and Lohvansuu (1991).

Another important issue is related to the potential effects of a changing climate on heavy rainfall intensity and frequency. Results from a recent study in the western region of the United States (Chagnon and Huff, 1992) indicated that warmer and wetter conditions, such as the GCMs predict, could be accompanied by increases of 10% to 15% in heavy rainstorm events with recurrence intervals of 5 to 10 years.

Despite the progress in the efficiency of GCMs and the research efforts spent in case studies in various places of the world, the detailed impacts of global warming on regional precipitation and evapotranspiration are still not well understood, and much uncertainty remains. Controversial aspects including the direct action of CO_2 as an antitranspirant — the phenomena of increased stomatal resistance and increased plant water use efficiency in high carbon dioxide environment - (Morison,

1989; Dooge, 1989; Kuchment et al., 1989; Allen, 1991), and the sensitivity of GCM-derived predictions to initial climate conditions (Tsonis and Elsner, 1989), still remain to be further investigated. The uncertainty concerning future changes in regional precipitation and evapotranspiration is the main reason for the diversity of various climatic change scenarios used as inputs to regional hydrologic models. These scenarios vary in their assumptions for the direction of changes in the climatic variables (see, for example, Table 3.1).

3.4.3.2 Runoff and soil moisture

Despite the uncertainties concerning future regional climatic changes, some of the studies mentioned previously have identified impacts on runoff and soil moisture that are rather insensitive to variations in input assumptions. In general, annual runoff is shown to be more sensitive to changes in precipitation than to temperature changes, in contrast to the seasonal distribution of runoff and soil moisture, which is more sensitive to temperature than to precipitation (Gleick, 1986b; 1987b; 1989; U.S. Environmental Protection Agency, 1988; Bultot et al., 1988b).

Regional studies of the sensitivity of annual runoff quantities to temperature increases and precipitation changes, have shown that temperature increases alone, with no change in precipitation, cause relatively small percentage reductions in the mean annual runoff, whereas changes in precipitation have a greater effect upon annual runoff. The effects of precipitation upon annual runoff appear to be characterized by a magnification factor, with respect to the change of the causative factor (precipitation change), which is believed to be positively associated with basin aridity (Gleick, 1986a,b; Dooge, 1989; Arnell and Reynard, 1989; Mimikou and Kouvopoulos, 1991). Thus, a given percentage change in precipitation is expected to trigger a greater change in runoff in drier areas, and a nearly proportional change in runoff in wetter areas. Results from different case studies indicate that changes in mean annual runoff can vary depending on the alternative climate scenario used. But, on the average, changes in mean annual runoff of about 30% as a result of climate change seem to be realistic, with the possibility of greater changes in the most arid regions (Arnell, 1989; Shiklomanov, 1989; Mimikou et al., 1991a; Nash and Gleick, 1991).

As previously discussed, warmer environments entail an increase in rain-on-snow events, and significant decreases in snow accumulation and melting with subsequent impacts on seasonal runoff and soil moisture. The increase of rainfall, in combination with the decrease of snowfall will increase winter runoff, making more soil moisture available during winter and early spring at the expense of late spring and summer soil moisture. Spring snowmelt and summer runoff will decrease with the time of occurrence of snowmelt runoff shifting from spring to winter. Thus, it is expected that the severity and frequency of winter floods will increase, whereas summer conditions will be drier. Some results indicating changes in seasonality of runoff and soil moisture in the Sacramento basin, California, resulting from eight GCM-derived climatic change scenarios, are schematically shown in Figure 3.4 (adapted from Gleick, 1987b; 1989). Results indicating a comparable character of changes have been obtained by other researchers in different geographical locations and climatic regimes. For example, in Figure 3.5 (adapted from Mimikou

et al., 1991a) one can see some results from a regional study in Greece. This figure presents the percent changes in winter and summer runoff as functions of hypothetical temperature and precipitation future changes for a medium-sized mountainous catchment in central Greece.

However, although different case studies suggest the general character of changes in regional runoff and soil moisture that would occur under a global warming, the specific hydrologic changes for any given catchment would depend on local climatic, physiographic, geologic and topographic characteristics, and should be investigated individually.

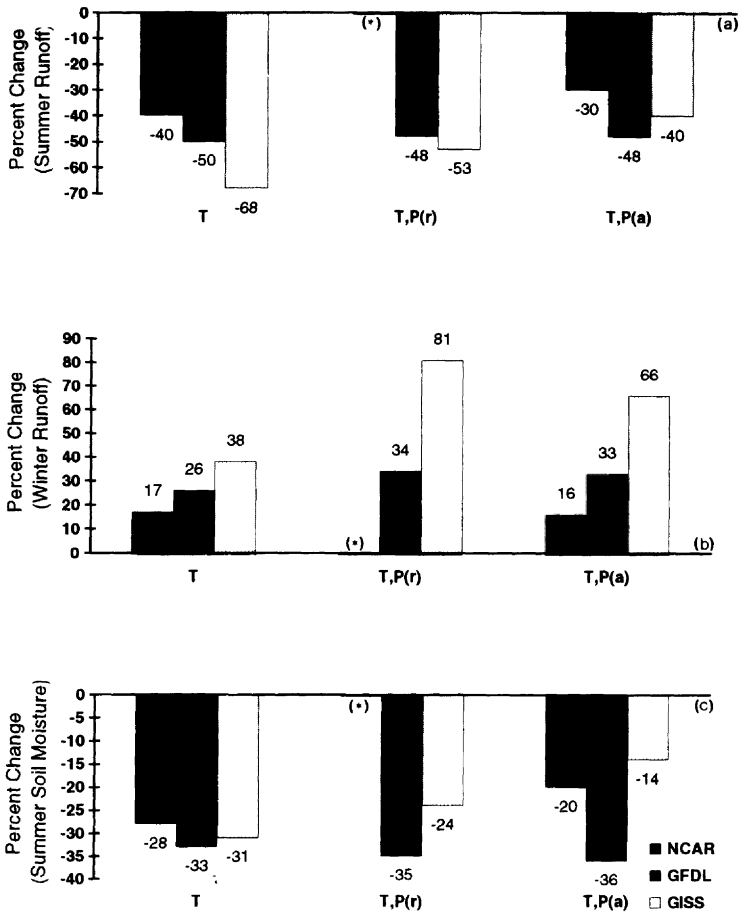
3.4.3.3 Groundwater

Groundwater is a very important source of freshwater for domestic, agricultural, and industrial use. In dry regions of the earth, it is the sole resource for water supply over large parts of the year, and, therefore, its sensitivity to climatic variations and changes is very critical. Groundwater resources play a significant role in the hydrologic cycle and water balance of a catchment since they act as a buffer, providing significant volumes of flow during years of low precipitation or absorbing large excess runoff during periods of high precipitation. Under a changing climate, long-term groundwater fluctuations resulting from significant perturbations of temperature and precipitation may be of critical importance for the water balance, and thus, for the water use.

Despite the importance of groundwater resources, only a few researchers have focused their studies on the consideration of the influence of climatic variation and change on groundwater (WMO, 1987). Regional studies of climatic change impacts on groundwater aquifers require changes in climatic and hydrologic inputs often averaged over long time periods, and additionally, hydrogeological and geological studies of the region. Since winter rainfall is the most important source of groundwater recharge, it is consequent that regions where winter rainfall is predicted by GCMs to increase, can expect an increase in groundwater recharge, and versa for regions where rainfall would decrease under global warming.

Thomsen (1989), in his review study of the effects of climate variability and change on groundwater in Europe, provides information concerning areas in Europe that can expect increase or decrease in groundwater recharge depending on the regional distribution of positive or negative changes in average precipitation evaluated by GCMs. Another experiment undertaken in Australia (Sharma, 1989) has shown that changes in rainfall produced amplified effects on groundwater recharge, a magnification effect which was greatly influenced by land use. For example, it was found that a $\pm 20\%$ change in rainfall resulted in a $\pm 30\%$ change in recharge beneath grassland, whereas the corresponding change in groundwater recharge beneath pine plantation was four times higher.

The effects of a reduction in groundwater recharge are many-sided: drying up of streams and lands, water shortages, salt water intrusions, as well as other consequent contamination effects. Basic research in the area of climatic change impacts on groundwater must be further developed because of its great and immediate practical importance for economic development.



LEGEND:
T only: temperature changes only
T,P(r): temperature and relative (percent) change in precipitation
T,P(a): temperature and absolute change in precipitation
(*): NCAR relative precipitation run not included

Fig. 3.4. Changes in seasonality of runoff and soil moisture in the Sacramento basin, California, due to climate change (adapted from Gleick, 1987b; Gleick, 1989). (a): The percent change in average summer runoff between the model historical run and eight GCM-derived doubled carbon dioxide scenarios. (b): The same plot for average winter runoff. (c): The same plot for average summer soil moisture.

3.4.3.4 Sea-level

Global warming will lead to a rise in world mean sea-level resulting from an actual increase in the volume of the world ocean due, both to the addition of water melting from glaciers and polar ice caps, as to the thermal expansion of the oceans (WMO/ICSU, 1990). Four factors mainly contribute to a change in sea-level (Oerlemans, 1989): Melting of glaciers, change in the net accumulation of

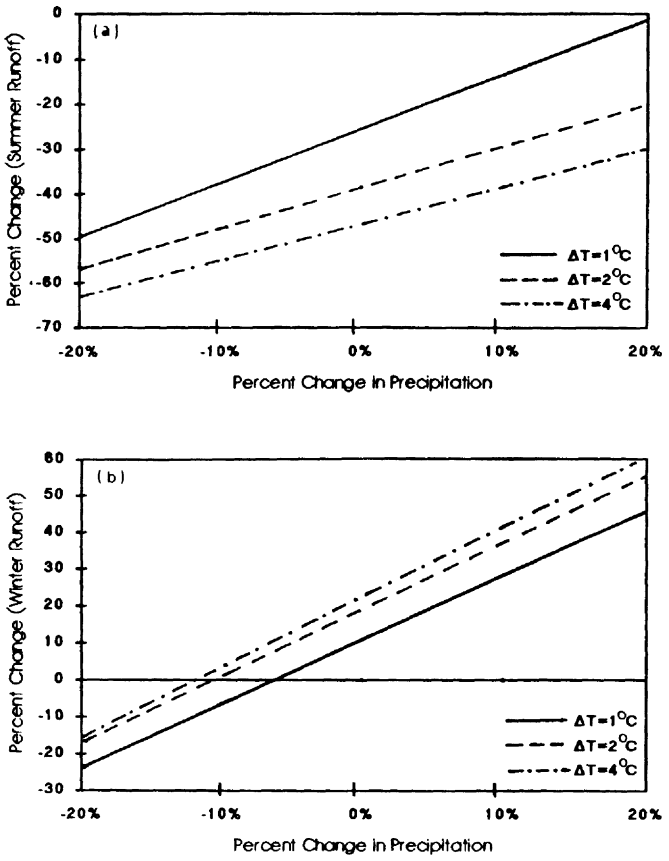


Fig. 3.5. Changes in seasonality of runoff in the Mesohora basin, central Greece (adapted from Mimikou et al., 1991a). (a): The percent change in average summer runoff between the model historical run and 15 hypothetical climate change scenarios. (b): The same plot for average winter runoff.

snow over the polar ice sheets of the Antatctic and Greenland, thermal expansion of ocean water, and possible disintegration of the West Antarctic Ice Sheet (WAIS). According to current research results, a global mean value of about 60 centimeters rise of mean sea-level, relative to the year 1980, can be expected by the end of next century (U.S. Environmental Protection Agency, 1983; Robin, 1986; Oerlemans, 1989). This is a rough estimate of the climate change effect on global mean sea-level. In order to make estimates for specific locations, local trends that can vary significantly have to be taken into account. The global mean value of estimate of future sea-level rise includes many uncertainties, a major part of which results from our poor knowledge of the reaction of the polar ice sheets to a changing climate. In fact, it is possible that global warming could result in net accumulation, rather than wastage, of the antarctic ice sheet with a consequent lowering of mean sea-level (Ambach, 1980; Oerlemans, 1982). Another important factor affecting the reliability of future protections of sea-level concerns the possible disintegration

of WAIS, for which there is not enough convincing evidence that it will have a substantial contribution to sea-level changes over the next one or two centuries (Van der Veen, 1985; 1986; Budd et al., 1986; Oerlemans, 1989).

Research, such as accurate simulations of the thermal expansion of the ocean by using transient-response models, studies of the WAIS, changes of the rate of snowfall in Antarctica, and investigation of the causes of sea-level change over the last 100 years, are currently under progress in order to improve our estimates of future sea-level rise.

The accelerated research work on sea-level rise partly reflects the fact that this climate change impact area is receiving a great deal of international attention because of its consequences on a variety of aspects. Such aspects include: property and life losses in coastal regions, aggravation of saline intrusion to groundwater aquifers affecting water withdrawals, changes in estuaries, and flooding along coastal rivers.

3.5 Climatic Change Impacts on Water Resource Planning and Management

3.5.1 Basic Considerations

The necessity of an assessment of the climatic change impacts on water resources becomes apparent when one realizes that, for the majority of the water resource systems which are to start operation early in next century, the planning has already started. Such an assessment would show the changes in the efficiency of the water resource systems under the new climatic conditions and indicate the necessary measures needed to mitigate the impacts of adverse changes or to take advantage of favorable changes. An assessment of the climate change impacts on water resources presupposes that quantitative estimates of changes in the long-term indices of the major climatic and hydrological variables can be converted into estimates of how the availability and the quality of freshwater resources may change. As discussed in the preceding sections, the state-of-the-art real time climate forecasting does not permit any specific estimate of future changes in temperature and precipitation to be considered as a deterministic forecast. For this reason, though it certainly would be preferable for water managers to have assessments based on reliable real-time forecasts, the assessment of climatic change impacts on water resources is usually made by using plausible climatic change scenarios. Therefore, the results of such assessments should be interpreted as a sensitivity analysis of water resource systems to alternative feasible climatic conditions rather than as predictions (Klemes, 1985).

The assessment of water-resource impacts is not only based on the perturbations to the hydrological variables, but also takes a use-oriented view focusing on water that is available for human exploitation; i.e. water at the right time and place and in the right quantity and quality (WMO, 1987). Changes in major hydrological variables will have serious impacts on the potential for water supply, entailing implications for domestic water use, irrigation, industrial water use and energy production. Since the key element of most large water supply projects is a storage reservoir, most of the studies are concentrated on the impacts on characteristics of a reservoir, and particularly, on the reliability of its performance

under changed climatic and hydrologic conditions. Moreover, the effects on the occurrence frequency of extreme flow events, such as floods and droughts, will have potentially devastating results, and cause problems mainly related to physical security and food production. A possible sea-level rise would affect navigation and the quantity and quality of freshwater resources in coastal zones. With respect to flooding, a combination of a sea-level rise and an increased upstream runoff could prove particularly dangerous for coastal lowlands. In addition to these primary water resource management activities, there are other critical issues that would be seriously affected by a climate change, such as a wide range of environmental services, fisheries, wetlands preservation, wildlife, conservation and recreation.

In the earliest literature most of the studies focused on more direct hydrological impacts, whereas there have been relatively few attempts to determine the sensitivity of water management activities to climatic changes. Even fewer were the studies that considered further socio-economic impacts of altered water resources. The very unusual climatic conditions and the severe drought and flooding problems that have characterized much of the post-1987 period in many places of the world, stimulated considerable scientific and public interest on the possible impacts of the greenhouse warming on the availability and quality of freshwater resources. Research efforts have been accelerated, and attempts are being made towards an integrated approach to study all aspects of the problem from basic physical issues to questions of environmental, political and socio-economic nature. An overview of some of the studies that may be described as related specifically to water resources, is given in Table 3.2.

In this table, each impact study is presented with the name(s) of author(s), the year of publication, and information concerning the specific water resource issue studied. It is notable that all studies concern mainly surface water resources, rivers, and lakes. In addition, one can find in the literature interesting review and state-of-the-art reports and papers (Novaky et al., 1985; Callaway and Currie, 1985; Beran, 1986; Glantz and Wigley, 1987; Parry et al., 1987; Warrick et al., 1987; Reibsame, 1988; Parry and Read, 1988; Adrahamson and Johansson, 1989; Duband et al., 1989; Nemec, 1989; Law, 1989; Arnell, 1989; Beran and Arnell, 1989; Kaczmarec and Kindler, 1989; Rhodes et al., 1991), which provide useful information, and a summary of activities undertaken by a variety of public and private agencies at either national or international level. For example, Arnell (1989) provides interesting information on the sensitivities of various water management activities in the UK to different dimensions of climatic change, such as changes in runoff, temperature increase, and sea-level rise. This information is summarized in Table 3.3 (from Arnell, 1989).

Finally, treatises concerning either specific aspects of the problem (Mac Cracken, 1990) or all aspects from modeling to socio-economic and political aspects (AAAS, 1990; Mintzer, 1992), have been published recently.

Despite the increasing research efforts that focus on the water-resource impact area, the coverage of this multiplex problem is not satisfactory yet since several issues must be further investigated. Such issues include: impacts on groundwater withdrawal and recharge, effects of sea-level rise on quantity and quality of coastal freshwater, and the establishment of strategies to mitigate impacts on water resource systems of national, and especially of shared international rivers. Many

TABLE 3.2
Overview of water-resource impact studies

Study	Water management issue
Schwarz (1977)	Reservoir reliability-water supply
Nemec and Schaake (1982)	Reservoir reliability-water supply-storage
Klemes (1982; 1985)	Reservoir reliability-water supply-storage
Henderson-Sellers (1987)	Reservoir - water quality
Fitzerald and Walsh (1987)	Reservoir - water quality
U.S. Environmental Protection Agency (1988)	Irrigation
Institute of Hydrology (1988)	Water quality
Lindh et al. (1989)	Irrigation
Coker et al. (1989)	Water resource management in coastal zones
Niemczynowicz (1989)	Water resource management in coastal zones
Dracup and Kendall (1990)	Sewage systems Floods and droughts
Byron and Goldman (1990)	Primary production of a lake
Peterson and Keller (1990)	Electric power production
Miller (1990)	Irrigation
Jacoby (1990)	Water quality
Cooper (1990)	Recreation-wildlife-aesthetics
Mimikou et al. (1991b)	Reservoir reliability-water and energy supply-storage
Allen et al. (1991)	Irrigation
Lettenmaier and Sheer (1991)	Reservoir reliability-water supply

other water resource problems will also arise related to impacts on environmental services, international-political security, and socio-economic development, which consequently require a lot more research.

3.5.2 Sensitivity of Some Critical Water Management Issues to Climatic Change

3.5.2.1 Reservoir reliability

Research conclusions strongly suggest that reservoirs designed and operated under current climatic conditions will be severely stressed by climatic changes (Schwarz, 1977; Nemec and Schaake, 1982; Dracup and Kendall, 1990; Lettenmaier and Sheer, 1991; Mimikou et al., 1991b). The relationship between the amount of water that can reliably be taken from a reservoir and the storage capacity of the reservoir, the so-called storage-yield relationship, is significantly affected by altered runoff due to changes in temperature, precipitation and evapotranspiration (Klemes, 1982; 1985; Law, 1989). Nemec and Schaake (1982) in their evaluation of reservoir reliability under conditions of altered runoff, concluded that increased

TABLE 3.3
An initial assessment of the sensitivities of water management issues in the UK to climatic change (*)

WATER MANAGEMENT ACTIVITY	DIMENSION OF CLIMATE CHANGE *HYDROLOGICAL*				
	Total runoff	Seasonal distribution	Extreme behavior	Increased temperature	Higher sea-level
Surface water supplies ("small storage")	*	**	**	**	–
Surface water supplies ("large storage")	**	*	**	*	–
Groundwater supplies	**	**	**	*	*
Agricultural water use	**	**	**	**	–
Flood management	–	*	**	–	**
Urban stormwater drainage	–	–	**	–	*
Reservoir safety	–	*	**	–	–
Drainage of low-lying lands	**	**	**	–	**
Power generation	**	**	*	–	–
Water quality	*	**	**	**	–
Effluent dilution	*	**	**	*	–
Inland navigation	*	**	–	–	–
Fisheries	–	**	*	**	–
Recreation	–	*	–	*	–
Saline intrusion into estuaries	*	*	–	–	**
Aquatic ecosystem management	–	**	**	**	–

(*) From *Arnell (1989)*

– little sensitivity
* some sensitivity
** very sensitive

storage may be needed to set off the effects of climatic changes, if reliable yields are to be maintained. Significant changes in the volume of storage required to maintain existing water yields, as greenhouse warming progresses, were also evaluated by Dracup and Kendall (1990) in their study of potential changes in floods and droughts due to climatic changes. More recently, Mimikou et al. (1991b) showed that dramatic increases of risk levels of the annual firm water and energy supply levels of reservoirs designed under current climatic conditions are exhibited under plausible alternative scenarios of climate change. Moreover, they concluded that radical increases of reservoir storage volume are required in order to maintain firm water and energy yields at tolerable risk levels. The effect of temperature increases on both risk level and storage volume appeared to be marginal as compared to the effect of precipitation reduction. In Figure 3.6 (adapted from Mimikou et al., 1991b) one can see the risk isolines for various combinations of annual firm water supply and energy production from a multi-purposed reservoir in Thessaly, Greece, under current climatic conditions (a), and the significant variation of design risk associated with the guaranteed annual production, under plausible changes in temperature and precipitation (b). Moreover, Figure 3.7 (adapted from Mimikou et al., 1991b) presents the risk of guaranteed annual production from another reservoir in central Greece as a function of storage volume increase under various climatic change scenarios. As can be seen from this figure, significant increases of

storage volumes, some times greater than 100%, are required in order to maintain production at tolerable risk levels.

3.5.2.2 Water supply for urban and industrial use

Urban and industrial activities need water to meet various requirements of population and plants, and besides that, to evacuate urban and industrial wastes. Such urban and industrial water supplies are influenced by many factors, i.e. climatic, environmental, social and economic. Current water resource planning and decision making for future water supply are based on "standard design" conditions, which assume stationarity of climate over time. Such an assumption is no longer believed to be realistic, since climatic conditions are expected to change due to the greenhouse effect. Altered runoff patterns, and more particularly changes in extreme hydrological conditions of both water shortages and water excess situations due to climatic changes, will affect urban and industrial water supplies and influence sewage disposal systems. In addition, possible sea-level rise may result in additional effects on both surface and groundwater aquifers of coastal urban areas. Moreover, the previously discussed extreme sensitivity of the storage-yield-risk relationships of reservoirs that play a key role in coping with climatic and hydrological variability, clearly indicates that certain measures have to be taken to mitigate impacts from adverse effects of climatic changes on planning and management of water supply systems.

An increase of mean air temperature is likely to result in significant changes in urban hydrologic cycle, which contain fewer storage components compared with the rural basins, and may respond with further decreases of storage capacity and runoff. Exceptionally dry climatic conditions and persistent low flows exhibit thus amplified effects on urban water supply and sewage disposal. Increasing demands, combined with decreasing yield and storage of reservoirs, can lead to serious implications for urban water supply and sewage disposal systems, associated with environmental social and economic dislocations. Expectations of long-term drought conditions must be taken into account in planning and designing urban water resource systems (Parry and Read, 1988). Also impacts on urban sewage and stormwater networks can be induced by increased rain-to-snow events during winter, which may bring significant flooding problems and reduce the efficiency of sewage treatment plants, thus causing significant losses and environmental damages. A number of measures to mitigate the above negative impacts can be implemented, e.g. detention facilities and infiltration basins located upstream of the urban catchment, which could relieve the urban networks and leave some extra storage capacity for future flooding (Niemczynowicz, 1989).

On the other hand, industrial water use is a significant component of urban water cycle, since industrial water demand is a large portion of national demand in many industrialized countries. The major part of the industrial water is used for cooling purposes, whereas the rest is needed for other purposes, such as boiling, processing, air conditioning, and cleaning. Therefore, an increase of mean air temperature will increase evaporative losses, lower efficiencies in cooling, and increase use of water for air conditioning. Measures to mitigate these negative effects can be implemented, e.g. decreasing industrial raw demands and increasing

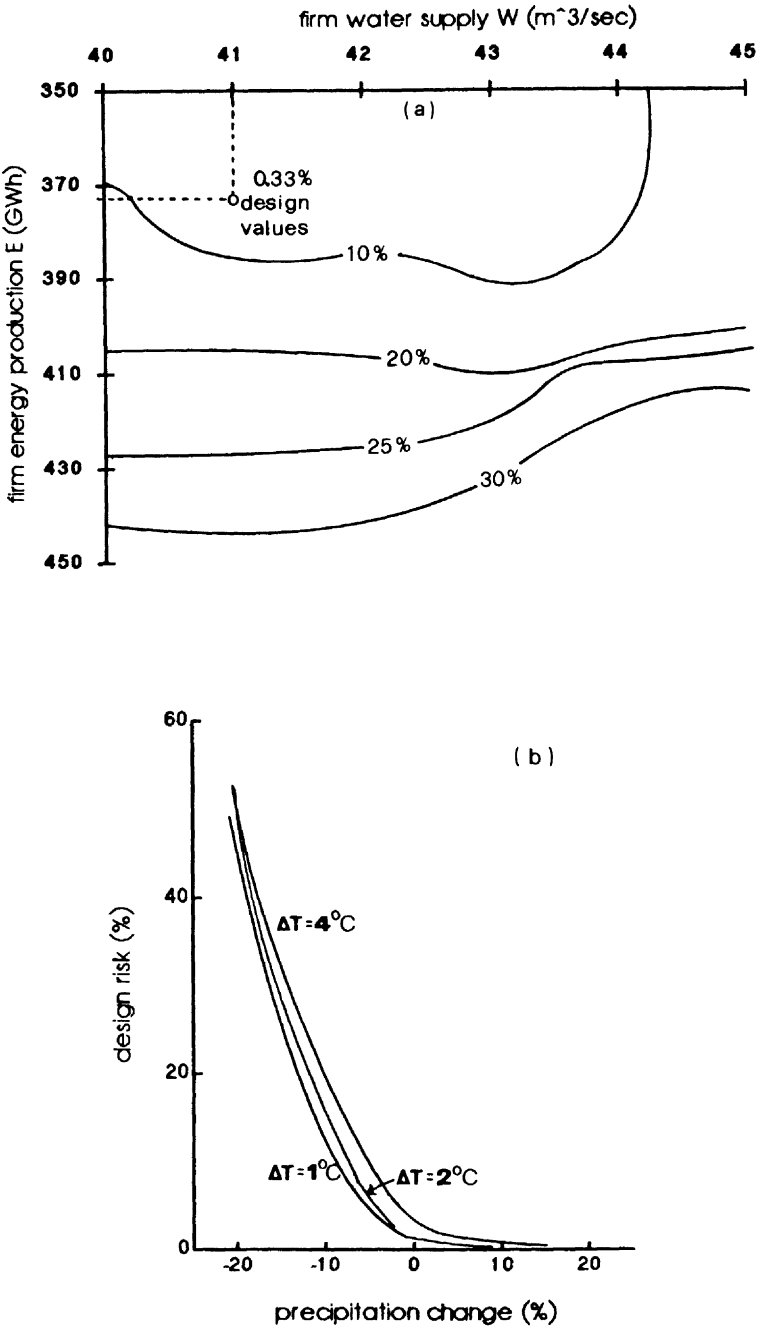


Fig. 3.6. Risk analysis of annual water and energy firm production from the Mouzaki reservoir, Thessaly, Greece (adapted from Mimikou et al., 1991b). (a): Isolines of risk under current climatic conditions. (b): Variation of design risk under plausible changes in temperature and precipitation.

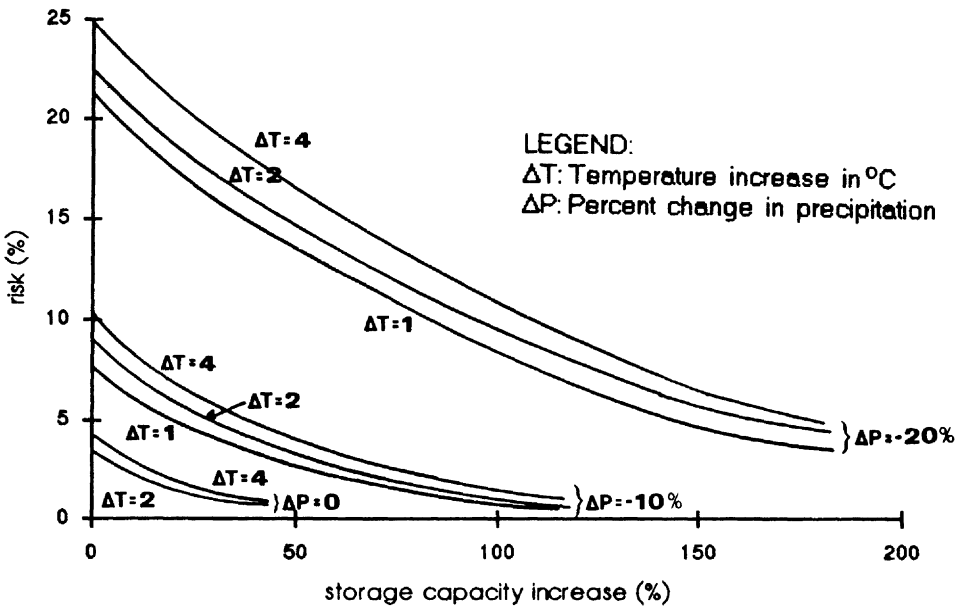


Fig. 3.7. Risk of guaranteed annual production as a function of storage volume increase for the Mesohora reservoir, central Greece, under various climatic scenarios (adapted from Mimikou et al., 1991b).

the efficiency of industrial water use (Parry and Read, 1988; Kaczmarek and Kindler, 1989).

3.5.2.3 Irrigation

Increases in atmospheric CO_2 concentrations affect crop production directly by changing the plant processes of photosynthesis and transpiration, and indirectly, by modifying the climate (Morison, 1989; Allen et al., 1991). CO_2-induced changes in climatic variables, such as temperature, net radiation, windiness and relative humidity, and consequent modifications of local hydrologic cycles, will affect both rainfed and irrigated agriculture and cause major impacts on agricultural activities in general. Higher air temperature results in increases in evaporative demands and changes in precipitation patterns, which may result in increasing irrigation requirements. Thus, the availability of water for irrigation becomes increasingly important to ensure the adequacy of water supplies for crop production and food supply, especially for countries in semi-arid and arid zones (Hargreaves, 1981; Parry et al., 1987; Nemec, 1989). The assessment of regional impacts on irrigation water is obviously a subset of the assessment of impacts on agriculture in general (Nemec, 1989). The latter is a more complex issue including other aspects of the problem as well, which are described by Warrick et al. (1987) as:

* Estimation of possible spatial shifts in cropping patterns;

* Analysis of the agricultural sector, which focuses on the estimation of the range of impacts within and between agricultural regions, and emphasizes the positive and negative feed-back mechanism;

* Assessment of impacts on agriculture of past climatic variations based on historical case studies.

Regional studies of impacts on irrigation are mainly based on sensitivity analysis of various irrigation features to alternative climate scenarios, which are either hypothetical or outputs from GCMs. For example, Fitzerald and Walsh (1987) studied the sensitivity of irrigated areas in New South Wales on altered availability of runoff. They found that a 20% decrease of rainfall, combined with an increase of evapotranspiration of 15%, resulted in the reduction of the irrigated area by 75%. A more sophisticated case study commissioned by FAO (Nemec, 1989) with the British Institute of Hydrology (1988) evaluated the impacts of plausible alternative climate scenarios, not only on irrigation areas, but also on the probability of failure in meeting irrigation requirements by using actual data from the basin of Malibamatsamo in Lesotho. Under conditions of the $2\times CO_2$ climate change scenario, the total yearly irrigation demand was increased by about 60% compared with the current value. The combination of increased irrigation demands and of decreased inflows in the reservoir used for water supply resulted in a rise of the frequency of failure in satisfying irrigation needs, which from the current 1 in 5 year failure became 1 in 2 years with 10% less precipitation. Peterson and Keller (1990) evaluated the implications of plausible climate change scenarios for the net irrigation requirement, the depth of irrigation needed to optimize production, and the future irrigated percent of cultivated land. They concluded that greenhouse warming will have major impacts especially in dry regions, where it will be very difficult to maintain current levels of irrigation. They found that a 3.0°C temperature increase and a 10% decrease in precipitation will result in a 26% increase of average net irrigation requirement and a 30% decrease of cultivated areas in the western United States. More recently, Allen et al. (1991) evaluated the impacts of various GCM-derived $2\times CO_2$ scenarios on irrigation-water requirements in the Great Plains region. They concluded that increases in irrigation requirements were caused primarily by increases in atmospheric evaporative demands, changes in precipitation pattern and lengths of crop life cycles, which were shifted to months with lower solar radiation.

Measures to offset the potential impacts of climate change on water use and management in agriculture must be taken, an example of which would be to increase overall irrigation efficiency of water use, and to develop new supplies.

3.5.2.4 Energy production

The energy sector will be strongly influenced by a global climatic change. The energy production systems, and especially hydroelectricity production, will be affected by changes in temperature, precipitation, evaporation and runoff patterns. Changes will also be induced in temperature-sensitive energy demands. Finally, necessary policies to reduce greenhouse gas emissions will severely limit future

energy policy options, especially the ones related to fossil fuels (Abrahamson and Johnson, 1989; Mac Cracken, 1990). Existing systems for electric power generation, transmission, and distribution facilities, have been designed and operated under the assumption of climatic stationarity, which includes the possibility of extreme variations (flood, droughts) based on past experience. Thus, any further modification of the climate due to possible global warming, which will be beyond the conditions that the earth has experienced in the past, could result in considerable implications for the energy domain (Duband et al., 1989).

Particularly, the hydroelectricity production, which is vulnerable to existing climatic variability, would be severely affected by a possible climate change as well as by consequent alterations in runoff and water availability. Such changes may lead to the need for expensive replacement of power, and cause serious implications for regions where electricity production is mainly based on hydropower (Gleick, 1988; 1990). Miller (1990) examined in detail the significant impacts of altered water availability, and more particularly of seasonal water availability, on the electric power system in the United States. She concluded that under such changed conditions the reliability of the system is severely threatened. The extreme sensitivity of the reliability of guaranteed energy supply levels from large storage reservoirs to temperature increases and precipitation changes has been also identified by Mimikou et al. (1991b), (see, for example, Figure 3.6). Thermal power plants are also affected, mainly due to limitations in the availability of cooling water. Alternative cooling technologies should be adopted to offset impacts from future water shortages.

To mitigate possible adverse climate change impacts on the energy domain and to limit greenhouse gas emission additionally, certain measures should be adopted with high priority. Such measures could include the application of more efficient energy conservation technologies, and the use of other alternative energy sources, such as solar and other renewables.

3.5.2.5 Water quality

Assessment of possible climatic change impacts on water quality and timely development of feasible strategies in order to maintain satisfactory water quality, are research areas of apparent and primary importance. Research has shown that rising temperature and increasing evaporation accompanied by reductions in streamflows could lead to considerable implications for water quality. For example, Henderson-Sellers (1987), investigating the effects of greenhouse warming on reservoir water quality, concluded that increased evaporative losses could result in significant enhancement of nutrient loading and acceleration of eutrophication. The U.S. Environmental Protection Agency (1988), and Jacoby (1990) examined the issue of identifying how to maintain satisfactory water quality under conditions of greenhouse warming. It was concluded that expensive water quality protection measures may be required to mitigate negative effects of possible reductions in freshwater supply for diluting waste water or recharging aquifers. Moreover, salt water penetration into coastal aquifers, due to possible sea-level rise, has been identified as an important potential impact of climate change on water quality, which threatens both water supply facilities and natural ecosystems in coastal regions (Lindh et al., 1989;

Coker et al., 1989; Jacoby, 1990). Possible climatic change impacts combined with impacts due to increasing urbanization and population density (Lettenmaier et al., 1991) could lead to serious implications for future water quality.

3.6 Future Research Directions

Research during the past few years has highlighted the potential effects of climatic changes on hydrology and water resources. Despite the progress in modeling of climatic changes caused by an increased concentration of atmospheric greenhouse gases, and despite the improvements in our understanding of appropriate methods for addressing regional hydrologic and water resource questions, there is still much work to be done. Future research directions, where further effort needs to be placed, include the following:

- Climate models must be improved in order to help with the derivation of realistic climate scenarios, and provide detailed information on the spatial and short time-scale characteristics of climatic changes. Detailed spatial information and information on changes in short time-scales, such as days or weeks rather than annual or monthly averages, are needed to be incorporated into regional hydrologic studies in order to improve their efficiency and usefulness for water resource planners.
- Results from regional studies indicate that it may be difficult to generalize impact assessments even within a relatively small area (Arnell, 1989). It is really important therefore to understand the current relationship between climate and hydrological response, and investigate the interrelationships of all basin components, before attempting to assess possible changes. Thus, realistic regional hydrologic assessments must be able to incorporate important regional complexities, such as snowmelt characteristics, vegetation changes, soil characteristics, geology, natural or artificial storage, etc.
- Assessment of impacts on water resource planning and management should be directed towards the needs of the user. Nevertheless, related activities may vary depending on the nature of the water management infrastructure and according to the specific needs of different types of water resource systems (WMO, 1987). Therefore, there is a need of detailed case studies based on typical water management schemes, which, though restricted by our present limited knowledge concerning potential changes in climatic inputs, will help to improve our understanding of the sensitivities of the water resource systems and indicate the appropriate measures necessary to mitigate the effects of change.

A search of the literature helps to identify wider frameworks of recommendations and courses of actions concerning future activities of research on climate change. As for example, the framework of activities that are included in a WMO report (WMO, 1987) on the sensitivities of water resource systems to climatic change and variability, where four categories of fruitful courses of action are distinguished as follows: *State-of-the-art projects*, such as completion of national and international data bases with data sets, information, and analyses concerning all aspects of the problem, production of realistic climate scenarios based on appropriate historical data sets, testing and validation of current models, and other

activities which are within our present scientific and technological capabilities; *Soft research projects* to create realistic climatic and hydrological scenarios, and improve modeling of process mechanisms in order to derive simulated results as provisional substitutes for the unknown real ones; *Hard research projects*, such as estimation of sub-grid scale climate characteristics from GCMs, examination of water and energy balances on regional and global scales, improvement of areal precipitation and vapor flux measurements over land, and other activities attempting to give answers to basic problems of our state-of-the-art knowledge, test new theories, design new field experiments, and improve existing monitoring and data collection systems; and finally, *Dissemination of information* through education, training, workshops, and symposia.

From the point of view of the water user, the actual impacts of climatic changes on water resources will be mostly determined by the ability of water management to respond timely to changing external conditions. With our present state of knowledge, even if it is assumed that global warming will take place on the scale and over the time period predicted by different GCMs, it cannot be translated into reliable prediction of specific changes in climatic and hydrological variables over a watershed for the design and operation of specific water projects. Thus, a major question is whether it is sensible to take action now despite uncertainty, or to wait until climate change and water resource impacts are better defined through the ongoing research effort. At present, no categorical answer can be given to this question, although the implementation of certain preventive measures can be recommended. Such measures could include increasing flexibility, adaptability and robustness of new designed water resource systems, decreasing vulnerabilities of existing water systems, and taking early advantage of water conserving technologies. Nevertheless, drastic steps in water resource planning and management to accommodate climate change will only be taken if planners, managers, and decision makers can be convinced that greenhouse warming is a real problem requiring immediate actions.

References

Abrahamson, D., and Johansson, T.B. (1989). 'Global climatic change: Implications for energy policy'. In: Conference on Climate and Water, The Publications of the Academy of Finland, Vol. 2, pp. 245 - 274.

Allen, R.G., Gichuki, F.N., and Rosenzweig, C. (1991). 'CO_2 - induced climatic changes and irrigation - water requirements'. Journal of Water Resources Planning and Management (ASCE), Vol. 117, No 2: 157 - 177.

American Association for the Advancement of Sciences, (AAAS) (1990). 'Climate Change and U.S. Water Resources' (Waggoner, P.E., Ed.), Wiley Series in Climate and the Biosphere, John Wiley and Sons, New York, 496 pp.

Ambach, W. (1980). 'Anstieg der CO_2 - konzentration in der atmosphare und klimaanderung: Mogliche answirungen auf den Gronlandischer eisschild'. Water and Leben, 32: 135 - 142.

Arnell, N.W. (1989). 'The potential effects of climatic change on water resources management in the UK'. Paper II 11, Proc. Int. Seminar on Climatic Fluctuations and Water Management, Cairo, Egypt, (11 - 14 Dec., 1989).

Arnell, N.W., and Reynard, N. (1989). 'Estimating the impacts of climatic change on river flows: Some examples from Britain'. In: Conference on Climate and Water, The Publications of the Academy of Finland, Vol. 1, pp. 426-436.

Bach, W. (1988). 'Development of climatic scenarios: A. From general circulation models' In: The Impact of Climatic Variations on Agriculture, (Parry et al., Eds.), Kluwer Acad. Publ., Dordrecht, Vol. 1, pp. 125 - 157.

Becker, A., and Nemec, J. (1987). 'Macroscale hydrologic models in support to climate research'. In: The Influence of Climate Change and Climatic Variability on the Hydrologic Regime and Water Resources, (Solomon et al., Eds.), IAHS Publ.. 168, pp. 431 - 445.

Beran, M.A. (1986). 'The water resource impact of future climate change and variability'. In: Effects of Changes in Stratospheric Ozone and Global Climate, Vol. 1, Overview, (Titus, J.G., Ed.), United States Environmental Protection Agency, Washington D.C., Vol. 1, pp. 299 - 330.

Beran, M.A., and Arnell, N.W. (1989). 'Effect of Climatic Change on Quantitative Aspects of UK Water Resources. Report to DOE Institute of Hydrology, Wallingford, UK.

Berger, A. (1989). 'Expected climatic change'. (Report on Session III). In: Carbon Dioxide and Other Greenhouse Gases: Climatic and Associated Impacts. (Fantechi, R., and Ghazi, A., Eds.), CEC - Kluwer Academic Publishers, Dordrecht, pp. 135 - 137.

Budd, W.F., Mc Innes, B.J., Jenssen, D., and Smith, I.N. (1986). 'Modelling the response of the West Antarctic Ice Sheet to a climatic warming'. In: Dynamics of the West Antarctic Ice Sheet. (Van der Veen, C. J., and Oerlemans, J., Eds.), Reidel, Dordrecht, pp. 321 - 358.

Budyko, M.I. (1989). 'Climatic conditions of the future', In: Conference on Climate and Water. The Publications of the Academy of Finland, Vol. 1, pp. 9 - 25.

Bultot, F., and Gellens, D. (1989). 'Simulation of the impact of CO_2 atmospheric doubling on precipitation and evapotranspiration - study of the sensitivity to various hypotheses'. In: Conference on Climate and Water. The Publications of the Academy of Finland, Vol. 1, pp. 73 - 92.

Bultot, F., Dubriez, G.L., and Gellens, D. (1988a). 'Estimated annual regime of energy-balance components, evapotranspiration, and soil moisture for a drainage basin in the case of a CO_2 doubling'. Climatic Change, 12: 39 - 56.

Bultot, F., Coppens, A., Dubriez, G.L., Gellens, D., and Menlenberghs, F. (1988b). 'Repercussions of a CO_2 doubling on the water cycle and on the water balance - A case study for Belgium'. J. Hydrol., 99: 319 - 347.

Byron, E.R., and Goldman, C.R. (1990). 'The potential effects of global warming on the primary productivity of a subalpine lake'. Water Resources Bulletin, Vol. 26, No 6: 983 - 989.

Callaway, J.M., and Currie, J.W. (1985). 'State-of-the-art report: Water resources'. In: State-of-the-Art Report of the Impact of CO_2 - induced Climatic Change. U.S. Dept. of Energy, Washington D.C. chapter 7.

Chagnon, S.A., and Huff, F.A. (1992). 'Potential effects of changed climates on heavy rainfall frequencies in the Midwest'. Water Resources Bulletin, Vol. 27, No 5: 753 - 759.

Clark, W.C. (1982). 'Carbon Dioxide Review: 1982'. Oxford University Press, New York.

Cohen, S.J. (1986). 'Impacts of CO_2 - induced climatic change on water resources in the Great Lakes Basin'. Climatic Change, 8: 135 - 153.

Cohen, S.J. (1987). 'Projected increases in municipal water use in the Great Lakes due to CO_2 - induced climatic change'. Water Resources Bulletin, 23: 91 - 101.

Coker, A.M., Thompson, P.M., Smith, D.I., and Penning-Rowsell, E.C. (1989). 'The impact of climate change on coastal zone management in Britain: a preliminary analysis'. In:

Conference on Climate and Water. The Publications of the Academy of Finland., Vol 2, pp. 148 - 160.

Commission of the European Communities (1989). 'Carbon Dioxide and Other Greenhouse Gases: Climatic and Associated Impacts'. (Fantechi, R., and Ghazi, A., Eds). Kluwer Academic Publishers, Dordrecht, 279 pp.

Conference on Climate and Water (Helsinki, 14 - 15 Sept., 1989). The Publications of the Academy of Finland, September 1989, Vol. 1, 520 pp., and Vol. 2, 393 pp.

Cooley, K.R. (1990). 'Effects of CO_2 - induced climatic changes on snowpack and streamflow'. Hydrol. Sc. J., 35, 5: 511 - 522.

Cooper, C.F. (1990). 'Recreation, and wildlife'. In: Climate Change and U.S. Water Resources. (Waggoner, P.E., Ed.), John Wiley and Sons, New York, pp. 329 - 339.

Dickinson, R.E. (1982). 'Modelling climate changes due to carbon dioxide increases'. In: Carbon Dioxide Review: 1982. (Clark, W.C., Ed.), Oxford University Press, New York, pp. 101 - 133.

Dooge, J.C.I. (1989). 'Effects of CO_2 increases on hydrology and water resources'. In: Carbon Dioxide and Other Greenhouse Gases: Climatic and Associated Impacts. CEC - Kluwer Academic Publishers, Dordrecht, pp. 204 - 213.

Dracup, J.A., and Kendall, D.R. (1990). 'Floods and droughts'. In: Climate Change and U.S. Water Resources. (Waggoner, P.E., Ed.), John Wiley and Sons, New York, pp. 243-267.

Duband, D., Rabut, P., Clavel, P., and Leveugle, J. (1989). 'The impact of climate on the operation of the French electric system'. In: Conference on Climate and Water. The Publications of the Academy of Finland, Vol. 2, pp. 275 - 289.

Fitzerald, B.M., and Walsh, M.A. (1987). 'Significance of greenhouse changes to irrigation water supplies in New South Wales - A case study of the Severn Valley'. Report, Dept of Water Resources, N. Sydney, N.S.W., Aust., 16 pp.

Flaschka, I.M., Stockton, C.W., and Boggers, W.R. (1987). 'Climatic variation and surface water resources in the Great Basin region'. Water Resources Bulletin, 23 (1): 47 - 57.

Glantz, M.H., and Wigley, T.M.L. (1987). 'Climatic variations and their effects on water resources'. In: Resources and World Development. (McLaren, D.J., and Skinner, B.J., Eds.), John Wiley and Sons Ltd, pp. 625 - 641.

Gleick, P.H. (1986a). 'Regional Water Availability and Global Climatic Change: The Hydrologic Consequences of Increases in Atmospheric CO_2 and Other Trace Gases'. Phd Dissertation, Univ. of California, Berkley, California, 615 pp.

Gleick, P.H. (1986b). 'Methods for evaluating the regional hydrologic impacts of global climatic changes'. J. Hydrol., 88: 97 - 116.

Gleick, P.H. (1987a). 'The development and testing of a water balance model for climate impacts assessment: Modelling the Sacramento Basin'. Water Resour. Res., 23: 1049-1061.

Gleick, P.H. (1987b). 'Regional hydrologic consequences of increases in atmospheric CO_2 and other trace gases'. Climatic Change, 10: 137 - 161.

Gleick, P.H. (1988). 'Climatic changes and California: Past, present, and future vulnerabilities'. In: Societal Responses to Regional Climatic Change: Forecasting by Analogy. (Glantz, M.H., Ed.), Westview Press, Boulder, Colorado, pp. 307 - 327.

Gleick, P.H. (1989). 'Climate change, hydrology, and water resources'. Reviews of Geophysics, 27, 3: 329 - 344.

Gleick, P.H. (1990). 'Vulnerability of water systems'. In: Climate Change and U.S. Water Resources. (Waggoner, P.E., Ed), John Wiley and Sons, New York, pp. 223 - 240.

Hansen, J., and Lebedeff, S. (1987).'Global trends of measured surface air temperatures'. J. Geophys. Res., Vol. 92: 13,345 - 13,372.

Hansen, J.E., Wang, W.-C., and Lacis, A.A. (1978). 'Mount Agung eruption provides a test of a global climatic perturbation'. Science, Vol. 199: 1065 - 1068.

Hansen, J., Russell, G., Rind, D., Stone, P., Lacis, A., Lebedeff, S., Ruedy, R., and Travis, L. (1983). 'Efficient three-dimensional global models for climate studies: Models I and II'. Mon. Weather Rev., 111 (4): 609 - 662.

Hansen, J.E., Lacis, A., Rind, D., Russell, G., Stone, P., Fung, I., Ruedy, R., and Lerner, J. (1984). 'Climatic sensitivity: Analysis of feedback mechanisms'. In: Climate Processes and Climate Sensitivity. Geophys. Monogr. Ser. (Hansen, J.E., and Takahashi, T., Eds.), AGU, Washington D.C., Vol. 29, pp. 130 - 163.

Hansen, J., Fung, I., Lacis, A., Rind, D., Lebedeff, S., Ruedy, R., Russell, G., and Stone, P. (1988). 'Global climate changes at forecast by Goddard Institute for Space Studies three-dimensional model'. J. Geophys. Res., 93: 9341 - 9364.

Hargreaves, G.H. (1981). 'Water requirements and man-induced climate change'. Proc. Am. Soc. Civ. Eng., 107 (IR3): 247 - 255.

Henderson - Sellers, B. (1987). 'The impact of increasing atmospheric carbon dioxide concentrations upon reservoir water quality'. In: The Influence of Climate Change and Climatic Variability on the Hydrologic Regime and Water Resources. (Solomon, S.I., Beran, M., and Hogg, W., Eds.), IAHS Publ., 168, pp. 571 - 576.

Houghton, J.T., Jenkins, G.J., and Ephraums, J.J. (Eds.) (1990). 'Climate change: The IPCC Scientific Assessment (Report of Working Group 1)', Cambridge Univ. Press, New York, 416 pp.

Institute of Hydrology (1988). 'Effects of Climatic Change on Water Resources for Irrigation, an Example from Lesotho'. Report prepared for FAO, Wallingford, UK.

Jacoby, H.D. (1990). 'Water quality'. In: Climate Change and U.S. Water Resources. (Waggoner, P.E., Ed.), John Wiley and Sons, New York, pp. 307 - 328.

Jager, J., and Ferguson, H. (Eds.) (1991). 'Climate Change. Science, Impacts and Policy". Proc. of the 2nd World Climate Conference, Cambridge Univ. Press, New York, 592 pp.

James, L.D., and Burges, S.J. (1982). 'Selection, calibration and testing of hydrologic models'. In: Hydrologic Modelling of Small Watersheds. (Haan, C.T., Johnson, H.P., and Brakensiek, D.L., Eds.), American Society of Agricultural Engineers, St. Joseph, Mich., pp. 437 - 472.

Jurak, D. (1989). 'Effect of climate change on evaporation and water temperature'. In: Conference on Climate and Water. The Publications of the Academy of Finland, Vol. 1, pp. 138 - 145.

Kaczmarek, Z., and Kindler, J. (1989). 'The impacts of climate variability and change on urban and industrial water supply and wastewater disposal'. In: Conference on Climate and Water, The Publications of the Academy of Finland, Vol. 2, pp. 161 - 176.

Kates, R.W., Ausubel, J.H., and Berberian, M. (Eds) (1985). 'Climate Impact Assessments'. SCOPE Rep., 27, 625 pp.

Klemes, V. (1982). 'Effect of Hypothetical Climatic Change on Reliability of Water Supply from Reservoirs'. Report, National Hydrology Research Institute, Environment Canada, Ottawa, 57 pp.

Klemes, V. (1985). 'Sensitivity of Water Resource Systems to Climate Variations'. World Clim. Appl. Programme WCP - 98, World Meteorol. Organ., Geneva, 17 pp.

Kuchment, L.S., Motovilov, Yu.G., and Startseva, Z.P. (1989). 'Sensitivity of evapotranspiration and soil moisture to possible climatic changes'. In: Conference on Climate and Water. The Publications of the Academy of Finland, Vol. 1, pp. 236 - 251.

Law, F.M. (1989). 'Identifying the climate-sensitive segment of British reservoir yield'. In: Conference on Climate and Water. The Publications of the Academy of Finland, Vol. 2, pp. 177 - 190.

Lettenmaier, D.P., and Gan, T.Y. (1990). 'Hydrologic sensitivities of the Sacramento - San Joaquim River Basin, California, to global warming'. Water Resour. Res., Vol. 26, No 1: 69 - 86.

Lettenmaier, D.P., and Sheer, D.P. (1991). 'Climatic sensitivity of California water resources'. J. of Water Resources, Planning, and Management, Vol. 117, No 1: 108 - 125.

Lettenmaier, D.P., Hooper, E.R., Wagoner, C., and Faris, K.B. (1991). 'Trends in stream quality in the Continental United States, 1978 - 1987'. Water Resour. Res., Vol. 27, No 3: 327 - 339.

Lindh, G., Hanson, H., and Larson, M. (1989). 'Impact of sea level rise on coastal zone management in Southern Sweden'. In: Conference on Climate and Water. The Publication of the Academy of Finland. Vol. 2, pp. 128 - 147.

Linsley, R.K. (1982). 'Rainfall - runoff models. An overview'. In: Rainfall - Runoff Relationship (Singh, V.S., Ed.), Water Resources Publications, Littleton, Colorado, pp. 3 - 22.

Mac Cracken, M.C. (Ed.) (1990). 'Energy and Climate Change'. Official report of the Department of Energy Multi-Laboratory Climate Change Committee, Lewis Publishers, Inc., 161 pp.

Mac Cracken, M.C., Ellis, T.S., Ellsaesser, H.W., Luther, F.M., and Potter, G.L. (1981). 'The livermore Statistical Dynamical Climate Model', Lawrence Livermore National Laboratory, UCID - 19060, Livermore, California.

Mac Cracken, M.C., Hecht, A.D., Budyko, M.I., and Izrael, Y.A. (1990). 'Prospects for future climate: A Special US/USSR Report on Climate and Climate Change'. Lewis Publishers, Inc., 300 pp.

Manabe, S. (1982). 'Simulation of climate by General Circulation Models with hydrologic cycles'. In: Land Surface Processes in Atmospheric General Circulation Models. (Eagleson, P.S., Ed.), Cambridge Univ. Press, pp. 19 - 66.

Manabe, S., and Wetherald, R.I. (1967). 'Thermal equilibrium of the atmosphere with a given distribution of relative humidity'. J. Atmos. Sci., Vol. 24: 241 - 259.

Manabe, S., and Wetherald, R.I. (1975). 'The effect of doubling the CO_2 concentration on the climate of a General Circulation Model'. J. Atmos. Sci., Vol. 32: 3 - 15.

Manabe, S., and Wetherald, R.I. (1985). 'CO_2 and hydrology'. Adv. Geophys., 28A: 131 - 157.

Manabe, S., and Wetherald, R.I. (1986). 'Reduction in summer soil wetness induced by an increase in atmospheric carbon dioxide'. Science, 232: 626 -628.

Manabe, S., and Wetherald, R.I. (1987). 'Large scale changes in soil wetness induced by an increase in atmospheric carbon dioxide'. J. Atmos. Sci., Vol. 44: 1211 - 1253.

Mather, J.R. (1978). 'The Climatic Water Budget in Environmental Analysis'. D.C. Heath CO., Lexington Books, Lexington, Massachusetts.

Mearns, L.P., Gleick, P.H., and Schneider, S.H. (1990). 'Climate forecasting'. In: Climate Change and U.S. Water Resources. (Waggoner, P.E., Ed.), John Wiley and Sons, New York, pp. 87 - 137.

Meehl, G.A. (1984). 'Modelling the Earth's climate'. Climatic Change, Vol. 6: 259 - 286.

Miller, D.H. (1977). 'Water at the Surface of the Earth: An Introduction to Ecosystem Hydrodynamics', Academic Press, New York

Miller, K.A. (1990). 'Water, electricity, and institutional innovation'. In: Climate Change and the U.S. Water Resources. (Waggoner, P.E., Ed), John Wiley and Sons, New York, pp. 367 - 393.

Mimikou, M.A., and Kouvopoulos, Y.S. (1991). 'Regional climate change impacts: I. Impacts on water resources'. Hydrol. Sc. J., 36,3,6: 247 - 258.

Mimikou, M., Kouvopoulos, Y, Cavadias, G., and Vayiannos, N. (1991a). 'Regional hydrological effects of climate change'. J. Hydrol., 123: 119 - 146.

Mimikou, M.A., Hadjissavva, P.S., Kouvopoulos, Y.S., and Afrateos, H. (1991b). 'Regional climate change impacts: II. Impacts on water management works'. Hydrol. Sc. J., 36,3,6: 259 - 270.

Mimikou, M.A., Hadjissavva, P.S., Kouvopoulos, Y.S., and Anagnostou, E.N. (1992). 'The influence of basin aridity on the efficiency of runoff predicting models'. Nordic Hydrology, 23: 105 - 120.

Mintzer, I. (Ed.) (1992). "Confronting Climate Change. Risks, Implications and Responses". Cambridge Univ. Press, New York, 324 pp.

Mitchell, J.F.B., and Warrilow, D.A. (1987). 'Summer dryness in northern mid-latitudes due to increased CO_2'. Nature, 330: 238 - 240.

Morison, J.I.L. (1989). 'Plant growth in increased atmospheric CO_2'. In: Carbon Dioxide and Other Greenhouse Gases: Climatic and Associated Impacts. (Fantechi, R., and Ghazi, A., Eds.), CEC-Kluwer Academic Publishers, Dordrecht, pp. 228 - 244.

Nash, L.L., and Gleick, P.H. (1991). 'Sensitivity of streamflow in the Colorado basin to climatic changes'. J. Hydrol., 125: 221 - 241.

National Research Council Panel on Water and Climate (Ed.) (1977). 'Climate, Climatic Change, and Water Supply'. National Academy of Sciences, Washington D.C., 132 pp.

Nemec, J. (1989). 'Impact of climate variability and change on water resources management in agriculture'. In: Conference on Climate and Water. The Publications of the Academy of Finland, Vol. 2, pp. 357 - 371.

Nemec, J., and Schaake, J. (1982). 'Sensitivity of water resource systems to climate variations'. Hydrol. Sc. J., 27 (3): 327 - 343.

Niemczynowicz, J. (1989). 'Impact of "greenhouse effect" on sewage systems. In: Conference on Climate and Water. The Publications of the Academy of Finland, Vol. 2, pp. 191 - 206.

North, G.R., Cahalan, R.F., and Coakley, J.A. (1981). 'Energy Balance Climate Models'. Rev. of Geophysics and Space Physics, Vol. 19: 19 - 121.

Novaky, B., Pachner, C., Szesztay, K., and Miller, D. (1985). 'Water resources'. In: Climate Impacts Assessment. (Kates, R.W., Ausubel, J.H., and Berberian, M., Eds.), SCOPE Rep. 27, pp. 187 - 214.

Oerlemans, J. (1982). 'Response of the Antarctic Ice Sheet to a climatic warming: a model study'. J. Climatology, 2: 1 - 11.

Oerlemans, J. (1989). 'Greenhouse warming and changes in sea level'. In: Carbon Dioxide and Other Greenhouse Gases: Climatic and Associated Impacts. (Fantechi, R., and Ghazi, A., Eds.), CEC - Kluwer Academic Publishers, Dordrecht, pp. 214 - 227.

Paluticof, J.P. (1987). 'Some possible impacts of greenhouse gas induced climate changes on water resources in England and Wales'. In: The Influence of Climate Change and Climatic Variability on the Hydrologic Regime and Water Resources. (Solomon, S.I., Beran, M., and Hogg, W., Eds.), IAHS Publ. 168, pp. 585 - 596.

Parry, M.L., and Read, N.J. (Eds.) (1988). 'The Impact of Climatic Variability on UK Industry'. Air report 1, Atmospheric Impacts Research Group, University of Birmingham, UK.

Parry, M.L., Carter, T.R., and Konijn, N.T. (Eds.) (1987). Report of UNEP/IIASA Project 'The Impact of Climatic Variations on Agriculture', Reidel, Dordrecht, Netherlands.

Pearman, G.I. (Ed.) (1988). 'Greenhouse: Planning for Climate Change'. Division of Atmospheric Research. Commonwealth Scientific and Industrial Research Organisation (CSIRO), CSIRO Publications, East Melbourne, Austr., 752 pp.

Peterson, D.F., and Keller, A.A. (1990). 'Irrigation'. In: Climate Change and the U.S. Water Resources. (Waggoner, P.E., Ed.), John Wiley and Sons, New York, pp. 269 - 306.

Rango, A., and Martinec, J. (1987). 'Large scale effects of seasonal snow cover and temperature increase on runoff'. In: Large Scale Effects of Seasonal Snow Cover. Proc. Vancouver Symp. Aug., 1987, IAHS, Publ. No 166, pp. 121 - 127.

Reck, R.A. (1974). 'Aerosols in the atmosphere: Calculation of the critical absorption/backscatter ratio'. Science, Vol. 186: 1034 - 1036.

Refsgaard, J.C., Alley, W.M., and Vuglinsky, V.S. (1989). 'Methodology for Distinguishing Between Man's Influence and Climatic Effects on the Hydrological Cycle', IHP-III Project 6.3, UNESCO, Paris, 1.1 - 7.14 pp.

Revelle, R.R., and Suess, H.E. (1957). 'Carbon dioxide exchange between atmosphere and ocean and the question of an increase of atmospheric CO_2 during the past decades'. Tellus, Vol. 9: 18 - 27.

Revelle, R.R., and Waggoner, P.E. (1983). 'Effects of a carbon dioxide - induced climatic change on water supplies in the Western United States'. In: Changing Climate. Nat. Academy of Sciences, National Academy Press, Washington, DC., pp. 419 - 432.

Rhodes, S.L., Miller, K.A., and Mac Donell, L.J. (1992). 'Institutional response to climate change: Water provider organisations in the Denver Metropolitan region'. Water Resour. Res., Vol. 28., No 1: 11 - 18.

Riebsame, W.E. (1988). 'Adjusting water resources management to climatic change'. Climatic Change, 13: 69 - 97.

Robin, G. de Q. (1986). 'Changing the sea level'. In: The Impact of an Increased Concentration of Carbon Dioxide in the Environment. John Wiley, New York.

Sanderson, M., and Wong, L. (1987). 'Climatic change and Great Lakes water levels'. In: The Influence of Climate Change and Climatic Variability on the Hydrologic Regime and Water Resources. (Solomon, S.I., Beran, M., and Hogg, W., Eds.), IAHS Publ. No 168, pp. 403 - 420.

Schaake, J.C. (1990). 'From climate to flow'. In: Climate Change and the U.S. Water Resources. (Waggoner, P.E., Ed.), John Wiley and Sons, New York, pp. 177 - 206.

Schlesinger, M.E., and Mitchell, J.F.B. (1985). 'Model projections of the equilibrium climatic response to increased carbon dioxide'. In: Projecting the Climatic Effects of Increasing Carbon Dioxide. (Mac Cracken, M.C., and Luther, F.M., Eds.), Report DOE-0237, U.S. Dept. of Energy, Washington D.C., pp. 81 - 148.

Schneider, S.H. (1990). 'The global warming debate heats up: Analysis and perspective'. Bulletin Amer. Meteor. Soc., 71: 1292 - 1304.

Schwarz, H.E. (1977). 'Climatic change and water supply: How sensitive is the Northeast?'. In: Climate, Climatic Change and Water Supply. National Research Council Panel on Water and Climate (Ed.), National Academy of Sciences, Washington, D.C., pp. 111 - 120.

Sharma, M.L. (1989). 'Impact of climate change on groundwater recharge'. In: Conference on Climate and Water. The Publications of the Academy of Finland, Vol. 1, pp. 511 - 520.

Shiklomanov, I.A. (1989). 'Anthropogenic changes of climate, water resources and water management problems'. In: Conference on Climate and Water. The Publications of the Academy of Finland, Vol. 2, pp. 310 - 347.

Siegenthaler, U., and Oeschger, H. (1987). 'Biospheric CO_2 emissions during the last 200 years reconstructed by the deconvolution of ice core data'. Tellus, Vol. 398: 140 - 154.

Sokolov, A.A., and Chapman, T.G. (1974). 'Methods for Water Balance Computations'. International Guide for Research and Practice, UNESCO Press, Paris.

Solomon, S.I., Beran, M., and Hogg, W. (Eds.). 'The Influence of Climate Change and Climatic Variability on the Hydrologic Regime and Water Resources'. IAHS Publ. 168, 640 pp.

Stockton, C,W., and Boggers, W.R. (1979). 'Geohydrological Implications of Climate Change on Water Resource Development'. Report, U.S. Army Coastal Eng. Res. Cent., Fort. Belvoir, Va.

Study of Man's Impact on Climate (SMIC) (1971). 'Inadvertent Climate Modifications', MIT Press, Mass, 308 pp.

Thomsen, R. (1989). 'The effects of climate variability and change on groundwater in Europe'. In: Conference on Climate and Water. The Publications of the Academy of Finland, Vol. 1, pp. 486 - 500.

Thornthwaite, C.W. (1948). 'An approach toward a rational classification of climate'. Geographical Review, Vol. 38: 55 - 94.

Thornthwaite, C.W., and Mather, J.R. (1955). 'The water balance'. Publications in Climatology, Drexel Inst. of Technology, Laboratory of Climatology, Vol. VIII, No 1, 104 pp.

Thornthwaite, C.W., and Mather, J.R. (1957). 'Instructions and tables for computing the potential evapotranspiration and the water balance'. Publications in Climatology, Drexel Inst. of Technology, Laboratory of Climatology, Vol.. X, No 3, 311 pp.

Tsonis, A.A., and Elsner, J.R. (1989). 'Testing the global warming hypothesis'. Geophys. Res. Lett., 16: 795 - 797.

U.S Army Corps of Engineers (1980). 'Guide Manual for Preparation of Water Balances, (By Hayes, R.J., Popko, K.A., and Johnson, W.K.,). The Hydrologic Engineers Centre, Davis, California.

U.S. Environmental Protection Agency (1983). 'Projecting Future Sea Level Rise'. Office of Policy and Resource Management EPA, 230 - 09 - 007, Washington D.C.

U.S. Environmental Protection Agency (1984). 'Potential Climatic Impacts of Increasing Atmospheric CO_2 with Emphasis on Water Availability and Hydrology in the United States'. Report, Strategic Studies Staff, Office of Policy Anal., Office of Policy, Plann. and Eval., Washington D.C.

U.S. Environmental Protection Agency (1988). 'The Potential Effects of Global Climate Change on the United States'. Draft Report to Congress, Office of Policy, Plann. and Eval., Washington D.C.

U.S. National Academy of Sciences (1977). 'Climate, Climatic Change and Water Supply. National Academy of Sciences, Washington D.C.

U.S. National Academy of Sciences (1979). 'Carbon Dioxide and Climate: A Scientific Assessment'. National Academy of Sciences, Climate Research Board, Washington, D.C.

U.S. National Academy of Sciences (1983). 'Changing Climate'. National Academy of Sciences, National Academy Press, Washington D.C.

Van der Veen, C.J. (1985). 'Response of a marine ice sheet to changes at the grounding line'. Quaternary Res., 24: 257 - 267.

Van der Veen, C.J. (1986). 'Ice Sheets, Atmospheric CO_2 and Sea Level'. Thesis, Univ. of Utrecht, Netherlands, 184 pp.

Vanderwiele, G.L., Chog-Yu Xu, and Ni-Lar-Win (1992). 'Methodology and comparative study of monthly water balance models in Belgium, China, and Burma'. J. Hydrol., 134: 315 - 347.

Vehvilainen, B., and Lohvansuu, J. (1991). 'The effects of climate change on discharges and snow cover in Finland'. Hydrol. Sc. J., 36, 2, 4: 109 - 121.

Warrick, R.A., Gifford, R.M., and Parry, M.L. (1987). 'CO_2, climate change and agriculture'. In: The Greenhouse Effect, Climate Change, and Ecosystems. SCOPE series, John Wiley, London.

Washington, W.M., and Meehl, G.A. (1983). 'General circulation model experiments on the climatic effects due to a doubling and quadrupling of carbon dioxide concentration'. J. Geophys. Res., 88: 6600 - 6610.

Washington, W.M., and Meehl., G.A. (1984). 'Seasonal cycle experiment on the climate sensitivity due to a doubling of CO_2 with an atmospheric general circulation model coupled to a simple mixed-layer ocean model'. J. Geophys. Res., 89: 9475 - 9503.

Webb, T., and Wigley, T.M.L. (1985). 'What past climates can tell us about a warmer world'. In: Projecting the Climatic Effects on Increasing Carbon Dioxide. (Mac Cracken,

M.C., and Luther, F.M., Eds.), DOE/ER - 0235, Carbon Dioxide Research Division, Washington, D.C., pp. 237 - 257.

Wigley, T.M.L., and Jones, P.D. (1985). 'Influences of precipitation changes and direct CO_2 effects on streamflow'. Nature, 314: 149 - 152.

Wilson, C.A., and Mitchell, J.F.B. (1987). 'A doubled CO_2 climate sensitivity experiment with a global climate model including a simple ocean'. J. Geophys. Res., 92: 13, 315 - 13, 343.

World Meteorological Organisation (1975). ' Intercomparison of Conceptual Models Used in Operational Hydrological Forecasting'. Oper. Hydrol. Rep., 7, Geneva, 172 pp.

World Meteorological Organisation (1987). 'Water Resources and Climatic Change: Sensitivity of Water resources Systems to Climate Change and Variability'. Rep. WCAP-4, WMO/TD 247, Geneva, 50 pp.

World Meteorological Organisation (report compiled by Beran, M.) (1989). 'Report of Conference on Climate and Water', Helsinki, Finland, 11 - 15 Sept. 1989, 23 pp.

World Meteorological Organisation and International Council of Scientific Unions (WMO/ICSU) (1990). 'Global Climate Change'. A scientific review presented by the World Climate Research Programme, (Mc Bean, G.A., and Mc Ewan, A.D.), 35 pp.

Yeh, T.C., Wetherald, R.T., and Manabe, S. (1983). 'A model study of the short-term climatic and hydrologic effects of sudden snow-cover removal'. Mon. Weath. Rev., 111 (5): 1013 - 1024.

CHAPTER 4

Understanding river hydrology

B.L. Finlayson and T.A. McMahon

4.1 Introduction

Environmental hydrology is a developing area within the science of hydrology which seeks to apply the knowledge, skills and techniques of scientific and engineering hydrology to the understanding and solution of environmental problems. In this sense environmental hydrology is complementary to a range of other disciplines, particularly aquatic ecology and biology, which seek to study freshwater aquatic systems for the purposes of their better management towards sustainable ecological goals. In this chapter we deal with river hydrology, an important subset of environmental hydrology.

The knowledge and techniques available for describing and analysing river hydrology have for the most part been generated for scientific and particularly engineering uses. They may not be the most appropriate tools to use in analysing the system for ecological purposes but at this stage they are all we have. The research task ahead for environmental hydrology is to develop methodologies and techniques of analysis specifically suited to the needs of ecologically sound management.

River hydrology, as we use the term here, refers to the flow of water in river channels, to the characteristics of the catchment and the channel network which influence runoff processes, and to the interrelationships between river flows and channel form. Over the time scales relevant to these analyses (10^0–10^2 yrs) the catchment and the channel network can be considered as independent variables, except as they may be changed artificially, while channel form is dependent on interactions between characteristics of the flow, the materials forming the bed and banks and the sediment load being transported by the river.

In this chapter we first discuss the types of hydrological data that are available and their sources. Subsequent sections deal primarily with methods of analysis of those data and the kinds of information that can be extracted from them for ecological purposes. Section 4.3 deals with the characteristics of the river hydrograph; Section 4.4 with flow variability; in Section 4.5 we investigate the interrelations

V. P. Singh (ed.), Environmental Hydrology, 107–135.
© 1995 *Kluwer Academic Publishers.*

between discharge and channel form and consider the case of streams for which no gauged discharge record is available; Section 4.6 deals with the hydrology of regulated streams; Section 4.7 covers the interrelations between streams and their riparian surroundings and the habitat aspects of channels. In Section 4.8 we discuss the important issue of management philosophy since this ultimately determines the use to which the available information will be put.

4.2 Data

4.2.1 Types of Data and Their Sources

In many hydrologic applications, information based on continuous discharge or flow measurements is the basis of analysis and decision making. However, when environmental flow issues and particularly aquatic habitat needs are under consideration, stream stage (water level) is also important. Except in unusual situations stage is measured continuously or at regular intervals and, through a rating curve or rating table which relates stage to discharge at the recording site, the continuous or interval stage values are converted to discharge. Generally, these basic data are converted to daily average flows (usually in m^3/s) for publication purposes. Gauging authorities also provide monthly data based on the summation of daily flows for each calendar month. Daily and/or monthly data are suitable for most investigation, design, operation and management needs. Instantaneous peak discharges are required for some analyses associated with flood investigations. Data for such investigations are taken as the instantaneous peak discharge in each year of record (sometimes maximum daily average flow is used as a compromise) or all instantaneous peaks above some threshold discharge. For low flow analysis, daily average data are required.

The basic data are held generally by national or state water authorities. For example in the U.S. the Geological Survey is the main repository of streamflow and related data whereas in Australia data are held by the respective state water authorities. Flow data may be supplied in calendar years or water years (hydrologic years) but for most purposes it is desirable to carry out analyses using water years. The water year for any location is defined such that the flood season is not split between successive years (see Srikanthan and McMahon, 1984).

4.2.2 Data Problems

In addition to errors in models (simple and complex) and analytical techniques, there are potentially four sources of error in making hydrologic estimates using streamflow data namely measurement, inhomogeniety, sampling and transposition errors. Measurement errors result from errors in stage measurement, in current meter discharge measurements, in interpreting and plotting stage-discharge data, and as a result of hydraulic regime variations.

Inhomogeneity errors result from systematic changes in runoff due to alteration of natural phenomena within the catchment or from human activity. Errors from this source can usually be detected by double mass curve analysis (Gordon et al.,

TABLE 4.1
Minimum length of record (years) required to estimate mean annual flow to a given accuracy as expressed by the standard error of the mean over a range of variabilities

Standard error of mean annual flow	Coefficient of variation of annual streamflows		
	0.25	0.50	1.00
0.05 (5%)	25	100	400
0.100 (10%)	7	25	225

1992) or by one of several statistical tests (Langford and Lewis, 1975) and data modified accordingly.

The third major source of error, known as sampling error, is a function of the length of the record - the longer a record, the smaller the sampling error.

The final source of error, transposition error, relates to situations where hydrologic parameters are required at stream locations away from gauged sites. Hence errors result from transposing data or estimating parameters at an ungauged site.

Assuming errors due to inhomogeneities can be eliminated, the total expected hydrologic error is given by the following (assuming errors to be independent):

$$E = (E_m^2 + E_s^2 + E_t^2)^{0.5} , \tag{4.1}$$

where E is total error and E_m, E_s and E_t are measurement, sampling and transposition errors.

To gain some appreciation of the magnitude of errors involved, Table 4.1 shows the minimum length of record in years to yield a given accuracy of mean annual flow as a result of sampling error. A detailed discussion of streamflow data errors can be found in McMahon (1985) and McMahon and Cronin (1980).

4.2.3 Synthetic Data

In considering synthetic streamflow data, we need to distinguish between two data types. First, there is the situation where streamflow records covering only a few years need to be extended to provide a streamflow sequence equal in length and concurrent with the available rainfall record. The procedure known as data estimation involves developing a relationship between concurrent rainfall (and evaporation and possibly other data) with the known streamflow data. The relationship could be a simple regression equation through to a complex rainfall-runoff model. Details of such procedures are available in McMahon (1979).

Streamflows generated by stochastic procedures are the second type of data to be distinguished. These procedures allow data to be synthesised using a technique that includes a random component. Models representing streamflow are used to

synthesise annual and monthly and sometimes daily data and provide sequences of flows that are statistically indistinguishable from the historical sequence upon which they are based. Stochastic data "... do not improve poor records but merely improve the quality of designs made with whatever records are available" (Fiering and Jackson, 1971). Most text books on hydrology now incorporate a major section on stochastic data generation (see, for example, McMahon and Mein, 1986, and Maidment, 1993).

4.3 Characteristics of the River Hydrograph

The river hydrograph is a plot of discharge through time, the timescale being determined by the purpose for which the data are to be used. Figure 4.1 shows an example of the hydrographs for the Buchan River in Victoria, Australia (catchment area 850 km^2) plotted at a range of time scales from daily (a), monthly (b) and annual (c). An instantaneous hydrograph for this river is shown later in Figure 4.3.

4.3.1 The Generation of the River Hydrograph

The river hydrograph is made up of a number of flow components which consist of water which has taken different pathways through the catchment. The two basic components are illustrated in Figure 4.3. The peak which immediately follows rainfall (the storm hydrograph) consists of water which has not infiltrated into the catchment soils (quickflow or surface runoff) and which has run into the stream at the relatively high velocities characteristic of open channel flow. The rest of the hydrograph consists of water which has taken a longer and slower path through the catchment soils and bedrock and is termed baseflow. The techniques commonly used to separate these two flow types from the hydrograph are described later in this section.

Two basic models are used to explain the generation of the storm hydrograph and their salient features are presented in Figure 4.2. Figure 4.2(b) illustrates the Horton or infiltration excess model of storm runoff generation (Horton, 1933) which assumes that all of the surface runoff, or quickflow, consists of part of the rainfall from the whole catchment. In this model, the catchment soils are considered to have a limiting infiltration capacity and when rainfall intensity exceeds this value the excess runs off over the surface.

The alternative model is illustrated in Figure 4.2(a). Here it is assumed that the infiltration capacity of the catchment soils is higher than rainfall intensity and that surface runoff will only occur when the soil becomes saturated from below. This is referred to as the partial contributing area or saturation excess model of storm runoff generation (Betson, 1964). In this model, the quickflow consists of all the rainfall on the saturated area of the catchment which includes the stream channel. As the catchment becomes wetter with prolonged rainfall or high antecedent moisture conditions the saturated area will expand in size.

While most rainfall-runoff models in common use assume that catchments behave according to the Horton model, the reverse is in fact usually the case. While the distinction probably does not matter for the purposes of modelling

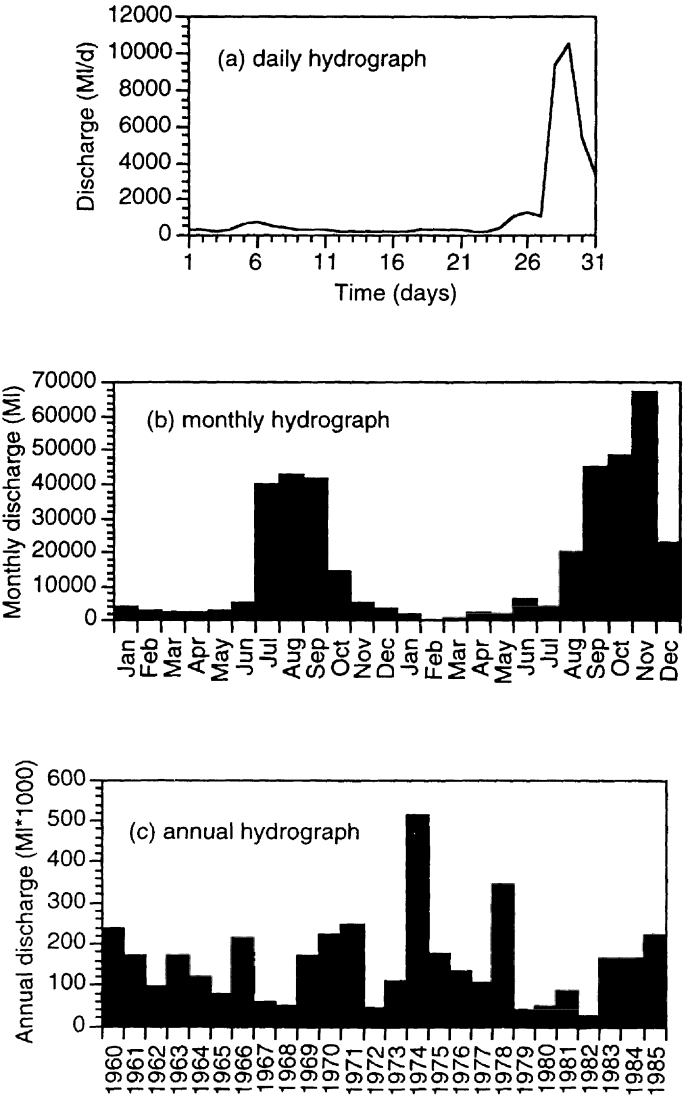
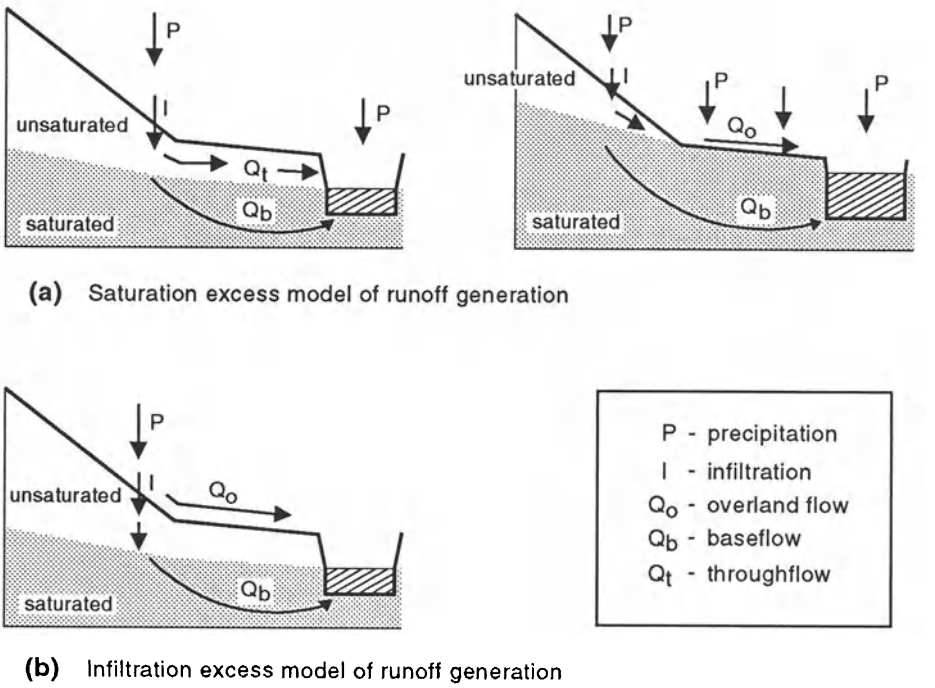


Fig. 4.1. Hydrographs for the Buchan River, Victoria, Australia, catchment area 850 km^2, at a range of time scales. (Data supplied by the Rural Water Corporation, Victoria)

a catchment response to rainfall only in terms of the hydrograph of runoff, it is vitally important when considering catchment erosion, pollutant pathways in runoff and catchment management (Finlayson, 1990). In the management of forested catchments, for example, it can be shown that provided the partial contributing area remains undisturbed during logging operations there is little detectable impact by the logging on stream water quality (Grayson et al, 1993a). Alternatively, agricultural chemicals used on the partial contributing areas may find their way quickly into the stream ecosystem.



(a) Saturation excess model of runoff generation

(b) Infiltration excess model of runoff generation

Fig. 4.2. Schematic illustration of the two models of generation of the storm hydrograph.

The base flow component of the hydrograph consists of water which has passed through the catchment soils and bedrock. The discharge of baseflow at any time is proportional to the volume of saturated storage in the catchment and, in the absence of new inputs of water from rainfall, the flow rate declines exponentially until the saturated storage area is empty and streamflow ceases (Boughton, 1986). Since quickflow passes rapidly through the system it does not have time or opportunity to take part in weathering reactions so that the concentration of dissolved solids is low relative to delayed flow. The converse is true for suspended solids and bedload. Quickflow travels over the surface of the soil at velocities high enough to entrain and transport soil particles so that high suspended solids concentrations and bedload transport rates are associated with the storm runoff peak (Finlayson, 1990).

4.3.2 Hydrograph Separation

Three techniques for separating the hydrograph into its two main components — quick flow or surface runoff and baseflow — are considered next. The first technique is based on the assumption referred to above that during the recession phase of the hydrograph the baseflow discharge is proportional to the amount of water remaining in the storage feeding the baseflow. In effect this means that

$$Q_t = Q_0 k^t, \tag{4.2}$$

where Q_t and Q_0 are baseflows at times t and $t = 0$ and k is the baseflow recession constant. Thus when the recession limb of a hydrograph becomes linear on a plot of $\log Q_t$ versus time t (point B in Figure 4.3), surface runoff is assumed to have ceased and further streamflow is defined as baseflow. In reality this is not a straight forward task because of the complication of other flow types, and analysts should refer to detailed discussions of the procedure (for example in Maidment, 1993, Sec. 9.2.3). The start of surface runoff is identified by the initial rise in the hydrograph - point A in Figure 4.3. The simplest method for the separation of quick flow and baseflow is to join A to B.

The second technique to separate baseflow from the hydrograph is to locate point B by a master recession curve developed for the gauging site in question. The curve is obtained by superimposing the recession limbs of many hydrographs so that the lower discharges coincide as closely as possible. The lower envelope of all the recessions is adopted as the master recession curve. This curve is used to determine for any particular hydrograph where quickflow ceases which is the point on the recession where the hydrograph becomes coincident with the master recession (IEA, 1987).

In the third technique, a recursive digital filter of the following form is used to disaggregate the hydrograph into quick flow and baseflow:

$$f_t = af_{t-1} + \frac{(1+a)}{2}(q_t - q_{t-1}), \tag{4.3}$$

where f_t and f_{t-1} are filtered quickflow at times t and $t-1$, q_t and q_{t-1} are original discharges and a is the filter parameter. Using daily data three passes of the filter are normally required for which the optimum value of a is 0.925 (Nathan and McMahon, 1990a). Chapman (1991) has proposed an alternative recursive filter algorithm:

$$f_t = \frac{3a-1}{3-a} f_{t-1} + \frac{2}{3-a}(q_t - aq_{t-1}). \tag{4.4}$$

In this case the filter parameter a is equivalent to the baseflow recession constant, k, in Equation (4.2) above. Both of these methods have the decided advantage that they enable processing of data automatically which the more traditional methods of flow separation do not.

4.3.3 Flow Duration Curves

The flow duration curve (FDC) is one of the simplest but most informative tools in hydrology. A flow duration curve shows the percentage of time during which specified flows were equalled or exceeded during the period of record.

As illustrated in Figure 4.4 the flow duration curve is a cumulative frequency distribution which is derived directly from the probability density function of the data. It is usual for the abscissa scale to be a normal distribution and the ordinate scale to be logarithmic values of flows. The following interpretations can be made from the FDC.

* Appropriately integrating the curve yields the mean annual flow.

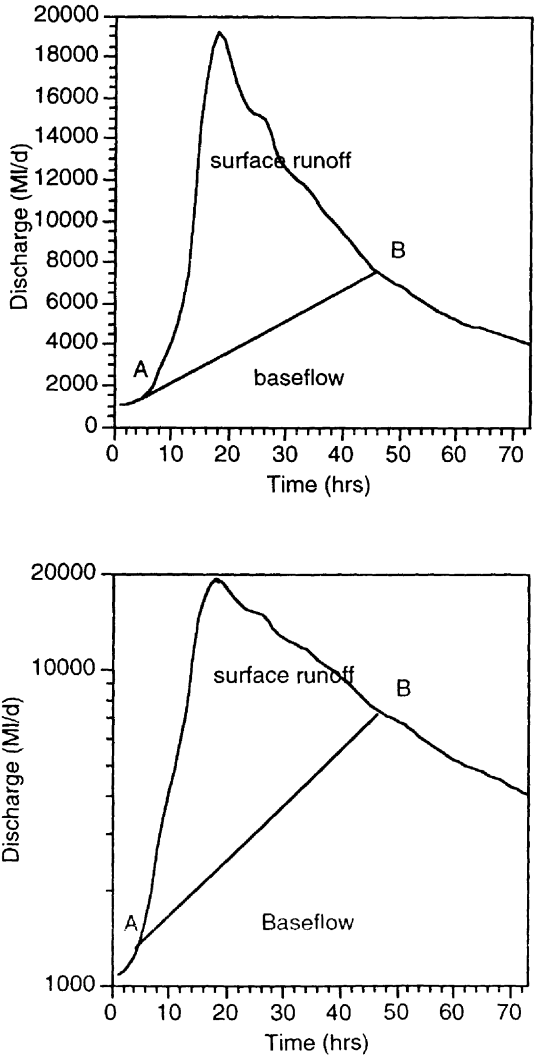


Fig. 4.3. River hydrograph illustrating flow separation on linear axes (top) and log-linear axes (lower).

* The median flow can be read directly from the curve.
* The slope of the FDC represents streamflow variability.
* Percentage time of cease-to-flow conditions is shown specifically.
* The shape of the curve during low flows provides quantitative information about baseflow and hence groundwater characteristics of the aquifer system feeding the stream.

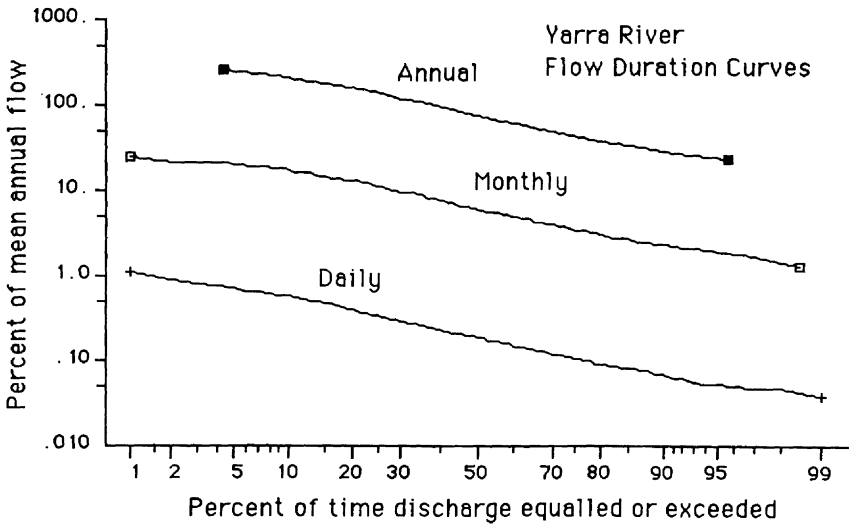


Fig. 4.4. Flow duration curves for the Yarra River for daily, monthly and annual flow records.

* The high flow end of the curve summarises the flood characteristics of the catchment.

4.3.4 High Flows

The peaks of flow in the hydrograph, the floods (see Figure 4.5), are important features of the runoff pattern for a whole range of purposes. The large floods are used by engineers in design calculations for bridges and dam spillways; floods are used in geomorphology in relation to characteristics of the channel such as width, depth and meander wavelength; floods are responsible for most of the sediment transport in rivers; and they are periods of stress and change in biological communities. Floods have properties of magnitude (peak discharge), frequency of occurrence and duration.

The most commonly used methods of analysis of the flood behaviour of a river are those that deal with magnitude and frequency. The data used for this are the peaks of discharge which follow a typical extreme value distribution such that there are many small floods and only a few large ones in any given time period. Using the example of the Yarra River, part of the record for which is shown in Figure 4.5, the annual maximum flood series is formed by selecting only the largest flood in each year. This approach has the advantage of ensuring statistical independence of the observations in the data set but the disadvantages that no estimates can be made of the magnitude/frequency of floods at time intervals of less than one year; that where more than one very large flood occurs in any one year only one will be included in the analysis; and that insignificantly small floods may be included if they are the only floods in any particular year.

The frequency distribution of the annual flood series is strongly skewed but its statistical manipulation can be facilitated using a log transformation of the data and

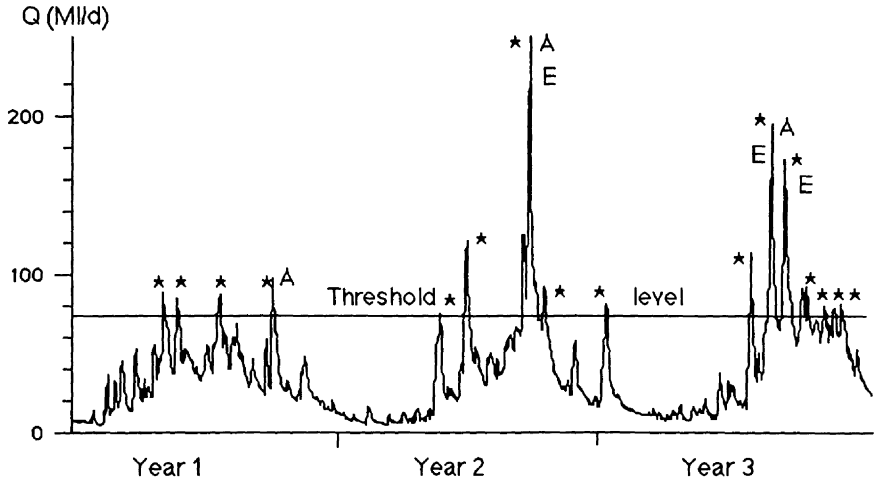


Fig. 4.5. Three years of daily flow record from the Yarra River, Australia. Flood peaks labelled A are those which would be used to form the annual maximum flood series, peaks labelled E would be used to form the annual exceedance series and the peaks labelled with an asterisk all lie above the threshold level for inclusion in the partial duration series.

for this reason most flood magnitude/frequency analysis is carried out in the log domain. In this discussion we will deal only with the average recurrence intervals of floods (rather than probability of exceedance which is the reciprocal of the average recurrence interval). For a more complete treatment of these issues see Gordon et al. (1992) and IEA (1987). Average recurrence interval (T), the average time interval between floods of a given size, is calculated from the ranks (m) of the floods, the largest event having rank 1, and the number of years of record (N):

$$T = \frac{(N+1) - 2a}{m - a}, \tag{4.5}$$

where a is a constant usually set at 0.40 (Cunnane, 1978; IEA, 1987).

Peak discharges can be plotted against average recurrence intervals as shown in Figure 4.6. The annual maximum series is used to estimate the magnitude of floods of long recurrence interval (> 10 yrs) and a frequency distribution is normally fitted to the plotted data points. While there are a number of extreme value distributions in common use in the analysis of floods, McMahon et al. (1992) have shown that the Log-Pearson Type III distribution fits the flood data of rivers around the world better than any other.

Two of the disadvantages of the annual maximum series, that it includes insignificantly small floods and excludes some large floods, can be overcome by using the annual exceedance series where the N largest floods are taken from a record N years long. This series can be plotted the same way as the annual maximum series and is shown for the Yarra River in Figure 4.6 where it can be seen that estimates of average recurrence interval for the annual maximum and annual exceedance series will be essentially identical for recurrence intervals

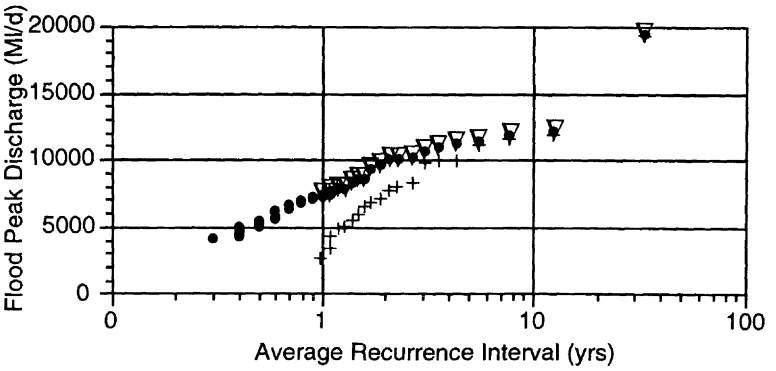


Fig. 4.6. The annual maximum flood series (+), the annual exceedance series (▽), and the partial duration series (•) for the floods of the Yarra River, Australia.

longer than about 10 years. At short recurrence intervals the annual exceedance series gives higher discharges for any given average recurrence interval.

The third disadvantage of the annual maximum series, that it does not deal with average recurrence intervals of less than one year, applies also to the annual exceedance series and can be overcome by using the partial duration series. In this case floods are chosen for inclusion in the series on the basis of some criterion of interest such as bankfull discharge, the discharge at which the substrate begins to become mobile or exceedance of some critical threshold. If the threshold is set so that there are more floods in the series than there are years in the record, then the partial duration series can be used to estimate average recurrence intervals of less than one year. Figure 4.6 also includes a plot of the partial duration series of the Yarra River calculated using three times the number of floods from the record (selected in rank order) as there are years of record. Note that the partial duration series is an extension of the annual exceedance series and that it gives very different average recurrence intervals for small floods than the annual maximum series does.

Flood frequency analysis using the annual series has been well developed in engineering practice since there the interest is usually in the floods of long recurrence interval. Use of the annual series helps to ensure statistical independence of the data. When the partial series is used it is much more difficult to ensure this statistical independence yet this form of analysis is probably of more relevance to river ecologists and geomorphologists who have an interest in the recurrence intervals of higher frequency events.

For many biological applications the height reached by the water (stage) may be of more interest than discharge. Stage frequency can be analysed instead of discharge frequency as described above (Dalrymple, 1960). It should be noted that stream gauging organisations usually do not publish stage data and a special request will need to be made to the hydrographic office to obtain it.

The discharge records available for most rivers are quite short, with few having good records of around 100 years or longer and most having records of 30 years or less. It is almost certain therefore that the gauged record will not include the

largest floods with very long recurrence intervals. Estimates of the magnitudes of these long recurrence interval floods can be obtained by indirect methods; from historical documentary records such as newspapers; or from the identification and dating of slackwater deposits left by palaeofloods in river valleys (Gillieson et al., 1991).

For some issues, such as flood plain management and the ecological interactions between floodplains and channels, the relevant characteristic of floods is time duration of the flow above a critical threshold. This is dealt with through spell analysis, a topic which we will deal with further in the discussion of low flow analysis below. Spell analysis has been little used in relation to floods though one example is given by Maheshwari et al. (1993) in which the analysis is carried out in terms of the duration of water level (rather than discharge) above ecologically significant threshold levels.

4.3.5 Low Flows

Low flows are also an important element of the river hydrograph for the management of water resources and have considerable ecological significance. Low flows have characteristics of discharge, duration and frequency of occurrence and are necessarily relatively long events — only highly regulated rivers drop to low levels and recover again quickly. Low flows can be analysed in terms of the lowest flow event of some specified duration occurring in each year as shown in Figure 4.7. These low flow frequency curves are based on data sequences that are independent (random) which allows the probability of occurrence of a specific low flow event to be determined.

Low flows can also be analysed using spell analysis in which the lengths of time the flow is below some critical threshold level (spell durations), or the volumes of flow which would be required to make up to that critical level (deficiency volumes), are analysed statistically. The derivation of these values is illustrated in Figure 4.8 and an analysis of deficiency volumes is shown in Figure 4.9 for the Yarra River.

In some countries stream water quality criteria have been defined with respect to a low flow event of given duration (often 10 days) and for an average recurrence interval (either 2 or 10 years) — see for example the Hydrologic Investigation Atlas published by the U.S. Geological Survey for the State of Wisconsin (Gebert, 1971). A low flow atlas has also recently been produced for the State of Victoria, Australia (Nathan and Weinmann, 1993). Protracted periods of low flow provide excellent opportunities to examine the spatial characteristics of water quality parameters in a catchment. Under these steady flow conditions water quality parameters can be compared from site to site across the catchment without the confounding influence of discharge fluctuations. An example of the application of this methodology has recently been published by Grayson et al. (1993b).

4.4 Variability

The variability of streamflow can be considered in three ways: as differences through the year related to the annual march of the seasons (the seasonal regime);

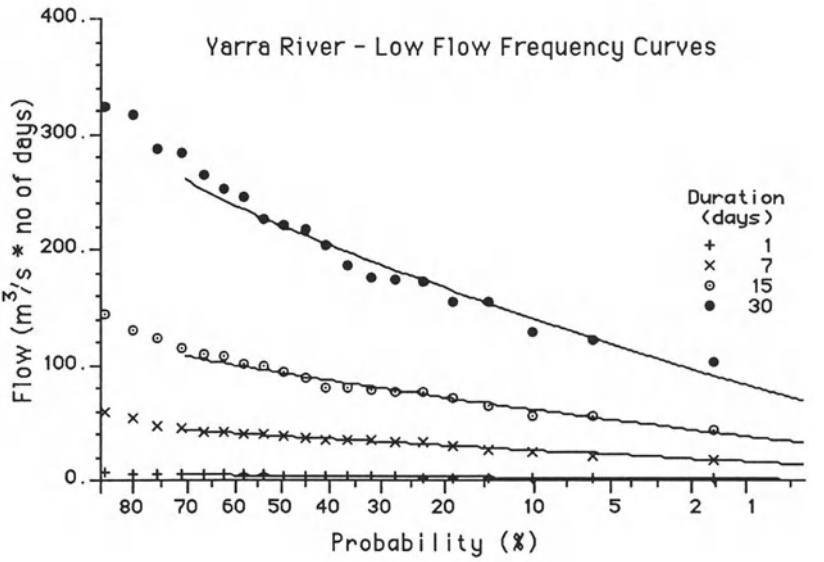


Fig. 4.7. Low flow frequency curves for the Yarra River for low flow events of durations of 1, 7, 15 and 30 days.

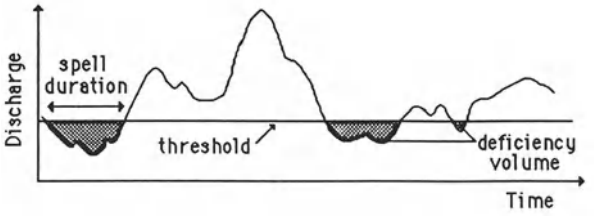


Fig. 4.8. Illustration of the derivation of values used in spell analysis. (Source: Gordon et al., 1992)

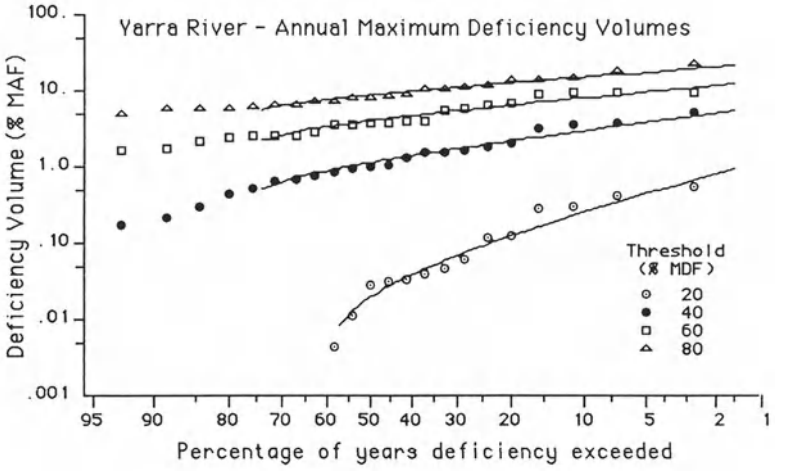


Fig. 4.9. Annual maximum deficiency volumes for the Yarra River. (MAF - mean annual flow; MDF - mean daily flow)

as variability in the stream hydrograph related to the catchment response to rainfall (flashiness); and as the variability between the annual flow of different years or between the flows the same month in different years (interannual variability).

4.4.1 Seasonal Regime

The seasonal regime is the expected pattern of flows through the year related to seasonal conditions. At its simplest this would mirror the seasonal distribution of rainfall though with some delay in the case of very large catchments. Where there is a strong seasonal pattern in evaporation this can substantially modify the pattern of streamflow and where freezing conditions occur in winter, the winter precipitation is stored as snow and becomes runoff only when the spring melt occurs. Given the range of combinations possible of seasonal precipitation and evaporation, snow and other catchment storages (eg lakes and groundwater bodies), many seasonal regimes are possible around the globe. Haines et al. (1988) have classified seasonal regimes using cluster analysis on a data set of nearly 1000 streams from all around the world and the seasonal patterns produced are shown in Figure 4.10. Seasonal regime regions can be mapped from these data (see Haines et al., 1988) and used as a basis for extrapolation to ungauged catchments or to determine the natural regime pattern in streams modified by regulation.

Seasonal regimes are normally calculated using mean monthly flows for the length of the available record. Of interest also is the variation from year to year in the regime type since for highly variable streams the regime type may not be constant. Colwell's indices (Colwell, 1974), developed originally for biological applications, have some potential as a tool in this area though they have been little used in this context so far (though see Gordon et al., 1992 and Gan et al., 1991). Three indices can be calculated, *predictability*, *constancy* and *contingency*, each having a range from 0 to 1 and where predictability is the sum of constancy and contingency. As applied to seasonal streamflow regimes, a *predictability* of 1 indicates exactly the same regime type and flow in each year; a *constancy* of 1 indicates that the flow is the same in every month of every year; and a *contingency* of 1 indicates that the flow is the same in each month from year to year.

4.4.2 Hydrograph Variability

This aspect of flow variability represents the outcome of interactions between rainfall characteristics and the physical properties of the catchment. Catchments with highly permeable soils and/or bedrock and large storage capacities (such as limestone) have flow hydrographs dominated by baseflow and show subdued responses to rainfall. In contrast, catchments with less permeable soils and bedrock have flow hydrographs with marked responses to rainfall events. Urban areas are particularly good examples of this behaviour since the extensive impermeable surfaces (roads and building roofs) generate a quick runoff response. Differences of this type are apparent in the flow duration curves as discussed previously in Section 4.3.

In addition to the effects of catchment materials, catchment shape and topography also influence this aspect of the hydrograph behaviour. Catchments with steep

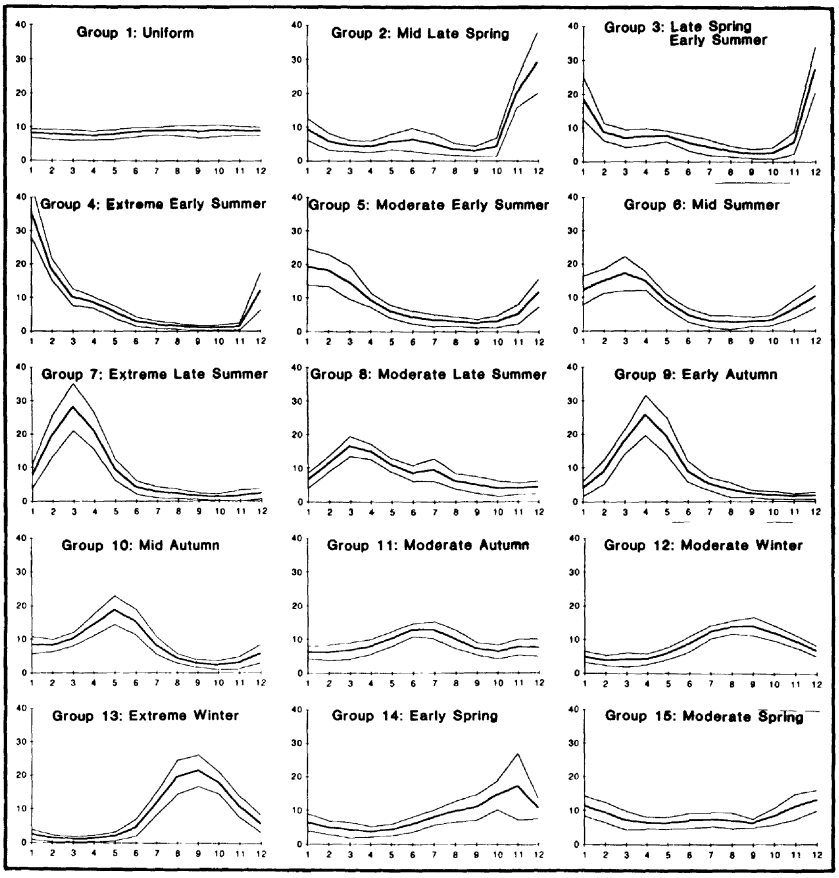


Fig. 4.10. A global classification of seasonal regime types. (Source: Haines et al., 1988)

slopes tend to shed water more rapidly than those with gently sloping topography and fan shaped catchments concentrate flow better than long narrow ones. Climate is also important here. Where rainfall comes in intense storms the hydrograph will be dominated by quickflow response. This is particularly noticeable in arid areas where the only runoff which occurs is that which follows storms and baseflow is largely absent from the hydrograph.

4.4.3 Interannual Variability

As noted earlier the slope of the flow duration curve provides a very good measure of the interannual variability of streamflow. However, another useful measure is the coefficient of variation of annual streamflows (C_v). This measure has been used as the basis for a number of studies of world hydrology (McMahon et al., 1992). That analysis shows that the mean C_v for 974 streams world wide is 0.43 with 20% having values greater than 0.7 and 20% less than 0.2.

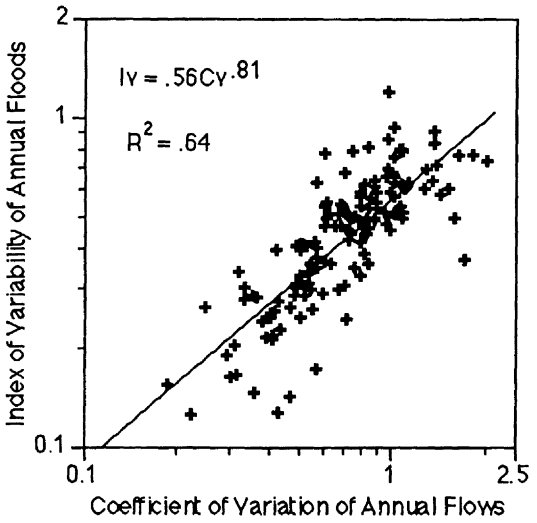


Fig. 4.11. Correlation between the variability of annual floods (I_v) and the variability of annual flows (C_v) for Australian rivers. (Source: Peel, unpublished data)

The interannual variability of floods is described by the index of variation (I_v) of annual floods, defined as the standard deviation of the logarithms of the annual peak discharges. Based on 931 streams world wide, the mean value of I_v is 0.28 with 20% being greater than 0.4. Another measure of the variability of floods is the ratio of the mean annual flood to the flood with an average recurrence interval of 100 years ($q_{100}/\bar{q}$). The world mean (based on 931 streams) is 4.4; for European and North American streams values are in the range 2–4 while for Australia and Southern Africa the range is 5–14 (McMahon et al., 1992).

The variability of annual flows and annual floods tend to be closely correlated. McMahon et al. (1992) found that the annual C_v explained 81% of the variance in $q_{100}/\bar{q}$. Peel (1993) has investigated the relationship between C_v and I_v for Australian streams and found a similarly close correlation (Figure 4.11). Peel also correlated the annual flow series and the annual maximum flood series for individual rivers in Australia and found these to be highly correlated with values of R^2 in the order of 80% and a tendency for the relationship to be better in more arid areas. These analyses indicate that the interannual variability of rivers tends to be dominated by the largest flood in each year.

The engineering impacts of high values of C_v and I_v are significant. Reservoir storage size is a function of C_v^2 (McMahon and Mein, 1986, p. 121) and consequently highly variable streams require relatively large storages. For example, in Australia a stream with a C_v of 0.7 (the median for Australia) would need a storage of more than five times that required for a stream of median variability in Europe, where the median C_v is 0.3, to achieve the same level of supply even before the higher evaporation levels in Australia are taken into account.

Flooding is a similar story. Because of the higher variability associated with Australian floods, on average the peak discharge per unit area of, say, the once in

100 years flood in Australia is much larger than that in Europe. Consequently, the impact of higher C_v and I_v values in Australia (and in southern Africa) translates into relatively higher costs for the engineering aspects of water projects than in other continents. As a consequence, decision makers are more reluctant to spend money on the related environmental water issues.

As a result of high hydrologic variability (and for other reasons too) there is the need for stream regulation by reservoirs and sometimes inter-basin transfers. Both these engineering activities normally have major consequences for the aquatic environment in the streams affected. To date stream regulation directly or through inter-basin transfers has been carried out with little concern for, and usually no knowledge of, its effects on the aquatic environment.

High streamflow variability directly impacts the aquatic environment. For example, Walker (pers. comm.) suggests that in Australian inland streams many fish species are opportunistic breeders and for reproduction take advantage of infrequent floods which cover the floodplain, where there is a large reserve of food, in contrast to less variable systems such as those of Europe where species exhibit a regular, predictable annual breeding cycle.

This opportunism characterises many Australian plants and animals, as they have evolved in conditions of high hydrological variability, and is a recurring theme in the literature. In dealing with aquatic plants in Tasmanian rivers, Hughes (1988) has this to say: " ... nonequilibrial or stochastic processes such as discharge fluctuations ... tend to favour opportunistic species that recover rapidly vegetatively".

Mitchell and Rogers (1985) identify long term wet and dry cycles in southern Africa and Australia as an important component of aseasonal fluctuations in environmental conditions. In this regard it is worth noting that Australian geomorphologists have evoked similar arguments (Erskine and Warner, 1988). They propose that flood-dominated and drought-dominated regimes each lasting several decades can be used in the explanation of the behaviour of river morphology in streams in eastern Australia.

In arid zones like the Lake Eyre basin in central Australia the effect of high streamflow variability on the aquatic environment is even more spectacular. Here on the rare occasions when substantial flood runoff reaches the lake, nomadic water birds quickly colonise it in large numbers and begin breeding. Serventy (1985) observes that the whole food chain undergoes a similar "explosion of life".

4.5 Catchments and Channels

4.5.1 Fluvial Geometry

In general, discharge increases as catchment area increases (this does not necessarily hold for arid zone catchments) and this relationship can be used in predicting the runoff from ungauged catchments (see Table 4.2). Larger flows require larger channels to carry them so that a close relationship can be observed between channel size (width and depth) and catchment area (see, for example, Gordon et al., 1992, Figure 7.9). The discharge which just fills the channel, termed the bankfull discharge, is commonly assumed to be responsible for the form of the channel.

Various recurrence intervals have been reported for bankfull discharge though it is obvious that this varies between rivers (Richards, 1982). One of the confounding factors in the search for this relationship in the past has been the use of the annual flood series to determine recurrence intervals for bankfull discharge. As shown in Figure 4.6, for average recurrence intervals less than 10 years the partial duration series will give lower values than the annual maximum series. The annual maximum series will only give average recurrence intervals of greater than one year for discharges which in fact occur more frequently than this.

Statistically significant relationships are also observed between discharge and other features of channel form such as meander wavelength but many other variables influence these aspects of channel form. Figure 4.12 is an attempt to portray graphically the relations between the various channel forms which are found in natural rivers and aspects of the fluvial system which are associated with them. Note that the variables on the axes of this diagram, sediment supply, sediment texture, channel gradient, channel sinuosity and stability are not necessarily independent of each other. It is important to note here that most of the available information which deals with the interrelationships depicted in Figure 4.12 has been derived for alluvial channels. Much less is known about the behaviour of channels in hard rock.

The importance of these relationships between discharge and the geometry of the river system for environmental hydrology is that the channel will respond to imposed changes in any of the controlling variables. Thus, flow modification will lead to adjustments in the channel, a factor which must be taken into consideration in the provision of environmental flows in regulated rivers. These interrelationships can be expressed using the qualitative relationship (Lane, 1955):

$$Q_s d \sim Q_w S, \tag{4.6}$$

where Q_s is the quantity of bed load carried by the stream, d is a measure of the particle size of the bed load, Q_w is the water discharge and S is the stream gradient. When a dam is constructed on a river, sediment is stored in the dam so that even if the discharge (Q_w) below the dam remained unchanged there is likely to be a reduction in slope (S) to compensate, achieved by incision of the stream bed. Increases in the size (d) or load (Q_s) of sediment will lead to aggradation as commonly occurs where mining waste is deposited in rivers. Increasing the discharge, as occurs where rivers are used to transport irrigation water, leads to bed degradation in the channel. Once adjustments of the system are set in train there is usually an ongoing demand for engineering works to attempt to prevent further undesirable consequences.

4.5.2 Ungauged Streams

The most useful field technique to determine streamflow characteristics at an ungauged site is based on a few low-flow discharge measurements at the otherwise ungauged site. These data can then be related to concurrent discharges at a nearby gauged station. Low flow frequency and duration data are transferred through the relationship to obtain appropriate flow characteristics at the ungauged site. For details and examples of the methodology see McMahon and Mein (1986).

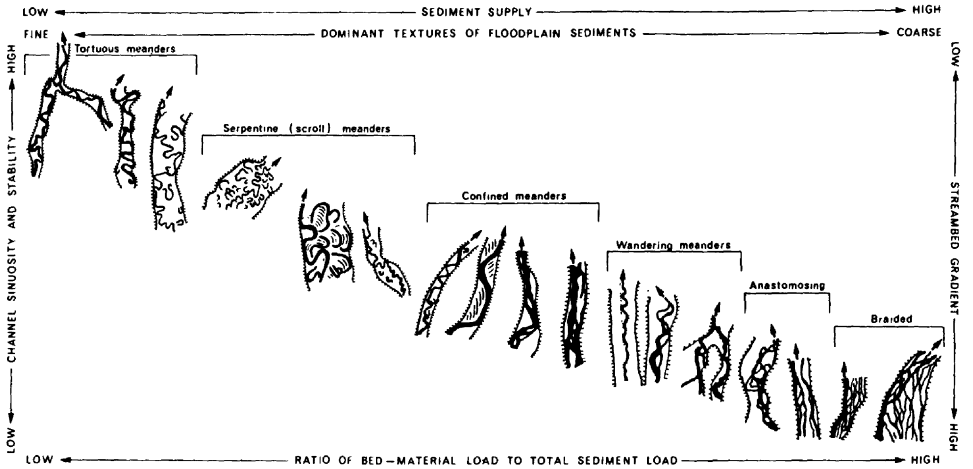


Fig. 4.12. River channel patterns and their relations with sediment supply and texture, streambed gradient and channel stability. (Source: Selby, 1985)

There are many simple empirical relationships that have been established to estimate streamflow characteristics at ungauged sites. Examples based on global, continental and regional data are given in Table 4.2.

Regionalisation is a technique that allows extrapolation of streamflow characteristics from a gauged location to an ungauged one. Essentially it is based on multivariate analysis in which the streamflow parameter is related to one or more climatic and catchment characteristics. An example from the Institute of Hydrology (1980) is:

$$Q_{95,10} = [a(\text{BFI})^{0.5} + b\{(P) \text{ or } L\}^{0.5} + c]^2, \tag{4.7}$$

where $Q_{95,10}$ is 10-day average flow exceeded 95% of time, BFI is base flow index, P is mean catchment rainfall, L is mainstream length and a, b, c are empirical constants.

In examples like the above, it is usual to assume that similar catchments will be located in a geographically contiguous region. However, techniques are now available in which regionalisation is achieved through similarity with respect to a number of measurable parameters using Andrew's Fourier Plots (Andrews, 1972; Nathan and McMahon, 1990b). For a site in question, an Andrew's Curve is prepared

$$F(T) = \frac{x_1}{2^{0.5}} + x_2 \sin(T) + x_3 \cos(T) + x_4 \sin(2T) + x_5 \cos(2T) + \cdots, \tag{4.8}$$

where $F(T)$ is the two dimensional curve which is plotted from $-T$ to $+T$ and x_1, x_2, ... represent the appropriate climate and catchment variables. Figure 4.13 is a set of master curves representing regions of similar hydrologic characteristics. The curve for an ungauged catchment can be matched with these in order to identify

TABLE 4.2
Examples of empirical relationships for estimating streamflow characteristics at ungauged sites

Empirical equation	Area	Reference
$C_v = 2.0\,\mathrm{MAR}^{-0.30}$	global	McMahon et al., 1992
$\bar{q} = 2.5A^{0.63}$	global	McMahon et al., 1992
$q_{100}/\bar{q} = 5.8C_{vf}^{0.92}$	global	McMahon et al., 1992
$\tau = [Z_p^2 C_v^2/4(1-D)] - dC_v^2$	global	Gould, 1964
$\mathrm{MAR} = 2300A^{-0.35}$	Australia	McMahon, 1976
$C_v = 3.9\,\mathrm{MAR}^{-0.33}$	Australia	McMahon, 1978
$\bar{q}_s = 3.3A^{-0.44}$	Australia	McMahon, 1985
$\bar{q}_s = 10^{-3}\,\mathrm{MAR}^{1.02}$	Australia	McMahon, 1982
$\mathrm{MAF} = 9.3 \times 10^{-6} A^{0.99}\,\mathrm{MAP}^{1.48}$	Southeast Australia	Gan et al., 1990
$C_v = 127A^{0.06}\,\mathrm{MAP}^{-0.63}C_{vp}^{0.87}$	Southeast Australia	Gan et al., 1990
$C_s = 4.3 - 1.2\log_{10}\mathrm{MAR}$	Northern Australia	McMahon, 1977
$q_{2.33} = 12A^{0.79}$	New Mexico	Leopold & Miller, 1956
$Q_{bf} = 8.9X^{1.3}(P/R)^{-0.27}S^{0.32}$	South Island New Zealand	Mosley, 1979
$Q = (a + bY)A$	New Hampshire	Dingman, 1978
$Q_{02} = k_1 Q$	(USA)	
$Q_{05} = k_2 Q$		
$Q_{30} = k_3 Q$		
$Q_{95} = e + fY = gY_2$		

Symbols:
a, e: constants
b, f, g, k_1, k_2, k_3: coefficients
A: catchment area (km^2)
C_s: coefficient of skewness of annual flows
C_v: coefficient of variation of annual flows
C_{vf}: coefficient of variation of annual peak flows
C_{vp}: coefficient of variation of annual precipitation
d: factor to adjust annual flows to gamma distribution
D: Draft from a reservoir as a ratio of annual streamflow
MAF: mean annual flow (m^3 $\times$ 10^6)
MAP: mean annual precipitation (mm)
MAR: mean annual runoff (mm)
P/R: wetted perimeter/hydraulic radius
$\bar{q}$: mean annual flood (m^3/s)
$\bar{q}_s$: specific mean annual flood (m^3/s/km^2)
$q_{100}/\bar{q}$: ratio of 100 year flood to mean annual flood
$q_{2.33}$: mean annual flood with 2.33 yr recurrence interval (ft^3/s)
Q: long term average flow
Q_{bf}: bankfull discharge (m^3/s)
Q_{02}, Q_{05}, Q_{30}, Q_{95}: flow equalled or exceeded 2%, 5% .... of time.
S: channel slope (m/m)
X: channel cross-section (m^2)
Z_p: standardised normal variate for p% risk of failure in any year
Y: elevation
τ: required reservoir size as a ratio of mean annual streamflow

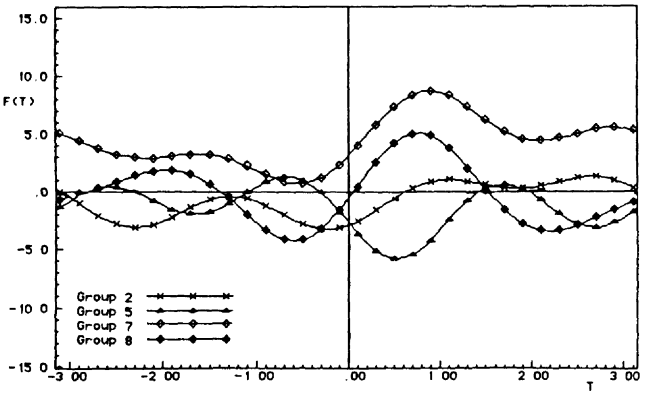


Fig. 4.13. Master Andrew's Curves for four groups of catchments. Curves for ungauged catchments can be compared with these master curves and the catchments assigned to one of the groups. (Source: Nathan and McMahon, 1990c)

the region with which the ungauged catchment should be identified. This approach adds a new dimension to regionalisation of hydrologic characteristics.

4.6 Regulated Rivers

River regulation is considered to be any alteration to the natural flow of the river by the removal or addition of water or by changing the timing of flows. Regulation is most commonly achieved by the construction of a dam on the river but pumping and other forms of flow diversion are also used. The effect of a dam on river flow depends on the purpose for which the dam was built and the way in which it is operated. In some cases dams are used to control the flow in the river below the dam so the total annual flow is little affected but the timing of flows is changed markedly to produce a different seasonal regime. In other cases the water from the dam is diverted into another catchment so that streamflow below the dam is markedly reduced. In climatically humid areas the effects of the dam or other form of regulation decreases downstream as unregulated tributaries enter the system. In arid areas, however, the regulation effects will be more persistent. In this section we provide two examples of regulation though it must be realised that a great variety of impacts are possible and each case needs to be examined individually.

The Eildon Reservoir on the Goulburn River in Victoria, Australia is the storage for a large irrigation scheme downstream of the dam. The river channel is used to transfer irrigation water to the irrigation scheme so that the total annual flow is little affected though the seasonal regime has been markedly altered (Gippel et al., 1991). This case can be used to illustrate the effect of dams on flood behaviour and the consequent impact this has on floodplain wetlands. This is shown in Figure 4.14 in which the flood frequency curves for a station just below the dam are presented for the unregulated condition, for regulation due to a small initial dam at the site and for the present large dam. The first small dam (Old Eildon in Fig. 4.14) reduced the size of flood peak for any given recurrence interval over the whole

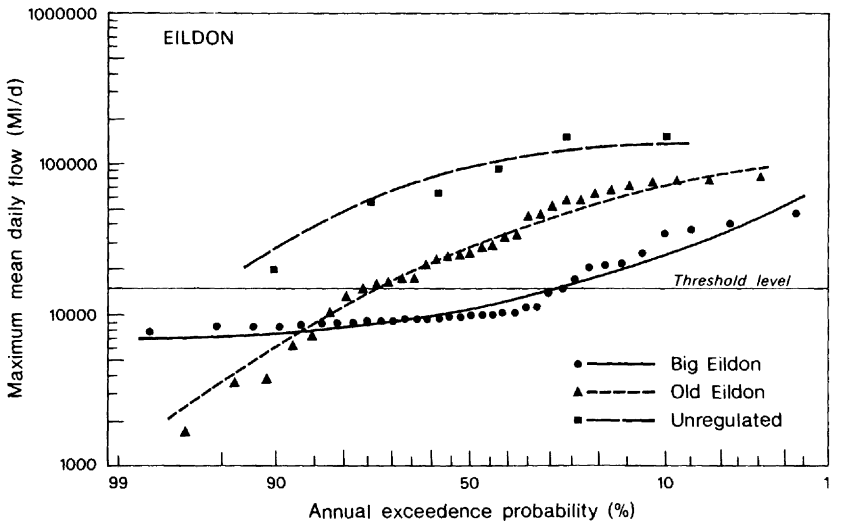


Fig. 4.14. Flood frequency curves for the annual maximum series for the unregulated condition and two phases of regulation of the Goulburn River, Victoria, Australia. (Source: Gippel et al., 1991)

range of floods. The two curves begin to close at the large flood end since under the conditions which produce extreme floods the storage will probably be full and the flood peak will be only slightly modified by its passage through the dam. The construction of the present dam (Big Eildon in Figure 4.14) was accompanied by a major expansion of the irrigation scheme and the additional water released for this has an effect on the lower end of the flood curve. At larger flood discharges the effects of the storage are obvious and here again the impact decreases under extreme conditions.

The floods which are now being stored in the Eildon Reservoir previously served to fill wetlands on the river floodplain below the dam. By analysing the behaviour of floods in relation to the threshold flood level at which the wetlands start to fill, Gippel et al. (1991) were able to show that the wetlands are now flooded about once in every three years whereas under natural conditions this occurred at least annually.

A second example (taken from Gippel et al., 1992a) relates to the inter-basin water transfer which occurs as a result of the Snowy Mountains Hydro-Electric Scheme in the mountains of southeastern Australia. Here water is transferred from the coastal Snowy River to the inland-flowing Tumut and Murray Rivers. The impact on the Tumut River, shown in Figure 4.15, is dramatic where the mean annual flow of the Tumut has nearly doubled. This is accompanied by a complete reversal of its seasonal regime. The Tumut River now flows near bank full for the summer six months; under natural conditions this was the annual low flow period. This has produced serious problems with bank erosion on the Tumut River and in order to maximise the channel capacity the channel is kept free of fallen trees and other debris. The fallen trees (snags or large woody debris) are important elements of in-stream habitat. The effect of the reversal of the seasonal

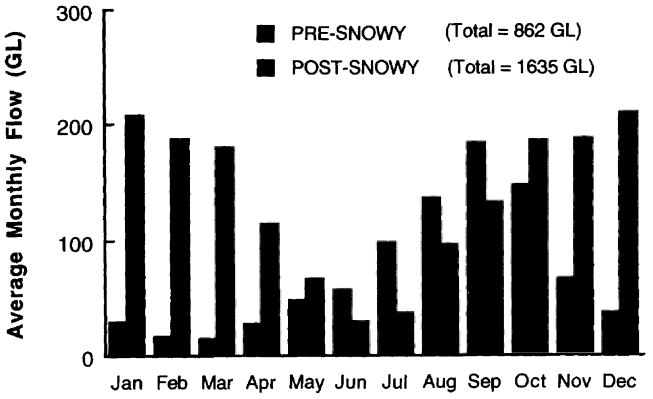


Fig. 4.15. Mean monthly flows in the Tumut River at Tumut (New South Wales, Australia) before and after completion of the Snowy Mountains Hydro-Electric Scheme. Water diverted from the east flowing Snowy River into the west flowing Tumut River has doubled the flow of the Tumut and reversed its seasonal regime. (Source: Gippel et al., 1992a)

flow regime on the river ecology has never been investigated though it is likely to be significant, particularly favouring introduced fish species such as trout at the expense of Australian native fish.

4.7 Waterways and Watersheds

4.7.1 Influent and Effluent Streams

An influent stream is defined as one where the stream water level is above the adjacent groundwater table resulting in the stream recharging the associated aquifer. This situation is sometimes termed transmission loss. Alternatively, an effluent stream is one where the groundwater feeds the stream.

Little quantitative information about the magnitude of transmission losses is available in the literature although Renard (1970) notes that the losses can be very high. For example, in the 150 km^2 Walnut Gulch experimental watershed in Arizona approximately 85% of the on-site runoff does not appear at the gauging station. In Australia, for a 900 km reach of the Darling River (located in the western arid zone of NSW), up to 17% of the inflow was considered to be transmission loss (Laurenson, 1962). In Queensland losses up to 0.03 $m^3/s/km$ have been observed (Hazel and Pappin, 1977). Another dramatic example of transmission losses is presented in Figure 4.16 which shows the monthly seasonal flows for the Niger River at three locations along its length.

Few daily rainfall-runoff models take into account influent stream effects yet most account for baseflow (effluent stream effect). MODHYDROLOG is one daily model that has the capability to simulate both processes. The model has been tested independently against an integrated surface-groundwater model for which there was an unusually good set of climate, runoff, watertable and piezometric data (Chiew and McMahon, 1991).

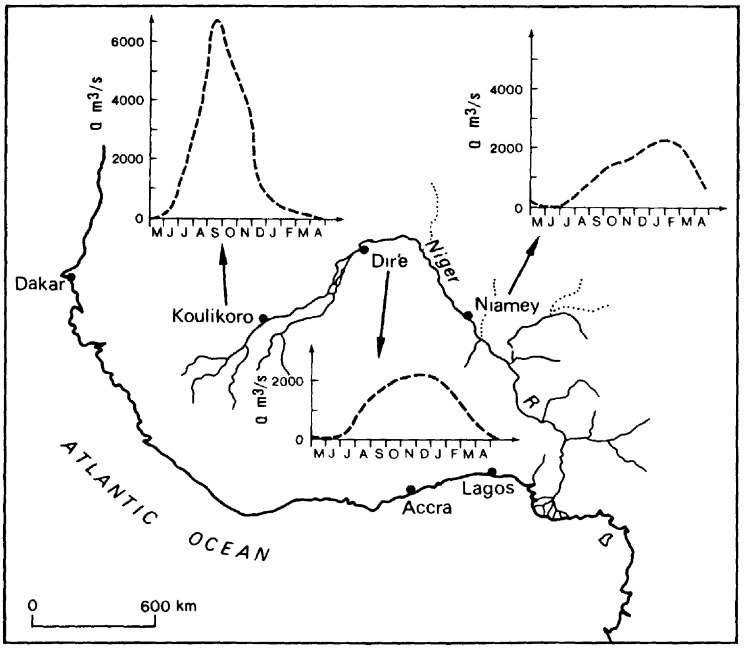


Fig. 4.16. Monthly flows for the Niger River at three locations showing the effects of downstream transmission losses. (Source: McMahon and Diaz Arenas, 1982)

4.7.2 Environmental Aspects

As shown in Section 4.6, the characteristics of flow in regulated rivers are substantially modified. In the past, the provision of environmental flows has consisted of little more that maintaining some arbitrary minimum flow level in the stream. A variety of techniques have been proposed for the determination of environmental flows which have been reviewed by Kinhill Engineers (1988). They group the available techniques as rule of thumb, transect or passage, and available habitat. The rule of thumb techniques do little more than nominate a predetermined proportion of the natural discharge to be retained in the channel to support in-stream activities and riparian users. These sometimes include provision for different proportions at different seasons of the year to provide for fish migration and spawning. Transect techniques use physical data from critical sites (cross sections) on the stream to determine the minimum flow at which critical activities such as migration and spawning can still occur.

The most elaborate methodologies are the available habitat techniques of which the instream flow incremental method (IFIM, Milhouse et al., 1984) is probably the best known. The IFIM methodology has as a central component a characteristic habitat area/discharge curve for the channel of the regulated stream. This curve is constructed on the basis of a combination of depth and velocity considered to be suitable habitat for the fish species of interest. Figure 4.17 shows the percentage of the maximum pre-regulation rearing habitat of the freshwater blackfish (Gadopsis mamoratus) over the range of discharges occurring in the Thomson River in

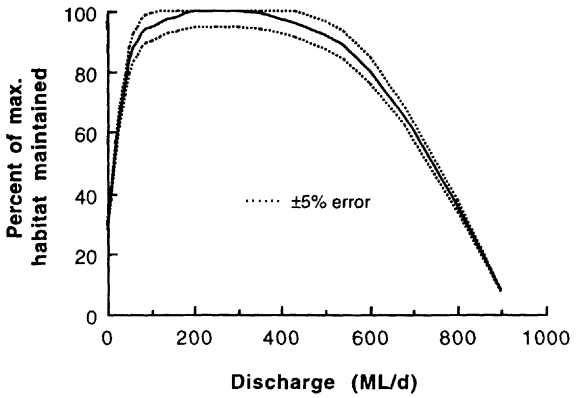


Fig. 4.17. Percentage of maximum available blackfish rearing habitat maintained across the range of observed flows in the Thomson River below the Thomson Dam, Victoria, Australia. (Source: Gippel et al., 1992b)

Victoria, Australia. At low flows the habitat is limited by shallow depths while at high discharges it is the velocity which becomes the limiting factor. Over a wide range of discharges, from 100 Ml/day to nearly 600 Ml/day, close to the maximum habitat is available and the implication is that 100 Ml/day is all that is necessary to provide maximum habitat for this species. More water does not necessarily increase the amount of habitat available.

The habitat/discharge curve in Figure 4.17 was constructed for the channel form which existed prior to regulation. If the channel form changes, this curve, and the flows specified for release on the basis of it, will no longer apply. As discussed in Section 4.5.1, there is a discharge to which the channel form is adjusted and the fluvial geometry relationships discussed in Section 4.5.1 can be used to predict the size and frequency of the channel forming flow. A flow of this magnitude and frequency can then be provided as part of the environmental flow release strategy to maintain the channel in its pre-regulation form. An additional consequence of failing to provide appropriate channel maintenance flows in regulated rivers is that on those infrequent occasions when a large flood comes over the spillway of the dam, substantial damage to the physical stream environment will occur if the channel has been allowed to contract in size.

4.8 Concluding Statement

Hydrology is an evolving science and the central concerns and emphases have been changing through time and continue to do so. The current importance of environmental hydrology in the discipline can be seen in the context of the development of hydrology over the past few decades.

Prior to the 1960s hydrology was primarily concerned with the land phase of the hydrological cycle and consisted of the strong practical component of engineering hydrology and a rather more academic field of scientific hydrology. Engineering hydrology in this period was concerned to develop practical methods

for the prediction of runoff from ungauged catchments and with flood prediction, both of these primarily for design purposes, and with the design and construction of water control systems (dams, irrigation schemes, hydro power stations, etc).

Scientific hydrology at this time was concerned with understanding hydrological processes such as evaporation and runoff generation. There were also close links with hydraulics in the study of sediment mobilisation and transport. In neither of these branches of hydrology was there any serious consideration given to the ecological relations of hydrology except as biological factors impacted on hydrological processes (as in forest hydrology) or on the operation of water systems (eg snags in rivers or weeds in irrigation canals). Few, if any, water development schemes in this period paid any attention to environmental impacts.

The nature of both engineering hydrology and scientific hydrology changed during the decades of the 1960s and 1970s with the advent and widespread use of computers. This facilitated the manipulation of large data sets, gave enormous impetus to hydrological modelling and led to the development of stochastic hydrology, still with relatively little concern for the wider environmental issues.

The 1980s saw the rise of environmentalism as both a public concern and a political movement and of increased public participation in policy development and decision making. Whereas previously the construction of yet another dam was welcomed by all except those whose land was to be flooded, there now began to be large scale public protests particularly against the construction of dams in pristine wilderness areas. At this time also, public concern developed over the single issue management of water regulation systems and the actual or perceived abuse of the natural landscape by the "engineering approach" to management.

This has coincided with both practical and economic constraints on the development of new dams and associated infrastructure. While governments began to scrutinise more carefully expenditure by the large public corporations involved in these construction projects it is also the case that, at least in the areas of greatest water demand, most of the good dam sites had already been built on and most of the available water was diverted for use.

These developments have combined to produce a climate in which hydrology as a science can no longer ignore the ecological aspects of water control systems. If water supply criteria are to be met while simultaneously maintaining an ecologically healthy river system then much better management procedures will be required than have been common in the past. In this context, the key issues and relevant variables are possibly different to those that have been required for the classical engineering approaches to hydrological analysis. There is, therefore, a need to adapt and to develop new methods and techniques in hydrology that are relevant to ecology. Despite this, the real limitation to better management is the lack of relevant ecological information and of an understanding of the relationship between flow, channel condition and in-stream ecology.

An important aspect of the changes now occurring in hydrology is the perception of what constitutes 'demand'. Whereas previously it had been the case that demand was calculated and the system designed and constructed to meet it, there is now the realisation that demand is a manageable quantity — irrigation demand can be reduced by better scheduling and more appropriate pricing policies; urban water demand can be reduced by more efficient water using appliances, advertising

campaigns and, of course, higher prices. System management, as perceived in this wider context, is now a high priority in engineering hydrology at both the research and management levels.

Ecological criteria need to be met in regulated systems so that these systems remain ecologically viable, but at the same time it is unlikely that highly regulated river systems can be returned to their 'natural' condition since the demands of water users (irrigation, hydro power, urban water supply) must continue to be met. The changes in attitude and practice that have occurred in engineering hydrology, especially the recognition of the the necessity for dealing appropriately with the wider aspects of the ecology of river systems, may well lead to the point where the terms hydrology and environmental hydrology will tend to become synonymous.

References

Andrews D.F. (1972): 'Plots of high dimensional data' Biometrics 28:125-136.

Betson R.P. (1964): 'What is watershed runoff?' J. Geophys. Res., 69:1541-1551.

Boughton W.C. (1986): 'Linear and curvilinear baseflow recessions' J. Hydrol. (N.Z.), 25:41-48.

Chapman T.G. (1991): 'Comment on "Evaluation of automated techniques for base flow and recession analysis" ' Wat. Res. Res., 27:1783-1784.

Chiew F.H.S. and T.A. McMahon (1991): 'Improved modelling of the groundwater processes in HYDROLOG' Hydrology and Water Resources Symposium, Perth. The Institution of Engineers, Australia, Nat. Conf. Pub. 91/22:492-497.

Colwell R.K. (1974): 'Predictability, constancy and contingency of periodic phenomena' Ecology 55:1148-1153.

Cunnane C. (1978): 'Unbiased plotting positions - a review', J. Hydrol. 37:205-222.

Dalrymple T., (1960): 'Flood frequency analyses', Manual of Hydrology: Part 3. Flood Flow Techniques, US Geol Surv. Water Supply Paper 1543-A (USGS, Washington, DC).

Dingman S.L. (1978): 'Synthesis of flow-duration curves for unregulated streams in New Hampshire' Water Resources Bull. 14:1481-1502.

Erskine W.D. and R.F. Warner (1988): 'Geomorphic effects of alternating flood- and drought-dominated regimes on NSW coastal rivers' Fluvial Geomorphology of Australia, ed. R.F. Warner (Academic Press, Sydney), 223-244.

Fiering M.B. and B.B. Jackson (1971): Synthetic Streamflow, American Geophysical Union, Water Research Monograph Series, No 1. (AGU, Washington, DC)

Finlayson B.L. (1990): 'The generation of suspended sediment in rivers and streams' In: R. Beckett (ed.) Surface and Colloid Chemistry in Natural Waters and Water Treatment. (Plenum Press, New York), pp. 57-70.

Gan K.C., T.A. McMahon and B.L. Finlayson (1990): 'Analysis of periodicity in streamflow and rainfall data by Colwell's indices', J. Hydrol., 123: 105 - 118.

Gan K.C., T.A. McMahon and I.C. O'Neill (1990): 'Errors in estimated streamflow parameters and storages for ungauged catchments' Water Resources Bull. 26:443-450.

Gebert W.A. (1971): 'Low flow frequency of Wisconsin streams, Hydrological Investigation Atlas, HA-390, (USGS, Washington, DC)

Gillieson D., D.I. Smith, M. Greenaway and M. Ellaway (1991): 'Flood history of the limestone ranges in the Kimberley region, Western Australia' App. Geog., 11:105-123.

Gippel C.J., B.L. Finlayson and B.A. Thompson (1991): 'Establishing the effect of river regulation on wetland flooding regimes' Rapid Techniques for Assessing Wetland Vegetation and Water Regime. Bureau of Rural Resources, Working Paper No. Wp/12/91, PartII

Gippel C.J., I.C, O'Neill and B.L. Finlayson (1992a) - The Hydraulic Basis of Snag Management, Project Completion Report, Land and Water Resources Research and Development Corporation, Project P88/33. 116pp

Gippel C.J., R. Marchant, M.J. Stewardson, S.O. Brizga, I.C. Campbell, J. Woodfull, B.L. Finlayson and T.A. McMahon (1992b) - A Review of Environmental Water Requirements of the Thomson River from Thomson Dam to Cowwarr Weir, (Melbourne Water Corporation.) 89pp.

Gordon N.G., T. A. McMahon and B. L Finlayson (1992): Stream Hydrology An Introduction for Ecologists (John Wiley & Sons, Chichester UK) 526pp.

Gould B.W. (1964): Statistical methods for reservoir yield estimation. Water Research Foundation of Australia Report No 8.

Grayson R.B., C.J. Gippel, B.L. Finlayson and B.T. Hart (1993b): 'A low-flow 'snapshot' of point and non-point sources loads in the Latrobe river catchment' Hydrology and Water Resources Symposium, Newcastle. The Institution of Engineers, Australia, Nat. Conf. Pub. 93/14:337-342.

Grayson R.B., S.R. Haydon, M.D.A. Jayasuriya and B.L. Finlayson (1993a): 'Water quality in mountain ash forests - separating the impacts of roads from those of logging operations', J. Hydrol., 150:459-480.

Haines A.T., B.L. Finlayson and T.A. McMahon (1988): 'A global classification of river regimes' App. Geog., 8:255-272.

Hazel C.P and L.B. Pappin (1977): Report on Mary River transmission loss study. Queensland Water Resources Commission, Hydrology Report No. 138000, PR/3.

Horton R.E. (1933): 'The role of infiltration in the hydrological cycle' Trans Amer. Geophys. Union, 14:446-460.

Hughes J.M.R. (1988): Lotic vegetation dynamics following disturbance along the Swan and Apsley Rivers, Tasmania, Australia. Unpublished PhD thesis, Department of Geography, University of Tasmania.

IEA (1987): Australian Rainfall and Runoff: A Guide to Flood Estimation, 1, (The Institution of Engineers, Australia, Barton, ACT)

Institute of Hydrology (1980): Low Flow Studies (Institute of Hydrology, Wallingford, U.K.)

Kinhill Engineers Pty Ltd (1988): Techniques for Determining Environmental Water Requirements - A Review Technical Report Series No 40, (Department of Water Resources, Victoria)

Lane E.W. (1955): 'The importance of fluvial morphology in hydraulic engineering' ASCE Proc., 81: No. 745.

Langford K.J. and R.A. Lewis (1975): 'Detection of change in sequences of hydrology data' Hydrology Symposium, Armadale. The Institution of Engineers, Australia, Nat. Conf Pub. 75/3:80-84.

Laurenson E.M. (1962): Hydrograph synthesis by runoff routing. University of New South Wales, Water Research Lab., Report No. 66.

Leopold L.B. and J.P. Miller (1956): Ephemeral streams - Hydraulic factors and their relation to the drainage net. U.S Geol. Surv. Prof. Paper 282A.

Maheshwari B.L., K.F. Walker and T.A. McMahon (1993): The Impact of Flow Regulation on the Hydrology of the Murray River and its Ecological Significance. Centre for Environmental Applied Hydrology, Department of Civil and Environmental Engineering, University of Melbourne and River Murray Laboratory, Department of Zoology, University of Adelaide, 154pp.

Maidment D.R.(1993): Handbook of Hydrology (McGraw-Hill Inc. New York).

McMahon T.A. (1976): 'Preliminary estimation of reservoir storage for Australian streams' Civil Engg. Trans., The Institution of Engineers, Australia, CE18:55-59.

McMahon T.A. (1977): Some statistical characteristics of annual streamflows of northern Australia. Hydrology Symposium, Brisbane. The Institution of Engineers, Australia, Nat.Conf. Pub. 77/5:131-135.

McMahon T.A. (1978): 'Australia's surface water resources: a potential development based on hydrologic factors' Civil Engg. Trans., The Institution of Engineers, Australia , CE20:155-164.

McMahon T.A. (1979): 'A review of data estimation procedures and associated errors' J. Hydrol. 41:1-10.

McMahon T.A. (1982): 'World hydrology: Does Australia fit?' Hydrology and Water Resources Symposium, Institution of Engineers Australia , Nat. Conf. Pub. 82/3:1-7

McMahon T.A. (1985): Flooding in Australia - yet another review. Proc. Ninth Invitation Symposium, Australian Academy of Technological Sciences (Aust. Acad. of Tech. Sci. Ltd.) 83-125.

McMahon T.A. and A. Diaz Arenas (1982): Methods of Computation of Low Streamflow, (UNESCO, Paris) 95pp.

McMahon T.A. and D.L.R. Cronin (1980): 'An approach to marginal economic analysis of hydrometric data collection' Water Resources Bull. 16:414-420.

McMahon T.A. and R.G. Mein (1986): River and Reservoir Yield (Water Resources Publications, Littleton, Co) 368pp.

McMahon T.A., B.L. Finlayson, R. Srikanthan and A.T. Haines (1992): Global Runoff: Continental Comparisons of Annual Flows and Peak Discharges. (Catena Verlag, Cremlingen-Destedt, West Germany) 166pp.

Milhouse R.T., D.L. Wegner and T. Waddle (1984): User's Guide to the Physical Habitat Simulation System (US Fish and Wildlife Service, Fort Collins).

Mitchell D.S. and J.A. Rogers (1985): 'Seasonality/aseasonality of aquatic macrophytes in Southern Hemisphere inland waters' Hydrobiologia 125:137-150.

Mosley M.P. (1979): 'Prediction of hydrologic variables from channel morphology, South Island Rivers' Jour. Hydrology (N.Z.) 18:109-120.

Nathan R.J. and P.E. Weinmann (1993): Low Flow Atlas for Victorian Streams, Drought Management Plan for Victoria's Water Resources, Long Term Planning Guidelines Background Report - L3 (Department of Conservation and Natural Resources, Victoria) 163pp.

Nathan R.J. and T.A. McMahon (1990a): 'Evaluation of automated techniques for base flow and recession analyses' Wat. Res. Res. 26 7:1465-1473.

Nathan R.J. and T.A. McMahon (1990b): 'Identification of homogeneous regions for the purposes of regionalisation' J. Hydrol. 121:217-238.

Nathan R.J. and T.A. McMahon (1990c): The Estimation of Low Flow Characteristics and Yield from Small Ungauged Rural Catchments. AWRAC Research Project Completion Report 85/105, (Department of Civil and Enviromental Engineering, University of Melbourne, Australia)

Peel M. (1993): Annual runoff variability in Australia. Unpublished BSc(Hons) Thesis, Department of Geography, University of Melbourne.

Renard K.G. (1970): The Hydrology of Semi-arid Rangeland Watersheds. (US Dept. of Agriculture, Agric. Res. Service), ARS41-162.

Richards K. (1982): Rivers: Form and Process in Alluvial Channels, (Methuen, London) 358pp.

Serventy V. (1985): The Desert Sea: The Miracle of Lake Eyre in Flood. (Macmillan, Melbourne).

Srikanthan R. and T.A. McMahon (1984): 'Hydrologic year: how should it be defined for analysis of annual streamflow volumes?' Fourth Congress, APD-IAHR, Chang Mai, Thailand, pp. 1235-1247.

CHAPTER 5

Transport of reacting solutes in rivers and streams

Robert L. Runkel and Kenneth E. Bencala

5.1 Introduction

In this chapter we discuss the major processes affecting solutes in rivers and streams. Here a *solute* is generally defined as any substance or entity that is transported downstream by the flowing waters. Under this definition, solutes may be pollutants, such as pesticides and hydrocarbons, or naturally occurring substances such as dissolved gases, nutrients, and trace elements. Study of the processes affecting solutes is an important undertaking, as pollutants may pose a threat to public health when the affected waterbody is used as a water supply or for recreational purposes. Another important issue is the effect of pollutants on the aquatic organisms that inhabit the stream ecosystem.

Naturally occurring solutes are also of interest as they often interact with pollutants, thereby intensifying or diminishing the toxicity of the pollutant. In addition, human activity may affect the concentrations of naturally occurring solutes, leading to the degradation of water quality. A classic example of this is the depressed dissolved oxygen levels that are frequently observed downstream from waste water treatment plants (WWTPs). These low oxygen levels are a direct result of the biochemical oxygen demand exerted by the nutrients contained in the WWTP effluent. Nutrients also play a role in eutrophication, the process wherein abnormally high nutrient concentrations cause excessive plant growth. Anthropogenic causes of eutrophication include failing septic systems, agricultural runoff, and WWTP discharges. A final example is the elevated concentrations of trace elements observed in streams that are acidic due to the effects of acid deposition and acid mine drainage.

A variety of processes determine the fate and transport of solutes within rivers and streams. To a casual observer, the most obvious processes are those which cause solutes to move downstream and to mix with other constituents in the water column.

V. P. Singh (ed.), Environmental Hydrology, 137–164.
© 1995 *U.S. Government.*

These *physical processes* are readily apparent to anyone who has ever observed the journey of a fallen leaf as it floats upon the water surface. Closer inspection of the problem indicates that other processes are at work. These additional processes are the chemical and biological reactions that influence solute concentrations. While it is interesting to study individual processes, it is important to consider the complex interaction of the physical, chemical and biological processes. Physical transport characteristics, for example, often determine the residence time of a given solute in the system under study. The residence time effectively determines the time scales over which chemical and biological processes have an effect. The various processes are therefore coupled, such that many complex interactions must be considered within the solute transport problem.

The basic solute transport problem may be formulated in one of two ways. One approach is to state the problem as that of prediction; given the solute concentration at a specific point in time and space, what are the solute concentrations at various locations downstream at some future time? A second approach is to state the problem in a descriptive manner; given a time series of solute concentrations along the stream, what are the dominant processes controlling solute concentrations? Consideration of either question requires a clear definition of the problem. Problem definition begins with a determination of the primary solute of interest. A second task is to define the secondary solutes that interact with the primary solute through chemical and biological processes such as co-precipitation, sorption and oxidation-reduction. The next step is to determine the potential sources of solute mass for the set of solutes under study. Given this information, the dominant processes are elucidated. A final task is to define the environmental forcing functions such as temperature, pH and meteorological variables that influence the rates of the dominant processes.

The goal of this chapter is to describe the concepts necessary to solve the problem described above. Many of these concepts are drawn from our research on the hydrology and geochemistry of small mountain streams. This is an appropriate focus as many of the concepts presented are equally applicable to larger rivers and streams. Our discussion also has relevance to larger systems in that small, upland streams may significantly affect the water quality of larger river systems located downstream. Finally, our focus on small streams highlights the scale dependence of the solute transport problem. In the discussion that follows, we are concerned with streams in which flow is predominantly uni-directional. This one-dimensional approach is frequently used for the analysis of solute transport in small streams, and is often applied to larger streams.

This chapter begins with a description of the physical mechanisms that influence solute transport. These physical processes affect both conservative, or inert, solutes, as well as nonconservative, or reactive, solutes that are affected by chemical and biological processes. This discussion is an appropriate starting point as a thorough understanding of physical transport is needed before one may consider the complex interaction between the hydrologic system and the biogeochemical regime. A second topic is the spatial and temporal distribution of the sources that contribute solute mass. This is important topic as spatial and temporal variability affects the manner in which we address the solute transport problem. Traditional analyses of solute transport have focused on point sources of pollution

such as outfalls from WWTPs. Many current environmental issues, however, concern distributed sources which are nonpoint in space and/or continuous in time. We conclude the chapter with a brief overview of some basic chemical processes controlling solute concentrations. For simplicity, biological processes are not considered herein.

Within the context of the specific topics described above, we wish to make two important points. First, study of solute transport in streams entails more than a study of the stream itself; a thorough study must consider the relationship of the stream to the surrounding watershed, and the associated hydrologic and geochemical implications. Second, streams are complex ecosystems in which a variety of processes control solute concentrations. Consideration of any one process tells only part of the story, such that the system must be considered as a whole. This holistic view requires explicit consideration of the time-scales over which the various processes influence solute behavior. Finally, we wish to state that our discussion of many of the topics described herein is necessarily brief. More complete descriptions may be found in one of several textbooks on water quality modeling and management (Fischer et al., 1979; Reckhow and Chapra, 1983; Chapra and Reckhow, 1983; Tchobanoglous and Schroeder, 1985; Thomann and Mueller, 1987).

5.2 Physical Transport Processes

Physical processes play an important role in determining the fate of solutes in surface water environments. For example, consider the release of dye into the center of a free-flowing stream. Upon entering the stream, the dye forms a cloud that moves downstream with the flowing waters. This process, known as *advection*, describes the downstream transport of solute mass at the mean flow velocity. As the dye moves downstream, small-scale mixing processes cause the dye to spread out, or disperse, thereby increasing the volume and decreasing the concentration of the dye-containing water. This spreading process is called *dispersion*.

Dispersion is attributed to both molecular diffusion and velocity variations caused by shear stress. Molecular diffusion occurs due to the random motion of particles, as illustrated in Figure 5.1. Figure 5.1a depicts a rectangular tank that is split in half by an impermeable divider. A soluble substance is added to the left half of the tank, while the right half of the tank contains distilled water. At some initial time, t_0, the molecules of the substance are uniformly distributed such that the concentrations at any two locations in the left half of the tank are identical. The remaining panels in the figure show the distribution of molecules after the divider is removed from the tank. Figure 5.1b shows the tank at an intermediate time ($t > t_0$), when 25 percent of the molecules have entered the right half of the tank. Given sufficient time ($t \gg t_0$), the random motion of particles results in a uniform distribution of molecules, as shown in Figure 5.1c.

Molecular diffusion may be a significant mixing process in waters with low flow velocities, such as quiescent lakes and groundwaters. The role of molecular diffusion is often diminished in streams and rivers, however, as velocity variation becomes the dominant mixing mechanism. This mixing arises due to the effects of friction or shear stress on the velocity profile. Figure 5.2a shows a vertical velocity

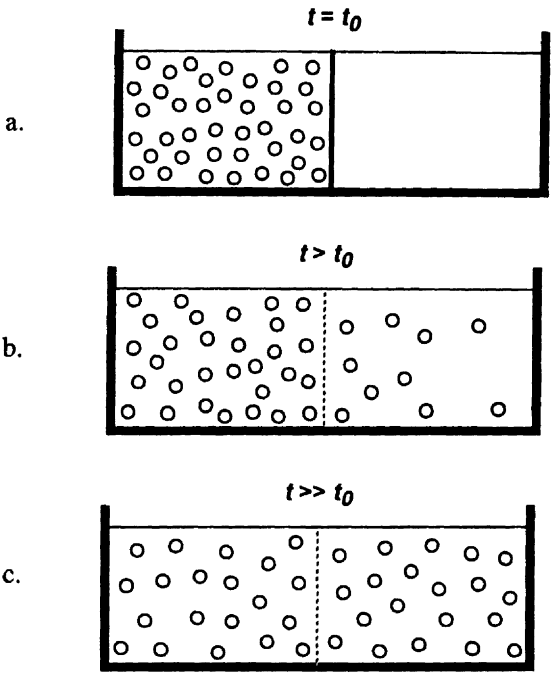


Fig. 5.1. Rectangular tank depicting molecular diffusion. a) At time t_0, an impermeable divider separates the left side of the tank, containing molecules of a soluble substance, and the right side of the tank, containing distilled water. b) After the divider is removed, random particle motion causes molecules to make their way into the right side of the tank. c) After a long period of time, random particle motion results in a uniform distribution of molecules throughout the tank.

profile of a typical stream. Due to the rough nature of the streambed, longitudinal velocity is at its minimum near the streambed. Velocity generally increases with distance from the streambed, but may decrease near the water surface due to the effect of wind shear. Velocity also varies in the horizontal direction, as given by the aerial view shown in Figure 5.2b. In this direction, velocity is at its maximum at the centerline of the stream and decreases as the stream bank is approached. These vertical and horizontal velocity variations cause considerable mixing as particles within the water column travel at different velocities.

The combined effects of advection and dispersion are depicted in Figure 5.3, where the size and location of our hypothetical dye cloud is given for three scenarios. Figure 5.3a shows the appearance of the dye to a stationary observer for the case of advection only. After the dye is placed in the stream at time t_0, the location of the dye cloud changes for times greater than t_0, but the volume of the dye cloud remains constant with time. When dispersion is the only transport process considered, as in Figure 5.3b, the center of the dye cloud remains stationary, but the volume of dye-containing water increases with time. The behavior of the dye in a real stream is most accurately shown in Figure 5.3c, where both advection and dispersion are considered. For this final scenario, the location of the dye changes

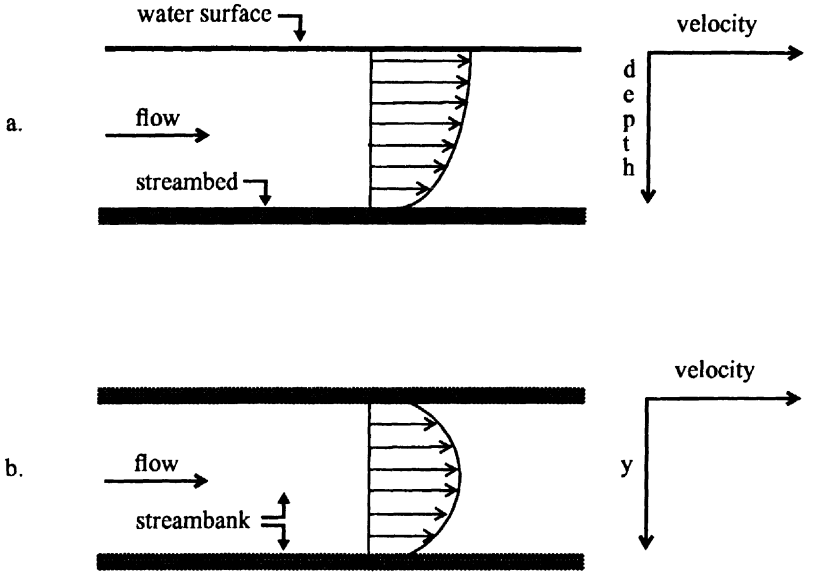


Fig. 5.2. Velocity variations due to shear stress. a) Vertical velocity profile in a stream. b) Horizontal velocity profile in a stream.

due to advection, and the volume of the dye-containing water increases due to dispersion.

An equation describing the spatial and temporal effects of advection and dispersion on solute concentration is presently developed using a fundamental concept, *conservation of mass*. Conservation of mass simply states that the accumulation of mass in a unit volume of water is equal to the difference between mass entering the volume and mass leaving the volume:

$$\text{Accumulation} = \text{Mass In} - \text{Mass Out}, \tag{5.1}$$

where each term in the equation is expressed in units of mass per time [M/T].[1]

The *mass balance equation* given above is formally developed by considering the fluxes in and out of a *control volume* shown in Figure 5.4. For simplicity, we assume that our stream is spatially uniform, i.e., the flow, cross-sectional area and mixing properties of the stream do not vary in space. Furthermore, we assume *steady flow*, such that the velocity and volume of water do not change with respect to time. Finally, we consider only inflows and outflows in the x direction, neglecting fluxes across the y and z planes. In doing so we implicitly assume that solute concentrations vary only in the downstream (x) direction and that solute mass is uniformly distributed over the stream's cross-sectional area (Fischer et al., 1979). Although this uniformity rarely occurs in nature, our one-dimensional approach is reasonable for streams of small to moderate width and depth.

[1]For the remainder of this chapter, physical quantities are described using the fundamental units of mass [M], length [L] and time [T].

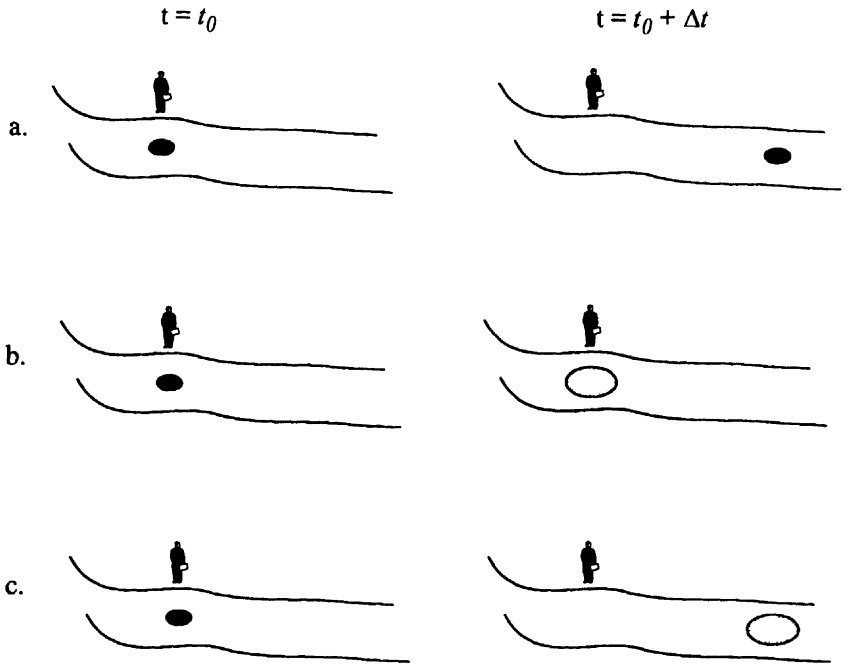


Fig. 5.3. Effects of advection and dispersion on a cloud of dye. a) the case of advection only; the location of the dye cloud changes with time but its volume is constant. b) the case of dispersion only; the center of the dye cloud is stationary but the volume increases. c) combined effect of advection and dispersion; the location and volume of the dye cloud changes with time.

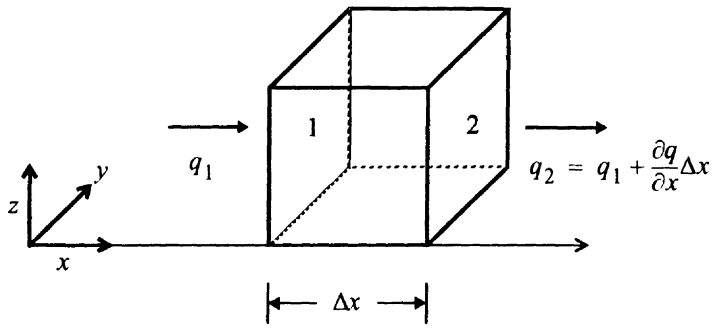


Fig. 5.4. Unit volume of water (control volume) used to develop the conservation of mass equation. A mass balance equation is developed by considering the fluxes into the volume across surface 1 and the fluxes out of the volume across surface 2. Fluxes in the y and z direction are not considered.

The first term in Equation (5.1) describes the accumulation of mass or the change in mass with respect to time. Using differential calculus, this is given by:

$$\text{Accumulation} = \frac{\Delta m}{\Delta t} = \frac{\partial m}{\partial t}, \tag{5.2}$$

where m is mass and t is time. Noting that mass equals concentration times volume and employing our constant volume assumption yields:

$$\text{Accumulation} = V\frac{\partial C}{\partial t}, \tag{5.3}$$

where V is volume [L^3] and C is solute concentration [M/L^3]. The right-hand side of Equation (5.1) is developed by considering the *flux* of solute mass across surfaces 1 and 2 denoted in Figure 5.4. The flux, q, is defined as the mass of solute crossing a unit area per unit time. The flux into the control volume is denoted by q_1 and flux out of the control volume is denoted by q_2. Note that q_2 equals the flux coming into the control volume (q_1) plus the change in flux through the control volume:

$$q_2 = q_1 + \frac{\partial q}{\partial x}\Delta x, \tag{5.4}$$

where Δx is the length of the control volume [L].

We now consider the individual fluxes due to advection and dispersion. The advective flux into the control volume (across surface 1) equals the product of the advective velocity, U [L/T], and the solute concentration at surface 1, C_1:

$$\text{Flux In}_{\text{adv}} = q_{1\,\text{adv}} = UC_1. \tag{5.5}$$

Employing Equation (5.4), the advective flux out of the control volume (across surface 2) is given by:

$$\text{Flux Out}_{\text{adv}} = q_{2\,\text{adv}} = UC_2 = UC_1 + U\frac{\partial C}{\partial x}\Delta x, \tag{5.6}$$

where C_2 is the solute concentration at surface 2.

Fluxes due to dispersion are developed by considering a fundamental law governing *diffusive* mass transport. *Fick's Law of Diffusion* states that mass flux due to molecular diffusion is proportional to the concentration gradient, $\partial C/\partial x$. By analogy, this law may be used to describe the dispersive mass flux. This flux is given by:

$$q_{\text{disp}} = -D_L\frac{\partial C}{\partial x}, \tag{5.7}$$

where D_L is a proportionality constant known as the *longitudinal dispersion coefficient* [L^2/T]. Dispersive flux into and out of the control volume is thus given by:

$$\text{Flux In}_{\text{disp}} = q_{1\,\text{disp}} = -D_L\left.\frac{\partial C}{\partial x}\right|_1, \tag{5.8}$$

$$\text{Flux Out}_{\text{disp}} = q_{2\,\text{disp}} = -D_L\left.\frac{\partial C}{\partial x}\right|_2 = -\left[D_L\left.\frac{\partial C}{\partial x}\right|_1 + D_L\frac{\partial^2 C}{\partial x^2}\Delta x\right]. \tag{5.9}$$

A differential equation corresponding to Equation (5.1) is now assembled using the accumulation and flux terms developed above. Equations (5.3), (5.5), (5.6), (5.8) and (5.9) are combined to give:

$$\underbrace{V\frac{\partial C}{\partial t}}_{\text{Accumulation}} = \underbrace{\left[AUC_1 - AD_L\frac{\partial C}{\partial x}\bigg|_1\right]}_{\text{Mass In}} - \underbrace{\left[AUC_1 + AU\frac{\partial C}{\partial x}\Delta x - AD_L\frac{\partial C}{\partial x}\bigg|_1 - AD_L\frac{\partial^2 C}{\partial x^2}\Delta x\right]}_{\text{Mass Out}}, \quad (5.10)$$

where A is the stream's *cross-sectional area* [L^2]. Because flux is specified on a per unit area basis, each flux is multiplied by A to obtain the units used in Equation (5.1) [M/T]. Using the relation $V = Ax$, Equation (5.10) is simplified to yield:

$$\frac{\partial C}{\partial t} = -U\frac{\partial C}{\partial x} + D_L\frac{\partial^2 C}{\partial x^2}. \quad (5.11)$$

This equation is the *Advection-Dispersion Equation* with constant coefficients (i.e., U and D_L are constant in space and time). The advection-dispersion equation describes the spatial and temporal variation in solute concentration, C, subject to the assumptions given previously. An explicit equation for C as a function of space and time is given in the section that follows.

Although advection and dispersion are the dominant physical processes affecting solutes in many surface waters, other physical processes may be of importance for a given water body. For example, many small streams and rivers gain significant amounts of water from *lateral inflows*. As defined here, lateral inflow is any water that is added to the stream due to groundwater inflow, overland flow, interflow or via small springs and seeps. As a result of lateral inflows, the volumetric flowrate of the stream, Q [L^3/T], increases in the downstream direction. It is important to note that these lateral inflows are not only a source of water to the stream, but also a source of solute mass. For a given stream reach, the lateral inflow concentration, C_L, represents the average solute concentration associated with the waters added by lateral inflow. If C_L is less than the instream concentration, C, lateral inflow has the effect of diluting the instream concentration. Conversely, if the lateral inflow concentration exceeds the instream concentration, the lateral inflow waters act to increase the instream concentration. As shown below, terms may be added to the advection-dispersion equation to account for the effects of lateral inflow.

Another physical process of importance in some systems is that of *transient storage* (Bencala and Walters, 1983). Transient storage has been noted in many small streams, where solutes may be temporarily detained in small eddies and stagnant zones of water that are stationary relative to the faster moving waters near the center of the channel. In addition, significant portions of the flow may

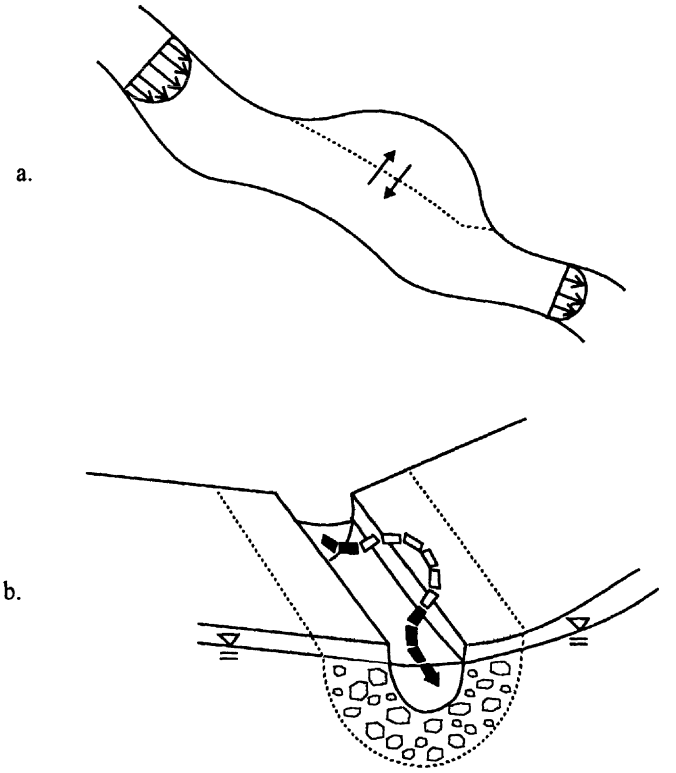


Fig. 5.5. Transient Storage occurs when solutes enter small pockets of slow moving water (a) and when solutes travel through the porous media that makes up the bed and banks of the channel (b).

move through the coarse gravel of the streambed and the porous areas within the stream bank. The travel time for solutes carried through these porous areas may be significantly longer than that for solutes travelling within the water column. These pockets of water and porous areas of the streambed are the two areas contributing to transient storage, as shown in Figure 5.5.

The effects of transient storage are illustrated by examining data from a tracer experiment shown in Figure 5.6 (Broshears et al., 1993). In a tracer experiment, a conservative tracer is injected into the stream and samples are taken at various downstream locations. Shortly after the injection begins, the concentration of the injected solute increases at each sampling site (Figure 5.6a). As the injection continues, the tracer concentration reaches a plateau level that is maintained until the injection is shut off. After the injection ceases, the tracer concentration gradually declines to its pre-injection level.

In Figure 5.6b, a rectangular area is superimposed over the tracer data. This rectangular pulse indicates the tracer concentration that would result if tracer mass was transported downstream by advective processes alone. Departures from this hypothetical concentration profile are due to the effects of mixing. Analysis of tracer data using the advection-dispersion equation derived previously accounts for some of the observed mixing through the specification of the longitudinal

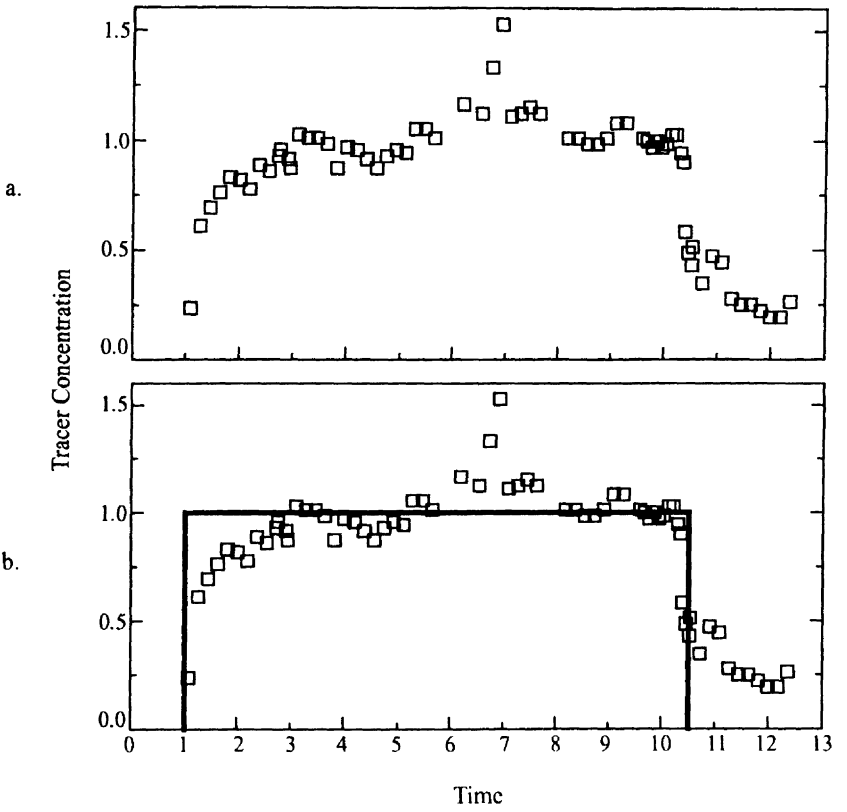


Fig. 5.6. Experimental tracer data illustrating the effect of transient storage. a) tracer concentration at a downstream sampling location. Tracer concentration increases after the injection begins, reaches a plateau and decreases after the injection terminates. b) rectangular pulse showing hypothetical tracer concentration in the absence of dispersion and transient storage. Observed concentrations differ from the hypothetical case due to dispersion and transient storage. (Adapted from Broshears et al., 1993).

dispersion coefficient. The general shape of the concentration profile cannot be described by this analysis alone, however, and additional mixing due to transient storage must be considered.

Two characteristic time periods of the observed profile are of interest. The first period relates to the arrival of the tracer pulse. As indicated by the rectangular region, the tracer concentration instantaneously increases to the plateau concentration (1.0) in a purely advective system. The arrival of the tracer pulse is delayed for the real stream, however, as indicated by the gradual increase in tracer concentration. This delayed arrival is due to dispersion and the movement of tracer mass into the transient storage areas. During this initial time period, the transient storage zones act to reduce tracer mass in main channel, as the tracer makes its way into the streambed and the slow moving pockets of water. A second time period of interest corresponds to the cessation of the tracer injection. Given a purely advective system, the tracer concentration drops instantaneously from the plateau concentration to background concentration, as indicated by the right

edge of the rectangular region. In the real stream, this decrease is more gradual, as indicated by the concentration tail observed in the data. During this time period, the tracer concentration in the storage zones generally exceeds that in the main channel, and tracer mass bleeds out of the storage zones producing the observed tail. In summary, the transient storage zones act as a sink during the initial stages of the tracer injection and a source of tracer during the latter stages.

To conclude our discussion on physical transport mechanisms, we briefly revisit the advection-dispersion equation given as Equation (5.11). Our derivation of this equation considers the processes of advection and dispersion, but neglects the effects of lateral inflows and transient storage. In addition, our assumptions result in an equation that is restricted to the case of spatially constant parameters and a steady flow regime. Most of us realize, however, that real streams are not spatially uniform, as the cross-sectional area varies over short distances and structural features such as boulders and tree stumps influence mixing processes. Furthermore, instream flowrates vary in time due to the effects of rainfall, snowmelt, evapotranspiration and the appropriation of water. Fortunately, a derivation similar to that provided above may be used to develop a general transport equation that encompasses the effects of advection, dispersion, lateral inflow and transient storage, while eliminating the restrictions of spatially and temporally constant parameters. This derivation produces:

$$\frac{\partial C}{\partial t} = -U\frac{\partial C}{\partial x} + \frac{1}{A}\frac{\partial}{\partial x}\left(AD_L\frac{\partial C}{\partial x}\right) + \frac{q_L}{A}(C_L - C) + \alpha(C_S - C), \quad (5.12)$$

where q_L is the lateral inflow rate per unit length of stream [L^3/T-L] and α is a first order mass-transfer coefficient controlling the rate of exchange between the main channel and the transient storage zone [/T]. Note that Equation (5.12) also introduces C_S, the average solute concentration in the storage zone [M/L^3]. Assuming waters in the storage zone are immobile relative to waters in the main channel, an equation describing C_S is given by (Bencala and Walters, 1983):

$$\frac{dC_S}{dt} = \alpha\frac{A}{A_S}(C - C_S), \quad (5.13)$$

where A_S is the cross-sectional area of the storage zone [L^2].

5.3 Sources of Solutes

Up until this point we have been concerned with the physical processes affecting solutes after they enter the stream. In this section we examine the pathways in which solutes enter the surface water environment. In our earlier example, a quantity of dye was added to a small stream or river. This dye was added at a specific point in space, and the addition occurred over a specific time period. Thus the addition of dye has both spatial and temporal components. Specifically, the dye addition represents a *point source* and an *instantaneous* release. As we see in this section, sources of solutes vary in both time and space. If the solutes of interest

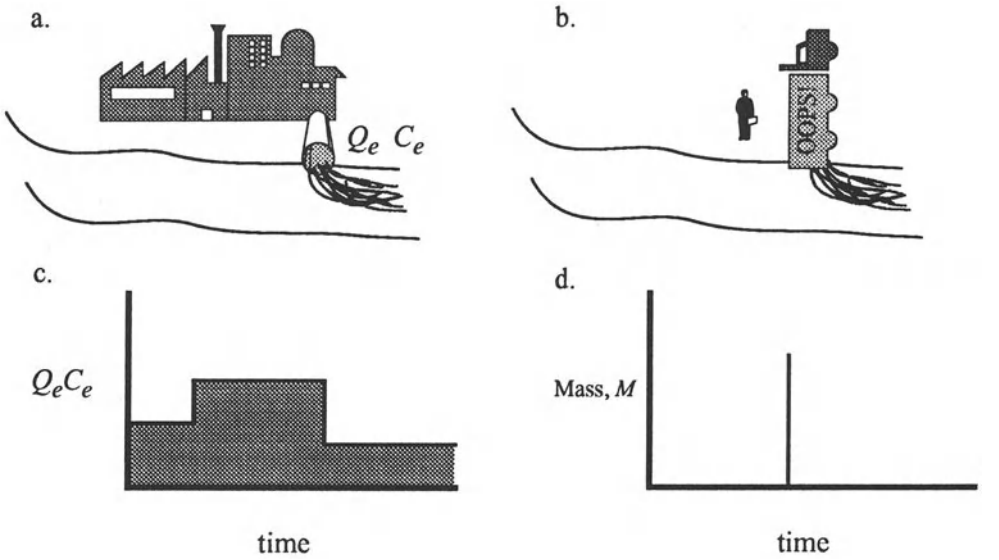


Fig. 5.7. Point sources of solutes. a) Discharges from municipal wastewater treatment plants and industrial facilities. b) Accidental spills from hazardous carriers. c) Point sources such as wastewater treatment plants (Figure 5.7a) may be continuous sources. d) Point sources such as accidental spills (Figure 5.7b) may be viewed as instantaneous sources.

are environmental contaminants, this variability may be an important factor in determining the ultimate impact of the contaminants on the stream environment.

From a spatial perspective, solutes enter surface waters through *point* and *nonpoint* sources. As the name implies, point sources introduce solutes at well defined locations along the watercourse. Examples of point sources include discharges from municipal wastewater treatment plants and industrial facilities (Figure 5.7a), and spills that occur due to accidents, when transporters of hazardous materials release contaminants into a waterbody (Figure 5.7b). From the standpoint of environmental monitoring, point sources are relatively easy to quantify. Mass loading, for example, may be estimated by measuring the flow and solute concentration associated with a plant's effluent (Figure 5.7a):

$$W = Q_e C_e, \tag{5.14}$$

where W is mass loading rate [M/T], Q_e is the volumetric flowrate for the point source and C_e is the solute concentration in the effluent. Calculations of mass loading rates are often more difficult than Equation (5.14) suggests, however, as Q_e and C_e frequently vary in time. The mass loading rate may be incorporated into the mass balance equations presented in the foregoing section.

Nonpoint sources are sources of solutes that are distributed along the watercourse. These sources are often diffuse, in that mass entering at any one point in space is relatively small, yet the aggregate mass loading rate is significant. An example of a nonpoint source is agricultural runoff that enters a stream as overland flow (Figure 5.8a). During precipitation events, runoff from cultivated fields may contain pesticide residues and fertilizer, as well as suspended sediments that sorb

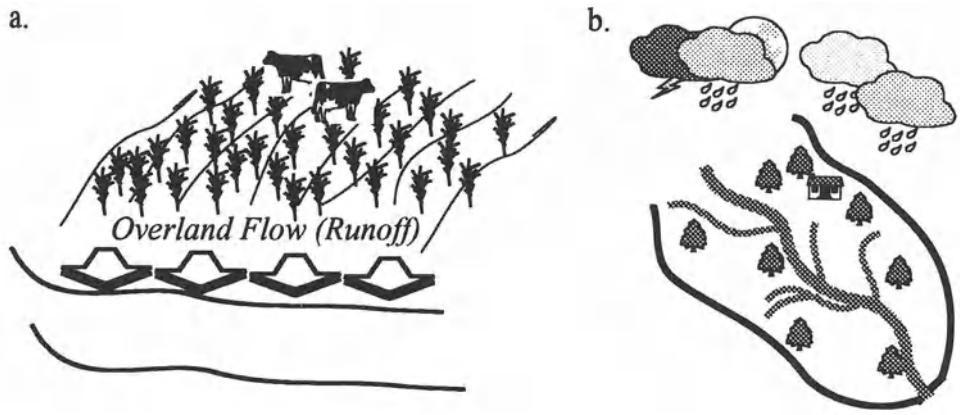


Fig. 5.8. Examples of distributed or nonpoint sources. a) Runoff containing pesticide residues and fertilizers from cultivated lands. b) Acid deposition.

contaminants. Loading due to this type of nonpoint source is represented by the lateral inflow terms in the general transport equation (q_L and C_L of Equation (5.12)). Another example of a nonpoint source is that of acid deposition (Figure 5.8b). Due to the burning of fossil fuels, rainfall in some areas of the United States is acidic (Drever, 1988). In addition, acidic particulate matter may deposit on the land surface. This loading may impact a watershed, leading to the acidification of surface waters.

Because they are not associated with a specific point in space, mass loading rates for nonpoint sources are often difficult to quantify. The lateral inflow solute concentration (C_L), for example, may vary dramatically between different sampling locations along the stream (Bencala and McKnight, 1987). For the case of acid deposition, rain gages distributed throughout the watershed may indicate different rates of mass loading. Due to these complications, care must be taken when designing sampling programs to monitor nonpoint sources of pollution.

Solute sources and loading rates also vary with respect to time. Sources may be roughly classified according to their duration. Some sources are *continuous* sources while others may be viewed as being *instantaneous*. Continuous sources introduce solutes to the stream for extended periods of time. An example of a continuous source is the wastewater treatment plant shown in Figure 5.7a. Although mass loading rates may vary in time (Figure 5.7c), many of these plants continuously discharge effluent into receiving waters, resulting in a continuous addition of solutes to the stream.

Instantaneous sources add solute mass to the stream over very short time periods. Although truly instantaneous sources do not exist, situations do arise in which solutes are added to the stream over time intervals that are short relative to the time-scales of interest. The addition of dye discussed in the previous section is one example of an instantaneous source. Another example is an accidental spill such as that depicted in Figure 5.7b, where contaminants enter a stream in a matter of minutes. For the purpose of analysis, this loading can be viewed as the instantaneous addition of mass at a discrete point in time (Figure 5.7d).

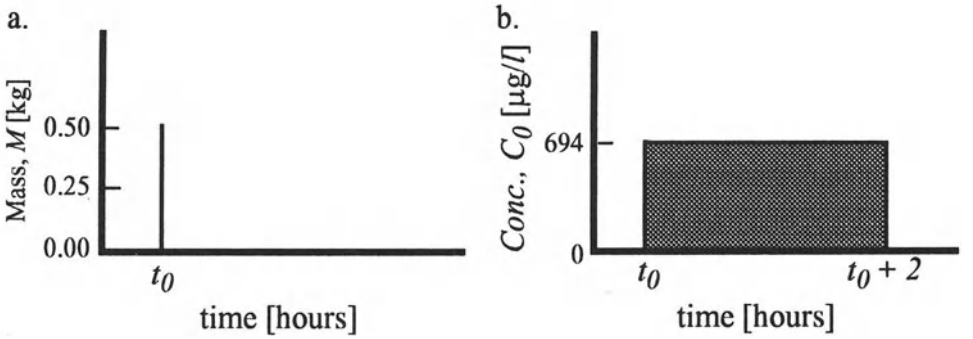


Fig. 5.9. Mass loading scenarios for a hypothetical release. a) Spill scenario is viewed as an instantaneous release of 0.5 kg of mass. b) Leak scenario; mass is released continuously over a two hour period. Concentration in the vicinity of the release is 694 µg/l.

In order to illustrate the significance of temporal variation, we consider the accidental release of a toxic substance into a small stream. Using our knowledge of physical transport mechanisms, we wish to determine the impact of the release on a fish population 500 meters downstream. The stream is fairly uniform throughout the 500 meter reach, having a cross-sectional area (A) of 0.75 m^2, a volumetric flowrate (Q) of 0.1 m^3/s, and a longitudinal dispersion coefficient (D_L) of 5 m^2/s.

Two cases are of interest: a 'spill scenario', where the toxic substance is instantaneously released, and the 'leak scenario', where the substance continuously enters the stream for a specified period of time. For the spill scenario, we consider the instantaneous release of 0.5 kg of the toxic substance (Figure 5.9a). For the leak scenario, we assume that the same amount of mass (0.5 kg) is uniformly released over a two hour time period. This continuous release results in an instream concentration of 694 µg/l (micrograms/liter) near the site of the release, that may be modeled as a rectangular input (Figure 5.9b).

These two cases may be analyzed using the advection dispersion equation (Equation (5.11)). Using the methods of calculus, *analytical solutions* are developed that provide the predicted solute concentration as a function of time and space. For a spill, the solution to Equation (5.11) is given by (Thomann and Mueller, 1987):

$$C(x,t) = \frac{M}{2A\sqrt{\pi D_L t}} \exp\left[\frac{-(x-Ut)^2}{4D_L t}\right], \tag{5.15}$$

where M is the mass of the contaminant in the spill and x is distance in the downstream direction. For a continuous rectangular input the solution to Equation (5.11) is given by Fischer et al., (1979):

$$C(x,t) = \frac{C_0}{2}\left[\text{erfc}\left(\frac{x-Ut}{2\sqrt{D_L t}}\right) + \exp\left(\frac{Ux}{D_L}\right)\text{erfc}\left(\frac{x+Ut}{2\sqrt{D_L t}}\right)\right] \quad \text{for } t < \tau, \tag{5.16}$$

where C_0 is the solute concentration after mixing over the cross-section of the stream, τ is the duration of the continuous input [T] and 'erfc' is the complimentary

error function. The complimentary error function may be evaluated using tabulated values (e.g., Eagleson, 1970; Freeze and Cherry, 1979) or numerical techniques (Press et al., 1986). Note that Equation (5.16) provides the solution during the duration of the continuous input, i.e., for all $t < \tau$. The principle of superposition may be used to develop the solution for all time periods after the termination of the continuous input (for $t > \tau$). This yields:

$$C(x,t) = \frac{C_0}{2}\left[\text{erfc}\left(\frac{x - Ut}{2\sqrt{D_L t}}\right) - \text{erfc}\left(\frac{x - U(t-\tau)}{2\sqrt{D_L(t-\tau)}}\right)\right.$$
$$\left. + \exp\left(\frac{Ux}{D_L}\right)\left\{\text{erfc}\left(\frac{x + Ut}{2\sqrt{D_L t}}\right) - \text{erfc}\left(\frac{x + U(t-\tau)}{2\sqrt{D_L(t-\tau)}}\right)\right\}\right]$$
$$\text{for } t > \tau. \tag{5.17}$$

These solutions may be applied to our problem to determine the effects of advection and dispersion on solute concentration. Note that these equations do not consider any reactions the substance may undergo after entering the stream. As such, we are considering the worst case scenario, in which the toxic substance is transported conservatively and not allowed to degrade. This solution therefore provides an upper bound on the solute concentrations that are likely to be realized within the stream.

The results of our analysis are shown in Figure 5.10. The solute concentrations at 500 meters for the instantaneous (Equation (5.15)) and continuous (Equations (5.16) and (5.17)) cases are given in Figure 5.10a. Two features of the concentration profiles illustrate the marked effect of the time component on solute concentrations and the corresponding implications for our fish population. First, the peak solute concentration for the instantaneous release is approximately double that of the peak for the continuous release (1.4 versus 0.7 mg/l). This difference may be important if the concentration at which fish are adversely impacted lies between 1.4 and 0.7 mg/l. Second, the duration of elevated toxicant concentrations is longer for the case of the continuous release. Fish are exposed to concentrations exceeding 0.3 mg/l for almost twice as long for the continuous release scenario. The relative impact of the release under the two different loading scenarios is therefore dependent on the acute and chronic toxicities of the particular substance for the fish species of interest.

Due to the effects of dispersion, the differences attributable to the duration of the release diminish in the downstream direction. This is shown in Figure 5.10b, where solute concentrations are given for a site 4000 meters downstream. At this location, peak concentrations and exposure durations are approximately equal for the instantaneous and continuous release scenarios.

This example illustrates the effects of source variability and physical processes on the fate of solutes in the stream environment. In the analysis above, we have assumed conservative conditions, such that only the physical processes of advection and dispersion are considered. In many situations, however, the concentration of a released solute may decrease due to sorption, oxidation/reduction, volatilization,

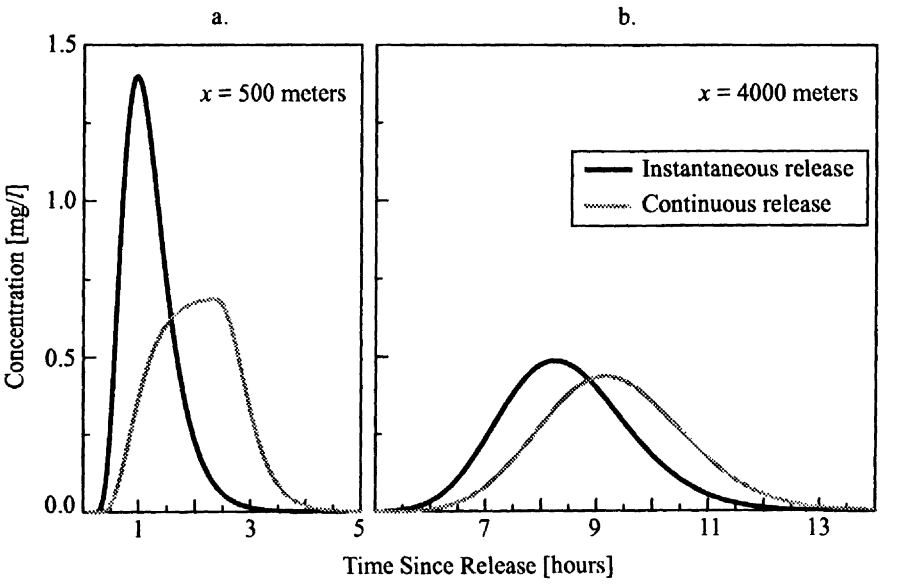


Fig. 5.10. Predicted solute concentrations at two downstream sampling sites. a) At 500 meters, the peak concentration for the instantaneous release is approximately double that of the peak for the continuous release. The duration of exposure is longer for the continuous release. b) At 4000 meters, there is little difference between the concentration profiles for the instantaneous and continuous releases.

hydrolysis, biodegradation and other biochemical processes. These processes are the topic of the following section.

5.4 Geochemical Processes

Thus far we have focussed our discussion on the strictly physical processes that affect solute transport. Although physical processes play a large role in determining the fate of solutes, chemical and biological processes may be equally important. In this section we turn our attention to some of the basic geochemical reactions affecting solutes in natural waters. Biological processes are not considered herein.

To illustrate the importance of geochemical reactions, consider the fate of a dissolved chemical species. Dissolved species may take part in both *homogeneous* and *heterogeneous* chemical reactions. Homogeneous reactions are those reactions in which dissolved species interact with other dissolved species. These reactions are homogenous as they involve two or more species of the same phase, i.e., all the species involved in the reaction are in dissolved form. Examples of homogeneous reactions include acid-base, complexation and hydrolysis reactions.

A dissolved species may also take part in heterogenous reactions. In contrast to homogenous reactions, heterogenous reactions involve species from more than one phase. In a heterogenous reaction, a dissolved phase species interacts with gas or solid phase species. Dissolved oxygen concentrations, for example, may be controlled by the oxygen-containing atmosphere overlying the surface water. If

dissolved oxygen concentrations are *undersaturated* relative to the gas phase, oxygen will diffuse across the air/water interface and dissolved oxygen concentrations will increase. Conversely, if dissolved oxygen concentrations are *oversaturated*, oxygen will be lost to the atmosphere and dissolved oxygen concentrations will decrease. This process is known as *degassing*. Other examples of heterogenous reactions include solid-phase reactions such as precipitation/dissolution and sorption.

In addition to the reactions noted above, solute uptake and transformation by aquatic organisms is also of importance. Thus we see that there are a variety of processes that may potentially affect the concentration of a given solute. The importance of any particular process is highly dependent on the solute of interest. Interaction with the atmosphere, for example, is often an important process controlling dissolved oxygen concentrations, whereas it is a relatively unimportant process for trace metals. Given this solute-dependence, it is necessary to limit the scope of our discussion to a particular class of solutes. The discussion that follows is therefore based on our experience with inorganic solutes (e.g., salts and trace metals). Organic solutes, including naturally occurring dissolved organic carbon, petroleum-based hydrocarbons, and pesticides, are not explicitly considered. Note, however, that many of the processes and concepts discussed below pertain to organic as well as inorganic constituents. Many excellent reviews of the processes specific to organic solutes may be found in the organic chemistry and groundwater contaminant transport literature (Lyman et al., 1982; Moore and Ramamoorthy, 1984; Princeton University Water Resources Program, 1984; Thurman, 1985; Devinny et al., 1990). An interesting framework for the analysis of organic fate and transport in surface waters is presented by Chapra (1991).

With these points in mind, we now consider some basic geochemical processes. We begin with a discussion of the rates at which reactions occur and two fundamental paradigms for classifying reactions. This is followed by a brief discussion of two reaction classes: precipitation/dissolution and sorption/desorption. The section concludes with a discussion of the role of pH and the coupling between hydrologic and geochemical processes. Discussion of each of these topics is necessarily limited. For additional information, the reader is referred to one of several texts on aquatic chemistry (Snoeyink and Jenkins, 1980; Stumm and Morgan, 1981; Drever, 1988; Pankow, 1991; Morel and Hering, 1993).

5.4.1 Rates of Reaction

Geochemical reactions in natural waters may be studied using one of two fundamental approaches. Selection of the appropriate approach for a given problem is often based on the rates at which the relevant reactions occur, relative to the time scale of interest. If reaction rates are 'slow' relative to the study time scale, a *kinetic* approach is employed, wherein the dynamics of the system are studied. If reaction rates are relatively 'fast', an *equilibrium* approach is implemented to determine the *steady-state* composition of the waters. As explained below, other factors, such as data availability, play a role in determining the appropriate approach.

All geochemical reactions require some finite amount of time, τ_c, to reach completion. The required amount of time is dependent on the rate of reaction, r,

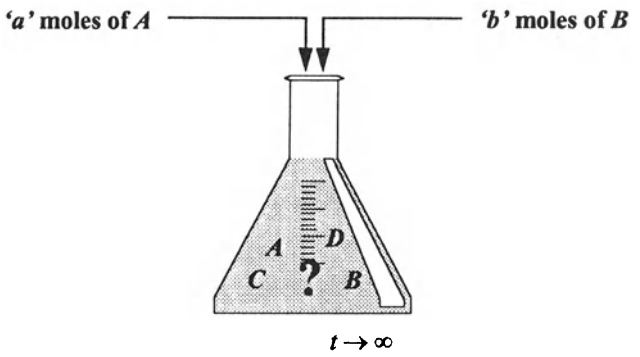


Fig. 5.11. Hypothetical chemical equilibrium problem. If a moles of A and b moles of B are added at an initial time, what are the species concentrations (A, B, C and D) as $t \rightarrow \infty$?

such that τ_c and r are inversely related (i.e., if the magnitude of r is small, the reaction proceeds slowly and τ_c is large). A kinetic approach is often implemented when τ_c exceeds the time period of observation, τ_o. For example, τ_c for many biologically-mediated oxidation/reduction reactions is on the order of days to years. If our interest is in the state of the system over the course of a day, a kinetic approach is needed to describe system dynamics, i.e., how the concentrations of the chemical species change over time. In this case, differential equations may be written to describe the temporal variation in species concentrations and kinetic rate constants (r) may be empirically determined. The equations may then be solved to yield the species concentrations over time. Examples of the kinetic approach are given by Morel and Hering (1993).

Although the kinetic approach is a useful paradigm for studying dynamic systems, it is not well-suited from many practical geochemical problems. As noted by Drever (1988), our knowledge of the kinetics governing many geochemical reactions is incomplete. As a result, kinetic rate information is unavailable for many reactions. In addition, some reactions (e.g., acid-base and complexation reactions) occur at very high rates, such that a study of kinetics is not relevant to the time scales of natural systems. Finally, our observational time period may exceed the time required for the reactions to reach completion ($\tau_o > \tau_c$).

For these reasons, an alternative approach is frequently used to study the system under conditions of chemical equilibrium. The *chemical equilibrium* approach differs from the kinetic approach in that the *steady-state* or equilibrium condition of the system is of interest, rather than the state of the system over time. Consider a reversible chemical reaction in which species A reacts with species B to form species C and D:

$$aA + bB \underset{r_b}{\overset{r_f}{\rightleftharpoons}} cC + dD, \tag{5.18}$$

where r_f is the forward rate of reaction, r_b is the backwards rate of reaction, and a, b, c and d are the number of moles of species A, B, C and D, respectively. Application of the equilibrium approach to this hypothetical reaction is depicted

in Figure 5.11. For this reaction, the chemical equilibrium problem may be stated as follows: "Given an initial state where a moles of A and b moles of B are added to a chemical system, what are the species concentrations (A, B, C and D) at equilibrium?". Here equilibrium is defined as the state of the system at steady-state (as $t \rightarrow \infty$), given constant environmental conditions such as temperature and pressure.

At equilibrium, the forward rate of reaction equals the backwards rate of reaction and thermodynamic principles may be used to formulate *mass action equations*. The mass action equation corresponding to Equation (5.18) is given by:

$$\frac{\{C\}^c\{D\}^d}{\{A\}^a\{B\}^b} = \frac{k_f}{k_b} = K_{eq}, \tag{5.19}$$

where k_f is the rate constant for the forward reaction, k_b is the rate constant for the backward reaction, K_{eq} is the equilibrium constant for the overall reaction, and the bracketed (i.e., $\{\}$) quantities denote chemical activities. Chemical activities may be considered 'effective concentrations' and are related to observed concentrations using *activity coefficients* (Drever, 1988). The equilibrium constant therefore defines the concentration ratio of the four species at equilibrium. Mass action equations may be written for all the reactions of interest. These equations are combined with mass balance equations and the resultant system of equations is solved for the equilibrium species concentrations.

Although natural systems are never truly in a state of equilibrium, the equilibrium approach is appropriate for many practical geochemistry problems. As noted above, many reactions occur at very high rates such that the assumption of equilibrium is warranted. For those reactions occurring at slower rates, the equilibrium approach provides valuable insight into system behavior. Equilibrium calculations may be compared with observed species concentrations to determine the degree of disequilibrium. As shown below, these comparisons allow us to determine the direction in which reactions are proceeding.

5.4.2 Precipitation/Dissolution

For inorganic solutes, one of the more dominant chemical reactions is that of precipitation/dissolution. Within a watershed, soil and rock formations near the land surface are subject to weathering reactions that are a source of solutes to surface waters. Limestone deposits, for example, react with water molecules to release both calcium and carbonate. Metallic ores are also subject to weathering reactions and are a source of trace metals such as iron, aluminum and zinc. These trace metals undergo hydrolysis reactions with water to form hydrous metal oxides. Examples of hydrous metal oxides include *ferrihydrite*, $Fe(OH)_3$, and *gibbsite*, $Al(OH)_3$.

In order to illustrate the process of precipitation/dissolution, we consider ferrihydrite, a solid phase that is often found on the bed of acidic mountain streams (McKnight and Bencala, 1989; Kimball et al., 1992). Ferrihydrite forms when free ferric iron, Fe^{3+}, combines with hydroxide, OH^-:

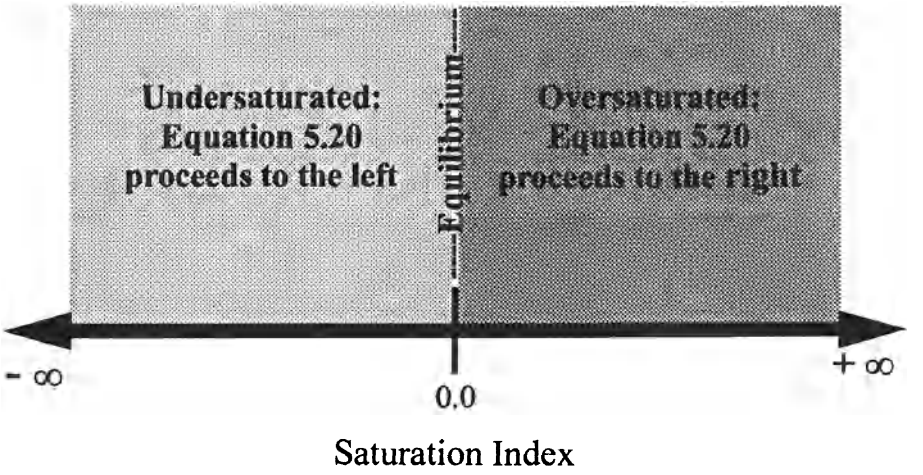


Fig. 5.12. Interpretation of the Saturation Index (SI). The solution may be undersaturated (SI < 0), oversaturated (SI > 0), or at equilibrium (SI = 0), with respect to the controlling solid.

$$Fe^{3+} + 3OH^- \Longleftrightarrow Fe(OH)_3(s). \tag{5.20}$$

At equilibrium, this reaction is represented using a mass action equation:

$$\{Fe^{3+}\}\{OH^-\}^3 = K_{sp}, \tag{5.21}$$

where K_{sp} is the *solubility product* for $Fe(OH)_3(s)$. Note that the solubility product is simply an equilibrium constant (K_{eq}, Equation (5.19)) describing the formation of a solid phase. Recall from our previous discussion that chemical activity is related chemical concentration using activity coefficients. Equation (5.21) therefore gives the concentrations of the dissolved species, Fe^{3+} and OH^-, that may be in equilibrium with the ferrihydrite solid phase. If the product of the Fe^{3+} and OH^- activities exceeds K_{sp}, the solution is said to be oversaturated and *precipitation* of ferrihydrite from the water column may occur. Conversely, if the product is less than K_{sp}, the solution is undersaturated. If a source of ferrihydrite is present (such as ferrihydrite coating the streambed), undersaturated conditions may result in the *dissolution* of ferrihydrite.

The degree of over or under saturation is quantified by calculating the *saturation index* (SI), a simple function of the *ion activity product* (IAP) and the solubility product. The ion activity product is given by:

$$\{Fe^{3+}\}\{OH^-\}^3 = IAP, \tag{5.22}$$

where the product on the left-hand side of Equation (5.22) refers to ambient conditions and the product on the left-hand side of Equation (5.21) reflects equilibrium conditions. As such, the ion activity product reflects the current state of the system and the solubility product represents the theoretical equilibrium state. The saturation index is given by:

$$SI = \log(\frac{IAP}{K_{sp}}). \tag{5.23}$$

Interpretation of the saturation index is depicted in Figure 5.12. If the ambient water concentrations (the IAP), exceed the equilibrium concentrations (the K_{sp}), the saturation index is positive, indicating oversaturation. Given oversaturated conditions, the reaction given by Equation (5.20) will tend to proceed to the right. Similarly, if the aqueous concentrations are less than those predicted at equilibrium, the saturation index is negative and the solution is undersaturated. Under these conditions, the reaction given by Equation (5.20) will tend to the left. Finally, a saturation index of zero indicates that the aqueous solution is in equilibrium with the solid-phase, and the forward and backward reactions given in Equation (5.20) are in equilibrium.

5.4.3 SORPTION/DESORPTION

Another important process controlling the concentrations of solutes in natural waters is that of *sorption*. Sorption may be generally defined as the process in which a dissolved species becomes associated with a solid surface. A dissolved species taking part in a sorption reaction is known as a *sorbate*, and the solid species or sorbing phase is known as the *sorbent*. When the dissolved species interpenetrates the sorbent, the process is termed *absorption*, whereas *adsorption* is used to describe the interaction of the dissolved species with the surface or interface of the sorbent (Weber et al., 1991).

Several physico-chemical mechanisms may be responsible for the sorption process. For example, a dissolved species may form a site specific complex with a sorbent phase. The formation of a chemical complex results in the previously dissolved species being strongly bound to the sorbent. Another potential mechanism is the *electrostatic attraction* that occurs between molecules of opposite charge. Some solid species, such as clays, have a net negative charge that attracts positively-charged metal species. A final mechanism is that of *hydrophobic binding*, wherein nonpolar molecules with a low solubility have an affinity for the solid phase. The specific mechanism responsible for sorption phenomena observed in a given system is a function of both the type of sorbate and the type of sorbent, and in many instances more than one mechanism is at work.

Within natural waters, solutes may sorb to both inorganic and organic sorbents. Inorganic sorbents include alumino-silicates (e.g., red clays), hydrous oxides of aluminum, iron and manganese, and carbonates (Morel and Hering, 1993). Organic sorbents include matter from dead or living organisms and natural organic matter (NOM) that may coat the surface of other solids. In addition to the distinction between inorganic and organic sorbents it is important to note the location of the sorbing phase. Sorption may occur within the water column as dissolved species interact with suspended particles. Sorption may also occur directly on the streambed where the solid phase is abundant. In this latter case, sorption removes solute mass from the water column and the dissolved species is no longer subject to hydrologic transport.

Sorption may have a significant effect on dissolved inorganic species within the water column. As noted by Morel and Hering (1993), trace elements such as zinc and copper are known to sorb to hydrous ferric oxides (e.g., ferrihydrite). Sorption onto hydrous oxides may be facilitated by the presence of NOM coatings

on the oxide surfaces that impart a negative charge to the sorbent. Other examples of sorption in natural waters include the research of Bencala et al. (1983) in which the sorption of strontium and potassium was observed, and the work of Kuwabara et al. (1984) wherein copper was found to sorb to the bed of a small stream.

Analysis of sorption reactions is often in the form of *sorption isotherms*. Sorption isotherms are mathematical expressions that relate the mass of solute sorbed to the dissolved solute concentration, for a chemical system at equilibrium. A common approach is to employ the *Freundlich isotherm*:

$$S_e = K_f C_e^n, \tag{5.24}$$

where S_e is the mass of solute sorbed per mass of sorbent [M/M], C_e is the dissolved solute concentration [M/L^3], K_f is the Freundlich isotherm constant [(L^3/M)^n], and n is a dimensionless exponent. Note that S_e and C_e are measured when the aqueous system is in equilibrium, i.e., when the rates of sorption and desorption are equal. The logarithmic form of Equation (5.24) is given by:

$$\log(S_e) = \log(K_f) + n \log(C_e). \tag{5.25}$$

The two model parameters, K_f and n, represent the sorption capacity and sorption intensity, respectively (Weber et al., 1991). Parameter estimates are obtained through the analysis of experimental data using Equation (5.25). In a typical experiment, a known quantity of the sorbent phase and solute are added to a completely mixed batch reactor (Tchobanoglous and Schroeder, 1985). After a sufficient amount of time has passed for the reactions to reach equilibrium, measurements of S_e and C_e are obtained. The experiment is repeated a number of times over a range of initial conditions to yield a set of S_e-C_e pairs that may be graphed on a log-log plot. If the sorption process under study adheres to the Freundlich model, plotted values will form a straight line with slope n and intercept $\log(K_f)$ (Equation (5.25)). Given the model parameters, Equation (5.24) provides a means of calculating the mass of solute sorbed at equilibrium.

For the special case of $n = 1$, Equation (5.24) reduces to the linear form of the Freundlich equation:

$$S_e = K_d C_e, \tag{5.26}$$

where K_d is known as the *distribution coefficient* [L^3/M]. The distribution coefficient is a physically meaningful parameter as it is a simple ratio of the amount sorbed (S_e) to the amount in the dissolved phase (C_e), at equilibrium. In addition to the Freundlich isotherm, other formulations are possible, such as the Langmuir isotherm.

Although isotherms provide a descriptive framework for the analysis of sorption process, recent research has focused on developing sorption models that are more strongly based in theory. An example of this research is the development of *surface complexation models* (Dzombak and Morel, 1990) that consider both the chemical reactions between solutes and functional surface groups (chemical bonding) and electrostatic interactions. Additional details on sorption isotherms and surface complexation modeling are provided by Morel and Hering (1993).

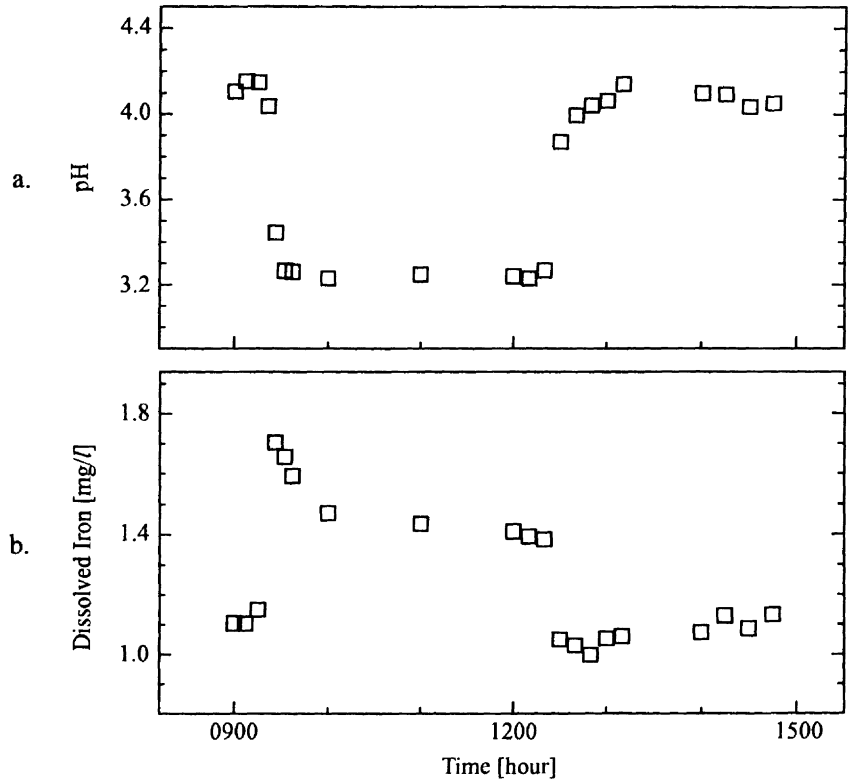


Fig. 5.13. pH modification experiment on the Snake River. a) pH of the stream was temporarily lowered from 4.2 to 3.2. b) dissolved iron concentrations increased in response to the depression in pH.

5.4.4 pH, THE MASTER VARIABLE

The pH of a solution is defined as the negative logarithm of the hydrogen ion activity, $pH = -\log(\{H^+\})$. All of the geochemical reactions discussed above are affected by pH. In order to illustrate the effect of solution pH on reaction chemistry, we now present some experimental data from the Snake River, a naturally acidic stream located in the Rocky Mountains of Colorado. Surface waters become acidic when additions of acid exceed the buffering capacity of the carbonate system. Anthropogenic sources of acidity include acid deposition and acid mine drainage. In addition, some sources of acidity arise naturally due to the oxidation of sulfide bearing ores.

Acidic waters are of concern as their low pH increases the solubility and mobility of trace metals. As described by McKnight and Bencala (1989), a pH modification experiment was conducted on the Snake River to assess the affect of pH on trace metal mobility. The experiment involved the injection of a strong acid to temporarily lower the pH of the stream. Figure 5.13a depicts the changes in pH in response to the acid injection. At an observation point located 628 meters downstream from the injection, instream pH values decreased from approximately

4.2 to 3.2. After 3 hours, the injection was terminated, and instream pH values returned to 4.2.

The effect of the pH decrease on dissolved iron concentrations is shown in Figure 5.13b. As shown in the figure, iron concentrations increased dramatically as the pH decreased. Much of this increase is attributable to the dissolution of hydrous iron oxides that coat the streambed. This increase in iron is to be expected in light of the mass action equation governing ferrihydrite precipitation/dissolution (Equation (5.20)). In this equation, free ferric iron (Fe^{3+}) activities are related to hydroxide (OH^-) activities to the third power. As a result, small changes in pH cause relatively large changes in dissolved ferric iron.[2]

The above example illustrates the effect of pH on the process of precipitation/dissolution. Many other reactions, including sorption/desorption and oxidation/reduction, are dependent on solution pH. Examples of the effect of pH on sorption are provided by Morel and Hering (1993). Due to the importance of pH, it is often referred to as 'the master variable'.

5.5 Coupling of Hydrologic and Geochemical Processes

Thus far we have discussed physical transport mechanisms and geochemical reactions as separate processes. In natural systems, however, these processes are closely coupled. Geochemical reactions determine the distribution of solute mass amongst dissolved, solid and gas phases and thereby affect the amount of mass available for physical transport. Physical transport mechanisms, meanwhile, affect the location of solute mass such that a given parcel of water is subjected to various geochemical regimes as it moves downstream. In this section we examine the coupling of the various processes through several examples.

The coupling between transport and chemistry may be illustrated by revisiting our conceptual classification of geochemistry based on homogenous and heterogenous reactions. This distinction between reaction types becomes important when considered within the context of physical transport. To begin, we divide the heterogenous reactions into three subclasses. Heterogenous I reactions are those reactions in which a dissolved species interacts with a mobile solid species. Examples of Heterogenous I reactions include the formation of a precipitate in the water column due to oversaturated conditions, and the sorption of a dissolved species onto a particulate solid within the water column. Heterogenous II reactions are those in which a dissolved species interacts with a stationary solid species. An example of a Heterogenous II reaction is the sorption of a dissolved species onto a stationary solid such as particles coating the streambed or debris in the channel. Finally, Heterogeneous III reactions are those in which a dissolved species interacts with a gas phase, such as the degassing of dissolved oxygen to the atmosphere.

The effect of the various reaction types on physical transport is depicted as a continuum in Figure 5.14. At one extreme of the spectrum, geochemical reactions do not affect the amount of mass that is subject to physical transport. Such is the case for homogenous reactions in which dissolved species interact with other

[2]pH is related to the hydroxide activity through the mass action equation describing the dissociation of water: $\{H^+\}\{OH^-\} = K_w$, where K_w is an equilibrium constant.

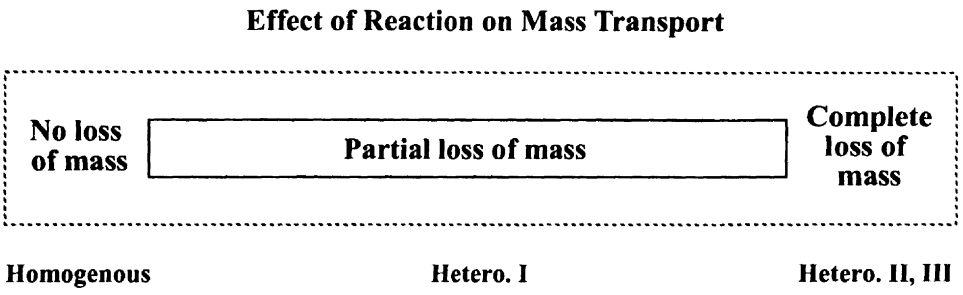


Fig. 5.14. Effect of reaction on mass transport, depicted as a continuum.

dissolved species. In this case, all of the reactants remain in the water column and the net amount of mass subject to transport does not change as a result of the reaction. At the other extreme, geochemical reactions result in a complete loss of mass such that all of the mass involved in the reaction is no longer transported downstream. Heterogenous II and Heterogenous III reactions fall at this end of the spectrum. For Heterogenous II reactions, mass involved in the reaction is deposited on stationary surfaces and is no longer transported. Likewise, Heterogeneous III reactions involve degassing; mass that leaves the stream via the air/water interface is no longer affected by the flowing waters.

The final reaction type, Heterogenous I, falls within the endpoints of the spectrum as it results in a partial loss of mass. For a Heterogenous I reaction, dissolved species interact with solids within the water column and continue to move in the downstream direction. Some of the mass involved in the reaction will be lost, however, as the force of gravity causes the solid to settle to the streambed over time. This loss is dependent on the nature of the solid. The surface area, surface charge and mass of some solids will be such that the losses due to settling are negligible. In this case, the reaction lies to the left end of the spectrum and the net effect is similar to that for a homogeneous reaction. Other solids are such that the rate of settling is significant, resulting in a substantial loss of mass.

The example presented above illustrates the effects of geochemical reactions on transport. As noted in the beginning of this section, this coupling between processes is bidirectional; transport processes also affect the degree to which chemical reactions take place. For example, consider an acidic stream such as the Snake River in which considerable concentrations of dissolved iron are present in the water column. As a parcel of water moves downstream, it may mix with waters of a higher pH, as with the case where tributary inflows add nonacidic waters to the stream. As a result of this mixing, the pH of the water will increase, and the solubility product for ferrihydrite may be exceeded, resulting in the formation of particulate iron in the water column. Given sufficient time and constant pH, these particulates will settle to the streambed and discontinue their downstream movement. Alternatively, the particulates may move farther downstream and encounter additional inflows that are acidic. If mixing results in a decrease in pH, the particulates may redissolve rather than settling to the streambed. Transport processes thus

result in a dynamic system wherein a parcel of water may be subject to a variety of geochemical regimes over time.

Quantitative analysis of the coupling between physical transport and geochemical reaction may be considered by adding terms to the advection-dispersion equation that represent the effects of chemistry. Numerous formulations of this approach are possible given the different reactions under study and the relative rates of chemical reaction and hydrologic transport. As an example, we consider the sorption of solutes on to the bed of a stream. Two modified formulations of the constant parameter advection-dispersion equation (Equation (5.11)) are presented. Both formulations assume reversible sorption reactions that are aptly described by the linear form of the Freundlich isotherm (Equation (5.26)). First, consider the case where the time scales of reaction are such that sorption may be modeled as an equilibrium process. The corresponding advection-dispersion-reaction equation is given by:

$$\frac{\partial C}{\partial t} = -U\frac{\partial C}{\partial x} + D_L\frac{\partial^2 C}{\partial x^2} - \rho\frac{\partial S}{\partial t}, \tag{5.27}$$

where S is the mass of solute sorbed per mass of sorbent [M/M], ρ is the mass of sorbent per volume of stream water [M/L^3] and,

$$\frac{\partial S}{\partial t} = K_d\frac{\partial C}{\partial t}. \tag{5.28}$$

The effect of sorption on the dissolved solute concentration (C) is embodied in the last term of Equation (5.27). If sorption is occurring, mass on the streambed will increase ($\partial S/\partial t > 0$) and mass will be lost from the dissolved phase. Conversely, if desorption is occurring, mass on the streambed will decrease ($\partial S/\partial t < 0$) and mass will be added to the dissolved phase. A similar formulation for modeling groundwater systems is presented by Pickens and Lennox (1976).

The above equations are applicable when the rate of sorption/desorption is high, relative to the flow velocity of the stream. As such, sorption/desorption is modeled as an instantaneous process. In many situations the assumption of instantaneous equilibrium is not warranted, due high flow velocities and the fact that not all of the solute mass in the water column has access to the sorbent surfaces on the streambed. This case requires a kinetic approach, as given by the following advection-dispersion-reaction equation:

$$\frac{\partial C}{\partial t} = -U\frac{\partial C}{\partial x} + D_L\frac{\partial^2 C}{\partial x^2} - \rho\lambda(K_d C - S), \tag{5.29}$$

where λ is a first-order rate constant for sorption [/T]. The last term in Equation (5.29) describes the effect of sorption/desorption as a function of the gradient between the mass sorbed at equilibrium, $K_d C$, and the mass present on the streambed, S. If the equilibrium amount exceeds the amount present on the bed ($K_d C > S$), sorption takes place and mass is lost from the dissolved phase at a rate that is proportional to λ. Desorption occurs when the amount present on the bed is greater than the amount that would be present at equilibrium ($S > K_d C$). An example of the kinetic approach for sorption/desorption is presented by Bencala (1983).

5.6 Summary

In this chapter we have discussed many of the processes that effect the fate and transport of solutes in stream ecosystems. We began our discussion with a quantitative look at the physical processes affecting solutes. To broaden our analysis, we also presented several potential geochemical processes that act to transform solute mass and alter species composition. We then described the complex interaction or coupling between the physical processes and the geochemical reactions. In doing so, our purpose was to point out that a variety of processes control solute concentrations, and that emphasis on any one process may result in inaccurate conclusions. Thorough understanding of the system requires an analysis of the coupled processes. Within this chapter, we have presented a primarily qualitative analysis of this coupling. A more rigorous and quantitative analysis requires the use of mathematical models such as those presented by several authors (Chapman, 1982; Bencala, 1983; Wood and Baptista, 1993; Runkel et al., 1995).

Another important consideration is the temporal and spatial variability in the sources of solutes to the stream. This variability contributes to the complexity and dynamics of the system and must be considered within a given solute transport problem. From our study of solute sources, it is evident that the stream is intricately linked with the surrounding watershed. As such, the watershed is a source of both water and solute mass. As pointed out by Bencala (1993), the stream is not merely a pipe conveying water downstream, but rather a complex system that interacts closely with the watershed. Consideration of the entire system is thus of paramount importance.

References

Bencala, K.E. and R.A. Walters, Simulation of solute transport in a mountain pool-and-riffle stream: a transient storage model, Water Resour. Res., 19(3), 718-724, 1983.

Bencala, K.E., A.P. Jackman, V.C. Kennedy, R.J. Avanzino and G.W. Zellweger, Kinetic analysis of strontium and potassium sorption onto sands and gravels in a natural channel, Water Resour. Res., 19(3), 725-731, 1983.

Bencala, K.E. and D.M. McKnight, Identifying instream variability: Sampling iron in an acidic stream, In Chemical Quality of Water and the Hydrologic Cycle, R.C. Averett and D.M. McKnight editors, Lewis Publishers, Ann Arbor, Ch. 15, 255-269, 1987.

Bencala, K.E., A perspective on stream-catchment connections, J. of the N. Amer. Benthological Society, 12(1), 44-47, 1993.

Broshears, R.E., K.E. Bencala, B.A. Kimball, and D.M. McKnight, Tracer-dilution experiments and solute-transport simulations for a mountain stream, Saint Kevin Gulch, Colorado, U.S. Geol. Surv. Water-Resour. Invest., 92-4081, 1993.

Chapman, B.M., Numerical simulation of the transport and speciation of nonconservative chemical reactants in rivers, Water Resour. Res., 18(1), 155-167, 1982.

Chapra, S.C., and K.H. Reckhow, Engineering Approaches for Lake Management, Volume 2: Mechanistic Modeling, Ann Arbor Science, Ann Arbor, Michigan, 1983.

Chapra, S.C., Toxicant-loading concept for organic contaminants in lakes, J. of Envir. Engrg., 117(5), 656-677, 1991.

Devinny, J.S., L.G. Everett, J.C.S. Lu, and R.L. Stollar, Subsurface Migration of Hazardous Wastes, Van Nostrand Reinhold, New York, New York, 1990.

Drever, J.I., The Geochemistry of Natural Waters, Second Edition, Prentice-Hall, Englewood Cliffs NJ, 1988.

Dzombak, D.A., and F.M.M. Morel, Surface Complexation Modeling, Wiley-Interscience, New York, New York, 1990.
Eagleson, P.S., Dynamic Hydrology, McGraw-Hill Book Company, New York, 1970.
Fischer, H.B., E.J. List, R.C.Y. Koh, J. Imberger and N.H. Brooks, Mixing in Inland and Coastal Waters, Academic Press, San Diego, 1979.
Freeze, R.A., and J.A. Cherry, Groundwater, Prentice Hall, Englewood Cliffs, New Jersey, 1979.
Kimball, B.A., D.M. McKnight, G.A. Wetherbee, and R.A. Harnish, Mechanisms of iron photoreduction in a metal-rich, acidic stream (St. Kevin Gulch, Colorado, U.S.A.), Chem. Geol., 96, 227-239, 1992.
Kuwabara, J.S., H.V. Leland, and K.E. Bencala, Copper transport along a Sierra Nevada stream, J. Environ. Engr., 110(3), 646-655, 1984.
Lyman, W.J., W.F. Reehl, and D.H. Rosenblatt, Handbook of Chemical Property Estimation Methods, Environmental Behavior of Organic Compounds, McGraw-Hill, New York, New York, 1982.
McKnight, D.M. and K.E. Bencala, Reactive iron transport in an acidic mountain stream in Summit County, Colorado: A hydrologic perspective, Geochim. Cosmochim. Acta, 53, 2225-2234, 1989.
Moore, J.W., and S. Ramamoorthy, Organic Chemicals in Natural Waters, Springer-Verlag, New York, New York, 1984.
Morel, F.M.M., and J.G. Hering, Principles and Applications of Aquatic Chemistry, Wiley-Interscience, New York, New York, 1993.
Pankow, J.F., Aquatic Chemistry Concepts, Lewis Publishers, Chelsea, Michigan, 1991.
Pickens, J.F. and W.C. Lennox, Numerical simulation of waste movement in steady groundwater flow systems, Water Resour. Res., 12(2), 171-180, 1976.
Press, W.H., B.P. Flannery, S.A. Teukolsky, and W.T. Vetterling, Numerical Recipes: The Art of Scientific Computing, Cambridge University Press, Cambridge, U.K., 1986.
Princeton University Water Resources Program, Groundwater Contamination from Hazardous Wastes, Prentice Hall, Englewood Cliffs, New Jersey, 1984.
Reckhow, K.H., and S.C. Chapra, Engineering Approaches for Lake Management, Volume 1: Data Analysis and Empirical Modeling, Ann Arbor Science, Ann Arbor, Michigan, 1983.
Runkel, R.L., K.E. Bencala and R.E. Broshears, An equilibrium-based simulation model for reactive solute transport in small streams, in U.S.Geological Survey Toxic Substances Hydrology Program – Proceedings of the technical meeting, Colorado Springs, Colorado, September 20-24, 1993, D.W. Morganwalp and D.A. Aronson editors, U.S. Geol. Surv. Water Resour. Invest. Rpt. 94-4015, 1995.
Snoeyink, V.L., and D. Jenkins, Water Chemistry, John Wiley and Sons, New York, New York, 1980.
Stumm, W., and J.J. Morgan, Aquatic Chemistry, Wiley-Interscience, New York, 1981.
Tchobanoglous, G., and E.D. Schroeder, Water Quality, Addison-Wesley, Reading, Massachusetts, 1985.
Thurman, E.M., Organic Geochemistry of Natural Waters, Nijhoff/Junk Publishers, Dordrecht, The Nederlands, 1985.
Thomann, R.V., and J.A. Mueller, Principles of Surface Water Quality Modeling and Control, Harper and Row, New York, 1987.
Weber, W.J., P.M. McGinley, and L.E. Katz, Sorption phenomena in subsurface systems: Concepts, models and effects on contaminant fate and transport, Water Research,499-528, 1991.
Wood, T.M., and A.M. Baptista, A model for diagnostic analysis of estuarine geochemistry, Water Resour. Res., 29(1), 51-71, 1993.

CHAPTER 6

Water and contaminant transport in the vadose zone

Peter J. Wierenga and Mark L. Brusseau

6.1 Introduction

Present interest in the vadose zone is largely a result of concern by the general public about the quality of ground water. People are beginning to understand that drinking water obtained from aquifers can easily become polluted, and that great care is needed to prevent such pollution. Examples of contaminant sources include agricultural chemicals, road salts, and municipal and industrial wastes which are intentionally or accidentally applied to soil from where they leach to the subsoil and into ground water. Contamination of ground water may also occur by contaminants leaking from landfills, disposal sites, leach fields, underground storage tanks, nuclear waste disposal facilities, holding ponds, etc.

The rate at which contaminants move to aquifers depends on the physical and chemical properties of the soil and of the contaminants, on prevailing rainfall or irrigation rates, and on the depth to ground water. The depth to ground water varies greatly between different regions of the country and the world, from a meter or less to several hundreds of meters or more. In areas with deep ground water, the unsaturated zone above the water table, also called the vadose zone, may serve as a temporary buffer between the soil surface and ground water. The effectiveness of this buffer in protecting ground water from contaminantion, even in areas with deep vadose zones, is often questionable.

In this chapter, we will discuss the basic principles of water storage, and of flow and contaminant movement through unsaturated soil. These principles will be discussed relative to characterization of soil properties, and their effects on monitoring water and contaminants in the vadose zone.

V. P. Singh (ed.), Environmental Hydrology, 165–191.
© 1995 *Kluwer Academic Publishers.*

6.2 Water Storage

6.2.1 Water Content

For most common soils, the soil solid phase occupies from 40 to 70% of the total space. The remaining 60 to 30% is pore space filled usually with water and air. In the vicinity of a spill of an immiscible liquid contaminant (e.g., gasoline, cleaning solvent), the pore space may also contain immiscible liquid. If the pore space is completely filled with water (no air present) the soil is called saturated. If the pore space is not completely filled with water, as is the case in the vadose zone, but contains air as well as water, then the soil is said to be unsaturated.

There are two methods to express the status of water in soil. The most common method is to determine the amount, i.e. the volume or the mass of water in soil. The other method is to determine the energy status of water in soil.

The standard method for determining the volume or mass water content of soil is taking samples, drying the samples in an oven, and calculating water contents by differences. (For details, see Klute (1986)). Where repetitive measurements are needed, indirect methods may be used advantageously. The most common indirect method for determining water contents in the vadose zone is the neutron probe method. This method involves lowering a radioactive source emitting fast neutrons through an access tube installed in soil. The presence of neutrons slowed by hydrogen is counted, and is a direct measure of the soil-water content (Klute et al., 1986). Newer methods for measuring soil water in the field employ the dielectric constant of water which is much larger (i.e. 60–80) than the dielectric constant of dry soil (i.e. 2–5). Most promising among these is Time Domain Reflectometry (TDR), which measures the soil-water content in layers of 0–5 cm thickness (Brisco et al., 1992). For a useful review of the TDR for measuring water content see Reeves and Elgezawi (1992).

6.2.2 Energy Relationships

Knowing the amount of water in soil is often not enough. For example, the water content by itself gives no indication whether a soil is saturated or unsaturated. Water content also provides no measure of availability for plant uptake, and is of no value in determining the direction of water flow. For many purposes percent saturation, i.e. the volumetric water content divided by the water content at saturation, is a more useful parameter. However, a complete evaluation of water in soil requires knowledge of the energy status of water in soil.

The energy status of water is expressed through the soil-water potential. The total soil- water potential is formally defined as the mechanical work required to transfer water from soil to a standard reference state, where the total soil-water potential is zero by definition. The units of soil-water potential are those of the reference state. In soil science and hydrology, soil-water potential is mostly expressed in energy/volume (J/m^3; bar) or energy/weight (head, cm or m) (cf., Jury et al., 1992).

The total soil-water potential (ψ_T, in energy/volume units; H, in head units) is more easily understood if it is broken down into its component potentials. For

the vadose zone three important component potentials are: gravitational potential ψ_z (with energy/volume units) or z (with head units); solute potential ψ_s (energy/volume units); and (tensiometer) pressure potential ψ_p (energy/volume units) or h (head units).

The gravitational potential is the potential resulting from the gravitational force field. The solute potential results from the difference in potential energy between soil solution and pure free water. Soil solutions contain a variety of dissolved minerals and thus its energy status is decreased relative to pure water. For example, a soil whose soil solution contains 1000 ppm of dissolved salts (not uncommon) has a solute potential of about -600 cm.

The tensiometer pressure potential (or pressure potential, ψ_p) results from the capillary and other forces in unsaturated soils. Its value is always negative in unsaturated soils, and usually expressed in bars, cm, or equivalent units. In saturated soils capillary forces are zero, but water below the water table surface is subject to pressure from the overlying water. This hydrostatic pressure is often expressed in head units, and is also called pressure head.

Because solute potential does not contribute much as a driving force for water flow (except at the soil-root interface where solute potential differences are most important for water uptake into roots), its contribution is often ignored. Adding the two remaining potentials, and expressing these in terms of head units, one obtains:

$$H = h + z. \tag{6.1}$$

Thus in unsaturated soil the total pressure head H equals the sum of the pressure head (h) and gravity or elevation head (z). The pressure potential or pressure head in unsaturated soils is always negative, becomes zero at the water table and positive below the water table. The absolute value of the negative pressure potential in unsaturated soils is also called tension.

There is a relationship between the amount of water held in soil and its energy status. This relationship is unique for each soil and is called the soil-water characteristic curve or soil-water retention curve. Examples of water retention curves for a loamy fine sand and for a clay loam are shown in Figure 6.1 (Hills et al., 1989). These data illustrate that as soil dries its tension increases (pressure-potential becomes more negative). Differences between the two curves in Figure 6.1 are significant. The clay loam has much higher water contents at saturation, and at intermediate tensions. The water retention curve for the fine sand is very steep at tensions above 10^3 cm.

Two points on the water retention curve are of special interest. They are Field Capacity (FC) and Permanent Wilting Point (PWP). Field capacity is the water content of soil after two days of redistribution following a heavy infiltration of water. In the older literature, field capacity is often defined as the water held at 333 cm tension. In reality, careful measurements of tension in well-watered field soils show that the tension at FC varies from 40 to 60 cm in coarse textured soils, and from 60 to 100 cm in fine textured soils. The permanent wilting point is the soil-water condition at which movement of soil water to plant roots is too slow to keep up with transpiration losses, and plants do not recover upon wilting. The

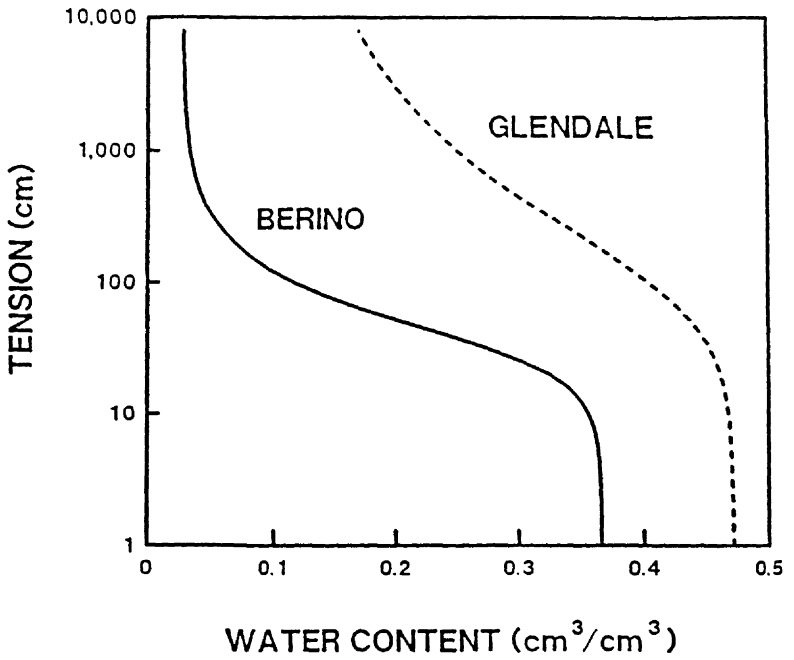


Fig. 6.1. Soil-water retention curves for Berino loamy fine sand and Glendale clay loam.

permanent wilting point is different for different soils and is generally taken around 15 000 cm, but may vary from 10 000 to 40 000 cm tension. Its precise value is of little significance because the water available from soil with a tension greater than 10 000 cm (see Figure 6.1) is usually small, although of sufficient magnitude for survival of many drought resistant plants.

Water retention curves are frequently determined in the laboratory on intact cores or using disturbed soil samples (Klute, 1986). The methodology allows for starting with wet or dry soil. Unfortunately, this makes a difference, and the final retention curves are different depending on whether the procedure was started with wet or dry soil. This behavior is called hysteresis in the water retention curve (Figure 6.2). In field studies, hysteresis is often ignored because water retention data, with hysteresis measured in the field, is generally not available.

Water storage in soils is often estimated by subtracting the water contents at PWP from the water contents at FC, and multiplying by the depth of interest. For the soils in Figure 6.1, the water stored in a 1 m deep profile is given in Table 6.1. It should be noted that field capacity moisture content values available from the literature often do not represent field measured values, but rather laboratory values which may be too low.

Water retention curves can be described by a number of empirical relationships (Brooks and Corey, 1964; van Genuchten, 1980). Recently the van Genuchten approach has found wide use in soil and hydrological studies. The relationship is

$$S_e(h) = \frac{\theta - \theta_r}{\theta_s - \theta_r} = [1 + (\alpha h)^n]^{-m}, \tag{6.2}$$

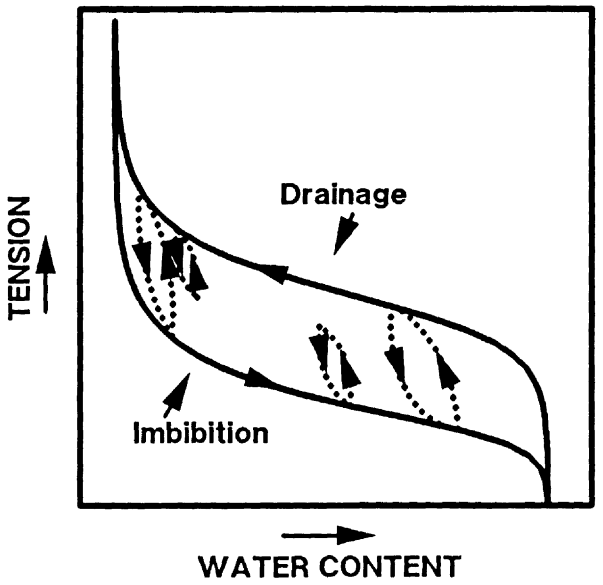


Fig. 6.2. Soil-water retention curves with hysteresis. The dotted lines represent scanning loops.

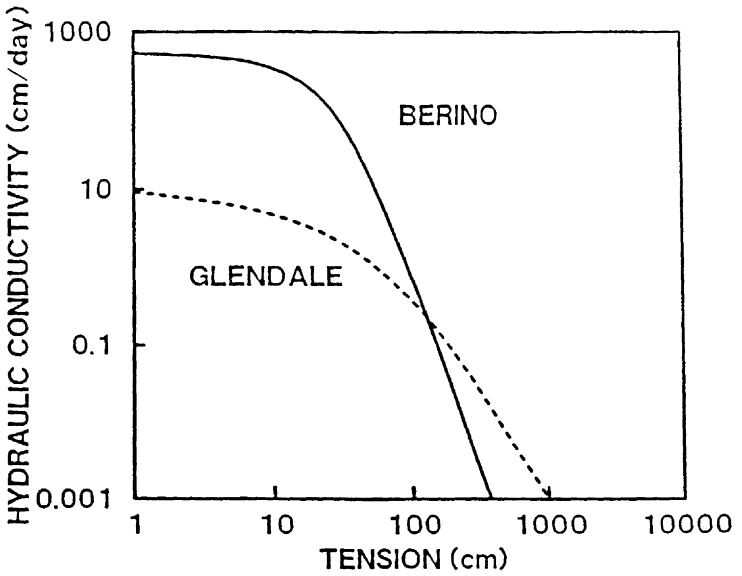


Fig. 6.3. Hydraulic conductivity versus suction for Berino loamy fine sand and Glendale clay loam.

where S_e is the effective saturation, θ is the volumetric water content (cm^3cm^{-3}). θ_r is the residual water content (cm^3cm^{-3}), θ_s is the saturated water content (cm^3cm^{-3}), and h is the tension; α (cm^{-1}), n (dimensionless), and m (dimensionless) are empirical parameters (van Genuchten, 1980). This function fits observed

TABLE 6.1
Volumetric water content (cm^3/cm^3) and water storage (cm per 100 cm soil profile)

	FC (100 cm)	PWP (15000 cm)	Storage (cm H_2O)
Berino fine sand	0.18	0.03	15.1
Glendale clay loam	0.39	0.14	24.9

TABLE 6.2
Van Genuchten parameters for Berino loamy fine sand, and Glendale clay loam. It is assumed then $m = 1 - 1/n$

	θ_r cm^3/cm^3	θ_s cm^3/cm^3	α cm^{-1}	n	K_s cm/day
Berino	0.03	0.37	0.03	2.24	541
Glendale	0.11	0.47	0.01	1.4	13

data well for a wide range of soils (van Genuchten and Nielsen, 1985). The curves in Figure 6.1 were calculated with Equation 6.2, using parameter values presented in Table 6.2 (Hills et al., 1989).

6.3 Water Movement

6.3.1 Steady Flow

Water flow through soil may be viewed microscopically by looking at water flow through individual pores, or macroscopically by looking at flow through the entire porous medium. This latter approach is the more common.

Macroscopic flow of water through soil is described with Darcy's equation

$$q = -K(\theta)\frac{\partial H}{\partial z} . \tag{6.3}$$

In this equation q is the Darcy flux or volume of water flowing through a unit surface area per unit time ($cm^3/cm^2/day$), $K(\theta)$ is the hydraulic conductivity (cm/day), H the total pressure head (cm) and z distance (cm). For saturated soils the hydraulic conductivity is assumed to be independent of θ.

6.3.2 Hydraulic Conductivity

In unsaturated soils, the hydraulic conductivity decreases sharply with decreasing water content. This is because as soil loses water (drains) pores empty, and the cross sectional area for flow decreases. Thus, for soils which drain rapidly, such as sands, the hydraulic conductivity decreases much more sharply with decreasing

water content, than for clay soils in which there are more intermediate-sized pores. These drain more slowly, and thus the hydraulic conductivity of clay soils decreases less sharply with decreasing water content. An example of changes in hydraulic conductivity with tension for the two soils of Table 6.2 is shown in Figure 6.3. Because water content is related to soil-water tension through the soil-water retention curve (Figure 6.1), the hydraulic conductivity is often shown as a function of soil-water tension. Note the exponential decrease in hydraulic conductivity with increasing tension.

There have been many efforts to relate the unsaturated hydraulic conductivity to water content, pressure head or tension (Gardner, 1958; Brooks and Corey, 1964; Campbell, 1974; van Genuchten, 1980). Because of its simplicity, the Gardner equation has been used quite frequently:

$$K(h) = K_{\text{sat}}\, \mathrm{e}^{-\alpha h}, \tag{6.4}$$

where K_{sat} is the saturated hydraulic conductivity (cm/day), h pressure head (cm), and α (cm^{-1}) a parameter, equal to the slope of the unsaturated log conductivity function. Values for K_{sat} and α must be determined by fitting Equation (6.4) to measured values of unsaturated conductivity. By plotting $\ln K(h)$ versus h one obtains α and K_{sat} from the slope and intercept, respectively. The empirical value of K_{sat} so obtained is not necessarily equal to the measured saturated conductivity. In practice, the log-linear relationship of Equation (6.4) is most applicable to the moist soil-water content range (e.g. $-100 < h\ (\text{cm}) < 0$).

A more versatile equation was proposed by van Genuchten (1980). The equation can be derived from the water content-pressure head relationship presented by Equation (6.2). The equation is:

$$K(S_e) = K_{\text{sat}}(S_e)^{\gamma}[1 - (1 - S_e^{1/n})^m]^2. \tag{6.5}$$

The symbols in Equation 6.5 were defined before, and are the same as for Equation (6.2). The parameter γ is often taken to be 0.5 (Mualem, 1976), although Hudson (1992) found that γ was greater than 2 for the Berino soil whose characterisitic curve is depicted in Figure 6.1. For Equation (6.5) to be valid it was also assumed that $m = 1 - 1/n$. The solid and dashed lines in Figure 6.3 were computed with Equation (6.5), using parameter values from Table 6.2.

Equations (6.4) and (6.5) show a strong dependence of $K(\theta)$ on water content. As a result of this dependency, a relatively small decrease in water content may reduce the hydraulic conductivity 10 to 100 fold, or more. This fact is important to consider when one wants to extract soil solution from an unsaturated soil. A high vacuum applied suddenly to a porous sampler may dry out the soil immediately around the sampler, and strongly reduce flow into the sampler.

6.3.3 TRANSIENT FLOW

Equation (6.3) is satisfactory for steady flow through soil. For transient conditions, i.e. when water contents or pressure heads change over time, Equation (6.3) needs to be combined with the equation of continuity. This equation is:

$$\frac{\partial \theta}{\partial t} = -\frac{\partial q}{\partial z}, \tag{6.6}$$

where t is time. The other symbols are as defined before. Combining Equations (6.3) and (6.6) yields the Richards equation:

$$\frac{\partial \theta}{\partial t} = \frac{\partial}{\partial z}\left[K(h)\left(\frac{\partial h}{\partial z} + 1\right)\right]. \tag{6.7}$$

Equation (6.7) is for vertical flow without sources or sinks (plant roots). By defining the inverse of the slope of the water retention function as the water capacity function,

$$C(h) = \frac{\mathrm{d}\theta}{\mathrm{d}h}, \tag{6.8}$$

one may write Equation (6.7) in its pressure head equivalent. Thus:

$$C(h)\frac{\partial h}{\partial t} = \frac{\partial}{\partial z}\left[K(h)\left(\frac{\partial h}{\partial z} + 1\right)\right]. \tag{6.9}$$

Analytical solutions of Equations (6.7) and (6.9) are available for some well defined conditions. However, given the complexity of most practical situations (variable upper boundary conditions, layered soils, uptake of water by roots), numerical solutions are most commonly used. Such solutions are available for one, two or three dimensions, with several of these coupled to solute transport models. A large number of the models have been described in the literature (van der Heijde, 1993), but unfortunately, very few, well documented unsaturated zone models are available in the public domain. Examples of public domain models are SWIM (CSIRO, Div. of Soils; Private Mail Bag, P.P. Aitkenvale, Queensland, Australia 4814) and HYDRUS (Kool and van Genuchten, 1991, U.S. Salinity Laboratory, 4500 Glenwood Drive, Riverside, CA 92502).

6.3.4 INFILTRATION

The rate at which water enters soil during rainfall or irrigation events has long been the subject of theoretical analyses and field and laboratory studies. For many years the Horton equation was used, i.e.

$$i = At^{-1/2}, \tag{6.10}$$

where i is infiltration rate (cm/day), A is a parameter and t is time (day). Following a series of theoretical analyses Philip (1958) suggested the following equation

$$i = St^{-1/2} + Kt. \tag{6.11}$$

The first part on the right hand side of Equation (6.11) represents early time infiltration while the second part represents long time infiltration behavior.

Equation (6.11) was derived by solving Equation (6.7) for given boundary conditions. Therefore, the parameters S and K can be computed from known

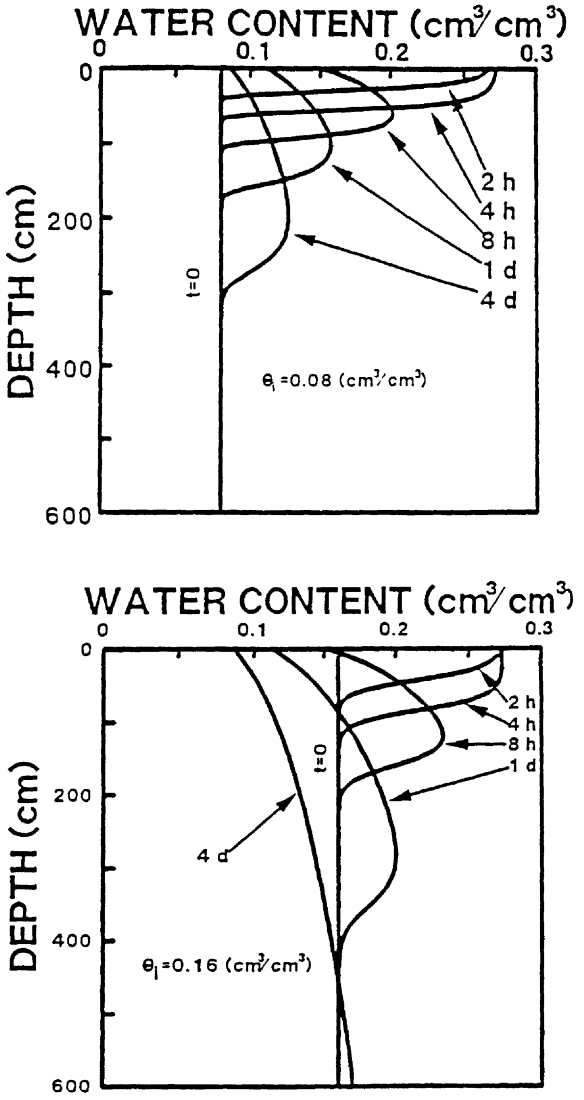


Fig. 6.4. Calculated water content distributions during and after infiltration of 10 cm H_2O in Berino loamy fine sand, at initial water contents of 0.08 cm^3/cm^3 (upper half) and 0.16 cm^3/cm^3 (lower half), respectively. The 10 cm water was added over a 4 hour period.

hydraulic conductivity and water retention functions (Haverkamp et al., 1977). However, the procedure is cumbersome, and also requires knowledge of $K(h)$ and $\theta(h)$ for the soil of interest. Because determining $K(h)$ and $\theta(h)$ is time consuming, it is often more efficient to determine S and K in Equation (6.11) directly by matching the equation to measured infiltration data.

Infiltration of water into soil may also be computed with a numerical solution of Equation (6.7). An example of computed water contents during infiltration of

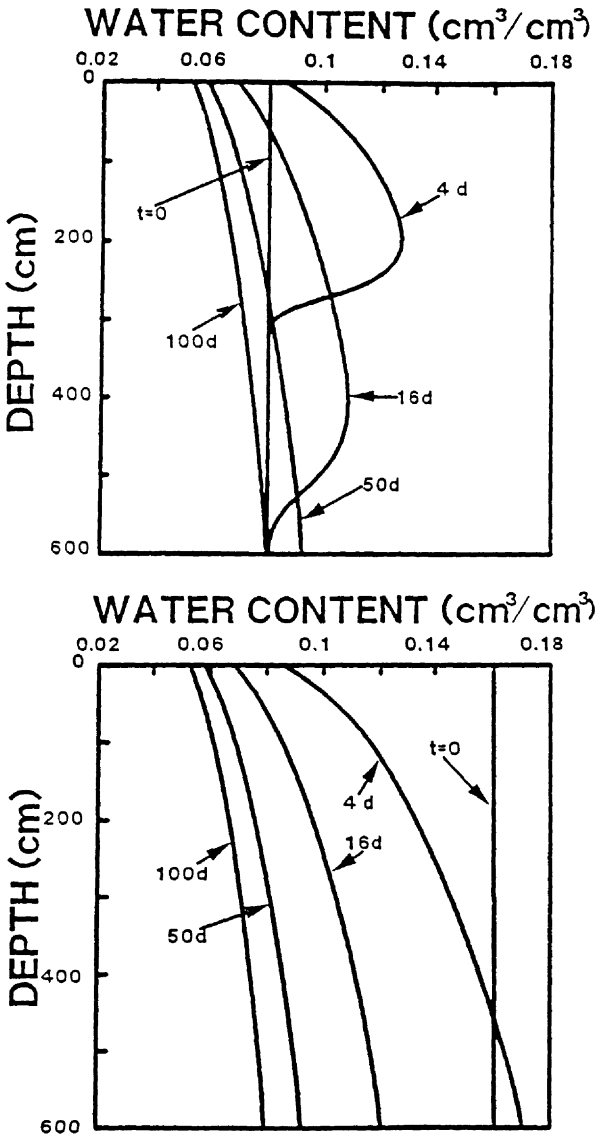


Fig. 6.5. Calculated water content distributions after infiltration of 10 cm H_2O in Berino loamy fine sand, at initial water contents of 0.08 cm^3/cm^3 (upper half) and 0.16 cm^3/cm^3 (lower half), respectively.

10 cm of water into a 6 m deep soil column of Berino loamy fine sand (Table 6.2) is presented in Figure 6.4. The example is presented for two initial water contents, e.g. 0.08 and 0.16 cm^3/cm^3, respectively. HYDRUS (Kool and van Genuchten, 1991) was used for the calculations. Evapotranspiration was assumed to be zero. The results clearly demonstrate the effects of initial water content on the rate at which water moves downward. The soil with the higher initial water content

(0.16 cm^3/cm^3) has a higher hydraulic conductivity and lower storage capacity for the infiltrating water, causing the water front to advance much more rapidly. Thus, if one needs to monitor subsoil moisture movement, one may expect much more rapid responses following rain in moist soil as opposed to dry soil. This is true, even though the initial infiltration rate into dry soil is higher than in moist soil.

6.3.5 REDISTRIBUTION

Redistribution of water in soil commences when infiltration stops. Figures 6.4 and 6.5 show computed water-content distributions after 4 hours of infiltration. Calculations were again made assuming no evaporation or root-water uptake. Note that decreases in water content are initially most pronounced near the soil surface. As time goes on, water contents decrease over the full length of the profile. Note that even though the initial water contents were quite different for the two simulations (8 and 16%), the water content distributions after 50 and 100 days are practically identical.

6.4 Contaminant Transport

6.4.1 EQUATIONS FOR NONREACTIVE SOLUTES

The transport of nonreactive solutes through soil is the result of three simultaneous processes: advective or mass transport, diffusion, and dispersion. Advective transport (J_m) refers to the passive movement of a dissolved solute with water. The equation for advective transport in one direction is:

$$J_m = qC, \tag{6.12}$$

where q is the Darcy flux and C is the volume-averaged concentration.

Diffusive transport (J_D) results from the natural thermal motion of dissolved ions and molecules. This process can be described with Fick's law of diffusion as follows:

$$J_D = \theta D_m \frac{\partial C}{\partial x}\,, \tag{6.13}$$

where θ is the volumetric water content of the medium, D_m the porous medium diffusion coefficient, and x the distance. The porous medium diffusion coefficient, D_m, is always smaller than the diffusion coefficient in water as a result of the tortuous nature of the porous medium.

The porous nature of soils and vadose zone materials results from the existence of pores, fissures, wormholes and other features, as well as from the loose stacking of the soil material itself. The different shapes, sizes, and orientation of the open spaces and pores result in fluid velocities which differ from place to place. In addition, fluid velocities differ in an individual pore, based on where the velocity is determined with respect to the pore wall. Velocities are lowest near the pore wall. The velocity variations cause solute to be transported at different rates, leading to mixing, macroscopically similar to molecular diffusion. This

mechanical spreading by local velocity variations is called mechanical dispersion (Bear, 1972).

The process can be described by an equation similar to (6.13), where J_h is dispersive transport, and D_h is the mechanical dispersion coefficient. This coefficient is strongly, and often linearly, related to fluid velocity. Thus:

$$J_h = -\theta D_h \frac{\partial C}{\partial z}, \tag{6.14}$$

$$D_h = \alpha v, \tag{6.15}$$

where v is the average pore-water velocity. The coefficient is frequently referred to as dispersivity. Because D_m and D_h are macroscopically similar, these terms are usually added. Thus,

$$D = D_m + D_h, \tag{6.16}$$

where D is the longitudinal hydrodynamic dispersion coefficient (Bear, 1972).

Combining (6.12), (6.13), (6.14) and (6.16) yields the solute flux J_s:

$$J_s = -\theta D \frac{\partial C}{\partial z} + qC. \tag{6.17}$$

The equation of continuity for soil solution is:

$$\frac{\partial \theta C}{\partial t} = -\frac{\partial J_s}{\partial z}, \tag{6.18}$$

where t is time. Combining (6.17) and (6.18) yields the general transport equation in one dimension:

$$\frac{\partial \theta C}{\partial t} = \frac{\partial}{\partial z}\left[\left(\theta D \frac{\partial C}{\partial z}\right) - qC\right]. \tag{6.19}$$

Defining pore water velocity v as q/θ, and assuming a uniform and constant water content, Equation (6.19) may be rearranged as follows:

$$\frac{\partial C}{\partial t} = D\frac{\partial^2 C}{\partial z^2} - v\frac{\partial C}{\partial z}. \tag{6.20}$$

Equation (6.20) shows that solute transport is driven by diffusion and dispersion (the first term on the right hand side of Equation 6.20), and mass transport (the second term on the right hand side of Equation 6.20). When pore water velocity is high, mass transport is dominating the process, but when velocities are low, diffusion becomes increasingly important.

The flux or velocity need to be known before Equation (6.20) can be used. If q or v change over time, as for example during infiltration of water into an unsaturated soil, the equation for water flow (Equation 6.7 or 6.9) needs to be solved along with Equation 6.20. This is most conveniently accomplished numerically, using one of several available computer models (Kool and van Genuchten, 1991).

6.4.2 EQUATIONS FOR REACTIVE SOLUTES

Many chemicals associate or sorb to the solid phase of soil. This interaction increases the retention capacity of the soil and reduces the net velocity of transport of the solute. For sorbing chemicals, the equation of continuity (6.18) is written as:

$$\frac{\partial}{\partial t}(\theta C + \rho S) = -\frac{\partial J_s}{\partial z}\,, \tag{6.21}$$

where ρ is the soil bulk density. Combining (6.17) and (6.21) yields:

$$\frac{\partial}{\partial t}(\theta C + \rho S) = \frac{\partial}{\partial z}\left(\theta D\frac{\partial C}{\partial z} - qC\right). \tag{6.22}$$

To reduce the number of unknowns in (6.22), a relationship is needed between S and C. The relationship between sorbed and solution concentrations is described by an isotherm. A widely used isotherm is the Freundlich, given by:

$$S = K_f C^n, \tag{6.23}$$

where S is the sorbed concentration, C is the solution concentration, K_f is the sorption coefficient, and n is a dimensionless parameter related to the energetics of sorption. In many cases n is assumed to be 1, based on either experiment or mathematical expediency.

Differentiation of (6.23), with $n = 1$, and substitution of the resulting equation in (6.22), eventually yields:

$$R\frac{\partial \theta C}{\partial t} = \frac{\partial}{\partial z}\left(\theta D\frac{\partial C}{\partial z} - qC\right). \tag{6.24}$$

Equation (6.24) is the general, one-dimensional convective-dispersive equation (CDE) for a sorbing chemical.

Defining pore water velocity v as q/θ, and assuming a uniform and constant water content, Equation (6.24) may be rearranged as follows:

$$\frac{\partial}{\partial t} = \frac{D}{R}\frac{\partial^2 C}{\partial z^2} - \frac{v}{R}\frac{\partial C}{\partial z}\,. \tag{6.25}$$

where R is the retardation factor equal to:

$$R = 1 + \frac{\rho K_f}{\theta}\,. \tag{6.26}$$

Equation (6.26) shows that the retardation factor can be defined in terms of relative velocities of water and solute. The retardation factor is the distance travelled by a solute front when $R = 1$, divided by the distance travelled by the solute front when R has the assigned value. Or, more simply, R = velocity of water divided by the velocity of the solute. Except for negatively charged ions such as chloride or bromide, R values are generally greater than 1.

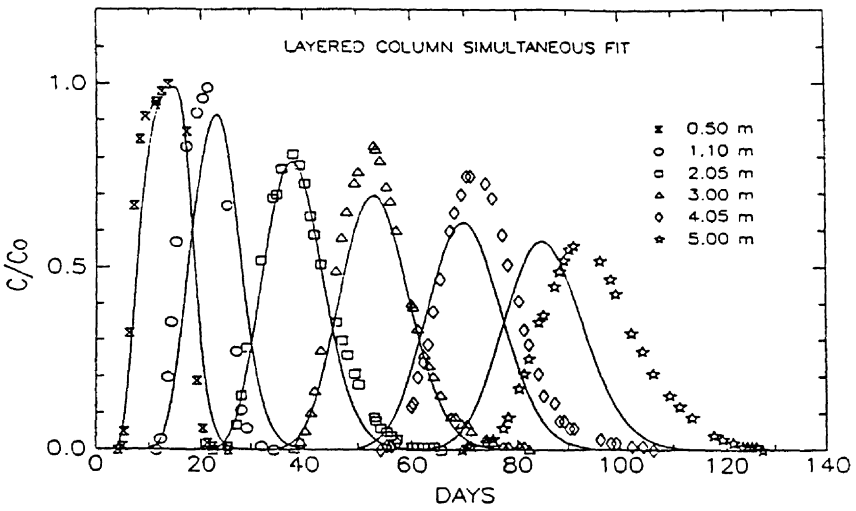


Fig. 6.6. Relative tritium concentration distributions in a layered soil column. The solid lines were obtained by fitting an analytical solution of Eq. 6.26 to the data.

Many contaminants of interest undergo transformation reactions such as biodegradation, hydrolysis, and radioactive decay. Thus, it is important to account for such reactions in transport equations. The typical method by which this is done is to represent the transformation reaction as first order with respect to solute concentration. Adding such reactions to (6.26) leads to:

$$R\frac{\partial C}{\partial t} = D\frac{\partial^2 C}{\partial z^2} - v\frac{\partial C}{\partial z} - k_a C - k_s(R-1)C, \tag{6.27}$$

where k_a and k_s are the first-order transformation coefficients for the solution and sorbed phases, respectively.

6.4.3 Transport in Uniform Soil: Steady Flow

Figure 6.6 is an example of solute flow through soil irrigated at a constant rate of 2.06 cm/day (Porro, Wierenga and Hills, 1993). A 6-m deep by 1-m diameter column was filled with alternate 20-cm thick layers of Berino loamy fine sand and Glendale clay loam. Water was added for several months at a slow but steady rate of 2.06 cm/day. The percentages of saturation once steady flow was reached were approximately 43% for the sand and 81% for the clay loam. Tritiated water was added with the water during the first 11 days of the experiment only. The symbols in Figure 6.6 are measured relative tritium concentration distributions at the various soil depths. The solid lines were obtained by fitting an analytical solution of Equation (6.24), with $R = 1$, $\theta = 0.31$ and $q = 2.06$ to the data.

Measured and fitted data in Figure 6.6 show what is typically observed in these situations, i.e., a decrease in the peak concentration and a spreading of the pulse, as the tritiated water moves down the 6 m deep layered soil profile. The curves were predicted with the same values for q, θ and R for the entire column. Using

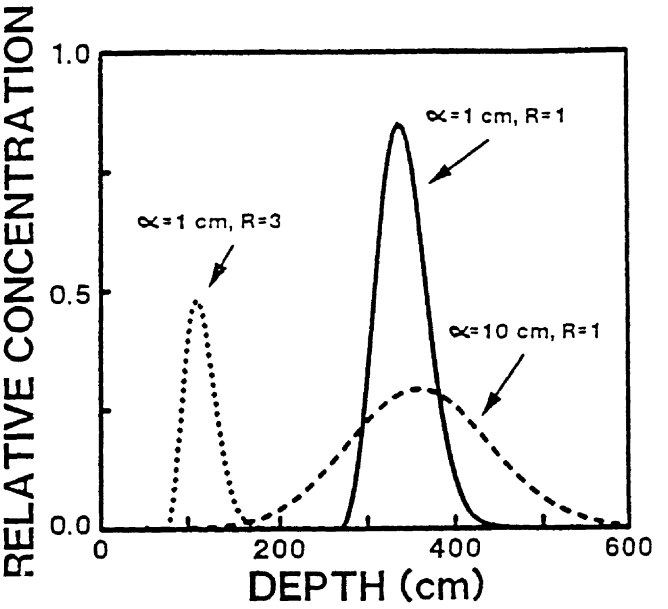


Fig. 6.7. Predicted solute concentration/depth distributions for different combinations of dispersivity and retardation factor (R). Flux 3.74 cm/day; water content 0.16 cm^3/cm^3; tension 70 cm H_2O; time is 15 days.

constant values for q, θ and R resulted in excellent fits between measured and observed curves at some depths (e.g. 2.05 m) but poorer fits at other depths (e.g. 5.0 m). This is the result of the variability of soil physical properties inside the column. Thus, even though the soil was hand packed with great care, there still is considerable variation in transport with depth. Even greater variation may be expected under field conditions. Except for the 5 m depth, the analytical solution of Equation (6.24) described the arrival times of the solute fronts quite well, indicating that simplified analytical solutions may be used for first-order estimates of arrival times. However, as is shown in the next example, a good estimate of the retardation factor is most important for such predictions.

The second example of computed solute transport with steady flow of water is shown in Figure 6.7. In this case, water was added continuously to a soil profile at a rate of 3.74 cm/day, but solute was added during the first two days only. The figure shows the relative solution concentrations with depth after 15 days, for two dispersivity values (1 and 10 cm) and two retardation values (1 and 3). The effect of chemical interaction (retardation) on solute transport is quite clear. A retardation factor of 3 instead of 1 significantly delays the downward movement. Thus, high retardation factors cause solute to move through soil or vadose zone material at much reduced rates. An R value of 3 represents a K_f value of 0.41 cm^3/g, assuming a bulk density of 1.4 and a water content of 20% by volume. R values greater than 3, and in many cases much greater than 3, are common for many of the contaminants introduced in the environment. Thus, when monitoring for solutes in the subsoil, one can expect the arrival of nonreactive chemicals ($R = 1$) much

sooner than of sorbing chemicals. This difference in arrival times can be very large, i.e. years. Figure 6.7 also shows that dispersivity mostly affects the shape of the concentration/depth distribution, but not the location of the peak concentration.

6.4.4 Transport in Uniform Soil: Transient flow

Water flow is seldom constant in the field, but changes over time due to variations in boundary conditions (e.g. rainfall, irrigation events). This has a significant effect on solute movement which occurs rapidly during periods of fast water flow. When the rate of water movement is slow, such as during the latter phases of the redistribution process (see Fig. 6.5), solute transport is mainly by diffusion, and very slow. Thus in the absence of significant water movement, solutes tend to stay where they are.

Figure 6.8 shows the advance of a solute front through the 6 m profile of Figures 6.4 and 6.5. No chemical interaction was assumed between the solute and the soil (i.e. $R = 1$). Calculations were again made with HYDRUS, and the physical properties of the soil were the same as before (Table 6.2). The distributions of water and solute in Figure 6.8 show that the solute fronts lag behind the water fronts. The magnitude of this lag is dependent on the initial water present in the soil, being greater for soil having a higher initial water content. The model results suggest that the initial water ("old water") is pushed ahead by the traced water ("new water").

6.5 Nonideal Contaminant Transport

6.5.1 Heterogeneity

The solute transport examples shown in Figures 6.6 through 6.8 represent close to ideal conditions such as may be found in carefully packed laboratory or field size columns. In the field, conditions are not nearly as ideal and one may expect different behavior. As found in a recent review, nonideal transport has been observed in most field-scale experiments of contaminant transport (Brusseau, 1994). One of the main reasons for deviations from ideal transport behavior in the field are heterogeneities in the physical, chemical, and biological properties of subsurface material. Soils and geologic materials are inherently non-uniform. This is a direct result of the formation of these materials and their exposure to various events during past geologic times. Thus, properties which affect solute transport can be expected to vary spatially. This is particularly true for soil hydraulic properties, such as the unsaturated hydraulic conductivity and the water retention properties.

Many studies have been conducted to determine the variability of hydraulic properties in the field. Examples are studies by Nielsen et al., 1973; Warrick and Nielsen, 1980; Gajem, et al., 1981; and Nash et al., 1989. Because a main driving force for solute transport in the field is advection, variation in hydraulic properties and thus advection has a large effect on contaminant transport in the subsoil.

Jury et al. (1989) reviewed chemical parameters in addition to hydraulic parameters affecting solute transport. While these authors paid special attention to dispersion coefficients, less is known about spatial variability of other chemical

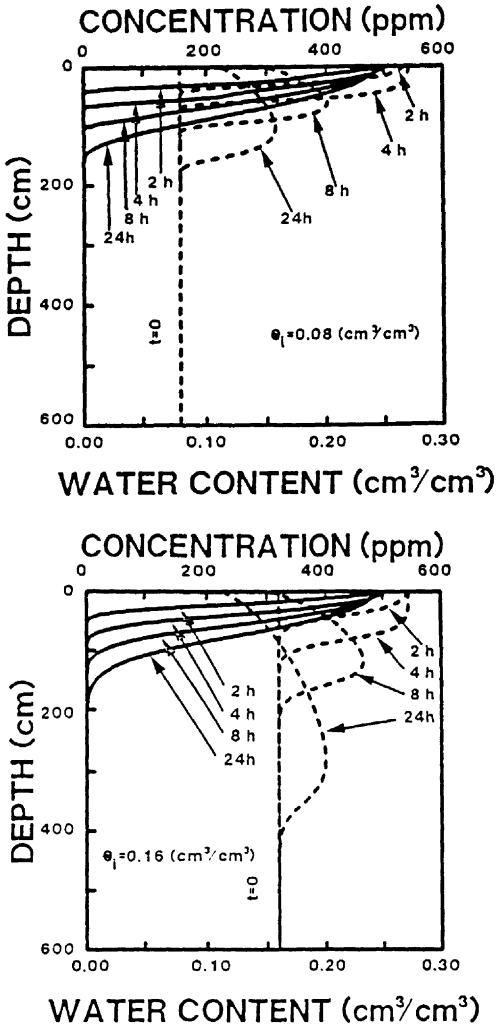


Fig. 6.8. Water content (dashed lines) and solute concentration (solid lines) distributions in Berino loamy fine sand irrigated with 10 cm water in 4 hours. Initial water contents were 0.08 cm^3/cm^3 (upper half) and 0.16 cm^3/cm^3 (lower half), respectively.

parameters such as sorption coefficients, or retardation factors. Even less is known about field heterogeneity of microbial properties.

Heterogeneity occurs laterally as well as vertically. However, soils tend to be more similar in the horizontal direction than in the vertical direction. In the horizontal plane changes in properties occur more gradually. Extensive layering is observed at most sites, and this causes sudden changes in texture and hydraulic properties. It is not clear how this layering affects solute transport in the vadose zone. However, one would expect that local differences in vertical fluid velocities are evened out by the layering. The result may be that over longer vertical

distances, across a number of distinct horizons, solute transport patterns may be similar to those observed in uniform soils.

Some evidence for this was shown experimentally by Porro and Wierenga (1992) who measured dispersivities during unsaturated transport through 6 m deep columns. They found that the dispersivity in a column uniformly filled with soil was twice as high as in a layered column. Similar observations were made by Roth et al. (1991). The implication is that heterogeneity in the vertical direction may have an averaging effect on solute transport in the vadose zone, and allow for application of simplified models based on Equation (6.26) for prediction purposes. The parameters in these models would have to represent the average behavior, such as average monthly or annual recharge for q, and an average retardation factor R calculated from values measured for each individual horizon (see Roth et al., 1991).

6.5.2 Preferential Flow

Solute transport in the vadose zone can also be affected by the structure of the soil or geologic material. In structured soils and in soils containing root holes, worm holes, cracks and other large pores, water and solute may move preferentially through these pathways. This phenomenon is variously called macropore flow, short-circuiting, bypassing, preferential flow, partial displacement, or transport with "mobile" and "immobile" domains (Brusseau and Rao, 1990).

Preferential flow is a well documented phenomenon known for over 100 years, and many publications, mainly in the soil sciences, deal with it (Beven and Germann, 1982; Elrick and French, 1966; Biggar and Nielsen, 1961). A recent publication by the American Society of Agricultural Engineers summarizes the role of preferential flow on ground water quality, processes affecting it, methods of measuring, and modelling for predictions (ASAE, 1990). According to Beven (1990) the important characteristic of preferential flow is that during wetting, local wetting fronts propagate into the soil to significant depths thus bypassing intervening matrix pore space. The result is movement of water and solutes to greater depths at faster speeds than would be predicted on the basis of Equation (6.26).

At the same time, there is diffusive mass transfer of solutes from the preferential pathways into the surrounding soil matrix and vice versa. The result is that in structured soil one observes early breakthrough and tailing of solute (van Genuchten and Wierenga, 1976; Brusseau and Rao, 1990). The effect of structured soil on solute transport is illustrated in Figure 6.9, which shows breakthrough curves for transport of a nonreactive tracer (pentafluorobenzoate) at three velocites through an aggregated soil (Brusseau et al., 1994).

The degree to which nonideal behavior is observed during transport in structured soil is dependent on several factors. The size of the soil structures, the size of the solute, and the pore-water velocity are major controlling factors (cf., van Genuchten and Wierenga, 1976; Brusseau, 1993a). The saturation status of the soil is also very important (Seyfried and Rao, 1987; Brusseau and Rao, 1990). At lower saturations the macropores will not contribute significantly to flow, which reduces the rapid transport behavior of the solute.

Although field observations on preferential flow are well documented for relatively shallow soils, the importance of preferential flow on transport through deep unsaturated profiles has not been documented. Because of layering in deep unsaturated profiles, preferential flow may be less important in such profiles than in surface soils.

6.5.3 NONLINEAR, RATE-LIMITED SORPTION

Most solute transport models, especially for field-scale applications, include the assumption that sorption is linear (i.e. $n = 1$ in Equation 6.23). Numerous laboratory experiments have shown that sorption is often linear for low-polarity organic compounds. However, this assumption should be evaluated for each case because non-linear isotherms have been reported. Non-linear isotherms are the norm rather than the exception for many polar organic, as well as for inorganic, chemicals. It is important to note that the nature of the isotherm is influenced by properties of the sorbent and the solution in addition to those of the solute. Several equations exist for describing sorption isotherms (see Travis and Etnier, 1981 and Kinniburgh, 1986 for reviews).

It is well known that non-linear sorption can cause asymmetrical breakthrough curves and concentration-dependent retardation factors (cf., Brusseau and Rao, 1989). This was clearly demonstrated by Rao and Davidson (1979), who reported breakthrough curves for transport of 2,4-dichlorophenoxyacetic acid through two soils. Retardation increased as the influent solute concentration decreased. In addition, the elution waves were much more dispersed than were the injection waves.

Most field-scale solute transport models include the local equilibrium assumption, which specifies that interactions between the solute and the sorbent are so rapid in comparison to hydrodynamic residence time that the interactions can be considered instantaneous. Based on laboratory experiments, it has long been known that sorption/desorption of many organic compounds by soils can be significantly rate limited. The mechanisms responsible for rate-limited sorption/desorption have been a focus of recent research. For low-polarity organic compounds, the current view is that nonequilibrium is caused by constrained intrasorbent diffusion (cf., Brusseau and Rao, 1989). While rate-limited specific chemical reactions (chemisorption) are not likely to be important for most low-polarity organic compounds, such reactions can be important for ionic and polar organic compounds. Rate-limited sorption/desorption of inorganic chemicals has been reported to be caused by intrasorbent diffusion (cf., Hodges and Johnson, 1987; Wood et al., 1990) and chemisorption (cf., Jardine et al., 1985).

Numerous experimental and theoretical studies have shown that rate-limited sorption/desorption can cause nonideal transport. This nonideality can take the form of asymmetrical breakthrough curves exhibiting early breakthrough and tailing, as well as decelerating plumes (temporally increasing R values). An example of the effect of rate-limited sorption/desorption on transport is presented in Figure 6.10, which shows breakthrough curves for transport of several solutes through a sandy soil. Note that the breakthrough curves for the nonreactive tracers (^3H_2O and pentafluorobenzoate) are sharp and symmetrical, signifying ideal behavior. In

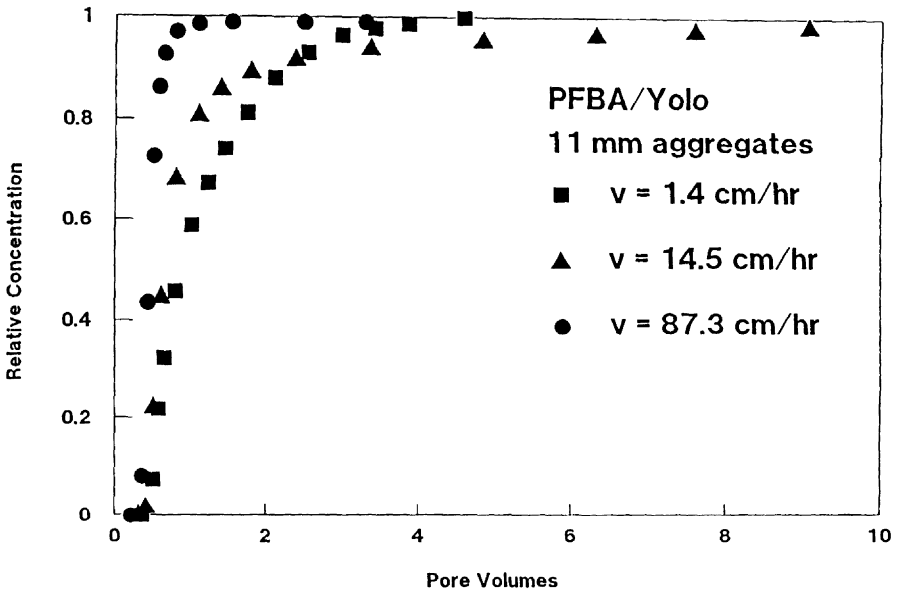


Fig. 6.9. Solute transport through an aggregated soil (from Brusseau et al., 1994).

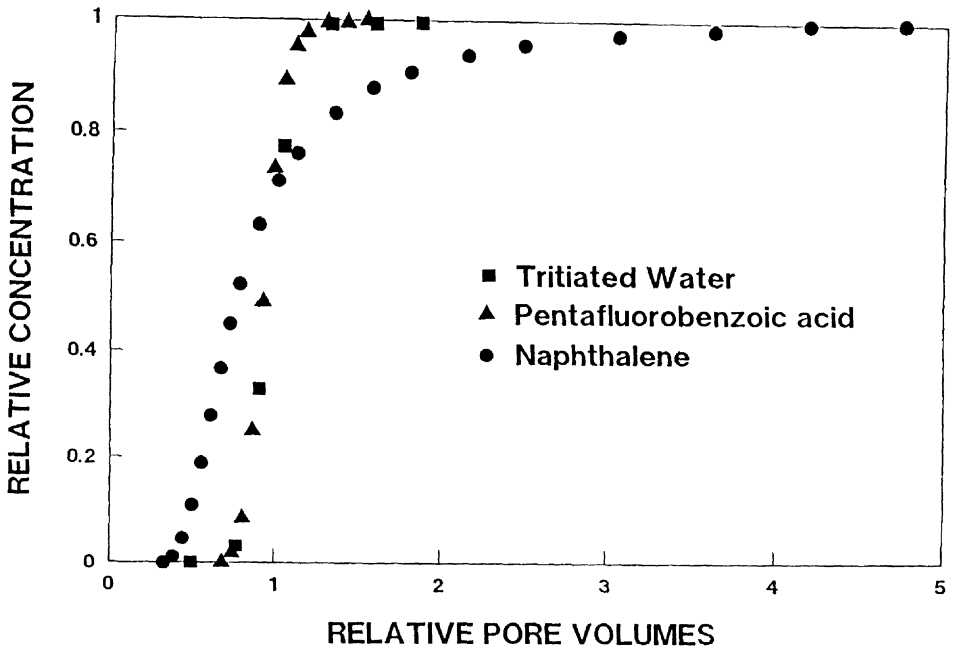


Fig. 6.10. Transport of nonreactive and sorbing solutes through a homogeneously packed column. Relative pore volumes is defined as pore volume divided by the retardation factor. Adapted from Brusseau et al., 1991.

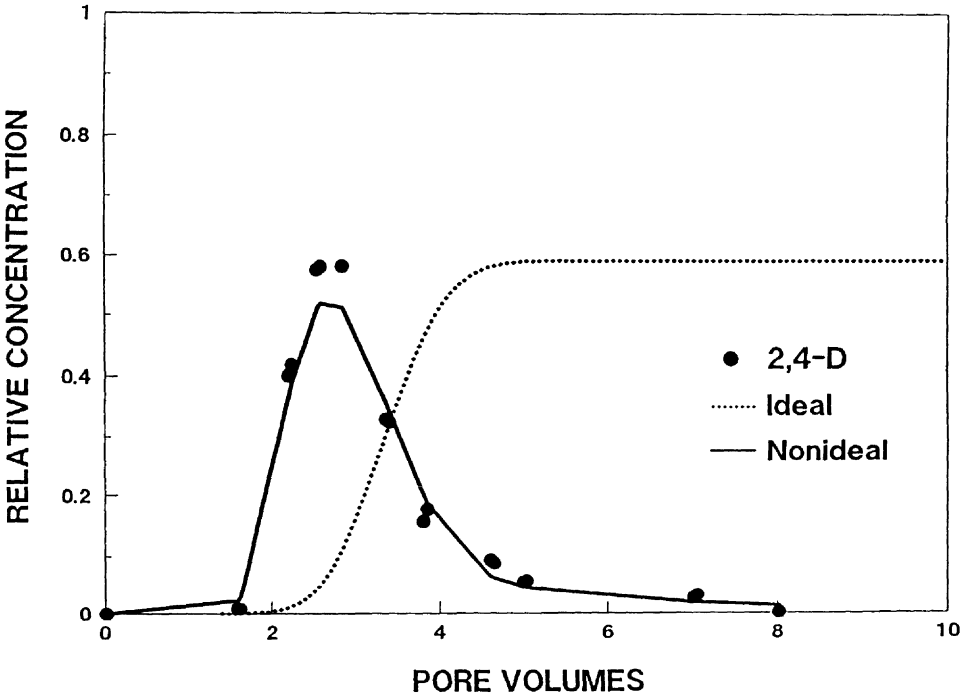


Fig. 6.11. Transport of 2,4-D through a column under (water) unsaturated conditions. Adapted from Estrella et al., 1993 and Brusseau, 1994.

contrast, the normalized breakthrough curve for the sorbing organic solute (naphthalene) is asymmetrical. This suggests that the nonideality factor involves the sorption process.

A phenomenon that, in many cases, is related to rate-limited sorption/desorption is contaminant aging, also refered to as "bound residues". Recent research has shown that contaminants that have been in contact with porous media for long times are much more resistant to desorption, extraction, and degradation. For example, contaminated samples taken from field sites exhibit solid:aqueous distribution ratios that are much larger than those measured or estimated based on spiking the porous media with the same contaminant (e.g., adding contaminant to uncontaminated sample) (Steinberg et al., 1987; Scribner et al., 1992). In addition, the magnitude of the desorption rate coefficients determined for previously contaminated media collected from the field have been shown to be much smaller than the values obtained for spiked samples (Steinberg et al., 1987). Over the time scale of years, physical, chemical, and biological processes can alter the environment in which the sorbed solute resides, thereby reducing the potential of the solute to desorb.

6.5.4 Facilitated Transport

Many contaminated sites have multiple contaminants in the aqueous phase. The existence of multi-component solutions leads to the possibility of nonideal behavior wherein one component may influence the behavior of another. Naturally occurring subsurface components associated with the fluid phase may also influence the behavior of contaminants. When multiple components are present in solution, a primary question to be addressed is the occurrence of antagonistic or synergistic interactions among the components, and between the components and the solid and aqueous phases (Brusseau, 1993b). The presence of co-components can affect the behavior of contaminants in several ways, resulting in the following three phenomena: 1. competitive solid-phase interactions; 2. cooperative solid-phase interactions; and 3. "solubilization". We will focus our brief discussion on the third, which can lead to "facilitated" transport, a subject of much interest.

For our purposes, solubilization is defined as an increase in the amount of contaminant that can reside in the aqueous phase beyond the solubility limit. Examples of solubilization include the effect of miscible solvents (i.e., the "cosolvent" effect), surfactants, and colloids (the "colloid" effect) on the solubility, sorption, and transport of low-polarity organic solutes. The presence of a cosolvent (e.g., ethanol), anionic or nonionic surfactant, or colloid (e.g., dissolved organic matter) has been shown to increase the apparent solubility of many low-polarity organic compounds. This, in turn, causes reduced sorption and retardation, thereby "facilitating" the transport of the low-polarity organic compound. Complexation by colloids, dissolved organic matter, or surfactants can lead to facilitated transport of inorganic compounds. Further discussion of these phenomena may be found in recent reviews (cf., McCarthy and Zachara, 1989; Brusseau, 1993b).

6.5.5 Vapor and Immiscible-Liquid Phase Transport

Immiscible organic liquids may exist in portions of the contaminated subsurface at many hazardous waste sites. Examples include gasoline, diesel fuel, and aviation fuel associated with various types of fuel stations, chlorinated solvents associated with cleaning facilities, and creosote associated with wood-treatment operations. The movement of immiscible liquids in porous media, which has become a major component of contaminant transport research, has received considerable attention in the petroleum engineering field. It is well known that when an immiscible liquid moves through a porous medium, a fraction of the liquid can become entrapped thereby forming a residual saturation. Very large pore-water velocities (i.e., hydraulic gradients) are usually required to displace residual quantities of immiscible liquids (cf., Wilson and Conrad, 1984; Hunt et al., 1988). Hence, the primary means of removal will be dissolution into water and volatilization into the soil atmosphere. The residual liquid, therefore, serves as a long-term source of contaminant.

While some of the most widely studied immiscible liquids are composed of a single component (e.g., trichloroethene), many others (e.g., gasoline, diesel fuel, coal tar) are multi-component liquids. Knowledge of the partitioning behavior of multi-component liquids is essential to the prediction of their impact on ground-

water quality. The partitioning of components into water is controlled by the aqueous solubility of the component and the composition of the immiscible liquid. The actual maximum concentration of a contaminant in a water phase in contact with a multi-component immiscible liquid will be less than the aqueous solubility. Hence, ignoring the multi-component effect will lead to erroneous estimates of aqueous-phase concentrations.

Many organic contaminants of interest are volatile (e.g., gasoline components, chlorinated solvents). This is of great significance for the vadose zone which, because it is usually not completely saturated with water, contains a gas phase. Consequently, a description of the transport of organic contaminants in the vadose zone will be incomplete without consideration of transport in the gas phase. Research in this area has increased greatly with the advent of "soil venting" or "vacuum extraction" as a popular means of remediating vadose-zone contamination.

One aspect of gas-phase transport that has been most often considered is gas-phase diffusion. Diffusive flux can be of great significance because gas-phase diffusion coefficients are much larger than aqueous diffusion coefficients. The relatively rapid transport by gas-phase diffusion can lead to larger than expected zones of contamination if such transport is ignored (cf., Sleep and Sykes, 1989). Advective transport in the gas phase, caused either by external pressure gradients (e.g., soil venting) or by gas-phase density differences (near immiscible liquid sources), can also be important.

6.5.6 Coupled Physical, Chemical, and Biological Processes

As previously discussed, many contaminants may undergo various types of transformation reactions during transport. A recent review of mathematical models designed to simulate transport of solutes influenced by transformation reactions revealed that the vast majority were based on the use of a first-order transformation reaction (Brusseau et al., 1992a). First-order kinetics, wherein the rate is a function only of solute concentration, accurately describes some transformation reactions (radioactive decay, pH-independent hydrolysis) but not others. Biodegradation is a good example of a transformation process that can be described as first order only for highly constrained conditions. Often, the rate of biodegradation of a contaminant will depend on many other factors (e.g., biomass growth, electron acceptor availability, sorption, environmental conditions) in addition to contaminant concentration. Contaminant transport under these conditions will not be describable with a model based on ideal transport (i.e., first-order reaction).

The effect of nonsteady-state microbial populations on contaminant transport is an example of transformation-related nonideality. This effect is illustrated in Figure 6.11, which shows the transport of a pesticide (2,4-dichlorophenoxyacetic acid, 24D) through soil packed in a column (Estrella et al., 1993). The soil contained a bacterial population capable of degrading the 24D as demonstrated by separate microcosm experiments. A simulated breakthrough curve obtained for ideal conditions, including first-order biodegradation, is also shown in Figure 6.11.

Clearly, the biodegradation and transport of 24D could not be simulated with the ideal model. One condition associated with the assumption of first-order

biodegradation kinetics is that there is no microbial growth. This condition is inconsistent with the data, which shows a steadily decreasing concentration of 24D in the effluent to a point of about 8 pore volumes, whereupon 24D was no longer detectable. It is important to note that the decline and disappearance of 24D in the effluent occurred while the solution containing 24D was still being injected into the column. This behavior indicates the occurrence of temporally non-uniform biological activity. A model that accounted for, among other factors, microbial growth could successfully simulate the data (see Figure 6.11).

The effects of interactions between sorption and biodegradation, two rate-controlled processes, on transport is another example of nonideal transport. Recent research performed with batch (i.e., non-flowing) systems has suggested that the rate of biodegradation can be controlled by the rate of desorption or mass transfer (cf., Rijnaarts et al., 1990; Robinson et al., 1990; Scow and Alexander, 1992), depending on the relative time scales of the two processes. When the rate of desorption or mass transfer is slower than that of biodegradation, the transport of the contaminant will not follow the behavior expected based on ideal conditions. This has been shown by Brusseau and colleagues, who have investigated the influence of rate-limited sorption/desorption (Angley et al., 1992; Estrella et al., 1993) and heterogeneous porous media combined with rate-limited sorption/desorption (Brusseau et al., 1992b) on the transport of contaminants undergoing biodegradation.

6.6 Conclusion

For the purpose of this introductory chapter, it suffices to say that as the understanding of the physics, chemistry and biology of vadose zone transport grows, models will get increasingly complex, containing many more parameters which will be difficult to evaluate independently. At the same time, averaging processes in field soils may allow the use of relatively simple models, with only a few bulk parameters. No matter which route is taken, it is important that physical, chemical and biological aspects of solute transport are taken into account. This applies to predictions of solute transport through unsaturated soils, as well as to characterization of vadose zone materials.

Acknowledgements

The authors wish to acknowledge Mr. Lehua Pan, Research Assistant, Department of Soil and Water Science, University of Arizona, who performed the Hydrus modelling and prepared many of the figures.

References

American Society of Agricultural Engineers, Preferential Flow, Proceedings of the National Symposium, Chicago, IL, 408, 1991.

Angley, J.T., Brusseau, M.L., Miller, W.L., and Delfino, J.J., Nonequilibrium sorption and aerobic biodegradation of dissolved alkylbenzenes during transport in aquifer material:

Column experiments and evaluation of a coupled-process model, Environ. Sci. Technol., 26, 1404-1410, 1992.
Bear, J., Dynamics of Fluids in Porous Media, American Elsevier Publishing Co., New York, 1972.
Beven, K., Modeling preferential flow: An uncertain future? in Proceedings National Symposium on Preferential Flow, Gish, T.J., Ed., American Society of Agricultural Engineers, St. Joseph, MI, 1-11, 1990,
Beven, K. and P. Germann., Macropore and water flow in soil. Water Resour. Res., 18, 311, 1982.
Biggar, J.W. and D.R. Nielsen, Miscible displacement: II. Behavior of tracers, Soil Sci. Soc. Am. Proc., 125, 1961.
Brisco, B., T.J. Pultz, R.J. Brown, G.C. Topp, M.A. Hares, and W.D. Zebchuk, Soil moisture measurement using portable dielectric probes and time domain reflectometry, Water Resour. Res., 28, 1339-1346, 1992.
Brooks, R.H. and A.T. Corey, Hydraulic Properties of Porous Media, Hydrol. Paper No. 3, Colorado State Univ., Fort Collins, 1964.
Brusseau, M.L., The influence of solute size, pore-water velocity, and intraparticle porosity on solute dispersion and transport in soil, Water Resour. Res., 29, 1071-1080, 1993a.
Brusseau, M.L., Complex Mixtures and Groundwater Quality, Environmental Research Brief, U.S. Environmental Protection Agency, Wash. D.C. 1993b.
Brusseau, M.L., Transport of Reactive contaminants in heterogeneous porous media, Rev. Geophys., 32, 285-314, 1994.
Brusseau, M.L. and Rao, P.S.C, Sorption nonideality during organic contaminant transport in porous media, CRC Critical Reviews in Environ. Control, 19, 33-99, 1989.
Brusseau, M.L. and P.S.C. Rao, Modeling solute transport in structured soils. A review, Geoderma, 46, 169-192, 1990.
Brusseau, M.L., Jessup, R.E., and Rao, P.S.C., Nonequilibrium sorption of organic chemicals: Elucidation of rate-limiting processes, Environ. Sci. Technol., 25, 134-142, 1991.
Brusseau, M.L., Rao, P.S.C., and Bellin, C.A., Modeling coupled processes in porous media: Sorption, transformation, and transport of organic solutes, pp. 147-184 in: Interacting Processes in Soil Science, Advances in Soil Science, J. Wagenet, P. Baveye, and B.A. Stewart, eds., Lewis Pub., Ann Arbor, MI, 1992a.
Brusseau, M.L., Jessup, R.E., and Rao, P.S.C., Modeling solute transport influenced by multi-process nonequilibrium and transformation reactions, Water Resour. Res., 28, 175-182, 1992b.
Brusseau, M. L., Gerstl, Z., Augustijn, D., and Rao, P. S. C. Simulating Solute Transport in an Aggregated Soil With the Dual-Porosity Model: Measured and Optimized Parameter Values. Journal of Hydrology, 163, 187-193, 1994.
Campbell, G.S., A simple method for determining unsaturated hydraulic conductivity from moisture retention data, Soil Sci., 1117, 311-314, 1974.
Chong, S.K. R.E. Green, and L.R. Ahuja, Simple in-situ determination of hydraulic conductivity by power function description of drainage, Water Resour. Res., 17, 1109-1114, 1981.
Elrick, D.E. and L.K. French, Miscible displacement patterns on disturbed and undisturbed soil cores, Soil Sci. Soc. Amer. Proc., 30, 153, 1966.
Estrella, M.R., Brusseau, M.L., Maier, R.S., Pepper, I.L., Wierenga, P.J., and Miller, R.M., Biodegradation, sorption, and transport of 2,4-Dichlorophenoxyacetic Acid (2,4-D) in a saturated and unsaturated soil, App. Environ. Microbiol., 59, 4266-4273, 1993.
Gajem, Y.M., A.W. Warrick, and D.E. Myers, Spatial dependence of physical properties of a typic torrifluvential soil, Soil Sci. Soc. Amer. J., 45, 709-715, 1981.
Gardner, W.R., Some steady-state solutions of the unsaturated moisture flow equation with application to evaporation from a water table, Soil Sci., 85, 228-232, 1958.

Haverkamp, R., M. Vauclin, J. Touma, P.J. Wierenga, and G. Vachaud, A comparison of numerical simulation models for one-dimensional infiltration, Soil Sci. Soc. Amer. J., 41, 285-293, 1977.

Hills, R.G., I. Porro, D.B. Hudson, and P.J. Wierenga, Modelling one-dimensional infiltration into very dry soils. Part 1. Water content versus pressure head based algorithms, Water Resour. Res., 25, 1259-1269, 1989.

Hodges, S.C. and Johnson, G.C., Kinetics of sulfate adsorption and desorption by cecil soil using miscible displacement, Soil Sci. Soc. Amer. J., 51, 323-331, 1987.

Hudson, D.B., A Transient Method for Estimating the Hydraulic Properties of Dry Unsaturated Soil Cores, Unpubl. Ph.D. thesis, New Mexico State University, Las Cruces, NM., 1992.

Hunt, J.R., Sitar, N., and Udell, K.S., Nonaqueous phase liquid transport and cleanup 1. Analysis of mechanisms, Water Resour. Res., 24, 1247-1258, 1988.

Jardine, P.M., Zelazny, L.W., and Parker, J.C., Mechanisms of aluminum adsorption on clay minerals and peat, Soil Sci. Soc. Amer. J., 49, 862-867, 1985.

Jury, W.A., W.R. Gardner, and W.H. Gardner, Soil Physics, 5th ed. John Wiley and Sons, Inc., pp. 328, 1991.

Kool, J.B. and M. Th. van Genuchten, Hydrus: One-Dimensional Variably Saturated Flow and Transport Model, Including Hysteresis and Root Water Uptake, U.S. Salinity Laboratory, Riverside, CA., USA, 1991.

Klute, A., Methods of Soil Analysis, Part 1. Monograph 9, American Society of Agronomy, Madison, WI, 1986.

McCarthy, J.F. and Zachara, J.M., Subsurface transport of contaminants, Environ. Sci. Technol., 23, 496-502, 1989.

Mualem, Y., A new model for predicting the hydraulic conductivity of unsaturated porous media, Water Resourc. Res., 12, 513-522, 1976.

Nash, M.S., P.J. Wierenga, and A. Butler-Nance, Variation in tension, water-content, and drainage rate along a 91 m transect, Soil Sci., 148, 94-101, 1989.

Nielsen, D.R., J.W. Biggar, and K.T. Erh, Spatial variability of field-measured soil-water properties, Hilgardia, 42, 215-260, 1973.

Nielsen, D.R., M.Th. van Genuchten, and J.W. Biggar, Water flow and solute transport processes in the unsaturated zone, Water Resour. Res., 22, 895-1085, 1986.

Philip, J.R., The theory of infiltration, 6, Soil Sci., 85, 278-286, 1958.

Porro, I. and P.J. Wierenga, Transient and steady-state solute transport through a large unsaturated soil column, Ground Water, 31, 193-200, 1993.

Porro, I., P.J. Wierenga, and R.G. Hills, Solute transport through large uniform and layered soil columns, Water Resour. Res., 29, 1321-1330, 1993.

Rao, P.S.C. and Davidson, J.M., Adsorption and movement of selected pesticides at high concentrations in soils, Water Res., 13, 375-342, 1979.

Reeves, T.L. and S.M. Elgazawi Time domain reflectometry for measuring volumetric water content in processed oil shale waste, Water Resour. Res., 28, 769-776, 1992.

Rijnaarts, H.M., Bachmann, A., Jumelet, J.C., and Zehnder, A.J.B., Effect of desorption and intraparticle mass transfer on the aerobic biomineralization of -hexachlorocyclohexane in a contaminated calcareous soil, Environ. Sci. Technol., 24, 1349-1354, 1990.

Robinson, K.G., Farmer, W.S., and Novak, J.T., Availability of sorbed toluene in soils for biodegradation by acclimated bacteria, Wat. Res., 24, 345-350, 1990.

Roth, K., W.A. Jury, H. Flühler, and W. Attinger, Transport of chloride through an unsaturated field soil, Water Resourc. Res., 27, 2533-2541, 1991.

Sardin, M., D. Schweich, F.L. Ley, and M.Th. van Genuchten, Modeling the nonequilibrium transport of linearly interacting solutes in porous media. A Review, Water Resour. Res., 27, 2287-2307, 1991.

Scow, K. and Alexander, M., Effect of diffusion on the kinetics of biodegradation: Experimental results with synthetic aggregates, Soil Sci. Soc. Am. J., 56, 128-134, 1992.

Scribner, S.L., Benzing, T.R., Sun, S., and Boyd, S.A., Desorption and bioavailability of aged simazine residues in soil from a continuous corn field, J. Environ. Qual., 21, 1115-1121, 1992.

Seyfried, M.S. and Rao, P.S.C., Solute transport in undisturbed columns of an aggregated tropical soil: Preferential flow effects, Soil Sci. Soc. Amer. J., 51, 1434-1443, 1987.

Sisson, J.B., A.H. Ferguson, and M.Th. van Genuchten, Simple method for predicting drainage from field plots, Soil Sci. Soc. Amer. J., 44, 1147-1152, 1980.

Sleep, B.E., and Sykes, J.F., Modeling the transport of volatile organics in variably saturated media, Water Resour. Res., 25, 81-92, 1989.

Steenhuis, T.S. and J.Y. Parlange, Preferential flow in structured and sandy soils, in Proceedings National Symposium on Preferential Flow, Gish, T.J., Ed., American Society of Agricultural Engineers, St. Joseph, MI: 11-21, 1990.

Steinberg, S.M., Pignatello, J.J., and Sawhney, B.L., Persistence of 1,2-dibromoethane in soils: Entrapment in intraparticle micropores, Environ. Sci. Technol., 21, 1201-1210, 1987.

van de Pol, R.M., P.J. Wierenga, and D.R. Nielsen, Solute movement in a field soil, Soil Sci. Soc. Am. J., 41, 10-13, 1977.

van Genuchten, M.Th. and D.R. Nielsen, On describing and predicting the hydraulic properties of unsaturated soils, Annals Geophysicae, 3, 616-628, 1985.

van Genuchten, M.Th. and P.J. Wierenga, Mass transfer studies in sorbing porous media: Analytical solutions, Soil Sci. Soc. Am. J., 40, 473-480, 1976.

van Genuchten, M.Th. and P.J. Wierenga, A closed-form equation for predicting the hydraulic conductivity of unsaturated soils, Soil Sci. Soc. Am. J., 44,892-898, 1980.

van Genuchten, M.Th. and P.J. Wierenga, Solute dispersion coefficients and retardation factors, in Methods of Soil Analysis, Part 1. Physical and Mineralogical Methods, Klute, A., Ed., Agron. Monograph No. 9, 1025-1054, 1986.

Warrick, A.W., D.E. Myers, and D.R. Nielsen, Geostatistical Methods Applied to Soil Science, in Methods of Soil Analysis. Part 1. Physical and Mineralogical Methods, Amer. Soc. of Agron. 53-80, 1986.

Warrick, A.W. and D.R. Nielsen, Spatial Variability of Soil Physical Properties in the Field, in Applications of Soil Physics, Hillel, D., Ed., Academic Press, NY, 319-343, 1980.

Wierenga, P.J., Solute distribution profiles computed with steady-state and transient water movement models, Soil Sci. Soc. Am. J., 41, 1050-1055, 1977.

Wilson, J.L., and Conrad, S.H., Is physical displacement of residual hydrocarbons a realistic possibility in aquifer restoration?, paper presented at Petroleum Hydrocarbons and Organic Chemicals in Ground Water, pp. 274-298, Nat. Water Well Assoc., Dublin, OH, 1984.

Wood, W.W., Kraemer, T.F., and Hearn, P.P., Intragranular diffusion: An important mechanism influencing solute transport in clastic aquifers?, Science, 247, 1569-1572, 1990.

CHAPTER 7

Transport of moisture and solutes in the unsaturated zone by preferential flow

Frank Stagnitti, J.-Yves Parlange,
Tammo S. Steenhuis, Jan Boll, Bruce Pivetz and
D.A. Barry

Abstract. Mathematical models are widely used in Soil Physics and Hydrology for predicting water percolation and water-aided transport of solutes and contaminants through the unsaturated zone. However, discrepancies between model results and actual field measurements are commonly found in the literature. In particular, several studies have detected high concentrations of pesticides in groundwater shortly after the first rainfall. Modelling and monitoring of moisture and solute transport through the unsaturated zone is difficult due to the complicated networks of interconnected pathways in the soil which can transmit water and its solutes at varying velocities. Preferred pathways in soils may result from biological and/or geological activity (e.g., macropores, earthworm burrows, channels consisting of highly conductive media) or from farm management practice (e.g., conservation tillage). Such paths may transmit water and its solutes at much higher velocities than those predicted by Darcy's theory. Further, many mathematical models assume uniform flow or do not include biological biodegradation below the root zone. However, recent research indicates that degradation of pesticides and other solutes may occur well below the root zone in the neighbourhood of macropores and cracks. Consequently, a better understanding of the factors affecting the percolation of water and the mobility, extent and nature of chemical transport and biological degradation of solutes is required to improve modelling efforts and management practices. In this chapter, the nature and impact of preferential flow on moisture and solute transport is discussed.

V. P. Singh (ed.), Environmental Hydrology, 193–224.
© 1995 *Kluwer Academic Publishers.*

7.1 Introduction

During recent years, worldwide concern has focussed on the potential for contamination of surface waters and groundwater by surface-applied chemicals in runoff and soil-water (Bergström and Jarvis, 1993). Regardless of the application mode, a portion of all surface-applied chemicals including pesticides eventually reaches the soil. After soil interception, the interaction of complex chemical and physical processes determine whether the chemical remains there or is transported downward to the groundwater. Below the soil surface, previously held belief that the root zone can act as a "living filter" capable of storing and biologically, chemically or radiologically degrading chemicals has given way to the reality that such chemicals often do not degrade in a sufficiently short time and that their eventual migration through the unsaturated zone to the groundwater is highly probable, particularly if the soil is well-structured consisting of macropores and channels of high hydraulic conductivity (Gee et al., 1991; Parlange et al., 1988; Smettem et al., 1991; Smettem and Ross, 1992; Edwards et al., 1993a).

Recent experimental research in soil structure and non-ideal chemical transport is forcing scientists to refine their analyses of the movement and fate of contaminants in the soil, and to consider processes which have become important because even small concentrations of highly toxic chemicals must be accurately accounted for in order to protect groundwater reserves from contamination. The mobility of chemicals and contaminants in soils is affected by the continuity as well as the size of the pores. Networks of interconnected, highly-conductive pathways which result from biological and geological activity, such as subsurface erosion, faults and fractures, shrink-swell cracks, animal burrows, worm holes, and decaying roots, may be responsible for transmitting moisture, solutes and contaminants to groundwater at much faster velocities than those predicted by current theory (Steenhuis et al., 1990; Stagnitti et al., 1991; Parlange et al., 1994a). Water and chemicals that travel in these pathways often bypass the bulk of the soil matrix. For example, Smettem and Collis-George (1985) found that a single, continuous 0.3 mm diameter macropore can conduct more water than a surrounding soil sample of 100 mm diameter. Some persistent and mobile chemicals (e.g., atrazine) can be transported long distances by soil water and may be detected far from the application site (Frank and Sirons, 1979; Glotfelty et al., 1984; Steenhuis and Parlange, 1990). To be acceptable for agricultural use, pesticides and other surface-applied chemicals must have mobility and persistence properties that fall within a narrow range. They must be sufficiently mobile to reach, and persist long enough to eliminate, the target organism. However, the physical heterogeneity of the natural environment makes accurate prediction very difficult.

Sorption of chemicals by the soil is also a major factor influencing their accessibility to the target organism and their potential to contaminate groundwater (Kookana et al., 1992a,b; Aylmore and Kookana, 1993). Since the soil-organic fraction governs the sorption potential, it will also govern the mobility. Sorption typically has two phases: early rapid adsorption, followed by slow diffusion of the dissolved chemical into the organic matrix (Gish et al., 1991). This suggests that the greatest potential for chemical movement is immediately after application. In addition to chemical parameters, chemical mobility is affected by surface-water

fluxes, i.e., the rate of application. For example, Rice et al. (1991) found that accelerated leaching of solutes occurred following intermittent flood and sprinkler application. However, under continuously flooded conditions, leaching was less rapid.

Even in relatively uniform sandy profiles, nonhomogeneous flow patterns are found as well. Typically, only 1 to 20% of the vadose zone is actually involved in the transport of the solutes and moisture (Starr et al., 1986; Parlange et al., 1994a). Rice et al. (1991) found that solutes and herbicides travel at velocities 1.6 to 2.5 times faster than traditional water balance models of piston flow in sandy soils with little or no structure. Steenhuis et al. (1994a) found that this rate can be as high as 6 times under certain irrigation regimes. Kladivko et al. (1991) found that relatively high concentrations of bromide and pesticides were detected very early in water samples collected from tile drains buried 75 cm in a silty loam soil. Kung, (1990a,b) found that water and solutes flowed through less than 10% of the total soil matrix between depths of 3.0 and 3.6 m and less than 1% between 5.6 and 6 m. He also found that the traced water had bypassed much of the soil matrix by flowing laterally over layers and vertically downwards through preferred paths.

The best way to account for all possible soil types, chemical environments and climatic conditions would be to conduct field experiments and monitor water and chemical movement on each site. However, this is clearly impractical. So, methods have been developed that simulate, measure and predict the way in which water transports solutes through the soil: This chapter reviews some recent developments in monitoring and modelling preferential moisture and solute transport and focuses particularly on physical transport of pesticides in the unsaturated zone. More traditional transport processes with uniform water flow across the soil layers are discussed in the preceding chapter (Wierenga and Brusseau, 1994).

7.2 Preferential Moisture and Solute Transport

Preferential flow is the rapid and nonuniform transport of soil moisture and solutes through macropores and subsurface channels. Many field studies have shown that preferential flow is a wide spread phenomenon (e.g., Starr et al., 1986; Steenhuis et al., 1994a). Flow in macropores can occur with little or no interaction with the surrounding soil matrix. On the other hand, matrix flow is the comparatively slower movement of water and solutes percolating though micropores. The relative importance of the two forms of percolation is dependent on the soil type and rainfall intensity (Kluitenberg and Horton, 1990; Shipitalo et al., 1990; Parlange et al., 1994a). For example, well-structured soils consisting of clay and loam typically experience low permeability rates. In such soils, less than 1% of the pore volume consists of cracks and subsurface channels. When it rains, water infiltrating from the soil surface often flows through these channels in preference to the surrounding soil matrix, whose small pores are penetrated comparatively slowly. Even though these channels consist of a relatively small percentage of the total pore volume, they may be responsible for the bulk of moisture and solute transport after an infiltration event (Steenhuis and Muck, 1988; Everts and Kanwar, 1988, 1992).

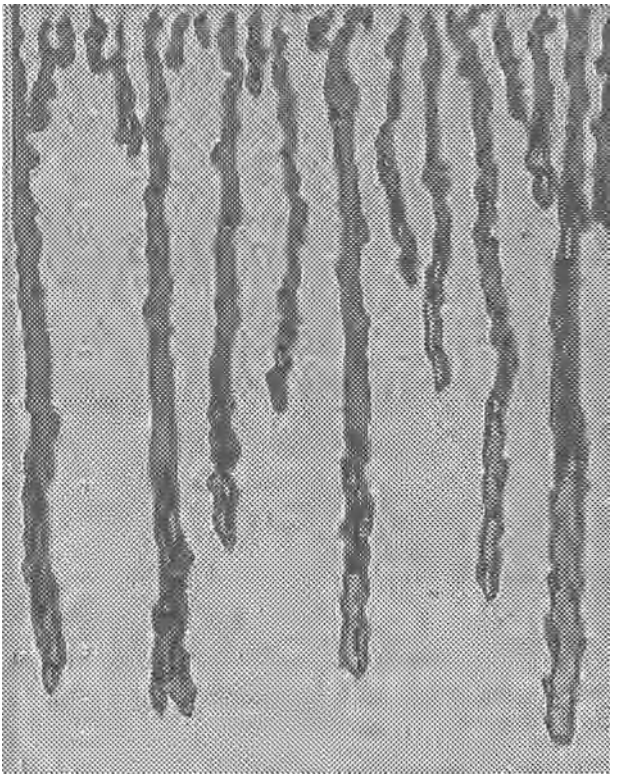


Fig. 7.1. The creation of soil water "fingers" in homogeneous sandy soils resulting from wetting front instability. The darker regions represent higher soil moisture content.

Further, preferential flow in structured soils can be initiated well below soil water saturation.

Solute concentrations in the soil matrix and in preferred paths also may be very different. The chemical composition of solutes in the macropores and subsurface channels often reflects the concentration of the water near the surface during an infiltration event. On the other hand, the solute concentrations in the soil matrix reflect an equilibrium concentration that was established prior to the infiltration event. Thus for certain infiltration events, high solute concentrations can leach to the groundwater before dispersion and/or biological or chemical degradation can occur (Hegert et al., 1981; Steenhuis and Muck, 1988; van Es et al., 1991; Richard and Steenhuis, 1988; Bouma, 1990). These studies also confirm that the distribution of the solutes is highly variable. Contrary to widely held belief, even in homogeneous soils, water does not necessarily take an average flow path. Hill and Parlange (1972) documented for the first time, preferential flow in homogeneous sand at low infiltration rates. Further research by others has led to an understanding of many of the mechanisms involved in this phenomena (Parlange and Hill, 1976;

Raats, 1973; Philip, 1975; Hoa, 1981; Starr et al., 1986; Hillel and Baker, 1988; Glass et al., 1989a,b,c; Glass et al., 1991; Selker et al., 1992a,b; Selker et al., 1994; Dekker and Ritsema, 1994; Liu et al., 1993, 1994a,b). This line of research has shown that water can find its way down through homogeneous, sandy soils in a number of channels, called fingers, which result from a wetting front instability (Figure 7.1).

The finger frequency is important for the estimation of the representative elementary volume used in solute sampling. For sandy and water-repellent soils, the finger frequency (f_f) can be found by dividing the wetted cross-sectional area (a_w) by the wetted area of each finger (A_f):

$$f_f = \frac{a_w}{A_f} . \tag{7.1}$$

The wetted portion of the soil and consequently the number of fingers per cross-sectional area is determined by the largest rainfall event into a dry soil (Selker et al., 1994). Once developed, the fingers remain in the same location with subsequent wetting and drying cycles, but more fingers can form if the carrying capacity of the existing finger network is exceeded. The cross-sectional area of a finger, as characterised by its diameter, d, has been extensively studied in laboratory experiments. Following the approach of Parlange and Hill (1976), the finger width is determined by balancing the forces of gravity and surface tension, and fingers form if the rainfall rate is much smaller than the saturated hydraulic conductivity (Liu et al., 1994a,b):

$$d = \frac{\beta S^2}{K_f \theta_f} , \tag{7.2}$$

where β is a constant, equal to 4.8 for three-dimensional fingers, K_f and θ_f are the hydraulic conductivity and moisture content at the fingertip, respectively, and S is the sorptivity at the fingertip for an initially dry soil and is given by:

$$S^2 = 2\theta_f^{1/2} \int_{-\infty}^{h_f} \theta^{1/2} K \, \mathrm{d}h, \tag{7.3}$$

where K is the unsaturated soil water hydraulic conductivity and h_f is the matric potential at the finger tip. Liu et al. (1994a) further simplified Equation (3) and obtained an expression for finger diameter which is equally valid for an initially dry or wet soil:

$$d \simeq \frac{2\beta\theta_f \dfrac{\mathrm{d}h}{\mathrm{d}\theta}}{\lambda + 1.5} , \tag{7.4}$$

where λ is related to the pore size distribution of the soil with values ranging from 3.3 to 4, and $\mathrm{d}h/\mathrm{d}\theta$ is the slope of the wetting branch of the soil characteristic curve at θ_f and is given by (Liu et al., 1994b):

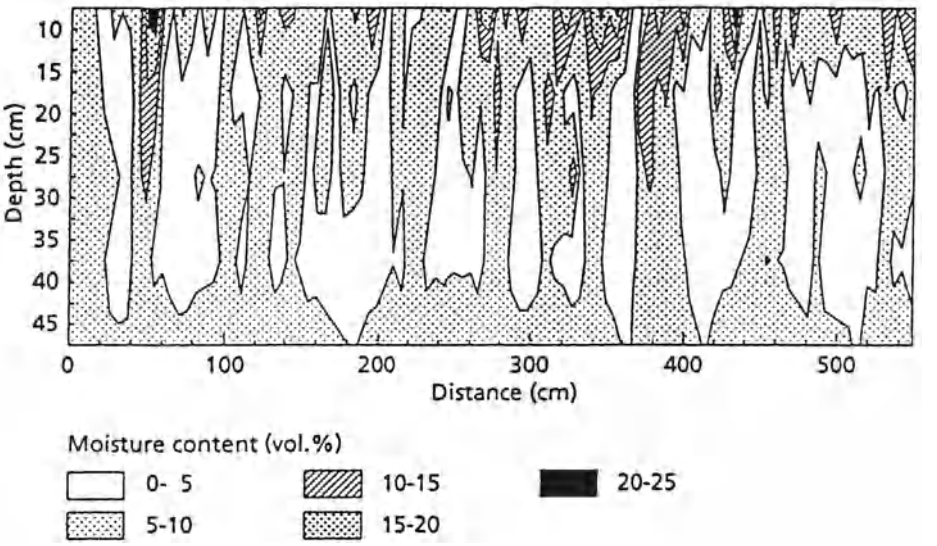


Fig. 7.2. Moisture distribution in soil transect, clearly showing fingers (hatched areas) for fine dune sand in Ouddorp, The Netherlands.

$$\frac{\mathrm{d}h}{\mathrm{d}\theta} = \frac{h_f}{\theta_f - \varepsilon}, \tag{7.5}$$

where ε is the soil porosity which can be found from the bulk density of the soil. The diameter, d, of each finger is essentially independent of the surface flux, q, and the amount of water carried by the fingers is a property of the soil. The distribution of water within a given finger is not uniform. Not only is there a sharp gradient at the finger's surface which is responsible for lateral diffusion, there is an increase of water content with depth as well. Although, the initial moisture content of the soil does not explicitly appear in the simplified mathematical formulation above, the finger diameter does differ with differing initial moisture contents. In general, the higher the initial moisture content, the wider the finger (Liu et al., 1994b). Thus under field conditions one expects that fingers, once established, will widen over time as the soil becomes wetter. This conjecture has been experimentally confirmed (Steenhuis et al., 1994a).

Wetting front instabilities can occur when there is a distinct difference in layering (e.g., Hillel and Baker, 1988) or when the surface application rate is much lower than the saturated hydraulic conductivity. Thus, fingering may occur widely in the field when soils are coarse or water-repellent (Ritsema and Dekker, 1993a,b). In homogeneous sand, the fingers are nearly vertical and generally do not merge with one another. On the other hand, in less homogeneous sand, the fingers can merge. In the field, where soil heterogeneities are far more pronounced than in laboratory experiments, most merging takes place in a shallow layer near the surface called the induction zone. Below the induction zone, merging is less common as long as the water flows through a relatively homogeneous medium. But, further down in the soil, sloping interfaces of texturally different layers may act as funnels (Kung, 1990a,b). Such structures can concentrate the water from a

large area into a single finger. Fingers may also split. This is the probable outcome when a finger reaches saturation and cannot carry the imposed flow. Figure 7.2 illustrates the development of wetting front instabilities for a fine dune sand in Ouddorp, The Netherlands (Steenhuis et al., 1994a).

7.3 Sampling for Preferential Flow in the Unsaturated Zone

Sampling the unsaturated zone can serve as an early warning system for groundwater contamination. However, one of the greatest uncertainties in monitoring groundwater contamination is the possibility that solutes flowing in preferred paths bypass samplers (Biggar and Nielsen, 1976; Barcelona and Morrison, 1988). Scaling point source measurements to field-scale estimates, therefore, may be unreliable and misleading even in the case when geostatistical methods are employed (Steenhuis et al., 1990). How sampling is affected by preferential flow patterns was studied by Beasly (1976) for forested watersheds, by Shaffer et al. (1979) and Shuford et al. (1977) for water sampling with suction cups in field soils. Methods for continuously sampling solutes in the unsaturated zone involve the collection of soil water drained by either the force of gravity (e.g., gravity pan samplers, agricultural tile lines, and shallow wells) or by applying a "capillary" suction (e.g., porous cup and wick pan samplers). Whilst all these sampling techniques result in collection of solutes and water in the unsaturated zone, only tile lines and pan samplers measure the amount moving to the aquifer (Steenhuis et al., 1994a). Gravity pan samplers collect the percolate from a saturated portion of the soil immediately in contact with the sampler possibly leading to "bypassing" (i.e., solutes and water flowing around the sampling device).

A modification of the gravity pan sampler is the Alundum tension plate sampler, in which the percolate is extracted from the unsaturated soil by suction applied across an alundum filter disc. Fibreglass wick samplers operate by the same principle (Rimmer et al., 1994). The wicks are self-priming and act as a hanging water column, providing a suction to the unsaturated soil. The wicks also are nonreactive and, therefore, can be used to sample solutes. Original wick sampling units consisted of a single fibreglass wick spread over a 30×30 cm area. The disadvantage of this design is that solutes entering the sides of the sampling unit had to travel a considerable distance to the center, whilst solutes near the center flowed without delay, giving rise to large instrument dispersion. Boll et al. (1991) improved this design using a multisegment percolation system as shown by Figure 7.3.

Multi-segment percolation systems have been applied successfully to study the transport of fertilisers and tracers (e.g., bromide) in unsaturated zone experiments (Boll et al., 1991). These units can be installed in field sites or, alternatively, undisturbed soil cores from selected sites can be extracted and transported to the laboratory for detailed study. In Boll's design (Figure 7.3), 25 individually sampled fibreglass wicks are placed in a 5×5 grid on the basal surface area of the sampling unit. The unit is then firmly pressed against the soil surface by springs. The length of the wicks provide a capillarity equivalent to that found in the soil and thus can be used to sample unsaturated flow. Being a porous medium, the wicks have been shown to provide boundary conditions which mimic those found in the undisturbed

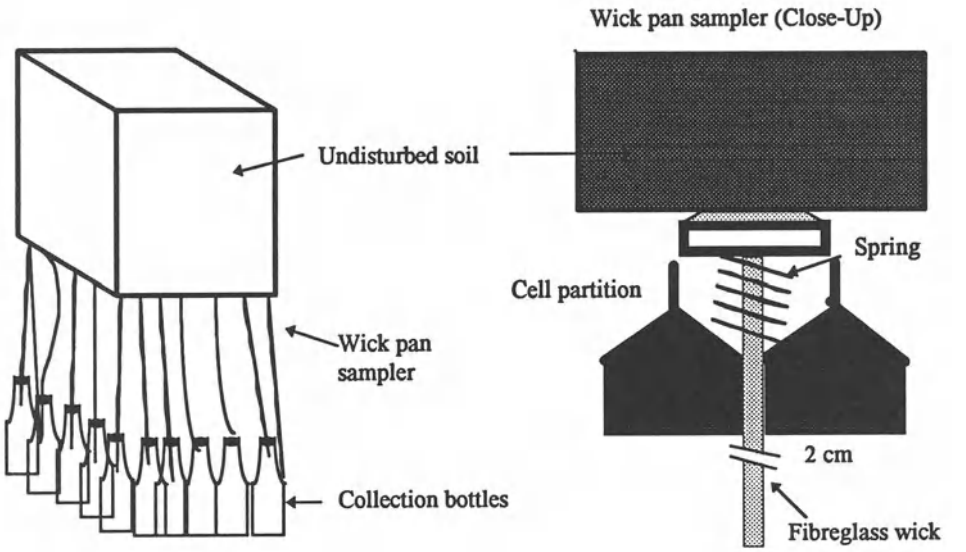


Fig. 7.3. Multisegment wick percolation system used for sampling moisture and solutes in the unsaturated zone (Boll et al., 1991).

soil. For instance, the capillary force in the wick decreases with increasing flux, thus eliminating flow field distortions created by suction cup samplers.

Boll et al. (1991) have conducted extensive field and laboratory investigations using a variety of sampling methodologies. They found that the effectiveness of each method largely depends on soil type and that no one device is best under all experimental situations. However, the multisegment wick sampling system better represented flow through the unsaturated zone when moisture and solutes were transported predominantly by matrix flow (e.g., in a loamy sand). On the other hand, in a clay soil, where the bulk of the transport was through macropores, both the gravity and wick pan samplers performed equally well. Porous cup samplers were not recommended for use in the unsaturated zone. These sampling devices spuriously indicated preferential solute movement in loamy soils and failed to register any solute flow in clay soils with macropores (Steenhuis et al., 1994a).

7.4 Characterisation of Soil Structure

Currently, one of the biggest hurdles in modelling and monitoring solute transport is an inability to properly characterise the macroporous structure of the soil. A knowledge of active transport zones in the soil matrix would provide a better understanding of the role and relative importance of macroporous channels in the transport process. Recently, X-ray Computer Assisted Tomography (CT-scans) has been used to characterise large biopores in unexposed interior portions of undisturbed soil samples (e.g., Withjack, 1988; Keller et al., 1989; Anderson et al., 1992; Peyton et al., 1992, 1994). This method is able to nondestructively measure macropore scale density in undisturbed soil cores (Peyton et al., 1994) and visualise three-dimensional flow patterns for solutes (Withjack, 1988). Hole porosity and

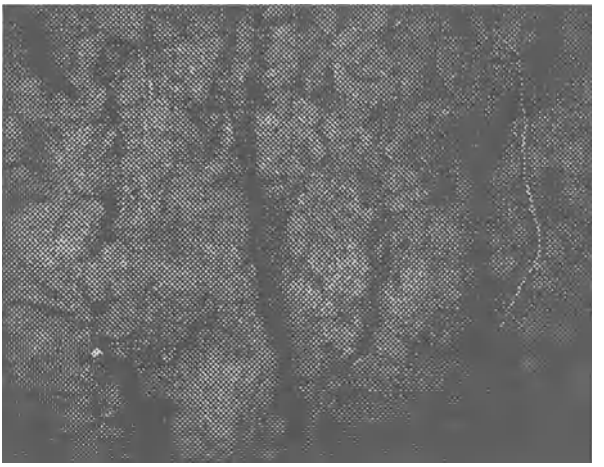


Fig. 7.4. Preferential flow paths in a structured soil near Ithaca, NY, are coloured by a blue dye, which infiltrated the soil from the surface.

equivalent hole perimeters and diameters to a resolution of 0.5 mm can also be identified by this method.

Whilst the CT-Scanning technique is an excellent nondestructive sampling methodology, it is very expensive and not readily available to most scientists. An alternative, inexpensive, but also less effective, method of characterising the soil matrix is to use food colouring dyes. Various dyes can be added to the soil surface and allowed to infiltrate to lower portions of the soil horizon. Later, a vertical cross-section of the soil profile can be obtained by excavation (e.g., Figure 7.4).

In a laboratory setting, the use of coloured dyes may be helpful in identifying flow paths for chemical and moisture transport in undisturbed soil columns from the field. For example, at the completion of a flow experiment, the undisturbed soil core can be stained with various coloured dyes and dismantled or serially sectioned and photographed using high resolution film or videography. These images can then be captured and processed using computer image enhancement software (e.g., Ghodrati and Jury, 1990; Vermeul et al., 1993). Once the images are digitised, software can be used to identify and group pixels into macropores. These parameters can be used to develop algorithms for determining equivalent hole diameters, porosity and circulatory indexes (Vermeul et al., 1993). Similar videographic and photographic methods have been developed to study moisture profiles in wetting front instabilities for homogeneous and layered soils (Yanuka and Elrick, 1985; Glass et al., 1989a,b,c; Glass et al., 1991; Selker et al., 1992a,b). An alternative method to using coloured dyes is to impregnate the soil core with paraffin containing anthracene, a fluorescent dye. After freezing and serial sectioning, each segment or layer of soil can then be photographed under UV light. The intensity of the illumination provides an indication of the relative rate of diffusion between the macropores and surrounding soil matrix. Unfortunately, all of these techniques are destructive and can only provide static information at the end of an experiment.

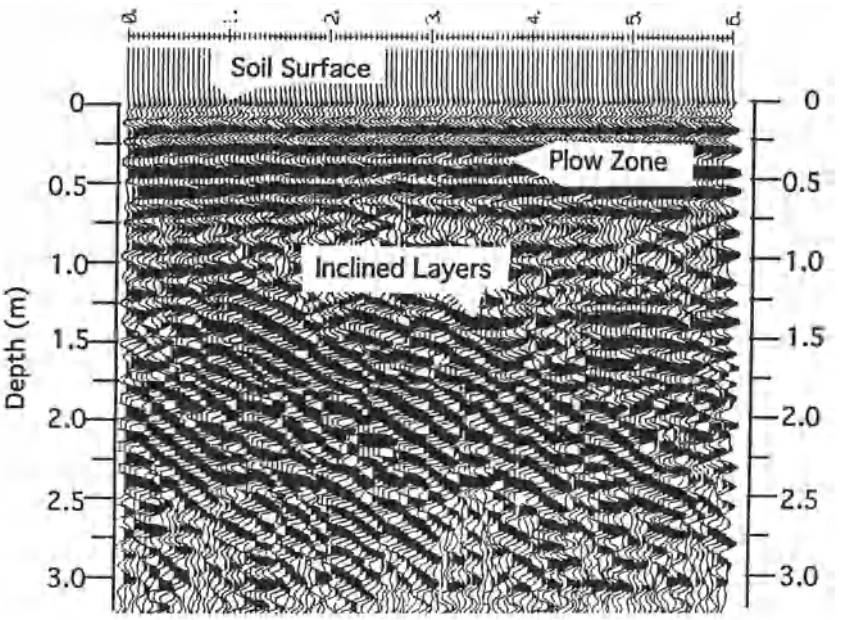


Fig. 7.5. Interpreted GPR section showing steeply inclined layers in Evesboro loamy sand, near Georgetown, DE (Boll et al., 1993).

There is also the problem of scaling; data obtained from soil characterisation at a local scale may not adequately represent the heterogeneous nature of the soil profile on a field scale. These techniques appear to be useful only for laboratory scale experiments.

On large field scale experiments the use of Ground-Penetrating Radar (GPR) as a nondestructive soil characterisation method has been proposed. GPR can be used to detect water table elevation, subsurface textural discontinuities, subsurface structures, wastes and waste migration. The technique has been largely used in drainage and construction engineering and operates much like sonar and reflection seismology. A short pulse of high frequency electromagnetic energy is transmitted into the ground. Changes in electrical properties of the soil, which are primarily controlled by textural characteristics (e.g., clay and moisture content) cause part of the transmitted signal to be reflected. The reflected signal is detected by a receiver and after amplification the amplitudes of the two-way travel time are stored in computer files for further analysis (Boll et al., 1993 for further details). Figure 7.5 is an example of the use of GPR for improving the siting of subsurface sampling devices. The investigation site was the University of Delaware Research and Education Center near Georgetown, DE. The field site consisted of a loamy sand. This figure shows a 5 m section. The upper 60 cm shows the direct arrival and reflection off the surface and plow layer, after which a small zone of homogeneous material follows to a depth of 1 to 1.2 m. Below 1.2 m, stratification is clearly visible showing inclined layers. The depth of penetration in this case was 3 m and

the depth of the water table was located at 2.5 m. GPR is a useful but difficult technique. It is also moderately expensive. The penetration depth of the radar wave depends on the antenna frequency, which in turn determines the resolution of the image. Greater penetration depths usually mean lower sampling resolution. The penetration depth also depends on soil properties. It is greatest for dry sandy soils.

7.5 Modelling Preferential Transport

Despite many experimental findings, the preferential flow phenomena has not been adequately incorporated into simple mathematical models for predicting moisture and solute flow in the unsaturated zone. The major difficulty of modelling this transport phenomena is primarily due to characterising the flow paths within the soil profile. The approach to predict water and solute transport is derived from the work of Richards who formulated a theory of water movement in unsaturated soils (Richards, 1931). van der Molen (1956) combined aspects of Richards' model with the theory of dispersive movement to predict desalinisation of land that had been inundated by seawater in The Netherlands. The resulting equation, the Convective-Dispersive Equation (CDE), assumes that water and solutes follow an *average* path through the soil. Others (e.g., Parker and van Genuchten, 1984; Sposito and Jury, 1988; Barry and Sposito, 1988) have improved the solute transport model by incorporating a stochastic description of soil heterogeneities (Barry, 1993; and Barry and Li, 1994 for further reviews). However, these results are usually limited to cases of constant or uniform mean velocity as well as constant dispersion (Gee et al., 1991). Transfer function models have received considerable field testing and have shown that scale dependent dispersion is present in the field (Li et al., 1994). The conclusion from these studies is that the CDE with a constant dispersion coefficient does not give as great a solute dispersion with depth than is observed (Gee et al., 1991).

Several more recent models separate matrix flow from macropore flow. These dual porosity models (e.g., Gerke and van Genuchten, 1993), typically write separate transport equations for each domain of flow (i.e., matrix and macropore). Mixing of water and solutes across the interface is permitted and described by transfer coefficients which are often functions of the pore size. Steenhuis et al., (1990) followed a similar conceptual framework and proposed a mathematical model that considers not two domains but any number of domains. Extensions and improvements to the model were later proposed by Steenhuis et al. (1991) and Stagnitti et al. (1991). In any modelling attempt, the characterisation of the soil profile is problematic. In the model proposed by Steenhuis and coworkers, preferential flow paths are simulated by taking piece-wise linear approximations of the hydraulic conductivity, resulting in two model parameters: N, the number of pore groups with mobile water and, n_p, the transport velocity of each pore group. These parameters can be related to physical properties but they usually require calibration to a particular field or laboratory experiment. Unlike the usual modelling assumptions applied in the CDE, the concentration of solutes in the percolating water is dependent on the varying rate of applied water and the time period between rainfall and chemical application. This is achieved by relating the solute flux to the water

flux. Thus transient field conditions can be simulated. The model can be applied to both large scale field experiments and small scale laboratory experiments. As the model has been fully described elsewhere (Steenhuis et al., 1990; Stagnitti et al., 1991; Parlange et al., 1994b), we shall present only a brief description here.

Water and solutes may travel through a soil by any number of possible routes and at differing velocities. Thus the residence times of water soluble contaminants in the vadose zone is highly variable. However, for purposes of modelling, even though these paths may be highly distributed throughout the soil, they may be collected into groups of pathways or 'capillary bundles' in which moisture and solutes travel at approximately the same flux. These 'capillary bundles' are also termed 'pore groups'. The question of how one determines these capillary bundles or pores groups was addressed in Stagnitti et al. (1991) and Parlange et al. (1994b). The total amount of moisture, $\theta(x,t)$ in the soil at time, t, and point, x, is the sum of all individual moisture contents for each capillary bundle, p:

$$\theta(x,t) = \sum_{p=0}^{N} \theta_p(x,t), \tag{7.6}$$

where θ_p is the individual moisture content for the pth pore group. The maximum amount of moisture that each group can hold and transmit is $\Delta M_p = M_p - M_{p-1}$, where M_p and M_{p-1} are various moisture contents representing upper and lower limiting values for the pth group and are a function of the size of pores in each group. When $\theta_p = \Delta M_p$, then all the pathways for the pth group are completely saturated with soil moisture. The soil is locally saturated when all flow paths in all groups are saturated, i.e., $\theta(x,t) = \theta_s$ and $\theta_p(x,t) = \Delta M_p$ for all p's where θ_s is the saturated moisture content of the soil. The group's moisture content, θ_p, is a function of the vertical percolation rate, q_p, and of the effects of precipitation, evapotranspiration and of loss or gain of moisture from interactions and exchanges with other groups. Therefore, from continuity, the transport equation may be written as:

$$\frac{\partial[\theta_p(x,t)]}{\partial t} + \frac{\partial[q_p(x,t)]}{\partial x} = A_p(x,t); \tag{7.7}$$
$$\text{for } t \geq 0;\ x \geq 0;\ 0 \leq \theta_p \leq \Delta M_p;\ p = 0,\ldots,N.$$

where $A_p(x,t)$ is a source/sink term representing the effects of precipitation, evapotranspiration and mixing between other groups; t is time, and x is a distance with $x = 0$ being the soil surface; N is the number of mobile pore groups. From consideration of Darcy's Law:

$$\begin{aligned} q_p(x,t) &= k(\theta_p)H(x,t); \quad \text{for } p = 1,\ldots,N, \\ q_0(x,t) &= 0; \qquad\qquad\quad \text{for } p = 0 \end{aligned} \tag{7.8}$$

and

$$0 \leq \theta_p \leq \Delta M_p; \quad q_0 \leq q_1 \leq q_2 \cdots \leq q_N,$$

where $H(x,t)$ is the hydraulic gradient and $k(\theta_p)$ is the hydraulic conductivity. For moisture in the unsaturated zone, a unit gradient is assumed, i.e., $H(x,t) = 1$. This is a useful and accurate approximation when gravity is the main driving force and is appropriate for percolation in macropores (Parlange et al., 1994b). Therefore, after substitution:

$$\frac{\partial \theta_p}{\partial t} + v_p \frac{\partial \theta_p}{\partial x} = A_p, \quad \text{where } v_p = \frac{\mathrm{d}k(\theta_p)}{\mathrm{d}\theta_p}, \tag{7.9}$$

where v_p is the velocity of percolating moisture in the pth group. Note that for $p = 0$, $v_p = 0$. The solution of Equation (9), obtained by the method of characteristics (e.g., Charbeneau and Street, 1979; Charbeneau, 1981, 1984), is given by:

$$x = v_p t + x_0, \tag{7.10}$$

and

$$\theta_p(x,t) = \int_0^t A_p \, \mathrm{d}t + \theta_p(x_0, 0); \quad x \geq v_p t, \tag{7.11}$$

where $\theta_p(x_0, 0)$ is an arbitrary function of x_0 when $t = 0$. Equation (11) cannot be integrated easily because A_p is generally not a simple analytical expression. Therefore, a fixed time step numerical solution for the moisture content θ_p was proposed. From Equation (10), in a small time interval, Δt, moisture will travel a distance, Δx_p, given by:

$$\theta_p(x,t) = \theta_p(x - \Delta x_p, t - \Delta t) + Ip(x,t); \quad \Delta x_p = v_p \Delta t, \tag{7.12}$$

where I_p is the integral of A_p between $(t - \Delta t)$ and t. Let Δx_p be a simple integer multiple γ_p of Δx_1, the distance travelled by moisture in the slowest moving group $(p = 1)$, i.e., $\Delta x_p = \gamma_p \Delta x_1$, with $\gamma_1 = 1$. Note that in the "smaller" time step, $\Delta t/\gamma_T$, the pth group moved a distance Δx_1. Therefore, define a grid of points in $(x - t)$ space based on Δx_1 and $\Delta t/\gamma_T$ where $T = N + 1$, is the label for the pore group with the greatest flux. Then:

$$x = i\Delta x_1; \quad i = 1, 2, \ldots, \frac{L}{\Delta x_1} \text{ and } t = j\frac{\Delta t}{\gamma_T}; \quad j = 1, 2, \ldots,$$

where L is the depth to groundwater. On this grid, only one parameter, either the spatial coordinate, Δx_1, or the time coordinate $\Delta t/\gamma_T$ needs to be specified. It is usually more convenient to fix time rather than distance and, therefore, the following equation for moisture transport results:

$$\theta_p^{(i,j)} = \theta_p^{(i - \delta_p \gamma_p, j-1)} + I_p^{(i,j)}, \tag{7.13}$$

where

$$\delta_p = \begin{cases} 0; & j \bmod \gamma_p \neq 0, \\ 1; & j \bmod \gamma_p = 0. \end{cases}$$

The solute concentration, $c(x,t)$, in the soil is given by:

$$c(x,t) = \frac{s(x,t)}{\theta(x,t)}\,, \tag{7.14}$$

where $s(x,t)$ is the total solute load given by:

$$s(x,t) = \sum_{p=0}^{T} s_p^{(i,j)} \tag{7.15}$$

and s_p is the solute mass in pore group p. Assuming, for the moment, that chemical processes like adsorption and solute precipitation are negligible, then the solute flux may be simply written as:

$$s_p^{(i,j)} = s_p^{(i-\delta_p \gamma_p, j-1)} + \psi_p^{(i,j)}, \tag{7.16}$$

where $\psi_p^{(i,j)}$ is the solute amount added to or subtracted from the pth group during $\Delta t/\gamma_T$. In each time step the concentration of solutes in each group is affected by three processes: evapotranspiration (moisture extracted from the soil), precipitation (moisture entering the soil after infiltrating the soil surface) and mixing or interchange between capillary bundles in order to attain chemical equilibrium. These three processes determine the value of $\psi_p^{(i,j)}$. Each process is described below. In the first step, moisture is extracted from each capillary bundle or pore group in the soil in various amounts as a result of evapotranspiration. Evapotranspiration also helps to redistribute moisture and, consequently, solutes between groups. Let $E^{(i,j)}$ represent the evapotranspiration loss from the soil at time j and depth i. Since this loss does not have to be uniform with depth nor with respect to the capillary bundle, let ξ_p represent the fraction of the total loss that was taken from group p during $\Delta t/\gamma_T$. If $\theta_p^{(i,j)}$ represents the moisture content just prior to accounting for evapotranspiration in the model, then the new moisture content after subtracting the evapotranspiration loss for the pth group is given by:

$$\theta_{p_E}^{(i,j)} = \theta_p^{(i,j)} - \xi_p \frac{E^{(i,j)}}{D}; \quad \text{for } 0 \le \xi_p \le \theta_p^{(i,j)} \frac{D}{E^{(i,j)}}\,, \tag{7.17}$$

where D is the depth to the root zone and the subscript E represents the new moisture content after evapotranspiration has been accounted for in the time step j. Although the concentration is affected, the amount of solutes remains unchanged.

In the second step, the net precipitation (precipitation less interception) is distributed in the soil to the various capillary bundles. Exactly how this distribution occurs in the soil is a complex process and depends on the infiltration rate and the moisture status of the soil. Since this distribution is often nonuniform, let μ_p represent the fraction of the total precipitation that is added to the pth group following rain. Therefore,

$$\theta_{p_{ER}}^{(i,j)} = \theta_{p_E}^{(i,j)} + \mu_p \frac{P^{(i,j)}}{D}; \quad \text{for } 0 \le \mu_p \le \Delta M_p - \theta_{p_E}^{(i,j)} \frac{D}{P^{(i,j)}}\,, \tag{7.18}$$

where $P^{(i,j)}$ is the amount of precipitation at (i, j) and the subscript ER represents the new moisture content after accounting for evapotranspiration and precipitation. Let $C_R^{(i,j)}$ be the concentration of solute in the precipitation and let $s_p^{(i,j)}$ be the amount of solute in the group just prior to rain. Therefore, the amount of solute in the soil after rain is:

$$s_{p_{ER}}^{(i,j)} = s_p^{(i,j)} + \mu_p \frac{P^{(i,j)}}{D} (\Delta t/\gamma_T) C_R^{(i,j)} , \tag{7.19}$$

where $s_{p_{ER}}^{(i,j)}$ represents the amount of solutes in the pth pore group after accounting for precipitation.

The network of flow paths in the soil are interconnected. Therefore, in the final step, there is an exchange of moisture and solutes between the various capillary bundles ranging from complete exchange of moisture and solutes to no exchange or mixing. Mathematically, this behaviour might be described by three strategies: (a) complete mixing, (b) partial mixing, and (c) no exchange between pore groups within the simulated time interval. These three strategies covering the continuum of possible behaviour should be implemented within the model. For example, under option (b), an appropriate partial mixing strategy might be one in which a substantial portion of the moisture and solutes in the slower pore groups are exchanged within the time interval whereas for faster moving pore groups or macropore groups very little exchange occurs in the simulated time step (Skopp et al., 1981). Various methods for implementing these mixing strategies in a computer algorithm were discussed by Stagnitti et al. (1991).

All strategies can be modelled easily by defining two sets of coefficients, $\lambda_p\theta_p$ and $\eta_p\theta_p$. Let λ_p represent the fraction of moisture extracted from group p and placed into a "common" pool ($0 \leq \lambda_p \leq 1$), and let η_p represent the fraction of the "common" pool that is moved back into the pth group after mixing. After mixing, the new moisture content will be given by:

$$\theta_{p_{ERM}}^{(i,j)} = (1 - \lambda_p)\theta_{p_{ER}}^{(i,j)} + \eta_p \sum_{g=0}^{T} \lambda_g \theta_{p_{ER}}^{(i,j)} . \tag{7.20}$$

The values for the mixing coefficients λ_p and η_p depend on the soil type and would normally have to be found by careful experimentation. In a similar fashion, after mixing is accounted for, we have:

$$s_{p_{ERM}}^{(i,j)} = (1 - \lambda_p)s_{p_{ER}}^{(i,j)} + \eta_p \sum_{g=0}^{T} \lambda_g s_{p_{ER}}^{(i,j)} , \tag{7.21}$$

where $s_{p_{ERM}}^{(i,j)}$ is the amount of solutes in group p after the mixing process has been accounted for and is equivalent to the right-hand side of Equation (16) after the processes of evapotranspiration, precipitation and chemical mixing and interchange have been accounted for. There are certain conditions and restrictions on the range of values that the mixing coefficients can have. These are fairly obvious and result from consideration of the conservation of mass. In particular, if a set of

coefficients for λ_p are adopted for a particular mixing strategy, then the values for η_p will be determined at each time step and depend on the amount of moisture in the pore group, the capacity of the pore group to store additional moisture and the value of λ_p. These conditions are described in Stagnitti et al. (1991) and Parlange et al. (1994b). Also, Parlange et al. (1994a) showed that there is a direct correlation between these mixing coefficients and dispersion coefficients obtained from the CDE for flow through homogeneous soils under steady moisture flux. Thus at the end of each time step, the concentration in each individual capillary bundle or pore group is:

$$c_p^{(i,j)} = \frac{s_{p_{ERM}}^{(i,j)}}{\theta_{p_{ERM}}^{(i,j)}}. \tag{7.22}$$

The preferential flow model has been successfully applied on a number of field and laboratory experiments (e.g., Stagnitti et al., 1991; Steenhuis et al., 1991; Nijssen et al., 1991; Parlange et al., 1994b). Figure 7.6, reproduced from Parlange et al. (1994b) illustrates application of the model to a series of experiments conducted on undisturbed soil cores by Anderson and Bouma (1977a,b). Anderson and Bouma conducted a series of laboratory experiments to determine the dispersion coefficients for soils consisting of blocky and coarse prismatic structures. In these experiments, three similar columns were subject to four different flow treatments using chloride as a tracer. Each soil core was 10 cm in diameter and the length ranged from 43 to 52 cm. Three replications of each experiment were studied. The first two experiments examined saturated flow conditions. The columns were initially saturated with water and then a 300 ppm solution of potassium chloride was ponded on the top of each column at shallow depth (Experiment 1). After the effluent concentration reached the input concentration, the columns were then flushed with water and drained, leaving the larger pores filled with air. Another continuous dose of 300 ppm of chloride concentration was applied by ponding (Experiment 2). This was then followed by another flushing and draining and then a pulse application of 1 cm per day irrigation by ponding (Experiment 3). The final experiment consisted of ponding water on a gypsum crust giving a steady, continuous infiltration of approximately 1 cm/d (Experiment 4).

Steenhuis et al. (1991) analysed these experiments by conducting two simulations. For the first simulation, $N = 7$ (i.e., seven mobile water pore groups and one macropore group), the velocity v_p (Equation (9)) of the slowest moving group was 1.85 cm/d and the fastest pore group was 118.6 cm/d. The simulation time.step was 3 minutes and solutes were blended at the following rates $\lambda_p = 0$, 0.1, 0.15, 0.5, 0.75, 1.5, 1.75, and 2.0% for every 3 minutes. Although the mixing coefficients were large for the larger pore groups, the overall mixing in the small pores was greater due to the longer retention time of solutes in those pores. In the second simulation, only the hydraulic conductivity function was modified. This had the effect of moving water and solutes in the fastest moving macropore group four times that in the first simulation. All other parameters remained unaltered. The predicted breakthrough curves for the two simulations plotted with the observed breakthrough curves are presented for each of the four experiments in Figure 7.6.

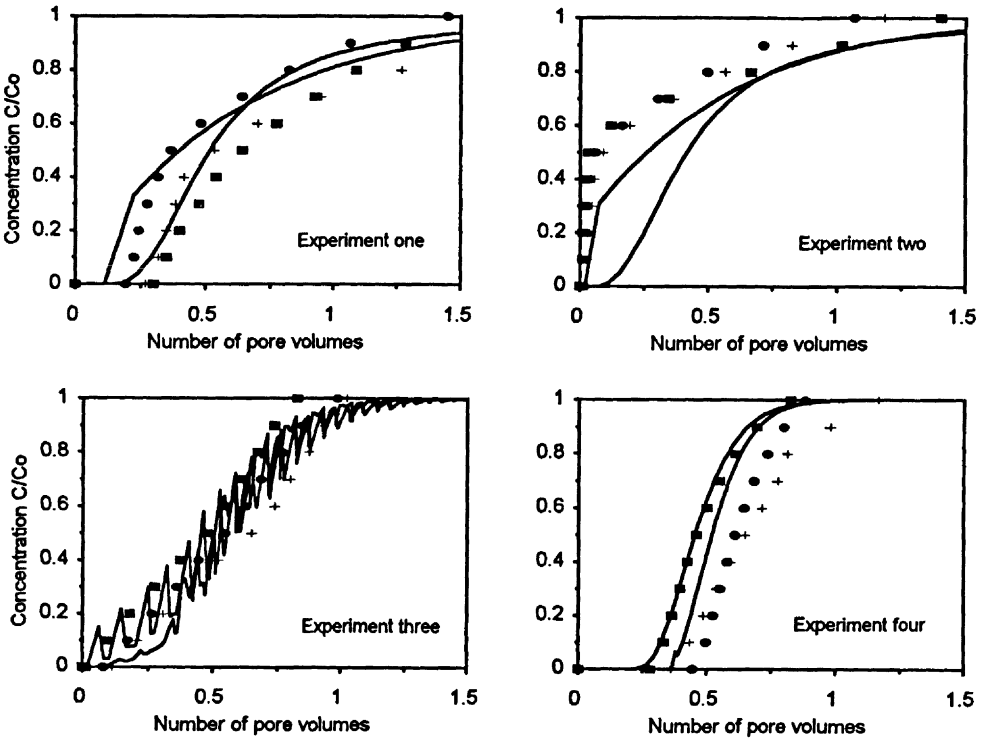


Fig. 7.6. Predicted and observed breakthrough curves for the Anderson and Bouma (1977a,b) data. The solid lines represent model predictions for the second simulation and the dashed lines represent model predictions for the first simulation. The symbols (squares, pluses and circles) represent the observed data for three different replications (soil columns).

The range of values for the observed breakthrough curves provides an indication of the experimental variation between the columns. In all cases, except for Experiment 2, the predicted breakthrough curves for both simulations fell within the range of observed values. For Experiment 3, the predicted pulse application showed large variation within a day. These were not observed in the original data set because the concentrations were only measured daily, whereas the model simulations were conducted on a three minute interval. If predicted concentrations were averaged on a daily basis a very similar breakthrough curve would result. The second experiment clearly shows the significant effect of macropore flow on the breakthrough curve. The almost instantaneous response in the effluent concentration was attributed by Anderson and Bouma to flow through the larger air-filled pores. The second simulation which enhanced the effect of the macropore flow more closely matched the observed values than the first simulation. It is very dif-

ficult to match the breakthrough curves for Experiments 1 and 2 with the same hydraulic conductivity function. It is likely that different conductivity functions for drained columns and saturated columns should be applied due to swelling of the soil and consequent narrowing of the cracks and voids (Steenhuis et al., 1991). The model, however, was able to simulate a wide variety of breakthrough curves without changing other physical properties of the soil. In contrast, Anderson and Bouma tried to fit dispersion coefficients to the experiments using the CDE but found that the coefficients varied considerably and were highly dependent on the flow regime.

7.6 Transport of Pesticides

In screening and registration programs, it is still common practice to estimate pesticide mobility by conducting short-term leaching tests on homogeneous soils in the laboratory. It is even more common to determine physical and chemical properties of the pesticides (e.g., adsorption constants, water solubilities and degradation rates) and simply predict leachability based on such information (Bergström and Jarvis, 1993; Jury et al., 1984). Environmental conditions in laboratory tests, however, can be quite different from natural soils and field conditions (Edwards et al., 1993a,b). Therefore, pesticide mobility in the field may be poorly predicted when based on laboratory experiments. Models of nutrient and fertiliser transport based on the CDE or the model presented in the last section have proven very satisfactory for the prediction of the overall transport of water and nutrients or tracers in soil columns and in field experiments. On the contrary, predictions of pesticide mobility with these models, except for very few cases, fail to represent the observed values. The predictive accuracy for pesticides and fertilisers is different because different quantities of applied chemicals will exceed health standards. For example, a very large percentage of fertilisers must reach the water table before toxic levels are exceeded. It is unlikely that the capacity of macropores in soils is large enough to transport such quantities in a sufficient amount or time to cause significant contamination. Only a comparatively small amount of pesticides needs to find its way to the water table through macropores to cause significant contamination of groundwater. Notwithstanding the considerable difficulties of modelling nonideal chemical transport in homogeneous soils, attempts to model and monitor the transport of small quantities of pesticides in heterogeneous soils is almost impossible, particularly on a field scale.

Recently, Steenhuis et al. (1994b) developed a simple conceptual model for predicting the concentration and time of arrival of pesticides in preferentially moving water. The conceptual framework of the model is shown in Figure 7.7 and is based on the observation that infiltration water is distributed in a thin layer at the soil surface before it is transported vertically (Figure 7.7a). In this layer, called the distribution zone, lateral moisture movement and mixing of surface-applied chemicals takes place prior to downward transport. During the time that the distribution layer becomes saturated, solutes can be absorbed into the soil matrix. After saturation, the pesticides are then desorbed and released into preferential flow paths or dispersed into the surrounding soil matrix. Water and pesticides flowing

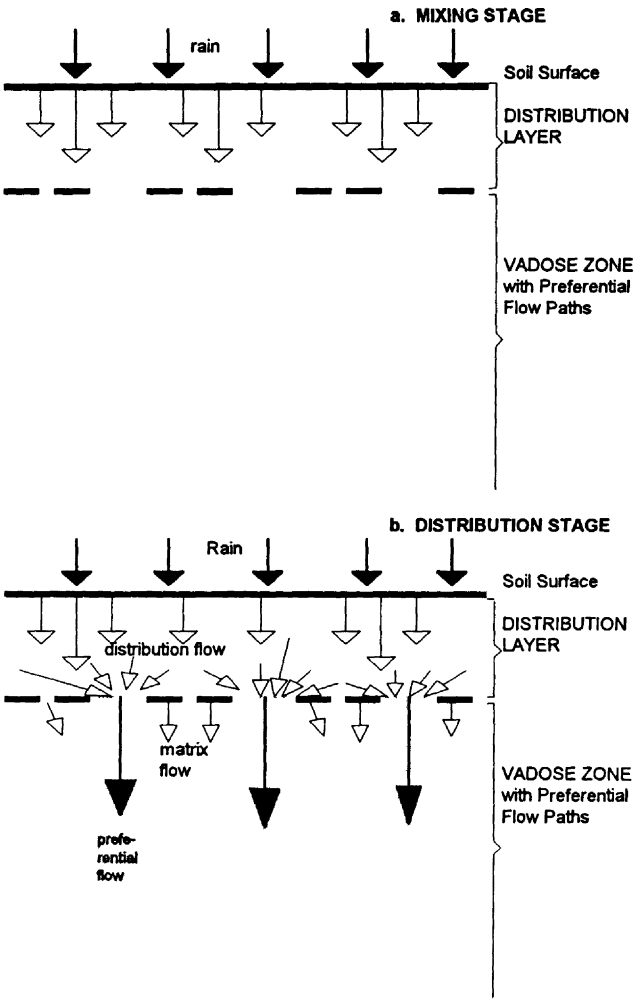


Fig. 7.7. Conceptual framework of preferential flow model. In the mixing stage (Figure 7.7a), water and solutes are absorbed in the distribution layer. After the distribution layer is saturated (Figure 7.7b), water and solutes are distributed to preferential flow paths (either macropores or fingers) (Steenhuis et al., 1994b).

through preferential pathways reach the groundwater with little modification or interaction with the surrounding matrix, and hence the concentration in this water is similar to the concentration of the surface-applied solution.

The calculation of the travel time and distance of pesticides in soils, requires the cross-sectional area of flow. Under convective flow, the average distance travelled by solutes in preferentially moving water, L, is expressed as:

$$L = \frac{qt_r}{\alpha_m \theta_e}, \tag{7.23}$$

where α_m is the fraction of the total pore space taken up by the mobile fractional of the soil ($\alpha_m = 1$ for the CDE), θ_m is the moisture in the mobile fraction of the soil after water is drained 24 to 48 hours after the rainfall event, q is the downward flux per unit area or specific discharge, and t_r is the duration of the rainfall event. For preferential flow, Steenhuis et al. (1994b) suggest that water and pesticides travelling in preferential paths move downward with a velocity close to the saturated hydraulic conductivity; an assumption that has been shown experimentally in finger flow experiments (e.g., Selker et al., 1992a,b). Also, the wetted portion of the cross-sectional area, a_w, adjusts itself such that the product of the saturated hydraulic conductivity and the wetted cross-sectional area equals the flux:

$$a_w = \frac{q}{k_s} . \tag{7.24}$$

This result was earlier proposed by Baker and Hillel (1990). Therefore, the distance can be expressed simply as:

$$L = \frac{k_s}{\theta_e}(t_r - t_d) \text{ for } t_r > t_d \text{ and } L = 0 \text{ for } t_r \leq t_d, \tag{7.25}$$

where t_d is the time to fill the distribution zone and k_s is the saturated hydraulic conductivity.

Assuming that the depth of the distribution zone is constant, the peak concentration of pesticide in the preferentially moving water may be expressed as:

$$C_w = C_{0w} \exp[-W_d^{-1} \int_{t_d}^{t_r} q \, \mathrm{d}t] , \tag{7.26}$$

where C_{0w} is the concentration at the end of the mixing phase in the distribution zone and before the distribution starts ($t = t_d$) and can be expressed as $C_{0w} = M_0/W_a$, where M_0 is the initial amount of pesticides applied. W_a and W_d are the apparent moisture contents per unit area in the distribution layer during adsorption and desorption phases and are equal to:

$$W_a = d_d(\theta_s + \rho k_a) \text{ and } W_d = d_d(\theta_s + \rho k_d), \tag{7.27}$$

where k_a is the adsorption-partition coefficient during the adsorption of water and solutes in the mixing stage, k_d is the adsorption-partition coefficient during the distribution stage, and d_d is the depth of the distribution zone. The distribution value, k_d is the tangent of the desorption isotherm and k_a is an average value over the entire range. For nonadsorbed solutes the apparent water contents per unit surface area are, as expected, equal to the amount of water in the distribution zone. For pesticides, the use of different partition coefficients is obviously justified, even through it is only approximate. Using this approach, Shalit and Steenhuis (1994) were able to predict the concentration of effluent in several field experiments. Figure 7.8 illustrates the predicted and actual concentration of chloride; 2,4-D and atrazine for a field plot under conventional tillage at a Cornell University Experimental Farm. It is interesting to note that all graphs experienced similar slopes,

indicating that the soil volumes contributing to this discharge were similar regardless of the solute and that preferential flow may be a more important mechanism in the delivery of pesticides to the groundwater in conventionally tilled soils.

7.7 Influence of Macropores on Biological Degradation of Pesticides

Even though several studies have found significant amounts of pesticides in groundwater much earlier than models predict, other studies have observed that concentrations of pesticide in macropores decrease rapidly due to biodegradation. Also, large populations of micro-organisms have been found below the root zone, suggesting that biodegradation at depth can occur. Since the type of chemical kinetics and the rate of biodegradation depend on the substrate concentration, initial microbial population, availability of electron acceptors and inorganic nutrients, all of which vary spatially in a soil, biodegradation cannot be considered uniform with depth and the role of macropores may be important. Models that assume uniform or even no biodegradation below the root zone in heterogeneous soils are unlikely to predict accurately the transport or fate of contaminants in the unsaturated zone. For more accurate modelling, the interaction of preferential flow and nonuniform biodegradation needs to be addressed.

Recently, Pivetz et al. (1994) and Pivetz and Steenhuis (1994) applied p-nitrophenol (PNP) and 2,4-D (2,4-dichlorophenoxyacetic acid) to the surface of soil columns containing preferential flow paths to study the influence of macropores on biological biodegradation. Both chemicals are rapidly biodegradable, only slightly absorbed, and easily analysed. Separate breakthrough curves were obtained from small areas, sampling matrix flow and macropore (or channel) flow. The velocity, retardation factor, hydrodynamic dispersion, and biodegradation of these chemicals were determined by using the difference in concentration obtained by fitting the theoretical solution of the breakthrough curves for "sterile" soils using CXTFIT (Parker and van Genuchten, 1984) and comparing them to actual breakthrough curves determined experimentally. We shall briefly discuss their results here.

7.7.1 Biodegradation of p-Nitrophenol (PNP)

Three types of columns were used in two different experiments to determine the influence of macropores and preferred channels on biodegradation of p-nitrophenol (PNP). The first set of experiments used one column that contained a single vertical artificial macropore with a diameter of 0.7 mm. The column was 5 cm in diameter and length. The macropore was an open cylindrical channel embedded in a homogeneous silt loam soil with a pore fraction smaller than 0.1 mm. Separate fibreglass wicks (0.6 mm diameter, 50 cm length) were used to sample the effluent near the macropore and in the surrounding soil matrix. The wicks were checked and found not to sorb organic chemicals. The biodegradation of PNP was tested using two different influent rates; a high application rate of 10 mg/L at 36 cm/day for the entire column resulting in apparent Darcy velocities of 2.5 cm/h for the macropore region and 1.3 cm/h for the matrix region. After 10 days the same column was subjected to a lower application rate of 10 mg/L at 16 cm/day, resulting

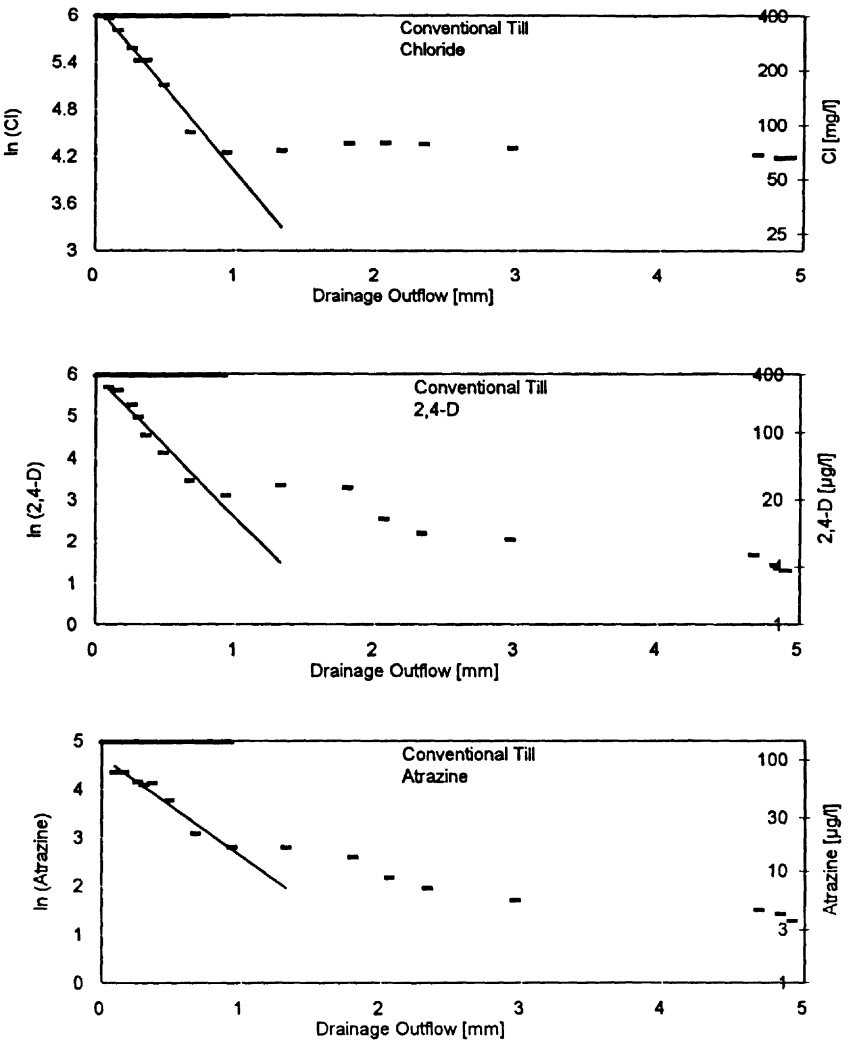


Fig. 7.8. The natural logarithm of predicted and actual solute concentration in the drainage outflow vs. accumulated drainage outflow for a conventionally tilled plot at a Cornell University Experimental Farm. The top plot is for chloride, the middle plot for 2,4-D and the bottom plot for atrazine. The straight lines indicate sections in which a linear regression was performed and the thick black line on the top of each graph indicates the duration of irrigation. (Shalit and Steenhuis, 1994).

in apparent Darcy velocities of 0.5 cm/h for the macropore region and 0.7 cm/h for the matrix region. The lower macropore velocity was attributed to the lack of preferential flow near the macropore at the lower application rate. Also, microbial acclimation for PNP biodegradation was tested at two temperatures, 25 degC and 21 degC. The breakthrough curves for the matrix and macropore region are shown in Figure 7.9 for high and low application rates of PNP. For the high application rate, the breakthrough curve for the macropore region occurred slightly before the

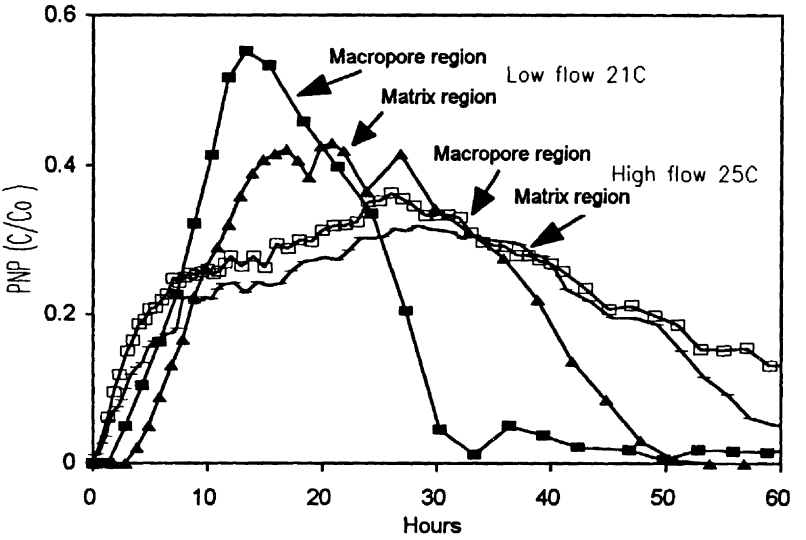


Fig. 7.9. PNP breakthrough curves in macropore and matrix regions.

matrix region and the peak concentration was obtained several hours earlier. For the low application rate, preferential flow did not occur. However, in the vicinity of the macropore, effluent was collected. Again, the peak concentration of PNP was slightly higher than in the matrix region, even though preferential flow was not generated.

Biodegradation of PNP in the matrix and macropore regions is illustrated in Figure 7.10. For the high application rate, the biodegradation rate increased faster in the macropore region than in the matrix region. However, the macropore region required greater time until PNP was no longer present due to the higher flux. The importance of temperature is suggested by the higher peak PNP concentration for the low temperature (21 degC). Presumably, the initial growth rate of microbes was lower at low temperature, allowing an initially higher concentration of PNP to pass through. At the high flow rate and high temperature, the biodegradation rates are greater and increase more rapidly than at low flow rates and temperatures.

A second set of experiments was conducted on paired sets of homogeneous and heterogeneous soil columns (4.1 cm diameter, 15 cm length). The homogeneous columns consisted of < 1.0 mm pore fraction silt loam. The heterogeneous columns contained a vertical channel of higher conductivity soil (1.7–2.0 mm fraction) surrounded by a soil matrix of the lower conductivity soil (0.6 to 1.0 mm fraction). Several different water velocities were used. Once again, the biodegradation rate was more rapid in columns with the channels than in the homogeneous soil columns. As the flow rate increased, the difference in biodegradation rates between the heterogeneous and homogeneous columns was more pronounced.

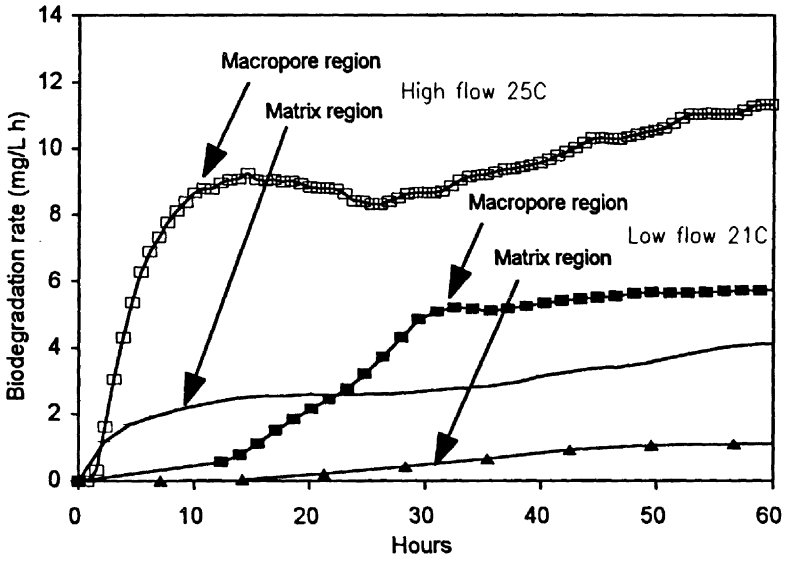


Fig. 7.10. Biodegradation rates in macropore and matrix regions.

7.7.2 Biodegradation of 2,4-D (2,4-dichlorophenoxyacetic acid)

Similar experiments using packed soil columns with an artificial macropore and undisturbed soil columns were conducted to measure the biodegradation of 2,4-D. The packed columns were 7.0 cm in diameter, 10 cm long, packed with silt loam (< 1.0 mm fraction) and had a 0.8 cm diameter cylindrical artificial macropore. The undisturbed columns were 10.0 cm in diameter, about 8 cm long, and consisted of a silty clay with earthworm burrows. The undisturbed columns were collected from a farm near Cornell University. A 50 mg/L solution of 2,4-D was continuously applied at 8.2 cm/d for packed soil columns and 5.5 cm/d for undisturbed soil columns for 480 hours and breakthrough curves indicating the loss of the pesticide in the effluent were obtained from both the macropore and matrix regions of the soil columns. In the packed soil columns with the artificial macropore, the biodegradation rate of 2,4-D in the matrix region was initially higher than that in the macropore region (Figure 7.11). With time, however, the biodegradation rate in the macropore region increased and surpassed that of the matrix region, presumably due to the increased microbial activity. The matrix and macropore regions had a larger difference in the rate of 2,4-D mass biodegraded when flow through the macropore was increased. Complete loss of the 2,4-D in both flow paths was observed after continuous application for 400 hours and biodegradation to nondetectable levels also was observed in both regions of the undisturbed soil columns (not shown here).

During the early phase of each experiment, the biodegradation rate in the macropore region was lower than the matrix region. This is probably due to the low initial population of 2,4-D degraders throughout the soil and to the slow movement of solutes in the matrix which allowed more biodegradation at low microbial

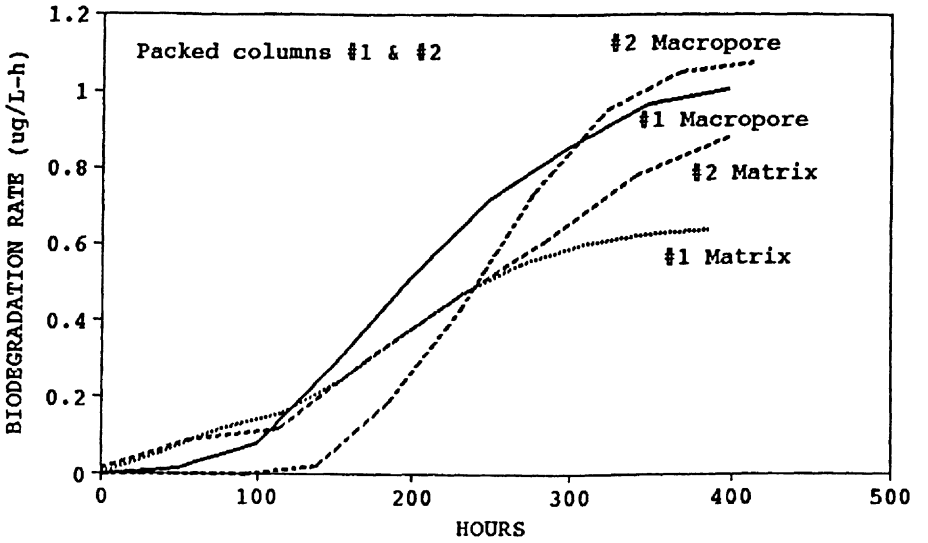


Fig. 7.11. 2,4-D biodegradation curves for packed soil columns with artificial macropores.

populations, whereas the rapid movement of the solute was not conducive to immediate biodegradation. However, as the microbial populations in the macropores increased with time, presumably as a result of increased oxygen availability and greater mass of substrate passing through, the biodegradation rate of 2,4-D in the macropore region exceeded the rates in the matrix regions.

Figure 7.12 shows the observed breakthrough curves for a packed soil column for matrix flow and the predicted breakthrough curves that would be obtained using the CDE for a range of first-order decay coefficients obtained from the literature. Clearly, the predicted breakthrough curves are different from the observed breakthrough curve and neither predict the peak concentration nor eventual loss of all 2,4-D. This indicates that in cases of continuous application of solutes, such as slow, steady leaks from chemical storage tanks, allowing microbial populations to increase significantly, the CDE with first-order decay coefficients cannot be used to predict accurately the peak concentration or the total mass delivered to groundwater.

These studies show that macropores and channels containing highly conductive soil may provide better conditions for biodegradation of pesticides. Microbial growth may occur along the entire length of the column and would result in microbial activity deeper in the soil profile. The coarser material in the channels might allow for a thicker mass of micro-organisms to grow. However, the fast flow may also have adverse effects because cells might not be able to use all of the substrate that rapidly moves past and might also be detached with the flow. Also, the walls of the macropore may only allow the development of a thin film of micro-organisms. Another apparent beneficial effect of macropores is that they may permit a greater diffusion of oxygen in the surrounding soil matrix at greater depth and this increased aeration may contribute to higher biodegradation rates.

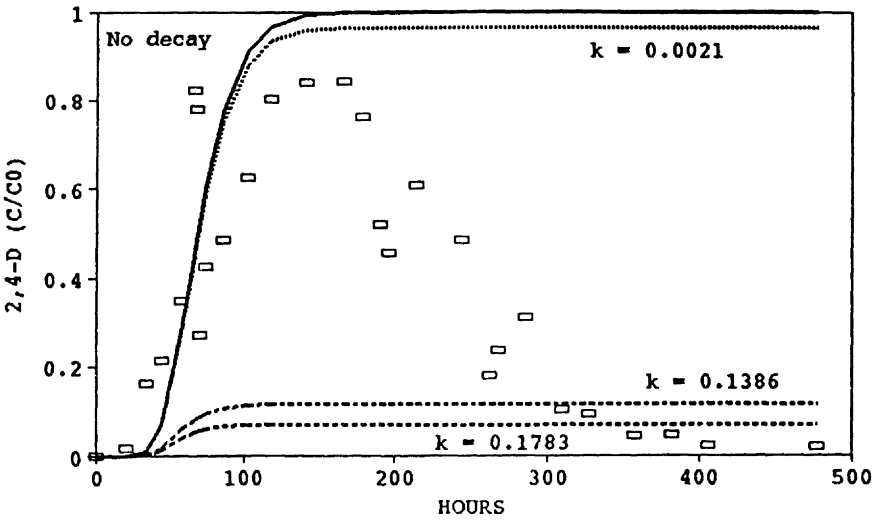


Fig. 7.12. Comparison of packed column #1 matrix BTC to BTCs predicted using the CDE with first-order decay coefficients, k (t^{-1}). Theoretical curves for $k = 0$, 0.0021, 0.1386, and 0.1783 are shown.

7.8 Summary and Conclusions

Local scale heterogeneities and any other factor leading to preferential flow in natural soils are a major concern in modelling solute movement to groundwater. The effect of preferential flow on groundwater quality is much the same whether it occurs in structured soils or homogeneous sandy soils. In fact, preferential flow seems to be most common in soils at both ends of the spectrum. In dense clay soils, cracks provide the only channels through which flow can occur, and in most coarse soils when rainfall intensity is low, wetting front instabilities result in preferential fingered flow. Preferential flow is especially important for chemicals that are toxic in the parts per billion or trillion range. For these chemicals, modelling contaminant transport with an averaged flux or averaged volume approach as advocated by the use of the CDE may result in gross inaccuracies. The presence of preferred paths in the soil presents not only significant difficulty for modelling but also for monitoring and measurement. Greater emphasis on experimental research directed to understanding the role of surface infiltration rates and chemical and physical processes in and between macropores is required to improve current models of unsaturated solute transport. Recent experimental findings show, unfortunately, that there are no simple means of estimating the size and the number of subsurface point measurements required to achieve representative samples.

Rapid biodegradation can occur in both macroporous and matrix flow regions. Preferential flow paths whether they are macropores or coarse grained channels, appear to increase the biodegradation rate of pesticides compared with a homogeneous soil. These paths may provide more favourable conditions for biofilm development and oxygen diffusion for microbial metabolism. In addition, the channels might transport the micro-organisms to greater depth in the soil profile.

Hence, they may aid in biodegradation. However, the initial peak concentration of pesticides and other chemicals in macropores appears to be greater and travel further in the profile than in the case of homogeneous soils. The highest loss rates occur in macropores and increasing the flow rates increases the loss rates. The rapid biodegradation in macropores after the initial quick breakthrough must be considered when modelling the transport and fate of contaminants moving through the soil.

A better understanding of preferential flow will lead to more realistic modelling in order to predict whether groundwater will be affected by agricultural chemicals, landfills, and industrial wastes. This, in turn, will lead to sounder management practices.

References

Anderson, J. L., and Bouma, J. (1977a). Water movement through pedal soils: I. Saturated flow. Soil Science of America Proceedings 37, 408 - 413.

Anderson, J. L., and Bouma, J. (1977b). Water movement through pedal soils: II. Unsaturated flow. Soil Science of America Proceedings 37, 419 - 423.

Anderson, S. H., Peyton, R. L., Wigger, J. W., and Gantzer, C. J. (1992). Influence of aggregate size on solute transport as measured using computed tomography. Geoderma, 53, 387 - 398.

Aylmore, L. A. G., and Kookana, R. S. (1993). Description and prediction of pesticide leaching. AGSO Journal of Australian Geology and Geophysics, 14(2/3), 287 - 296.

Baker, R. S., and Hillel, D. (1990). Laboratory tests of a theory of fingering during infiltration in layered soils. Soil Science Society of America Journal, 54, 20 - 30.

Barcelona, M. J., and Morrison, R. D. (1988). Sample collection, handling and storage: Water, soils and aquifer solids. In D. W. Nelson and R. H. Dowdy (Ed.), Methods for Groundwater Quality Studies. Proceedings of the National Workshop, Agricultural Research Division, University of Nebraska, Lincoln, NE. 49 - 62.

Barry, D. A. (1993). Modelling contaminant transport in the subsurface: Theory and computer programs. In H. Ghadiri and C. W. Rose (Ed.), Chapter 3. Modelling Chemical Transport in Soils: Natural and Applied Contaminants. Lewis Publishers. London. 105 -144.

Barry, D. A., and Li, L. (1994). Physical basis on nonequilibrium solute transport in soil. In 15th International Congress of Soil Science, Acapulco, Mexico. 20 pp.

Barry, D. A., and Sposito, G. (1988). Application of the convective-dispersion model to solute transport in finite soil columns. Soil Science Society of America Journal, 52(1), 3 - 9.

Beasly, R. S. (1976). Contribution of subsurface flow from the upper slopes of forested watersheds to channel flow. Soil Science Society of America Journal, 40, 955 - 957.

Bergström, L. F., and Jarvis, N. J. (1993). Leaching of Dichlorprop, Bentazon, and CI in undisturbed field lysimeters of different agricultural soils. Weed Science, 41, 251 - 261.

Biggar, J. W., and Nielsen, D. R. (1976). Spatial variability of the leaching characteristics of a field soil. Water Resources Research, 12, 78 - 84.

Boll, J., Selker, J. S., Nijssen, B. M., Steenhuis, T. S., Van Winkle, J., and Jolles, E. (1991). Water quality sampling under preferential flow conditions. In R. G. Allen, T. A. Howell, W. O. Pruitt, I. A. Walter, and M. E. Jensens (Eds.), Lysimeters for Evapotranspiration and Environmental Measures. Proceedings American Society of Civil Engineers. International Symposium on Lysimetry. American Society of Civil Engineers. New York City. 290 - 298.

Boll, J., Kung, K.-J. S., Ritter, W. F., Hendrickx, J. M. H., Herbert, S. J., Daliparthy, J., Tomer, M. D., and Steenhuis, T. S. (1993). Detection of textural interfaces using ground-penetrating radar. In H. J. M. Seytoux (Ed.), Proceedings of the 13th American Geophysical Union Hydrology Days, Colorado State University, Fort Collins, CO. 197 - 207.

Bouma, J. (1990). Using morphometric expressions for macropores to improve soil physical analysis of field soil. Geoderma, 46, 3 - 11.

Charbeneau, R. J., and Street, R. L. (1979). Modeling groundwater flow fields containing point singularities, streamlines, travel times, and breakthrough curves. Water Resources Research, 15(6), 1445- 1450.

Charbeneau, R. J. (1981). Groundwater contaminant transport with adsorption and ion exchange chemistry: Method of characteristics for the case without dispersion. Water Resources Research, 17, 705 - 713.

Charbeneau, R. J. (1984). Kinematic models for soil moisture and solute transport. Water Resources Research, 20(6), 699 - 706.

Dekker, L. W., and Ritsema, C. J. (1994). Fingered flow: The creator of sand columns in dune and beach sands. Earth Surface Processes and Landforms, 19, 153 - 164.

Edwards, W. M., Shipitalo, M. J., and Owens, L. B. (1993a). Gas, water and solute transport in soils containing macropores: A review of methodology. Geoderma, 57, 31 - 49.

Edwards, W. M., Shipitalo, M. J., Owens, L. B., and Dick, W. A. (1993b). Factors affecting preferential flow of water and atrazine through earthworm burrows under continuous no-till corn. Journal of Environmental Quality, 22(3), 453 - 457.

Everts, C. J., and Kanwar, R. S. (1988). Quantifying preferential flow to a tile line with tracers. 88-2635. American Society of Agricultural Engineers, St Joseph, MI. 15 pp.

Everts, C. J., and Kanwar, R. S. (1992). Interpreting tension-infiltrometer data for quantifying soil macropores: Some practical considerations. Transactions of American Society of Agricultural Engineers, 36(2), 423 - 428.

Frank, R., and Sirons, G. J. (1979). Atrazine: Its use in corn production and its loss to stream waters in southern Ontario. The Science of the Total Environment, 12, 223 - 239.

Gee, G. W., Kincaid, T., Lenhard, R. J., and Simons, C. S. (1991). Recent studies of flow and transport in the vadose zone. (U. S. National Report of the International Union of Geodesy and Geophysics. 1987 - 1900). Reviews of Geophysics, 29, 227 - 239.

Gerke, H. H., and van Genuchten, M. T. (1993). A dual-porosity model for simulating preferential movement of water and solutes in structured porous media. Water Resources Research, 29, 305 - 319.

Ghodrati, M., and Jury, W. A. (1990). A field study using dyes characterize preferential flow of water. Soil Science Society of America Journal, 54, 1558 - 1563.

Gish, T. J., Helling, C. S., and Mojasevic, M. (1991). Preferential movement of atrazine and cyanazine under field conditions. Transactions of American Society of Agricultural Engineers, 34(4), 1699 - 1705.

Glass, R. J., Oosting, G. H., and Steenhuis, T. S. (1989a). Preferential solute transport in layered homogeneous sands as a consequence of wetting front instability. Journal Hydrology, 110, 87 - 105.

Glass, R. J., Steenhuis T. S., and Parlange, J.-Y. (1989b). Mechanism for finger persistence in homogeneous unsaturated, porous media: Theory and verification. Soil Science, 148, 60 - 70.

Glass, R. J., Steenhuis, T. S., and Parlange, J.-Y. (1989c). Wetting front instability, 2. Experimental determination of relationships between systems parameters and two-dimensional unstable flow field behaviour in initially dry porous media. Water Resources Research, 25, 1195 - 1207.

Glass, R. J., Parlange, J.-Y., and Steenhuis, T. S. (1991). Immiscible displacement in porous media: Stability analysis of three-dimensional, axisymmetric disturbances with

application to gravity driven wetting front instability. Water Resources Research, 27, 1947 - 1956.

Glotfelty, D. E., Taylor, A. W., Esensee, A. R., and Glenn, S. (1984). Atrazine and simazine movement to Wye River estuary. Journal of Environmental Quality, 13, 115 - 121.

Hegert, G. W., Bouldin, D. R., Klausner, S. D., and Zwerman, P. J. (1981). Phosphorus concentration-water flow interaction in the effluent from manured land. Journal of Environmental Quality, 10, 338 - 344.

Hill, D. E., and Parlange, J.-Y. (1972). Wetting front instability in layered soils. Soil Science Society of America Journal, 36, 697 - 702.

Hillel, D., and Baker, R. S. (1988). A descriptive theory of fingering during infiltration into layered soils. Soil Science, 146(1), 51 - 56.

Hoa, N. T. (1981). A new method allowing the measurement of rapid variations on the water content in sandy porous media. Water Resources Research, 17, 41 - 48.

Jury, W. A., Spencer, W. F., and Farmer, W. J. (1984). Behavior assessment model for trace organics in soil: 111. Application of screening model. Journal of Environmental Quality, 13, 573 - 579.

Keller, J. M., Chen, S., and Crownover, R. M. (1989). Texture description and segmentation through fractal geometry. Computer Vision, Graphics and Image Processing, 45, 150 - 166.

Kladivko, E. J., Van Scoyoc, G. E., Monke, E. J., Oates, K. M., and Pask, W. (1991). Pesticide and nutrient movement into subsurface tile drains on a silt loam soil in Indiana. Journal of Environmental Quality, 20, 264 - 270.

Kluitenberg, G. J., and Horton, R. (1990). Effect of solute application method on preferential flow transport of solutes in soil. Geoderma, 46, 283 - 297.

Kookana, R. S., Aylmore, L. A. G., and Gerritse, R. G. (1992a). Time-dependent sorption of pesticides during transport in soils. Soil Science, 154(3), 214 - 225.

Kookana, R. S., Gerritse, R. G., and Aylmore, L. A. G. (1992b). A method for studying nonequilibrium sorption during transport of pesticides in soil. Soil Science, 154(5), 344 - 349.

Kung, K.-J. S. (1990a). Preferential flow in a sandy vadose soil: 1. Field observations. Geoderma, 46, 51 - 58.

Kung, K.-J. S. (1990b). Preferential flow in a sandy vadose soil: 2. Mechanisms and implications. Geoderma 46, 59 - 71.

Li, L., Barry, D. A., Hensley, P. J., and Bajracharya, K. (1994). Mass transfer in soils with local stratification of hydraulic conductivity. Water Resources Research. (In press.)

Liu, Y., Bierck, B. R., Selker, J. S., Steenhuis T. S., and Parlange, J.-Y. (1993). High intensity X-Ray and tensiometer measurements in rapidly changing preferential flow fields. Soil Science Society of America Journal, 57, 1188 - 1192.

Liu, Y., Steenhuis, T. S., and Parlange, J.-Y. (1994a). Closed-form solution for finger width in sandy soils at different water contents. Water Resources Research, 30(4), 949 - 952.

Liu, Y., Steenhuis, T. S., and Parlange, J.-Y. (1994b). Formation and persistence of fingered flow fields in coarse grained soils under different moisture contents. Journal of Hydrology, 159, 187 - 195.

Nijssen, B. M., Steenhuis, T. S., Kluitenberg, G. J., Stagnitti, F., and Parlange, J.-Y. (1991). Moving water and solutes through the soil: Testing of a preferential flow model. In T.J. Gish and A. Shirnohammadi (Eds.), Preferential Flow. American Society of Agricultural Engineers. St. Joseph, MI. 223 - 232.

Parker, J. C., and van Genuchten, M. T. (1984). Determining transport parameters from laboratory and field tracer experiments. Bulletin, 84-3. Virginia Agricultural Experimental Station, Blacksburg.

Parker, J. C., and van Genuchten, M. T. (1984). Flux-averaged and volume-averaged concentration in continuum approaches to solute transport. Water Resources Research, 20(7), 866 - 872.

Parlange, J.-Y., and Hill, D. E. (1976). Theoretical analysis of wetting front instability in soils. Soil Science, 122, 236 - 239.

Parlange, J.-Y., Steenhuis, T. S., Glass, R. J., Pickering, N. B., Waltman, W. J., Bailey, N. O., Andreini, M. S., and Throop, J. A. (1988). The flow of pesticides through preferential paths in soils. New York's Food and Life Science Quarterly, 18, 20 - 23.

Parlange, J.-Y., Steenhuis, T. S., and Stagnitti, F. (1994a). Percolation. In C. Finkl (Ed.), The Encyclopedia of Soil Science and Technology. Van Nostrand Reinhold. (In press.)

Parlange, J.-Y., Steenhuis, T. S., Stagnitti, F., Simmelink, E., and Nijssen, B. (1994b). Recent advances in modelling vadose zone transport. In V. P. Singh and B. Kumar (Eds.), International Conference on Hydrology and Water Resources. Proceedings. New Delhi. National Institute Of Hydrology, India. (In press.)

Peyton, R. L., Haeffner, B. A., Anderson, S. H., and Gantzer, C. J. (1992). Applying X-Ray to measure macropore diameters in undisturbed soil cores. Geoderma, 53, 329 - 340.

Peyton, R. L., Gantzer, C. J., Anderson, S. H., Haeffner, B. A., and Pfeifer, P. (1994). Fractal dimension to describe soil macropore structure using X-ray computed tomography. Water Resources Research, 30(3), 691 - 700.

Philip, J. R. (1975). Stability analysis of infiltration. Soil Science Society of America Proceedings, 39, 1042 - 1049.

Pivetz, B. E., Kelsey, J. W., Steenhuis, T. S., and Alexander, M. (1994). A procedure to calculate biodegradation during preferential flow through heterogeneous soil columns. Soil Science Society of America Journal. (In press.)

Pivetz, B. E., and Steenhuis, T. S. (1994). Soil matrix and macropore biodegradation of 2,4-D. Journal of Environmental Quality. (In press.)

Raats, P. A. C. (1973). Unstable wetting fronts in uniform and nonuniform soils. Soil Science Society of America Proceedings, 37, 681 - 685.

Rice, R. C., Jaynes, D. B., and Bowman, R. S. (1991). Preferential flow of solutes and herbicide under irrigated fields. Transactions of American Society of Agricultural Engineers, 34(3), 914 - 918.

Richard, T. L., and Steenhuis, T. S. (1988). Tile drain sampling of preferential flow on a field scale. Rapid and Far Reaching Hydrologic Processes in the Vadose Zone (Special Issue). Journal of Contaminant Hydrology, 3, 307 - 325.

Richards, L. A. (1931). Capillary conduction of liquids through porous mediums. Physics, 1, 318 - 333.

Rimmer, A., Steenhuis, T. S., and Selker, J. S. (1994). One-dimensional model to evaluate the performance of wick samplers in soils. Soil Science Society of America Journal. (In press.)

Ritsema, C. J., and Dekker, L. W. (1993a). Preferential flow mechanism in a water repellent sandy soil. Water Resources Research, 29(7), 2183 - 2193.

Ritsema, C. J., and Dekker, L. W. (1993b). Soil moisture and dry bulk density patterns in bare dune sands. Journal of Hydrology, 154, 107 - 131.

Selker, J. S., Leclerq, P., Parlange, J.-Y., and Steenhuis, T. S. (1992a). Fingered flow in two dimensions, 1. Measurement of matrix potential. Water Resources Research, 28(9), 2513 - 2521.

Selker, J. S., Parlange, J.-Y., and Steenhuis, T. S. (1992b). Fingered flow in two dimensions, 2. Predicting finger moisture profile. Water Resources Research, 28(9), 2523 - 2528.

Selker, J. S., Steenhuis, T. S., and Parlange, J.-Y. (1994). An engineering approach to fingered vadose pollutant transport. Journal of Environmental Engineering. (In press.)

Shaffer, K. A., Fritton, D. D., and Baker, D. E. (1979). Drainage water sampling in a wet dual pore system. Journal of Environmental Quality, 8, 241 - 246.

Shalit, G., and Steenhuis, T. (1994). A simple mixing layer model predicting solute flow to drainage lines under preferential flow. Journal of Hydrology. (In press.)

Shipitalo, M. J., Edwards, W. A., Dick, W. A., and Owens, L. B. (1990). Initial storm effects on macropore transport of surface-applied chemicals in no-till soil. Soil Science Society of America, 54, 1530 - 1536.

Shuford, J. W., Fritton, D. D., and Baker, D. E. (1977). Nitrate nitrogen and chloride movement through undisturbed field soil. Journal of Environmental Quality, 6, 255 - 259.

Skopp, J., Gardner, W. R., and Tyler, E. J. (1981). Two-region model with small interaction. Soil Science Society of America Journal, 45, 1147 - 1152.

Smettem, K. R. J., and Collis-George, N. (1985). Statistical Characterization of Soil Biopores using a Soil Peel Method. Geoderma, 36, 27 - 36.

Smettem, K. R. J., and Ross, P. J. (1992). Measurement and prediction of water movement in a field soil: The matrix-macropore dichotomy. Hydrological Processes, 6, 1 - 10.

Smettem, K. R. J., Chittleborough, D. J., Richards, B. G., and Leaney, F. W. (1991). The influence of macropores on runoff generation from a hillslope soil with a contrasting textural class. Journal of Hydrology, 122, 235 - 252.

Sposito, G., and Jury, W. A. (1988). The lifetime probability density function for solute movement in the subsurface zone. Journal of Hydrology, 102, 503 - 518.

Stagnitti, F., Steenhuis, T. S., Parlange, J.-Y., Nijssen, B. M., and Parlange, M. B. (1991). Preferential solute and moisture transport in hillslopes. Institution of Engineers, Australia, 22, 919-924.

Starr, J. L., Parlange, J.-Y., and Frink, C. R. (1986). Water and chloride movement through layered field soil. Soil Science Society of America Journal, 50, 1384 - 1390.

Steenhuis, T. S., and Muck, R. E. (1988). Preferred movement of non-adsorbed chemicals on wet, shallow, sloping soils. Journal of Environmental Quality, 17, 376 - 384.

Steenhuis, T. S., Nijssen, B. M., Stagnitti, F., and Parlange, J.-Y. (1991). Preferential solute movement in structured soil: Theory and application. Institute of Engineers, Australia, 22, 925 - 931.

Steenhuis, T. S., and Parlange, J.-Y. (1990). Preferential flow in structured and sandy soils. Engineering Quarterly, (Cornell University), 25(1), 7 - 14.

Steenhuis, T. S., Parlange, J.-Y., and Andreini, M. S. (1990). A numerical model for preferential solute movement in structured soils. Geoderma, 46, 193 - 208.

Steenhuis, T. S., Ritsema, C. J., Dekker, L. W., and Parlange, J.-Y. (1994a). Fast and early appearance of solutes in groundwater by rapid and far-reaching flows. Transactions of International Congress of Soil Science, 2a, 184 - 203.

Steenhuis, T. S., Boll, J., Shalit, G., Selker, J. S., and Merwin, I.A. (1994b). A simple equation for predicting preferential flow solute concentrations. Journal of Environmental Quality, 23(5), 1058 - 1064.

van der Molen, W. H. (1956). Desalinisation of saline soils as a column process. Soil Science, 81, 19 - 27.

van Es, H. M., Steenhuis, T. S., Geohring, L. D., Vermeulen, J., and Boll, J. (1991). Movement of surface-applied and soil-embodied chemicals to drainage lines in a well-structured soil. In T. J. Gish and A. Shirmohammadi (Eds.), Preferential Flow. American Society of Agricultural Engineers. St. Joseph, MI. 59 - 67.

Vermeul, V. R., Istok, J. D., Flint, A. L., and Pikul Jr., J. L. (1993). An improved method for quantifying soil macroporosity. Soil Science Society of America Journal, 57(3), 809 - 816.

Wierenga, P. J., and Brusseau, M. L. (1994). Water and contaminant transport in the vadose zone. (This volume.)

Withjack, E. M. (1988). Computed tomography for rock-property determination and fluid-flow visualization. Society of Petroleum Engineers. Formation Evaluation, 3, 969 - 704.

Yanuka, M., and Elrick, D. E. (1985). Applications of microcomputer-based image digitization in soil and crop sciences. Computers and Electronics in Agriculture, 1, 59 - 73.

CHAPTER 8

Groundwater contamination modelling

A. Ghosh Bobba and Vijay P. Singh

Abstract. In the development of groundwater protection and rehabilitation strategies, mathematical models play an important role. This chapter discusses the role of groundwater contamination models in planning, management, and regulation of groundwater systems, with a focus on generic and site specific contamination. The various approaches for modelling groundwater contamination are reviewed. The applicability of various kinds of models to groundwater contamination is discussed and an overview of available models presented. A case history from Canada is used to demonstrate and illustrate current modelling methodology.

8.1 Introduction

The increasing demand for water to meet agricultural, industrial, and municipal needs is placing greater emphasis on the development of groundwater resources. Yet, the very uses for which the water is utilized, e.g., agricultural, industry and human needs, are adding contaminants to groundwater reservoirs at an increasing rate. It is generally accepted that groundwater contamination is 'irreversible', i.e., once it is contaminated, it is difficult to restore the original water quality of the aquifer over a short span of time.

Groundwater contamination can occur from several sources. These include industrial wastes, solid waste disposal sites, waste water treatment lagoons, agricultural areas, cattle feed lots, artificial recharge sites using waste water, mine spills, septic tank tile fields, etc. In some cases wastes are directly put underground by means of shallow and deep wells and this could result in the contamination of adjoining aquifers. Currently, there are no generally accepted limits for contaminants in groundwater. However, the substances which are of main concern in a drinking water supply can also be considered as contaminants to a groundwater reservoir. Almost all these substances are soluble in water and to make the discussion applicable to any of these substances, the soluble contaminants will be subsequently referred to as contaminants. Contaminants and pollutants are used synonymously

V. P. Singh (ed.), Environmental Hydrology, 225–319.
© 1995 *Kluwer Academic Publishers.*

in this chapter and refer to the soluble substances that impair groundwater quality in some manner.

In studying groundwater contamination, scientists, engineers and others are often confronted with the problem of predicting the concentration of a particular contaminant in the aquifer with respect to time and distance. Such information is needed: (a) to forecast the water quality problems that could arise in the aquifer, (b) to locate wells whether for municipal, industrial or agricultural needs, and (c) to design methods for rehabilitating the aquifer from a water quality point of view.

Except injection by deep wells, all potential sources of groundwater contamination add contaminants to the aquifers by percolation. As such, these contaminants can generally be expected to end up in unconfined aquifers. Once the contaminants are added to the aquifer, their movement is governed by transport and mixing phenomena in porous media and the flow patterns occurring in the aquifer. The process that occurs when one fluid with a particular concentration of a contaminant mixes with and displaces the fluid with a different concentration is referred to as miscible displacement. The mixing and movement of contaminants in groundwater aquifers are an example of miscible displacement.

The transport of contaminants in groundwater aquifers is described by a set of partial differential equations, which constitute the mathematical model of the contaminant transport system. A solution to the set of these equations with appropriate initial and boundary conditions provides the calculated concentration values for the contaminant in the aquifer with respect to time and distance.

In the study of groundwater contamination problems, values of the parameters in the equation for the transport of contaminant in the aquifer are needed to predict the movement and distribution of contaminant in an extended time domain. These values are also needed for studying the behavior of similar aquifers under different boundary conditions. The use of a model to study the behavior of a system as it operates over time is referred to as simulation. When the state of the system (e.g. groundwater system) is defined by a set of differential equations, the mathematical process whereby the parameters embedded in the differential equations are determined from observations of system input and output is referred to as 'parameter identification'.

Any studies of the movement of contaminant in a groundwater system should consider a regional scale so that the studies have practical applications. A regional aquifer is generally one that extends over an appreciable area, with significant thickness and with the potential of yielding or storing significant quantities of water. When groundwater contamination problems are considered on a regional basis, the scale of such problems normally precludes detailed measurement of either concentration of contaminant or flow patterns at sufficient points within the system. The computer then is a useful tool in dealing with such problems.

8.2 Classification of Groundwater Contamination Models

8.2.1 MODELS

Groundwater contamination models used to predict contaminant transport can be classified into three categories: (a) advection models, (b) advection-dispersion models, and (c) advection-dispersion-chemical/biological reaction models.

(a) *Advection Models*
Advection models define the movement of contaminants as a result of groundwater flow only. A slug of water carrying contaminants moves through the soil system along with groundwater flow. Contaminants are transported with no change in concentration with distance.

(b) *Advection-Dispersion Models*
When the concept of dispersion is introduced into the model, a term is included which provides for dispersion related mixing and spreading and leads to time-related changes in contaminant concentration. The dispersion term takes into consideration molecular diffusion, microscopic dispersion, and macroscopic dispersion. Generally, because of the scale of applications in terms of land area involved and relatively high flow velocities, molecular diffusion is of small consequence compared to micro- and macro-dispersion.

(c) *Advection-Dispersion-Chemical/Biological Reaction Models*
Another step in model sophistication is the inclusion of effects of reactions which change the concentration of transported contaminants. The reactions may be chemical or biological and can be incorporated into advection-only models or advection-dispersion models. Because of the current lack of knowledge regarding subsurface reaction kinetics only chemical processes such as ion exchange and adsorption have been considered in most applications.

An additional class involves coupling of geochemical models with groundwater flow models. Such models are complex and were developed for studying the chemistry of natural waters and are not designed for application to contaminant transport problems. Their applications have been limited to simulating the evolution of groundwater quality along regional groundwater flow paths in systems dominated by calcium-magnesium-sulphate reactions. Application of this modelling approach to meet industry needs appears to be of limited value.

8.2.2 MODEL DEVELOPMENT

The recent concern over the contamination of groundwater systems and resulting environmental hazards has pushed groundwater modelling to the forefront in many cases as the method to provide all the needed answers. There is an apparent trend towards heavy reliance on modelling in a variety of regulatory programs. Modelling is a best judgment method because it is a powerful tool and can be used for a variety of purposes, including (i) prediction of contaminant transport, (ii) selection of new waste disposal site facilities, (iii) development of groundwater monitoring systems for new and existing waste disposal site facilities, and (iv) development of remedial action plans. It must be kept in mind that each model is subject to inherent limitations and their predictions contain uncertainty. Therefore

a proper use of groundwater contamination models should be as aids in making regulatory, management, planing, and policy decisions.

Groundwater contamination models are mathematical approximations of the complex natural phenomena which attend a particular situation. In making the approximation, certain assumptions and judgements must be made which makes model development possible. A clear understanding of these assumptions is important in any application. In general, the current generation of groundwater quality models is not completely reliable. Most are not suited without considerable judgment for application in regulations, planning and management.

8.3 Groundwater Modelling

Groundwater modelling is concerned with the behavior of subsurface systems. Essentially all models are simplified representations of these subsurface systems. Modelling, therefore, may be considered as an exercise in systems analysis whereby theories concerning the behavior of groundwater systems are organized into models which are used for their predictive capabilities.

A groundwater system is composed of interacting parts. While recognizing the parts or subsystems and their functions, the ultimate concern of modelling is with the operation of the groundwater system as a whole in relation to its surrounding environment. Models integrate fragmented knowledge of the system's component parts and develop a comprehensive conception of the entire system.

Some degree of simplicity or abstraction in modelling is required in attempting to represent or simulate groundwater systems. Approximations are factored into the analysis via the assumptions incorporated into the model after considering 1) the model purpose, 2) the status of available model theory, and 3) the data base to be used in developing and testing the model. Although a model by design may be less complex than the real system it represents, oversimplifying a system is not always justified.

Complete data is generally lacking for specific groundwater systems, and the gap between data needs and data availability increases with the complexity of the groundwater system. The effective application of models to field problems requires the ability to fill in data gaps with estimated, interpolated, or extrapolated values. Considerable scientific judgement of a subjective or intuitive nature is often necessary for any degree of success in modelling. Attempts at modelling without a measure of experienced judgement can often be counterproductive.

Adequate acknowledgment and documentation of data base limitations is an important aspect of modelling. A model should be in tune with the data base. Sophistication beyond data availability is generally not warranted and may be misleading. The reliability of model results cannot exceed the reliability and accuracy of the data base.

Models represent continuous and time dependent processes. The major mechanisms considered in models for our needs are fluid flow and contaminant transport. Groundwater flow is modelled without consideration of contaminant transport phenomena. Modelling of contaminant transport requires the meshing of contaminant

transport algorithms with either simultaneous solution with, or results from, a flow model.

Flow models simulate some aspects of groundwater flow such as direction and rate of flow, changes in water levels, stream-aquifer interactions, and interference effects of production wells. Most flow models are of the distributed type in that they have spacial components.

Contaminant transport models simulate movement and concentration in groundwater systems of various contaminants, in particular pollutants such as leached contaminants from landfills. These models generally contain a flow submodel which provides flow directions and velocities. A quality submodel utilizes these velocities to simulate advective transport, allowing for dispersion and reactions. Mass transport models include both conservative and non-conservative transport by containing such factors as chemical adsorption and ion exchange.

8.3.1 Principles and Concepts used in Groundwater Modelling

Contaminant models are developed to provide a simplified and easily understandable version of reality. The analytical as well as numerical groundwater models are built upon a number of principles and concepts which describe or are a best estimate of suspected physical, chemical, and biological events. The ability to access these events by mathematical relationships provides the basis for both analytical and numerical model development.

The processes that control the transport of contaminants are flow, hydrodynamic dispersion, and geochemical and biochemical reactions. Advection involves transport down gradient from the contaminant source by flowing groundwater with contaminants normally spreading as a result of dispersion into and occupying an increasing volume of the groundwater system. In the case of a conservative contaminant, reactions that alter the contaminant concentration do not occur between the contaminant and the soil matrix in the aquifer system. As a result the total mass of the contaminant in the flow regime does not change, but the mass occupies an increasing volume of the aquifer system. The transport of non-conservative substances is more complex. In addition to the described effect of advection and dispersion, the total mass in transport is reduced by chemical and biological activity.

Principles and concepts of note which are the foundation of groundwater contamination modelling include (a) Darcy's law, (b) Hubbert's force potential, (c) conservation of mass, (d) hydrodynamic dispersion, and (e) chemical and biological activity.

8.3.2 Darcy's Law

In 1856, a French engineer named Henry Darcy published results on the flow of water through sand filter beds. This experimental work resulted in the establishment of the basic law of groundwater movement, termed Darcy's Law. Darcy concluded that the rate of flow of water through sand beds is directly proportional to the head loss over the bed, and inversely proportional to the thickness of the bed. Mathematically, his findings can be stated in the following way:

TABLE 8.1
Hydraulic Conductivities of Unconsolidated Sediments

Material	Hydraulic Conductivity (cm/s)
Clay	10^{-9}–10^{-6}
Silt, sandy silts, clayey sands, till	10^{-6}–10^{-4}
Silty sands, fine sands	10^{-5}–10^{-3}
Well-sorted sands, glacial outwash	10^{-3}–10^{-1}
Well-sorted gravel	10^{-2}–1

$$V_D = Q/A = -K(\Delta H/\Delta x), \tag{8.1}$$

where V_D is the Darcy velocity (L/T), Q is the flow rate (L^3/T), A is the cross-sectional area perpendicular to the flow direction (L^2), K is the hydraulic conductivity (L/T) and H is the hydraulic head loss (L) over the distance x (L). The negative sign signifies that flow is in the direction of decreasing head.

The constant, K, hydraulic conductivity, is also referred to as the coefficient of permeability. It is a function of both liquid and soil characteristics. The soil characteristics which influence hydraulic conductivity values include soil matrix geometry, soil porosity, pore size distribution, and tortuosity. Liquid characteristics of importance include density and viscosity. For the general application of groundwater modelling, the fluid properties of density and viscosity are generally assumed constant.

Involvement of immiscible fluids is an additional complicating factor in a modelling exercise. Because of the complexities associated with immiscible fluids modelling, this topic is not included in this chapter. For modelling purposes, the common approach is to obtain a best estimate of hydraulic conductivity from laboratory and/or field tests. This best estimate may undergo reasoned adjustment during various stages of the predictive process.

There are a number of methods for obtaining hydraulic conductivity values. These methods are grouped into laboratory and field tests. Laboratory tests are conducted using soil samples from the study site and one of two types of apparatus. The two types of apparatus are the constant head permeameter and the falling head permeameter. Apparatus selection for a particular test is based on the general characteristics of the soil. The constant head system is best suited to samples with hydraulic conductivities greater than 0.01 cm/min while the falling-head system is best suited to samples with lower conductivity.

There are several test methods available for field determination of hydraulic conductivity. Slug tests and bail tests are initiated by causing an instantaneous change in well water level. Hydraulic conductivity values are obtained from observations of the recovery of the water level with time. Pumping tests are also used and can be important because they provide in-situ values that are averaged over a large aquifer volume.

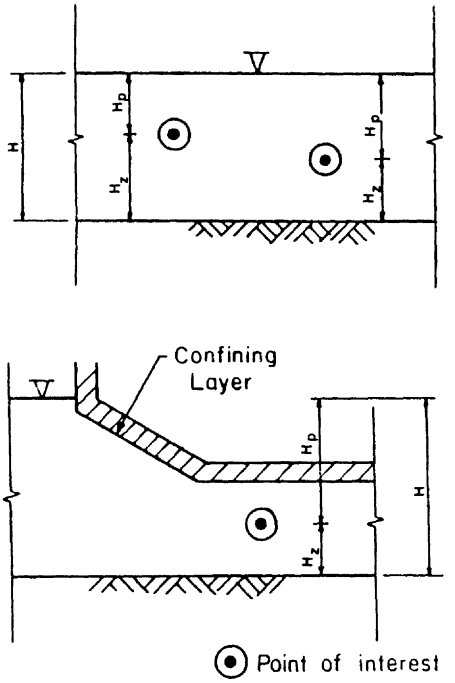


Fig. 8.1. The relationship between hydraulic head and elevation: (a) Confined aquifer, and (b) Unconfined aquifer.

Hydraulic conductivities vary with soil type. Table 8.1 is a typical listing of hydraulic conductivity which provides a general range of expected values for the soil shown. Whatever the hydraulic conductivity measured at some point in time, this value can change due to physical, chemical or biological processes related to the contaminants present in the groundwater.

The potential for validation of hydraulic conductivity with time as a result of reactive processes makes the use of modelling a very valuable tool for estimating changes in contaminant transport. Because models can be adjusted easily to reflect possible hydraulic conductivity changes, arrangement of information may be quickly and inexpensively made available for management purposes. Anticipating and accounting for these changes makes modelling more meaningful.

8.3.3 Hubbert's Force Potential

Darcy's law is an expression of the fact that groundwater moves in the direction of decreasing energy from higher to lower hydraulic head. The hydraulic head represented (Hubbert, 1940) in Equation (8.2) as H, incorporates two terms, pressure head and elevation head. The pressure head is that pressure potential (H_p) of the groundwater, generally expressed in terms of feet or meters of water, that results from a column of water which sits above the point in a groundwater system under consideration or indirectly impacts this point because of the action of a confining

layer. Elevation head (H_z) is simply the elevation of the point under consideration in the groundwater system, whether confined or unconfined, above the zero datum. A simple equation expressing this concept is:

$$H = H_p + H_z, \tag{8.2}$$

where H is Darcy's H (L). The two diagrams of Figure 8.1 schematically represent the concept. In Figure 8.1(a) the hydraulic head-pressure head plus elevation head relationship of any chosen location within the aquifer is clear. However, Figure 8.1(b) contains an intervening confining layer that displaces the water column which generates the pressure head from the point of interest. A well drilled at the point of interest would be artesian. The level of groundwater in this would rise to an elevation equal to the pressure head plus elevation head. Hubbert is given credit for clarifying the concept of groundwater potential and its relationship to Darcy's head by deriving it from basic physical principles.

8.3.4 Conservation of Mass

The law of conservation of mass is the basis of the development of mathematical relationships generally referred to as governing equations (Freeze and Cherry, 1979). From the concept that mass is always conserved, continuity equations are developed for both advection — only models (conservative transport) and contaminant transport models (non-conservative transport). This law is applicable whether the desired model is one, two or three dimensional, or whether describing steady state or transient conditions. In these considerations, water is usually considered as an incompressible liquid.

8.4 Flow Models

For regional groundwater flow problems, two dimensional horizontal flow is considered. The governing equations are well established (Bear, 1972).

8.4.1 Basic Assumptions

The following assumptions are valid for regional groundwater flow: (1) The flow is essentially horizontal in a two dimensional plane. This assumption is valid when the variation of thickness of the aquifer is much smaller than the thickness itself. This approximation fails in regions where the flow has a vertical component. (2) The fluid is homogeneous and slightly compressible. (3) The aquifer is elastic and generally non-homogeneous and anisotropic. The consolidating medium deforms during flow due to changes in effective stress with only vertical compressibility being considered. (4) For the two dimensional horizontal flow assumption, an average piezometric head is used where the average is taken along a vertical line extending from the bottom to the top of the aquifer,

$$H_{av}(x_1, x_2, t) = \frac{1}{B} \int_{x_3=0}^{B} H(x_1, x_2, x_3, t)\, dx_3, \tag{8.3}$$

where B is the thickness of the aquifer.

8.4.2 Two Dimensional Horizontal Flow

The combined equation of motion and continuity for flow in a two dimensional horizontal plane may be written as

$$\frac{\partial}{\partial x_i}\left[T_{ij}\frac{\partial H}{\partial x_j}\right] - P + I = S\frac{\partial H}{\partial t}, \qquad i,j = 1,2, \tag{8.4}$$

where T_{ij} is the transmissivity tensor equal to the aquifer thickness multiplied by the hydraulic conductivity K_{ij}, S is the storage coefficient, t is the time, I is the vertical recharge or infiltration into the aquifer, and P is strength of a sink or source function defined by

$$P = \sum_{m=1}^{M} P_w[(x_1)_m(x_2)_m]\,\delta[x_1 - (x_1)_m][x_2 - (x_2)_m], \tag{8.5}$$

where P_w is the discharge or recharge from the aquifer, M is the number of nodes in one element, and δ is the Dirac delta function.

8.4.3 Definition of Boundary and Initial Conditions

In order to obtain a unique solution of a partial differential equation corresponding to a given physical process, additional information about the physical state of the processes is required. This information is described by boundary and initial conditions. For steady state problems only boundary conditions are required, whereas for unsteady state problems both boundary and initial conditions are necessary. The initial conditions are simply the values of the dependent variable specified everywhere inside the boundary. Computation time can be lessened generally by choosing initial conditions which are approximately equal to the final conditions. Mathematically, the boundary conditions include the geometry of the boundary and the values of the dependent variable or its derivative normal to the boundary. In physical terms for groundwater applications, the boundary conditions are generally of three types: (1) specified value, (2) specified flux, or (3) value dependent flux, where the value is head or concentration depending on the equation. The boundary condition details are given in Table 8.2a.

8.4.3.1 Boundary conditions

In order to solve a partial differential equation describing a physical phenomenon, it is necessary to choose certain additional conditions imposed by the physical situation at the boundaries (S) for the domain (D) under consideration. In general the equation for the boundary condition can be written as

$$\beta_1 T_{ij}\frac{\partial H}{\partial x_j}l_i + \beta_2 H + \beta_3 = 0, \quad i,j = 1,2, \tag{8.6}$$

TABLE 8.2a
Boundary Condition

Type	Description
Specified value	Values of head or concentration specified along the boundary. (Dirichlet boundary condition)
Specified flux	Flow rate of water or concentration is specified along the boundary and equated to the normal derivative. For example, the volumetric flow rate per unit area for water in an isotropic media is given by $q_n = -K(\frac{\partial H}{\partial n})$, where the subscript n refers to the direction normal (perpendicular) to the boundary. A no-flow (impermeable) boundary is a special case of this type in which $q_n = 0$. (When the derivative is specified on the boundary, it is called a Neumann condition.)
Value-dependent flux	The flow rate is related to both the normal derivative and the value. For example, the volumetric flow rate per unit area of water is related to the normal derivative to head and head itself by $-K\frac{\partial H}{\partial n} = q_n(H_b)$. where q_n is some function that describes the boundy f rate given the head at the boundary (H_b)

TABLE 8.2b
Mathematical Equations of Boundary Conditions

The general equation for the boundary condition can be written as
$\beta_1 T_{ij}\frac{\partial H}{\partial x_j}l_i + \beta_2 H + \beta_3 = 0,$
where l_i are the directional cosines, β_1, β_2 and β_3 are known functions of space and time.

Type	Name	β_1	β_2	Equation	Remark
1	Dirchlet	0	#0	$H=-\frac{\beta_3}{\beta_2}=f(x_1,t)$	
2	Neumann	#0	0	$T_{ij}\frac{\partial H}{\partial x_j}l_i=-\frac{\beta_3}{\beta_1}$	If β=0, the flow on the boundary is zero
3	Cauchy	#0	#0	$T_{ij}\frac{\partial H}{\partial x_j}l_i+\frac{\beta_2}{\beta_1}H=\frac{\beta_3}{\beta_1}$	

where l_i are the directional cosines, and β_1, β_2 and β_3 are given functions of position and possibly time. For flow through an aquifer, three different boundary conditions are applicable.

8.4.3.2 Dirichlet or prescribed boundary condition

In this case the potential is specified for all points along the boundary:

$$H = -\frac{\beta_3}{\beta_2}, \qquad \beta_1 = 0, \qquad \beta_2 \neq 0. \tag{8.7}$$

8.4.3.3 Neumann or prescribed flux boundary condition

Along a boundary of this type, the flux normal to the boundary surface is prescribed for all points of the boundary as a function of position and time:

$$T_{ij}\frac{\partial H}{\partial x_j} l_j = -\frac{\beta_3}{\beta_1} \text{ on S}, \qquad \beta_1 \neq 0. \tag{8.8}$$

A special case of the Neumann condition is the impervious boundary where the flux vanishes everywhere on the boundary, i.e.,

$$\beta_3 = 0. \tag{8.9}$$

8.4.3.4 Cauchy Boundary

This problem occurs when the potential and its normal derivative are prescribed on the boundary in the combined form, and the entire Equation (8.7) is used. Different forms of Equation (8.8) for three types of boundary conditions are summarized in Table 8.2b. In general, for a flow problem one will have mixed boundary conditions in which the Dirichlet condition will apply over a part of the boundary and the Neumann condition will be specified for the remaining portion.

8.4.4 Initial conditions

At the initial time, either the piezometric heads are known in the entire domain (D) or the hydrologic stresses (such as pumping and recharge) are specified and boundary conditions are known. For the second case the system has reached the steady state, so the solution of the equation

$$\frac{\partial}{\partial x_i}\left[T_{ij}\frac{\partial H}{\partial x_j}\right] - P + I = 0, \quad i, j = 1, 2 \tag{8.10}$$

will yield piezometric heads for the initial time.

The governing equations account for all groundwater flow into and out of the soil region under consideration and can include terms for additions due to a variety of recharge options, withdrawals, or water which is held or released from storage.

Darcy's law is incorporated into the equation to provide a relationship between measured or estimated physical parameters and flow patterns.

Groundwater flow equation in two dimensions (vertical cross-section):

$$\frac{\partial}{\partial x}\left(K_{xx}\frac{\partial H}{\partial x}\right)+\frac{\partial}{\partial z}\left(K_{zz}\frac{\partial H}{\partial z}\right)=\frac{\partial H}{\partial t}. \tag{8.11}$$

Darcy velocity vectors:
Along x-direction:

$$v_x=-\frac{K_{xx}}{\theta}\frac{\partial H}{\partial x}. \tag{8.12}$$

Along z-direction:

$$v_z=-\frac{K_{zz}}{\theta}\frac{\partial H}{\partial z}. \tag{8.13}$$

8.5 Contaminant Transport Models

Contaminant transport models are a step beyond flow models. These models include all the considerations incorporated in flow models plus relationships which are designed to track the contaminants of interest and determine the change in their concentration with time. The continuity equations for contaminant transport simulations include not only terms for dispersion and flow, but also other processes such as chemical and biological reactions which quantify the expected changes in contaminant concentrations with time as this material travels through the soil system of interest. The controlling concept is that total mass is always accounted for.

In this section different mathematical models (Table 8.3) will be formulated to represent transport of contaminants generated from a source site into a groundwater flow system. Three major contaminant transport mechanisms are included separately or simultaneously in each model.

8.5.1 Advective Dispersion Phenomena

In this chapter the transport of a contaminant in a saturated flow through a porous medium is considered. This contaminant will be referred to as a "contaminant". The symbol C will be used to denote the concentration of a contaminant, i.e., mass of contaminant per unit volume of the solution. The term contaminant will be used to represent any species of interest in a solution.

8.5.2 Basic Assumptions

(a) It is assumed that the porous medium is homogeneous and isotropic with respect to dispersivity. (b) The flow regime is laminar. (c) In general, variations in contaminant concentration cause changes in the density and viscosity of the liquid.

TABLE 8.3
Contaminant Transport Models, Functions, Input Data Requirements and Outputs

Models	Function	Input	Output
Flow Models	Simulation of subsurface flow movement		
Saturated and unsaturated flow model	Simulation of saturated and unsaturated flow model	a) Hydraulic characteristics b) Boundary conditions; c) Initial conditions	Time-dependent groundwater: (a) potentials, (b) velocities; (c) flow rates; (d) flow paths; and (e) travel times.
Transient groundwater flow model.	Simulation of saturated flow	a) Transmissivity distribution; b) Initial conditions; c) Boundary conditions: includes quantities and location of recharge locations.	Time-dependent groundwater: (a) potentials; (b) velocities; (d) flow paths; (e) flow rates; and (f) travel times.
Contaminant Transport Models	Simulation of subsurface contaminant transport.		
Macro-ion transport model	Simulates macro-ion movement in saturated and unsaturated porous media.	a) Cation exchange capacities; b) Initial concentrations on the soils; c) Groundwater velocities from hydraulic models; d) Boundary conditions.	Macro ion concentrations in liquid and solid phases as a function of time and space.
Micro-ion transport model	Simulation of micro-ion movement in saturated and unsaturated soils.	a) Sorption coefficient distribution b) Initial concentrations on the soils c) Macro-ion concentrations as a function of time and space from the macro-ion model d) Groundwater velocities from flow models. e) Boundary conditions	Micro ion concentrations in liquid and solid phases as a function of time and space.

These, in turn, affect the flow regime (i.e., velocity distribution). At relatively low concentrations it is assumed that the concentration does not affect the liquid properties. This assumption leads to the following conclusions: (1) The viscosity is constant, and (2) the concentration does not affect the velocity distribution.

8.5.3 Advective Dispersion Equation in Cartesian Coordinates

The equation describing the mass transport and dispersion of dissolved chemical constituents in a saturated porous medium may be written as

$$\left[\frac{\partial C}{\partial t} + \frac{\partial}{\partial x_i}(V_i C)\right] - \left[\frac{\partial}{\partial x_i}\left(D'_{ij}\frac{\partial C}{\partial x_j}\right) + q'_c\right] = 0, \quad i,j = 1,2,3, \tag{8.14}$$

where C is the mass concentration of the contaminant, D'_{ij} is the coefficient of hydrodynamic dispersion, V_i is the component of seepage velocity, q'_c is the mass flux of source or sink, and x_i are the Cartesian coordinates. The theoretical basis and the derivation of the diffusion-convection equation are discussed in detail by Bear (1979). In Equation (8.14) the first term represents the time rate of change of the contaminant concentration. The second term describes the advective transport of C in the x_i-direction, which is proportional to the seepage velocity. The third term is the transport (redistribution) of C due to dispersion and the molecular diffusion. Finally, the last term represents the time rate of production or decay of C.

The advective dispersion equation is a nonlinear partial differential equation of parabolic type. The relation is nonlinear because of the advective term, and because of the transport coefficient which is a function of the dependent variable V. The advective term is nonsymmetric and has been a principal source of difficulty in the numerical solution of the advective dispersion equation.

8.5.4 Boundary and Initial Conditions

8.5.4.1 Boundary conditions

The general equation of the boundary conditions for the mass transfer equation is similar to the flow equation. As discussed earlier, it can be written as:

$$\alpha_1 D'_{ij}\frac{\partial C}{\partial x_j} l_i + \alpha_2 C + \alpha_3 = 0, \tag{8.15}$$

where α_1, α_2 and α_3 are known functions. Three different boundary conditions are:

(1) Dirichlet or prescribed concentration boundary condition:

$$C = -\frac{\alpha_3}{\alpha_2}, \qquad \alpha_1 = 0, \qquad \alpha_2 \neq 0. \tag{8.16}$$

(2) Neumann or prescribed flux boundary condition:

$$D'_{ij}\frac{\partial C}{\partial x_j}l_i = -\frac{\alpha_3}{\alpha_1}, \qquad \alpha_1 \neq 0, \qquad \alpha_2 = 0, \tag{8.17}$$

where $\alpha_3 = 0$, one has the no-flow boundary.
(3) Cauchy boundary condition:

$$D'_{ij}\frac{\partial C}{\partial x_j}l_i + \frac{\alpha_2}{\alpha_1}C = -\frac{\alpha_3}{\alpha_1}, \qquad \alpha_1, \alpha_2 \neq 0. \tag{8.18}$$

Again, as in the flow situation, usually along the boundary one has mixed boundary conditions, i.e., the Dirichlet condition applies over a part of the boundary and the Neumann condition over the remaining part.

8.5.4.2 Initial conditions

As an initial condition, the concentration distribution at some initial time t=0 at all points of the flow domain must be specified:

$$C(x_i, 0) = f_1(x_i), \tag{8.19}$$

where f_1 is a known function of x_i.

8.6 Hydrodynamic Dispersion

Flow models are used as conservative estimators of contaminant transport. This is a practical approach because flow models are generally well developed and satisfy the particular modelling requirements while additionally providing an extra measure of safety by assuming no reduction in contaminant mass. An important aspect of this simulation approach is dispersion. This physical process is a natural phenomenon which causes reductions in the concentration of a transported contaminant.

Dispersion is one of many processes which can reduce the concentration of contaminants transported by groundwater. It is a physical phenomenon of major importance which affects contaminant concentration as these materials travel in groundwater systems. This process will not only tend to mix contaminated flows with uncontaminated groundwater leading to reductions in concentration by dilution, but will also result in contaminants spreading longitudinally and transversely forming a typical plume. This process also results in the contaminants arriving at a distant location earlier than predicted by flow models which do not account for dispersion. It is important to note that the concentration of the contaminant that 'arrives early' will be less than the concentration reported by the flow models.

The general term dispersion refers to both the process of mechanical mixing during fluid advection and molecular diffusion due to the thermal-kinetic energy of the contaminant material. Diffusion, which is driven by concentration differences, is a dispersion process of importance only at low velocities. Dispersion due to mechanical mixing during fluid advection is referred to as hydraulic dispersion.

Hydraulic dispersion is generally separated into microscopic and macroscopic levels as shown in Figure 8.2. Hydraulic dispersion is the spreading of distribution

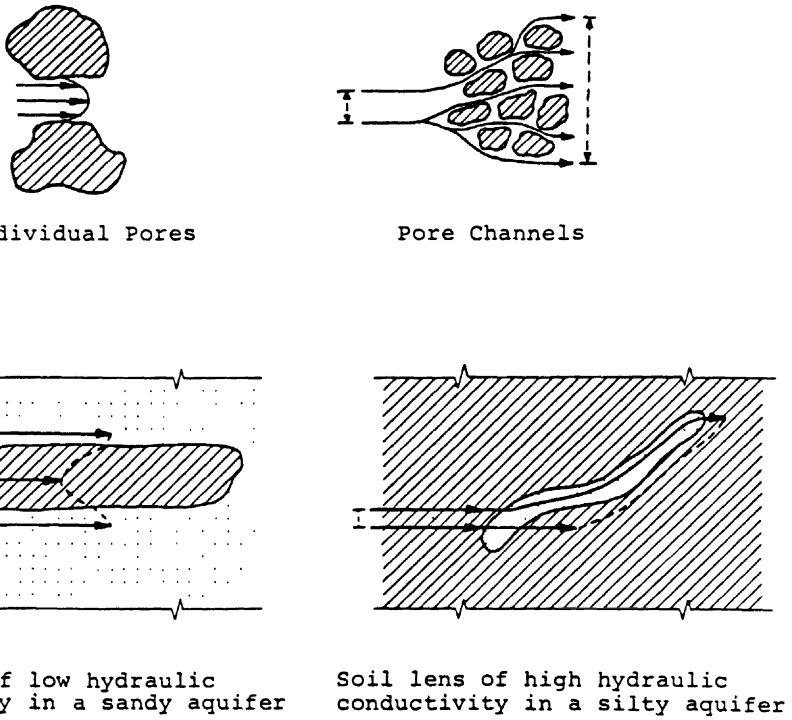


Fig. 8.2. Types of dispersion: (a) Microscopic dispersion, and (b) Macroscopic dispersion.

of contaminant material in groundwater systems and results from inherent heterogeneity of soil matrix geometry in all natural soil systems. Microscopic dispersion results from the numbers of pore pathways available to a slug of groundwater as well as the hydraulics of flow around individual soil particles. Macroscopic dispersion applies to the impact of small soil bodies or lenses present in larger soil bodies of significantly different hydraulic conductivity. This condition can lead to wide ranges in time of travel for contaminants depending on what route the groundwater carrier takes. Hydrodynamic dispersion is the macroscopic outcome of the actual movements of individual tracer particles through the pores and includes two processes (Bear, 1972; Reddel and Sunada, 1967). One mechanism is mechanical dispersion, which depends on both the flow of the fluid and the characteristics of the porous medium through which the flow takes place. The process of water movement through saturated porous media involves both transport and adsorption of fluid. Advective and dispersive transport are the mass movement mechanisms associated with hydrodynamic dispersion. It is generally assumed that the amount of material transferred parallel to any given direction is the sum of the advective and dispersive mass transport components. Advective or mechanical mass transport is attributed to the variation of local microscopic velocity in the porous medium matrix. The dispersive transport phenomenon, or the so-called physicochemical dispersion or molecular diffusion, is caused by the existence of a concentration

gradient in the fluid or liquid phase. It should be emphasized at this point that hydrodynamic dispersion includes the effects of both mechanical dispersion and molecular diffusion. Any separation of these two processes is artificial.The second process is molecular diffusion which basically results from variations in tracer concentration within the liquid phase, and is more significant at low velocities. Thus, the coefficient of hydrodynamic dispersion D'_{ij} includes the effect of both the mechanical dispersion D_{ij} and molecular diffusion $(D_m)_{ij}$. Hence,

$$D'_{ij} = D_{ij} + (D_m)_{ij}, \tag{8.20}$$

In Equation (8.20), $(D_m)_{ij} = T_{ij}D_m$, where D_m is the molecular diffusivity, and T_{ij} is the medium's tortuosity. For homogeneous and isotropic media the value of T_{ij} is approximately equal to 2/3 (Bear, 1972). For most situations the contribution of molecular diffusion to hydrodynamic dispersion is negligible when compared to the mechanical dispersion. For a gravel with seepage velocity ranging from 0.1 to 0.45 cm/sec, the magnitude of the dispersion coefficient varies from 0.01 to 0.08 cm^2/sec (Rumer, 1962). The molecular diffusivity for contaminants in water is very small and is in the range of 0.5 to 4.0 $\times$ 10^{-5}.

Many investigators have attempted to model the dependence of the hydrodynamic dispersion coefficient on media, fluid properties, and flow characteristics, in order to understand the dispersion process in porous media. A comprehensive discussion of the factors affecting the dispersion coefficient can be found in Bear (1972, 1979).The mechanical dispersion coefficient for an isotropic medium in cartesian coordinates can be written as:

$$D_{ij} = a_{II}d_{ij}V + (a_I + a_{II})V_iV_j/V, \quad i,j = 1,2,3. \tag{8.21}$$

In Equation (8.21), a_I and a_{II} are the longitudinal and transversal dispersivities of the medium, respectively, V_i and V_j are components of the seepage velocity in the i and j directions, V is the magnitude of the velocity, and d_{ij} is the Kronecker delta; its value is one when $i = j$ and is zero, otherwise, Equation (8.21) is commonly used to calculate the mechanical dispersion coefficient and hence is utilized in the chapter. It includes the major parameters causing the mechanical dispersion, and for practical purposes it is assumed adequate.

8.6.1 Effects of Dispersion

To illustrate the effects of dispersion, consider the contaminant movement between an injection well and a pumping well in a confined homogeneous aquifer, where the continuous injected concentration is $C = 1$. It is assumed that the initial aquifer concentration is $C = 0$ and that steady-state flow exists between the wells. The situation can be approximated by a simple one dimensional model as shown by Figure 8.3. The concentration front without dispersion would appear as a sharp front plug flow that moves out at the average fluid velocity from the injection well. With dispersion the front is no longer sharply marked at a particular point in time but rather elongated and shows a gradual concentration variation with distance from the contaminant source.

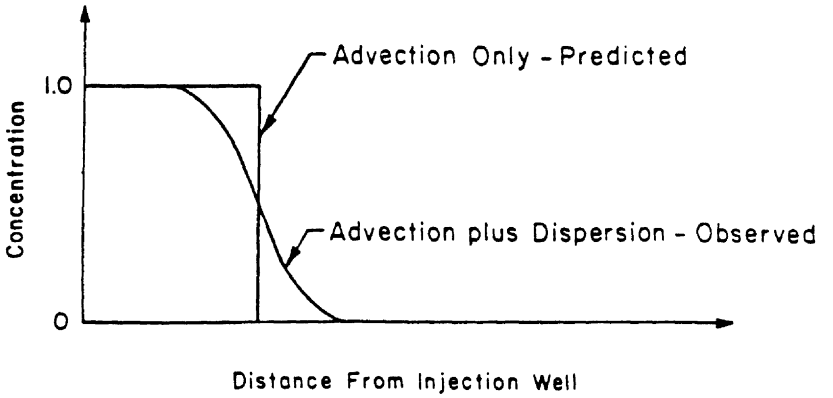


Fig. 8.3. Effects of dispersion on concentration of injected water.

8.6.2 Quantification of Dispersion

Dispersivity is defined to be a characteristic mixing length which is a measure of the mechanical dispersion of a contaminant. A common test for obtaining dispersivity values consists of injecting a tracer into the porous medium and measuring the rate of dispersion by monitoring concentrations. Using these test data, dispersivity is determined by calculation. It is important that because of the effect of soil heterogeneities on the magnitude of dispersivity, these heterogeneities are identified and their impact incorporated to the greatest extent possible into any method for dispersivity determination.

The magnitude of measured dispersivities changes with the scale at which measurements are taken. Laboratory measurements yield values in the range of 10^{-2} cm to 1 cm, while dispersivities of 10 to 100 m have been obtained for field experiments. It has been recommended that field measurements be made at one to four levels of scale depending on specific needs. The four suggested levels based on mean travel distance are: 2 m to 4 m, 4 m to 20 m, 20 m to 100 m and greater than 100 m. Studies continue on effective ways of obtaining time and space dependent dispersivity values and of introducing scale and time dependence of dispersivity for model application.

In modelling applications, dispersion is quantified by using the coefficient of dispersion (also called the coefficient of hydrodynamic dispersion). This coefficient includes factors for groundwater flow, the nature of the aquifer, and diffusion effects. For one dimensional flow, the longitudinal coefficient of hydrodynamic dispersion is expressed in terms of two components as

$$D_l = \alpha \bar{v} + D_m, \tag{8.22}$$

where D_l is the coefficient of dispersion (L^2T^{-1}), α is the dynamic dispersivity (L), $\bar{v}$ is the average linear groundwater velocity (LT^{-1}), and D_m is the coefficient of molecular diffusion (L^2T^{-1}). The term α (L) is a characteristic property of the porous medium known as the dynamic dispersivity, or simply dispersivity, and D_m is the coefficient of molecular diffusion for the contaminant in porous medium. The coefficient of molecular diffusion is an empirical value that takes into account

the effect of solid phase on the diffusion process. At low flow velocities, diffusion is an important contributor to the dispersion process; as a result, the coefficient of hydrodynamic dispersion is set equal to the diffusion coefficient ($D_l = D_m$). In the case of high velocity, mechanical mixing is the dominant dispersive process, and the coefficient of hydrodynamic dispersion is set equal to the hydrodynamic dispersion ($D_l = \alpha\bar{v}$).

Hydrodynamic dispersion in the transverse direction is generally weaker than dispersion in the longitudinal direction. As a result contaminant plumes usually develop an elliptical shape even though the aquifer system may be isotropic. At low velocities, the coefficients of longitudinal and transverse dispersion are nearly equal.

8.6.3 Determination of Coefficient of Molecular Diffusion

Molecular diffusion is the transport of contaminants in their ionic state due to the difference in concentration levels in a given species in the aquifer. Values for coefficient of molecular diffusion D_m, also referred to as the apparent diffusion coefficient, are empirically determined and based on the diffusion coefficient (D) from Fick's first law governing diffusion in bulk solutions. Coefficients of molecular diffusion are proportionately reduced from diffusion coefficient values because in saturated porous media, as compared to a bulk liquid solution, ions must follow longer paths of diffusion because of the presence of the solid matrix. The coefficient of molecular diffusion for nonadsorbed species in porous media (D_m) is related to the diffusion coefficient (D) by

$$D_m = XD, \tag{8.23}$$

where D_m is coefficient of molecular diffusion (L^2T^{-1}), X is empirical coefficient (dimensionless), D is Fick's diffusion coefficient (L^2T^{-1}). X, which is less than 1, takes into account the effect of the solid phase particles of the porous media on the diffusion process. In laboratory studies of diffusion of non-adsorbed ions in porous media, X values between 0.01 and 0.50 are commonly observed. Often in contaminant transport model development, the effects of diffusion are considered negligible because of anticipated velocity levels through the soil and are, as a result, dropped from further consideration.

In view of the difficulties in measuring dispersivity and because of uncertainties over its physical definition in model development and application, both longitudinal and transverse dispersion are generally assumed to be unknown and are determined during model calibration. That is, dispersivity values are adjusted until the model reproduces an observed concentration pattern. Recently, dispersivities have been computed by inverse methods (Bobba and Joshi, 1989).

It is commonly assumed that the soil medium is isotropic with respect to dispersivity. The assumption of isotropy is a convenient one, because as yet there is no standard technique for differentiating between longitudinal and transverse dispersivities. However, it should be remembered that the assumption of isotropy with respect to dispersivity implies that the medium is also isotropic with respect to hydraulic conductivity.

8.7 Chemical and Biological Activity

Chemical and biological processes can have a significant impact on the presence of transported contaminants. Adsorption of a contaminant on the porous medium depends on the physicochemical relationships existing between the contaminant and the solid matrix. This physicochemical relationship may be described by a transport step and an attachment. The transport step involves primarily physical phenomena, whereas the attachment step, basically a chemical process, can be influenced by the chemical and physical parameters. The development of a mathematical mass transport model is explained below.

The numerous chemical and bio-chemical reactions that can alter contaminant concentrations in groundwater flow systems can be grouped into six categories, (1) adsorption-desorption reactions, (2) acid-base reactions, (3) solution-precipitation reactions, (4) oxidation-reduction reactions, (5) ion paring or complexation, and (6) microbial cell synthesis. Although these reaction types are easily identified, the actual effects and reaction kinetics of many processes are still unknown largely because of the complexities associated with groundwater systems. A clear understanding of each process based on current knowledge is needed.

8.7.1 Chemical Processes

The chemical process which is often an element in non-conservative contaminant transport models is adsorption-desorption (Table 8.4). Because of the process similarities, ion exchange generally is included in this category. One reason for the interest in adsorption-desorption is the general understanding of the phenomena involved and the ability to mathematically define the reaction kinetics.

Adsorption - desorption has been identified as a significant sink or source for constituents derived from sanitary landfill leachate, spray irrigation of sewage, wastewater lagoons, septic tank liquid discharge, injection of pre-treated waste water into aquifers, agricultural fertilization use, and pesticide movement in soils. The effects of adsorption-desorption are incorporated into groundwater models using relationships based on adsorption isotherms. Common adsorption isotherms include (1) linear, (2) Langmuir, (3) Freundlich, and (4) BET (Brunauer, Emett and Teller). For modelling purposes, the adsorption capacity of the soil is determined and the isotherm relationship obtained using laboratory procedures. This information is included in non-conservative models in the form of a sink term in the mass transport equation.

Ion exchange reactions can be instrumental in reducing the concentration of certain contaminants. Ionic species are removed from solution via a replacement reaction with ions held by electrostatic forces to charged functional groups on the surface of soil particles. The process is quantified based on the exchange capacity of the soil matrix which is termed the cation exchange capacity (CEC).

Incorporation of ion exchange processes into contaminant transport models is an area of current development. Often the chemical equilibria involved are complex particularly if adsorption is also included in the modelling process. The effects on contaminant transport resulting from ion exchange reactions are included in non-conservative transport models as sink terms.

TABLE 8.4
Models of transport equations describing transport, and adsorption of contaminants in groundwater

Phenomenon	Model	Equation
Conservation of mass	Buildup= + convergence (or) - divergence + source (or) - sink	$\frac{\partial \theta C+S}{\partial t} = -\frac{\partial J_x}{\partial x} - \frac{\partial J_y}{\partial y} - \frac{\partial J_z}{\partial t} + Q$
Molecular diffusion	Fick's Law	$J_x = -D_{mx}\frac{\partial C}{\partial x}, \quad J_y = -D_{my}\frac{\partial C}{\partial y}, \quad J_z = -D_{mz}\frac{\partial C}{\partial z}$
Simple convection	Piston flow	$J_x = v_x C, \quad J_y = v_y C, \quad J_z = v_z C$
Hydrodynamic dispersion	Proportionality law	$J_x = -D_L\frac{\partial C}{\partial x}, \quad J_y = -D_T\frac{\partial C}{\partial y}, \quad J_z = -D_T\frac{\partial C}{\partial z}$
Adsorption or exchange (Equilibrium)	Linear equilibrium	$S = k_1 C + k_2$
Adsorption or exchange (Equilibrium)	Freundlich equilibrium	$S = k_3 C^{k_4}$
Adsorption or exchange (Equilibrium)	Competition between ions 'a' and 'b'	$S_a = S_b(\frac{k_5 C_a}{C_b})$
Adsorption or exchange (Equilibrium)	Modified Kjelland	$S = S_{\max}\frac{C/C_0}{C/C_0+(1-C/C_0)}\exp[\ln k_6 + k_7(1-2C/C_0)]$
Adsorption or exchange (Nonequilibrium)	Linear	$\partial S/\partial t = k_8(k_1 C + k_2 - S)$
Adsorption or exchange (Nonequilibrium)	Freundlich	$\partial S/\partial t = k_9(k_3 C^{k_4} - S)$
Adsorption or exchange (Nonequilibrium)	Competition between ions 'a' and 'b'	$\partial S_a/\partial t = k_{10}(C_a S_b - (1/k_5)S_a C_b)$
		$\partial S/\partial t = k_{11}(S_{\max} - k_8 C)\sinh k_{12}(1 - k_6 C/S_{\max})$
		$\partial S/\partial t = k_{13}\,e^{k_{14}} S(k_{15} C\,e^{-2k_{14}S} - S)$

Chemical equilibria reactions, of which precipitation/dissolution is one type, can play a significant role in moderating the groundwater transport of contaminants. The conditions of time and presence of interactive chemical species can result in reduced environmental impact. The law of mass action is commonly used for describing the process. This relationship states that the driving force of chemical reaction is related to the concentrations of the constituents of reaction and the concentration of reaction products. However, the law is only an equilibrium statement and establishes nothing about the kinetics of the chemical processes involved. Development of models incorporating these concepts is presently in the research stage.

8.7.2 Linear Adsorption

The simplest and most widely used of the equilibrium sorption isotherms is that given by the linear relationship. That is, it is assumed that the amount S of contaminant adsorbed by the soil matrix and the concentration C of contaminant in the soil solution are related by a linear relationship (Table 8.4). The linear isotherm model, in conjunction with an advective-dispersive contaminant transport model, has been used frequently to describe the transport of radioactive material through porous media. Bobba and Joshi (1988) used such an isotherm in a model for radioactive contaminant transport from a waste disposal site situated entirely above the water table to the nearby Lake Ontario.

8.7.3 Freundlich Isotherm

The Freundlich isotherm is defined by the nonlinear relationship (Table 8.4). This isotherm is the oldest of the nonlinear sorption isotherms and has been used widely to describe the sorption of contaminants in soils. It should be kept in mind, however, when applying the Freundlich isotherm that the flexibility of the two constants allows for easy curve fitting, but does not guarantee accuracy if the data are extrapolated beyond experimental points. One limitation of the Freundlich isotherm is that, like the linear isotherm model, it does not imply a maximum quantity of adsorption. Numerous examples exist in the literature where the Freundlich isotherm has been used to describe the adsorption of contaminants in the soil matrix.

8.7.4 Langmuir Isotherm

The Langmuir adsorption isotherm was developed by Langmuir (1918) to describe the adsorption of gases by solids. Langmuir assumed that the surface of a solid possesses a finite number of adsorption sites. If a gas molecule strikes an unoccupied site, it is adsorbed, whereas if it strikes an occupied site, it is reflected back into the gas phase. This model leads immediately to the concept of an upper limit of adsorption. The maximum amount of adsorption occurs when the surface of the solid is covered with a closely packed adsorbed layer of gas molecules.

The two standard forms of the Langmuir isotherm are

$$C/S = (1/kb) + (C/b) \tag{8.24}$$

and

$$S = kbC(1 + kC), \tag{8.25}$$

where k is a measure of the bond strength holding the sorbed contaminant on the soil surface, b is the maximum amount of contaminant that can be adsorbed by the soil matrix (μg/g), and C and S are as previously described. The Langmuir adsorption isotherm has been used extensively in the literature to describe the sorption of contaminants by the soil.

8.7.5 Biological Processes

Microorganisms are present in groundwater systems and have the potential for significantly reducing contaminant levels. Their activity is responsible for nearly all the important redox reactions that occur in groundwater. Bacteria are the microorganisms that are most important in groundwater zone. Different varieties of bacteria can withstand high fluid pressures, wide ranging pH and temperature conditions, and high salinity levels. Some forms of bacteria, the aerobes, require dissolved oxygen while others, the aerobes, require anoxic conditions. Another group, called facultative bacteria, can thrive with or without oxygen. Bacteria are small compared to the pore sizes in most unconsolidated geological materials and can migrate through porous geological materials. In unfavorable environments, many species can evolve into resistant bodies that may be activated at a later time. The catalytic capability of bacteria is produced by the activity of enzymes that normally occur within the bacteria cells.

Bacteria and their enzymes are involved in redox processes to acquire energy for synthesis of new cells and maintenance of existing cells. The main source of energy for bacteria in the groundwater zone is from the degradation of organic matter which requires the presence of certain essential nutrients for growth. The metabolic byproducts of organic degradation can contribute to soil matrix plugging with a resultant decrease in hydraulic conductivity.

A good example of the capability of biological systems to reduce aquifer contamination can be illustrated using petroleum products as the test material. There exist species of aerobic and anaerobic bacteria that grow rapidly in the presence of oil or gasoline if the necessary nutrients are available with the net result of aquifer contaminant reduction. Although examples such as the point of the important role bacteria play in the groundwater environment exist, the study of these organisms in this domain is in its infancy. Further research is needed before suitable mathematical relationships can be developed for inclusion of biological reactions in contaminant transport models.

8.8 Development of Contaminant Transport Models

Although the models developed for approximate subsurface transport of contaminants are simplifications of reality, they should still provide valid estimates of

the process occurring. In these models a set of mathematical relationships are combined with certain assumptions and boundary conditions to obtain the desired solution.

8.8.1 Analytical and Numerical Models

Analytical models like numerical models consist of equations together with appropriate initial and boundary conditions that express conservation of mass, momentum and energy. Additionally, each modelling approach entails phenomenological relationships such as Darcy's law for flow, Fick's law for chemical diffusion, and Freundlich's isotherm for adsorption (Table 8.4). However, analytical models differ from numerical models in that all functional relationships are expressed in closed form with fixed parameters so that equations can be solved by classical methods of analytical mathematics.

Because groundwater systems are complex and cannot be completely comprehended or defined in their entirety, they are imagined to be simpler than they really are considering only those aspects which pertain to the problem at hand. Such a simplification may allow for the application of analytical models. On the other hand, when analyzing a specific event, significant complexities may arise where the capabilities of analytical models are exhausted and the application of more complex numerical models is required.

An important aspect of analytical modelling is the acknowledgment of the approximate nature of these models based on a clear understanding of model assumptions and limitations. Adequate documentation and appreciation of analytical model assumptions greatly assists the modeller and model user in keeping resultant expectations within a realistic perspective.

Analytical and numerical models require different amounts and types of data. Generally, as the modelling approach becomes more sophisticated in order to closely conform to reality, the associated data requirement increases. Often a point is reached when benefits of applying a more complex and hence more realistic model are weighted against the difficulty of defining such a model. This decision process regarding data ultimately has bearing on whether analytical or numerical modelling methods are used in analyzing the subsurface conditions at hand.

8.8.2 Types of Models

Mathematical models can be statistical or deterministic. Statistical models provide a range of solutions based on probabilities of occurrence. Deterministic models are based on cause and effect relationships of known systems and processes. Models currently available generally are deterministic, thus, this chapter is limited to numerical deterministic contaminant transport models.

Figure 8.4 summarizes the steps for developing a numerical deterministic model. The first step is to carefully determine the physical concepts that play an important role in the behavior of the system to be modelled. These include the previously discussed processes of advection, dispersion, and chemical and biological reactions. The coordination of these concepts into a single general relationship meeting the problem's predictive needs leads to the conceptual model.

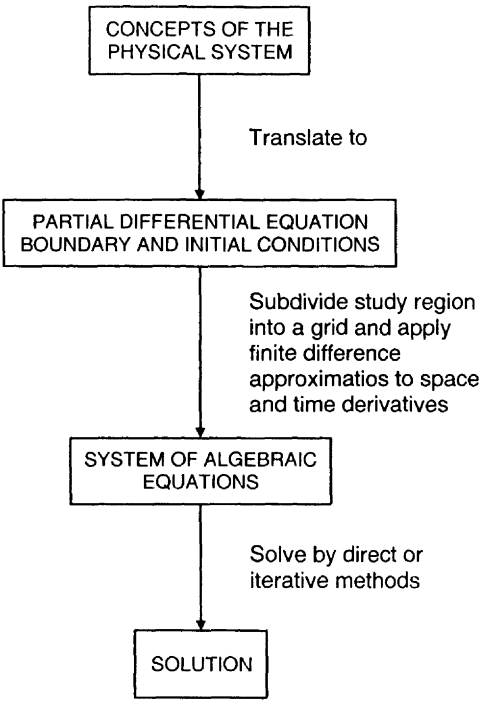


Fig. 8.4. The steps for developing a numerical deterministic model.

The next step is to translate the conceptual relationships in mathematical terms. This development, combined with certain simplifying assumptions, becomes the governing equation and constitutes the basic core of the mathematical model. For groundwater flow, the model consists of a partial differential equation together with appropriate boundary and initial conditions, and expresses conservation of mass and describes continuous variables over the region of interest. Additionally, certain laws describing rate processes are included. Darcy's law for fluid flow through porous media is an example. This law is generally used to express the conservation of momentum. Lastly, various assumptions are included such as those for one or two dimensional flow and those involving artesian or water table conditions.

When models are developed to include changes in transported contaminants, additional partial differential equations with appropriate boundary and initial conditions are required to express conservation of mass for the contaminant of interest. Mathematical terms are included which describe the effects of diffusion and hydrodynamic dispersion. To include the changes resulting from chemical and biological processes, relationships such as adsorption isotherms or the law of mass reactions are used.

Following the formulation of the basic mathematical model, the next step is to obtain a solution using one of two general approaches. If the model can be simplified further it may be amenable to analytical solution. Equations and

solutions of this type are referred to as analytical models. These are exact solution methods for the mathematical models developed.

When analytical models, because of their nature, no longer adequately describe the events to be modelled, numerical modelling techniques are used. The Partial Differential Equations (PDE) of the mathematical models are approximated numerically using finite difference or finite-element techniques. By using these approaches, continuous variables are replaced with discrete variables that are defined at nodal points. In this way, as an example, the continuous differential equation defining hydraulic head everywhere in the problem domain is replaced by a finite number of algebraic equations that define hydraulic head at specific points. This system of algebraic equations is solved by iterative methods using matrix techniques. The digital computer has made this solution approach practical because of the high speed and reasonably low cost at which the calculations can be performed.

There are important differences between the finite difference method and the finite element technique which is a relatively new development. A major benefit of the finite element method is its flexibility in formulating and generating a problem solution. Increasing use of this method is being made because of its advantages in analyzing situations involving irregular boundaries and for problems in which the medium is heterogeneous or anisotropic. The flexibility of the finite element method is also useful in solving coupled problems such as contaminant transport and in solving moving boundary problems, such as a moving water table. However, in the end, the selection of which method to apply generally depends on such factors as the complexity of the problem and the user's familiarity with each method.

8.8.3 Analytical Models

Although the number of analytical solutions to the differential equations of groundwater flow problems is limited, together they constitute a rather wide class. Because an analytical solution has the obvious advantage that it does not contain any approximation, the availability of such an analytical solution for certain problems makes it an attractive approach. The class of analytical solutions can be described in two ways: by defining the members of class, or by indicating the common restrictions. Analytical solutions exist for the following types of groundwater flow problems.

(a) Systems of wells in homogenous aquifers of infinite extent. The aquifer can be confined, unconfined, or semi-confined. By using image wells one can also construct solutions for regions in the form of a half-plane, a circle, or an infinite strip. All these solutions are of a two-dimensional character. In the aquifer head differences in vertical direction are neglected. Uniform infiltration into the aquifer can also be taken into account.

(b) Systems of wells in a homogeneous porous medium of infinite extent. By using image wells the solution can be extended towards a half space, or even a layer of constant thickness.

Although analytical solutions have the obvious advantage of being exact and thus error-free, they are restricted to relatively simple problems. The main restrictions are of the following form.

(i) The porous medium must be homogeneous. This applies to all properties of the porous medium: transmissivity, storativity, vertical resistance, and also to some of the hydraulic conditions: The infiltration rate must be uniformly distributed.

ii) The region occupied by the porous medium must be of mathematically simple shape: a half plane, a circle, an infinite strip, etc. Regions of arbitrary shapes cannot be accommodated.

It may be illustrative to present a solution, as an example. This is the solution for a homogeneous confined aquifer of infinite extent, with a number of wells, and a uniform flow at infinity. The solution is

$$H = -\frac{\theta}{K}(X\bar{v}_x + Y\bar{v}_y) + \sum_{i=1}^{n}\left[\frac{Q_i}{2\pi kB}\ln\left(\frac{r_i}{R}\right)\right]. \tag{8.26}$$

Here $\bar{v}_x$ is the velocity in the x-direction at infinity, and $\bar{v}_y$ is the velocity in the y-direction at infinity. The permeability of the aquifer is denoted by K, its porosity by θ, and its thickness by B. The parameter r_i denotes the distance from the i-th well to the point considered, Q_i is the discharge of the well, and R is a reference constant to be determined from the given value of the head at some point.

On the basis of formulas such as Equation (8.26) mass transport models have been developed. The basic solution for a well in a confined aquifer in Equation (8.26) is the logarithmic function. For other aquifers this function has to be replaced by some other function, for instance, a Bessel function in case of a semi-confined aquifer. From the analytical solution for the hydraulic head H the velocity components can be easily be determined with use of Darcy's law. Thus at every point of the field the velocity is known, and then it is a simple matter to construct stream lines by integrating the equations.

$$v_x = x/t \quad \text{and} \quad v_y = y/t, \tag{8.27}$$

where x and y are the coordinates of a material point.

8.8.4 Analytical Element Method Models

A generalization of the analytical model has been developed by Strack (1989) in the form of the analytical element method. The basic idea of this method is that solutions of the basic equations can be obtained by superposition of standard solutions for various problems, which may include singular solutions for inhomogeneities in the permeability, the infiltration rate, etc. Because the solution is in principle an analytical solution, it has all the advantages related to the analytical character. The solution is exact, for instance, also in a very small region, and this means that a small part of a large regional problem can be studied in minute detail. Also, the stream lines can be constructed on the basis of a continuous velocity field, which means that the stream lines are smooth. In this method all basic solutions are sought in the form of solutions of the differential equation

$$\frac{\partial^2 H}{\partial x^2} + \frac{\partial^2 H}{\partial y^2} = -I, \tag{8.28}$$

where I is an infiltration function. This equation applies to flow in a homogeneous region, with uniform infiltration. The infiltration function is also used to simulate leakage, and storage in non-steady problems, which then are considered constant over the region considered. The method has been elaborated for a single aquifer, and also for a multilayered aquifer. It should be mentioned that the solution becomes rather complex when considering a multi-layered aquifer with zones of variable permeability.

The main disadvantage of the method seems to be that it is primarily suited for homogeneous regions in steady flow conditions. The generalizations to a layered system, and to non-steady flow, are approximations, which lead to errors depending on the scale of the size of the elements. Perhaps another disadvantage is that it requires a considerable level of expertise on the part of the user. The method is more elegant, however, and should be considered as a serious competitor for the numerical models to be described later, primarily because of the possibility to include small details in a large scale model.

8.8.5 Finite Difference Models

Perhaps the most widely used type of model is the one based on a finite difference approximation of the spatial derivatives. The use of finite difference methods to solve partial differential equations was first introduced by Richardson in 1910. Although his paper was not directed to a problem in hydrology, he described a method by which diffusion equation might be solved. Shaw and Southwell (1941) were the first to apply the finite difference method to the steady state seepage problem in the field of hydrology. Since then, finite difference methods have been widely used to solve heat flow problems and the reservoir behavior problem in petroleum engineering. Douglas et al. (1959) employed an **A**lternating-**D**irection-**I**mplicit **P**rocedure (ADIP) to solve a two-dimensional, two-phase, incompressible flow model. Blair and Peaceman (1963) further extended this method to include compressibility. Quon (1965) utilized an **A**lternating-**D**irection-**E**xplicit **P**rocedure (ADEP) to solve two-dimensional mathematical models of petroleum reservoirs. Fagin and Stewart (1966) developed a two-dimensional, three phase reservoir simulator. They utilized the ADIP technique. A full volumetric account of three phases was performed simultaneously throughout the flow domain. Quon (1966) furthered the use of the ADEP technique by solving the natural gas reservoir problems which involved the nonlinear partial differential equation. They concluded that ADEP had an advantage over the conventional forward-difference-explicit procedure from the stability point of view and had an advantage over ADIP from a computational point of view.

Numerical methods and digital computer models had not been used as widely in the solution of groundwater flow problems as they had been used in petroleum engineering. Remson (1965) applied a digital computer model utilizing the finite difference method to predict the effect of a proposed reservoir on the water levels in the Old Bridge Sand aquifer of the Raritan formation in New Jersey. The aquifer was considered to be homogeneous and isotropic. The aquifer was partially confined and partially unconfined. For nodes representing confined aquifer, transmissivity remained constant but for unconfined aquifers the transmissivity

changed with time. Longenbaugh (1967) developed a generalized computer model to simulate the aquifer in a 25-mile reach of the Arkansas River Valley between La Junta and Las Animas, Colorado. The Gauss-sidel technique was used to solve the finite difference approximations of the governing groundwater flow equation. A comparison between the historical stream flows and computed return flows was used to check the validity of the model. Pinder and Bredehoeft (1968) employed an implicit finite difference scheme to build a digital computer model of the Musquodoboit Harbour in Nova Scotia, Canada. For hypothetical homogeneous isotropic aquifers of simple geometry, the results of the digital model were found to compare favourably with the results from the analytical solutions of the problems. The hydrogeologic parameters in the model were modified until the model duplicated the pumping test data. Then, as a check for the model, results from the digital model were compared with results from an electronic analog model of the aquifer. Hornberger et al. (1970) used a predictor corrector numerical technique to obtain groundwater recession and groundwater flow in response to changes in stream stage. The Dupuit-Forchheimer assumptions were utilized in developing the mathematical model. Also, the system was idealized by considering the stream as deep as the aquifer and by assuming the aquifer did not receive any recharge. It was observed that the model predicted the water table position very poorly; however, the predictions for flows were of reasonable accuracy. Utilizing an alternating direction implicit technique, Bredehoeft and Pinder (1970) solved a three-dimensional mathematical model for multiple aquifers in terms of two-dimensional aquifer models having essentially horizontal flows. The two aquifers considered were coupled through leakage. Taylor and Luthin (1969) developed computer models to obtain simultaneous solutions for flow in the saturated and unsaturated zones. The Gauss-Sidel iterative method was used to obtain a solution of finite difference approximations of the governing differential equation. The model was applied to the problem of withdrawal from the aquifer through a well. The results obtained from the computer model were compared with the results of the sand tank model. It was observed that at equilibrium the difference in water level derived from the computer model and the sand tank model was no more than 2 percent. Hornberger et al. (1970) used a finite-difference approximation technique to develop computer models for transient moisture movement in the composite soil moisture and the groundwater system. No comparison of computer model derived results was made against field or experimental results. The authors concluded that in view of the objective to solve field size composite groundwater soil moisture systems the model was not considered satisfactory due to practical and theoretical limitations. Shamir and Dagan (1971) developed a set of simultaneous partial differential equations to define motion of saltwater body in the aquifer. An implicit scheme was developed to solve these equations simultaneously. The solution resulted in an easily solvable seven-diagonal matrix. The results from the numerical solution were found to compare favourably with Hele-Shaw model of the system. Reddell (1969) employed an implicit numerical technique to obtain a simultaneous solution of partial differential equations describing dispersion in groundwater flow systems. Others utilizing numerical methods based on finite difference approximations are Stone and Brian (1963), and Hoops and Harleman (1967).

Having reviewed the previous work done in the field of groundwater modelling, utilizing numerical methods based on finite difference approximation, the subsequent sections will deal with the development of digital computer models for the system described by the following equation:

$$S\frac{\partial H}{\partial t} = \frac{\partial}{\partial x}\left[T\frac{\partial H}{\partial x}\right] + \frac{\partial}{\partial y}\left[T\frac{\partial H}{\partial y}\right] \pm Q(x, y, t), \tag{8.29}$$

where S, H, T, t, x and y are described previously and $Q(x, y, t)$ is a known function and refers to any withdrawal from or recharge to the aquifer. It is to be emphasized here that in Equation (8.29) T is equal to the product of hydraulic conductivity (K) and thickness of the saturated zone of the confined aquifer (h). Throughout the flow domain, it is assumed that the boundary condition can be given in the form of a Dirichlet or Neumann problem.

8.8.5.1 Finite Difference Approximations

The numerical solution entails replacement of partial derivatives by finite difference approximations. The finite difference approximations are directly derived from the fundamental definition of the derivative and Taylor series expansion of a function about a point. If $f(x)$ represents a function, then the Taylor series expansion of the function at point x_0 can be written as

$$f(x) = f(x_0) + (x - x_0)f'(x) + \frac{(x - x_0)^2}{2!} f''(x^*) + \cdots, \tag{8.30}$$

where $x_0 < x^* < x$. Therefore, the first derivative can be approximated by

$$f'(x) = \frac{f(x) - f(x_0)}{x - x_0}. \tag{8.31}$$

The error in this approximation of the first derivative is

$$E_T = -\frac{(x - x_0)^2}{2!} f''(x^*), \tag{8.32}$$

where E_T is a truncation error. Now if it is assumed that $x = x_0 + h$, then Equation (8.31) becomes

$$f'(x) = \frac{f(x_0 + h) - f(x_0)}{h}. \tag{8.33}$$

The above is a forward difference approximation of the first derivative. The backward difference approximation can be obtained by substituting $x = x_0 - h$ in Equation (8.31). The substitution gives the following backward difference approximation of the first derivative:

$$f'(x) = \frac{f(x_0) - f(x_0 - h)}{h}. \tag{8.34}$$

Similarly, an expression for the second derivative of $f(x)$ can be written as

$$f''(x) = \frac{f'(x_0 + h) - f'(x_0)}{h} . \tag{8.35}$$

In order to avoid a bias for any one approximation scheme, the forward difference scheme is used for $f'(x_0 + h)$ and the backward difference scheme is used for $f'(x_0)$. Substituting for $f'(x_0 + h)$ and $f'(x_0)$ in Equation (8.35) gives:

$$f''(x) = \frac{f(x_0 + h) - 2f(x_0) + f(x_0 - h)}{h^2} . \tag{8.36}$$

To utilize finite difference approximations for the solution of partial differential equations governing groundwater flow, the region of integration is overlaid by an x, y grid. The groundwater flow equation has one dependent variable (H) and three independent variables (x, y, t). Let i, j and n be indices of the x, y and t coordinates, and Δx, Δy and Δt be the size of increments in the x, y and t coordinates, respectively. By similarity to Equation (8.36), the partial derivatives in the groundwater flow equations for any point i, j in the region can be written as:

$$\left[\frac{\partial^2 H}{\partial y^2}\right]_{i,j} = \frac{H_{i,j+1} - 2H_{i,j} + H_{i,j-1}}{\Delta y^2} . \tag{8.37}$$

Regarding the values of T in Equation (8.29), Saulyev (1964) showed that even if the variable is discontinuous in the region under consideration, midway values can be used without causing any significant error. Bredehoeft and Pinder (1968) agreed with Saulyev's conclusion. Thus, in terms of finite difference approximations, Equation (8.29) can be written as follows:

$$\begin{aligned} S_{i,j}\left[\frac{\partial H}{\partial t}\right]_{i,j} = {} & T_{i-\frac{1}{2},j}\frac{(H_{i-1,j} - H_{i,j})}{\Delta x^2} - T_{i+\frac{1}{2},j}\frac{(H_{i,j} - H_{i+1,j})}{\Delta x^2} + \\ & + T_{i,j-\frac{1}{2}}\frac{(H_{i,j-1} - H_{i,j})}{\Delta y^2} - T_{i,j+\frac{1}{2}}\frac{(H_{i,j} - H_{i,j+1})}{\Delta y^2} + \\ & + Q_{i,j}. \end{aligned} \tag{8.38}$$

A proper average value of $(\partial H/\partial t)_{i,j}$ over the time interval t is found by integrating between $t = n\Delta t$ and $t = (n+1)\Delta t$, and thus the following expression results:

$$H_{i,j}^{n+1} - H_{i,j}^{n} = \left[\frac{\partial H}{\partial t}\right]_{i,j} \Delta t. \tag{8.39}$$

Substituting Equation (8.39) into Equation (8.38) gives

$$\begin{aligned} S_{i,j}\frac{(H_{i,j}^{n+1} - H_{i,j}^{n})}{\Delta t} = {} & T_{i-\frac{1}{2},j}\frac{(H_{i-1,j} - H_{i,j})}{\Delta x^2} - T_{i+\frac{1}{2},j}\frac{(H_{i,j} - H_{i+1,j})}{\Delta x^2} + \\ & + T_{i,j-\frac{1}{2}}\frac{(H_{i,j-1} - H_{i,j})}{\Delta y^2} - T_{i,j+\frac{1}{2}}\frac{(H_{i,j} - H_{i,j+1})}{\Delta y^2} + \\ & + Q_{i,j}. \end{aligned} \tag{8.40}$$

8.8.5.2 *Comparison of Different Methods*

Different numerical methods can be used to solve Equation (8.40). The basic difference between these methods is the weight given to the right-hand side flux terms in Equation (8.40) at time $t = n\Delta t$ and at time $t = (n + 1)\Delta t$. In general, these methods can be categorized under two headings: the implicit methods and the explicit methods. Implicit methods are interpolative and require a solution of a set of simultaneous equations for each time interval, or in terms of the matrix algebra find an inverse of the matrix for each time interval. Simultaneous equations can also be solved by iteration, relaxation or the Gauss-Sidel elimination technique. Computations are complicated and time consuming. Explicit methods are extrapolative and require a simple solution of an algebraic equation for one unknown for each point and for each time interval. The computation time required is much smaller than for the implicit technique. Also, computations are easy to program. However, there are many types of implicit and explicit methods.

All of the numerical methods to be discussed here can be obtained as a special case of the following equation obtained by multiplying Equation (8.40) by $\Delta x\,\Delta y$ and writing it for both time $t = n\Delta t$ and $t = (n + 1)\Delta t$:

$$
\begin{aligned}
&S_{i,j}\Delta x\,\Delta y\frac{(H_{i,j}^{n+1} - H_{i,j}^{n})}{\Delta t} = \\
&\beta_{i-\frac{1}{2},j}^{n+1}T_{i-\frac{1}{2},j}\frac{\Delta y}{\Delta x}(H_{i-1,j}^{n+1} - H_{i,j}^{n+1}) + \beta_{i-\frac{1}{2},j}^{n}T_{i-\frac{1}{2},j}\frac{\Delta y}{\Delta x}(H_{i-1,j}^{n} - H_{i,j}^{n}) + \\
&+\beta_{i+\frac{1}{2},j}^{n+1}T_{i+\frac{1}{2},j}\frac{\Delta y}{\Delta x}(H_{i+1,j}^{n+1} - H_{i,j}^{n+1}) + \beta_{i+\frac{1}{2},j}^{n}T_{i+\frac{1}{2},j}\frac{\Delta y}{\Delta x}(H_{i+1,j}^{n} - H_{i,j}^{n}) + \\
&+\beta_{i,j-\frac{1}{2}}^{n+1}T_{i,j-\frac{1}{2}}\frac{\Delta x}{\Delta y}(H_{i,j-1}^{n+1} - H_{i,j}^{n+1}) + \beta_{i,j-\frac{1}{2}}^{n}T_{i,j-\frac{1}{2}}\frac{\Delta x}{\Delta y}(H_{i,j-1}^{n} - H_{i,j}^{n}) + \\
&+\beta_{i,j+\frac{1}{2}}^{n+1}T_{i,j+\frac{1}{2}}\frac{\Delta x}{\Delta y}(H_{i,j+1}^{n+1} - H_{i,j}^{n+1}) + \beta_{i,j+\frac{1}{2}}^{n}T_{i,j+\frac{1}{2}}\frac{\Delta x}{\Delta y}(H_{i,j+1}^{n} - H_{i,j}^{n}) + \\
&+Q_{i,j}^{n+\frac{1}{2}}. \qquad (8.41)
\end{aligned}
$$

Each of the terms in the above equation on the right-hand side is a volumetric flux of water and the β's are the associated weighting factors placed on each flux term over the time interval.

8.8.5.3 *Forward Difference Method*

The forward difference method is an explicit method. It extrapolates the value of $H_{i,j}^{n+1}$ from values of known $H_{i,j}^{n}$ terms, i.e., in this method all the weight is placed on the flux term at $t = n\,\Delta t$ and none on $t = (n + 1)\,\Delta t$, i.e., all β's with superscript n are one and with superscript $n + 1$ are zero. Stability problems are encountered since the error of extrapolation may be accumulated. However, the computations are very easy to perform.

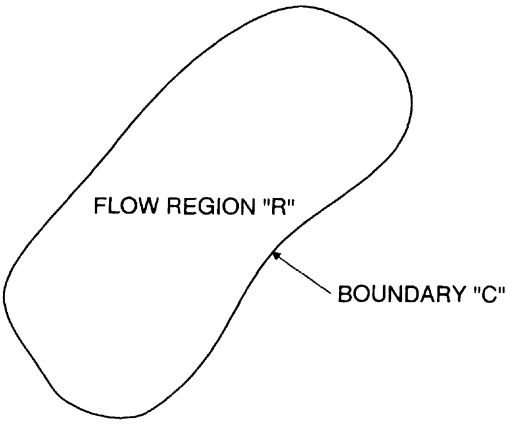


Fig. 8.5. A hypothetical two dimensional groundwater flow region.

8.8.5.4 Backward Difference Method

In the backward difference method, all the weight is placed on flux terms at $t = (n+1)\,\Delta t$ and none on $t = n\Delta t$, i.e., all β's in Equation (8.41) with superscript $n + 1$ are 1 and with superscript n are zero. Since the method is interpolative, no stability problems are encountered. As none of the potentials at $t = (n + 1)\Delta t$ are known, the computation involved is complicated and time consuming. Suppose there are N and M grid nodes in x and y directions, respectively. This method requires that an $NM \times NM$ matrix be solved for each time step. Since NM may be in the hundreds, the method requires considerable computer time.

8.8.5.5 Crank Nicholson Method

The Crank Nicholson method puts equal weight on all flux terms, i.e., all β's in Equation (8.41) are equal to 1/2. The method is conceptually sounder than the above two methods since all fluxes are given equal weight. However, it poses the same kind of computational difficulty as the backward difference method.

8.8.5.6 Alternating Direction Explicit Method

The Alternating Direction Explicit Method (ADEM) was first reported by Saulyev (1964) and was further improved by Larkin (1964). In ADEM, a cycle of calculation consists of two stages. Each stage, like the previous method, has two flux terms at $t = (n + 1)\Delta t$ and two flux terms at $t = n\Delta t$. However, the two flux terms at any time t either at $(n + 1)\Delta t$ or at $n\Delta t$ are 90 degrees apart instead of 180 degrees apart as in the previous method. Apparently, it would seem that the resultant equations would be implicit, since the equation for each grid node will contain three presumably unknown potential heads at $t = (n + 1)\Delta t$. However, if the starting point for the calculation is chosen at the corner, the appropriate boundary conditions would permit the determination of two of these potentials at $t = (n + 1)\Delta t$, making the resulting equation explicit. This method is also unconditionally stable.

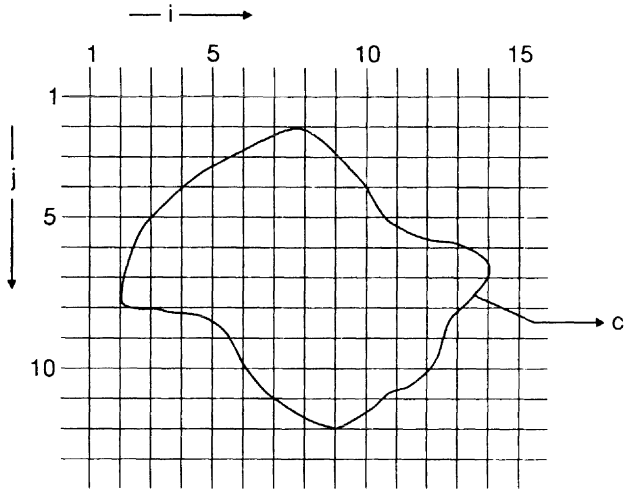


Fig. 8.6. Example of finite difference representation of an aquifer region.

The basic advantage of ADEM is that in spite of being explicit the stability conditions do not place any restrictions on the size of time and spatial increments. In addition, the computation time required by this method is considerably smaller than the other more conventional methods. For the purpose of illustrating ADEM, assume that the region R in Figure 8.5 represents an aquifer. The flow domain of an aquifer is bounded by the curve C. Conditions on the curve are assumed known either in the form of the Dirichlet problem (hydraulic, H, is known) or in the form of the Neumann problem (hydraulic head gradient $\partial H/\partial r$, where r is normal to C, is known) or any combination of these two. Also known are the hydraulic head H at every point in the region R at time $t = 0$ and any recharge or withdrawal from any point in the basin ($Q(x, y, t)$) for all time, t. Obviously, transmissivity or hydraulic conductivity and aquifer thickness and storage coefficient values are known at all points within the region, R. Figure 8.6 shows the hypothetical aquifer, under consideration, overlaid by a grid network. The index i varies from 1 to 15 and lies along the x-axis of the rectangular coordinate system while the index j varies from 1 to 13 and lies along the y-axis of the rectangular coordinate system. The distance between nodes need not be the same over the entire region. Here it is assumed that the distance between the nodes along the x-axis remains constant and is equal to Δx, and also, the distance between the nodes along the y-axis remains constant and is equal to Δy.

Using appropriate values of β in Equation (8.41) will give the necessary equations of ADEM. Since ADEM is a two stage process, two equations have been solved alternatively. They are:

$$S_{i,j}\frac{H_{i,j}^{n+1} - H_{i,j}^{n}}{\Delta t} =$$
$$= T_{i+\frac{1}{2},j}\frac{(H_{i+1,j}^{n} - H_{i,j}^{n})}{\Delta x^2} + T_{i,j+\frac{1}{2}}\frac{(H_{i,j+1}^{n} - H_{i,j}^{n})}{\Delta y^2} +$$

$$+ T_{i-\frac{1}{2},j}\frac{(H^{n+1}_{i-1,j} - H^{n+1}_{i,j})}{\Delta x^2} + T_{i,j-\frac{1}{2}}\frac{(H^{n+1}_{i,j-1} - H^{n+1}_{i,j})}{\Delta y^2} + - \frac{Q^{n+\frac{1}{2}}_{i,j}}{\Delta x\,\Delta y} \tag{8.42}$$

and

$$S_{i,j}\frac{H^{n+1}_{i,j} - H^{n}_{i,j}}{\Delta t} = = T_{i-\frac{1}{2},j}\frac{(H^{n}_{i-1,j} - H^{n}_{i,j})}{\Delta x^2} + T_{i,j-\frac{1}{2}}\frac{(H^{n}_{i,j-1} - H^{n}_{i,j})}{\Delta y^2} + + T_{i+\frac{1}{2},j}\frac{(H^{n+1}_{i+1,j} - H^{n+1}_{i,j})}{\Delta x^2} + T_{i,j+\frac{1}{2}}\frac{(H^{n+1}_{i,j+1} - H^{n+1}_{i,j})}{\Delta y^2} + - \frac{Q^{n+\frac{1}{2}}_{i,j}}{\Delta x\,\Delta y}. \tag{8.43}$$

The computations start from a corner point where all but one term for time $t = (n+1)\Delta t$ are known by virtue of being at boundaries. One by one all the rows are evaluated using Equation (8.42). In the second stage, Equation (8.43) is used and the computations start from a corner point diagonally opposite to the corner point of the first stage. This is done to balance errors and preserve symmetry.

As an example, calculations for point ($i = 2, j = 2$) are given here. Equation (8.42) is used and for point ($i = 2, j = 2$) it becomes

$$S_{2,2}\frac{H^{n+1}_{2,2} - H^{n}_{2,2}}{\Delta t} = T_{2.5,2}\frac{(H^{n}_{3,2} - H^{n}_{2,2})}{\Delta x^2} + T_{2,2.5}\frac{(H^{n}_{2,3} - H^{n}_{2,2})}{\Delta y^2} + + T_{1.5,2}\frac{(H^{n+1}_{1,2} - H^{n+1}_{2,2})}{\Delta x^2} T_{2,1.5}\frac{(H^{n+1}_{2,1} - H^{n+1}_{2,2})}{\Delta y^2} - \frac{Q^{n+\frac{1}{2}}_{2,2}}{\Delta x\,\Delta y}. \tag{8.44}$$

In the above equation, every term except $H^{n+1}_{2,2}$ is known and, thus, $H^{n+1}_{2,2}$ can be evaluated explicitly. In a similar fashion $H^{n+1}_{3,2} \cdots H^{n+1}_{16,14}$ can be evaluated. In the second stage, Equation (8.43) is written for point ($i = 16$, $j = 14$). In that equation it will be observed that all terms attached to are known. So, can be evaluated directly. In a similar way, computations are done for all points from ($i = 16, j = 14$) to ($i = 2, j = 2$) and there the second cycle starts. Thus it is clear that by choosing the two flux terms at a particular time, t, at right angles rather than in line, thereby making a clever use of the boundary condition, Equations (8.42) and (8.43) are transformed from implicit to explicit.

A detailed stability analysis of an equation very similar to Equations (8.42) and (8.43) is given by Larkin (1964). Following a procedure very similar to Larkin's, it can be shown that the stability restriction on Equations (8.42) and (8.43) is

$$S_{i,j} \frac{\Delta x^2}{\Delta t} > 0, \tag{8.45}$$

i.e., the equations are stable for all positive time increments.

It is required that the solution of a finite difference approximation scheme converges to the solution of the partial differential equation as the increment sizes of the independent variables are reduced. Larkin (1964) gives a detailed analysis of the truncation error for an equation very similar to the groundwater flow equation. He concludes that the alternating direction explicit method yields a favourable truncation error.

8.8.5.7 Alternating Direction Implicit Method

In the Alternating Direction Implicit Method (ADIM), a cycle of calculation consists of two stages. In each stage, weight is placed on two flux terms at $t = (n+1)\Delta t$ and two flux terms at $t = n\Delta t$. In the first stage, the flux terms in the x direction are taken at $t = (n + 1)\Delta t$ and in the y direction at $t = n\Delta t$. Thus, in solving for the potential head at $t = (n + 1)\Delta t$, each row of unknown potentials is effectively decoupled from all other rows. In practical terms, this means that instead of solving an $NM \times NM$ matrix, an $N \times N$ matrix is solved M times, a much simpler procedure. In the second stage, the two flux terms in the y direction are taken at $t = (n + 1)\Delta t$ and two flux terms in the x direction at $t = n\Delta t$. There is no stability problem in this method.

For a transport model the transport equation may also be approximated by ADIM. The details of the method are described by Bobba (1989). The method was applied to contaminant transport through fracture-porous system.

8.8.6 Finite Element Models

A second very powerful numerical method is the finite element method (Wang and Anderson, 1982). An elementary way of presenting this method is used in structural mechanics, where the elements are the actual parts of a structure, say the beams and columns in the framework of a building, or the grid of beams in the floor of a bridge. The deformation of each element is expressed in terms of the forces acting upon it at the two ends. This enables to express the displacement of each nodal point in terms of those of the neighbouring nodes, and the deformation of the connecting elements. The final system of equations is obtained from the conditions of equilibrium at each node.

In groundwater flow problems one could imagine that a region is subdivided into small elements, each with its own physical properties, in such a way that for each element the flow is described in terms of the head in the nodal points, using Darcy's law in each element, and that a system of algebraic equations is obtained from the condition that the flow must be continuous at each node.

The usual way of presenting the finite element method does not employ such a physical reasoning, however. Instead, a mathematical argument is used, in which the system of equations is obtained by requiring that the differential equation be satisfied 'on the average', using certain weight functions. The system of equations obtained in the finite element method has the same structure as in the finite

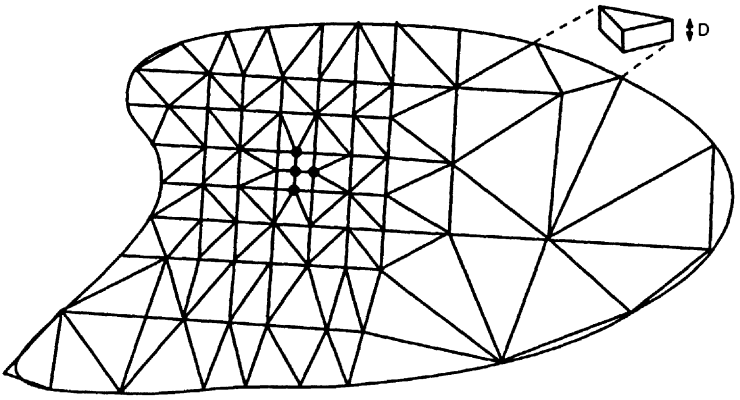


Fig. 8.7. Example of finite element representation of an aquifer region.

difference method. Actually, the two methods are very similar, and, for certain problems, it has been shown that they can be considered as two representations of the same mathematical model. However, the usual way of deriving and developing the equations reveals certain differences. For instance, the natural and simplest type of element is the triangle, which enables a flexible representation of the field, whereas the simplest and most natural mesh in the finite difference method is a mesh of squares or rectangles, which is less flexible. It can be seen from Figure 8.7 that in a mesh of triangles it is easily possible to have elements of different sizes, and also to follow smooth curved boundaries. Another advantage of the finite element method is that in its standard formulation it has the immediate property that each element can have its own values for the physical parameters, such as transmissivity and storativity.

8.8.6.1 Galerkin's approximation

To solve a differential equation

$$L(U) = 0 \tag{8.46}$$

by Galerkin's method, a trial solution of the form

$$U \approx \hat{U}(x, y, t) = \sum_{n=1}^{n} C_i(t)\, N_i(x, y) \tag{8.47}$$

is assumed. Here $N_i(x, y)$ $(i = 1, 2, \ldots, n)$ is system of functions (referred to as basis functions) selected before hand and which satisfy the essential boundary conditions of the original differential equations. The functions $N_i(x, y)$ $(i = 1, 2, 3, \ldots, n)$ are assumed to be linearly independent and to represent the first N functions of some system of functions that is complete in the given region. The functions $C_i(t)$ are the undetermined coefficients.

The approximating functions $\hat{U}(x, y, t)$ will be an exact solution to Equation (8.46) only if $L(U)$ is identically equal to zero, otherwise a residual $R_E = L(\hat{U})$

will result. In the method of weighted residuals an attempt is made to force this residual to zero in an average sense through the selection of constants C_i, $(i = 1, 2, \ldots, n)$. To do this the residual is distributed over the spatial domain Ω by using a weighting function $W_i(x, y)$ such that

$$\iint_{\Omega} R_E W_i(x, y)\, dx\, dy = 0, \quad i = 1, 2, \ldots, n \tag{8.48}$$

or recalling the definition of the inner product

$$\langle R_E(X, Y), W_i(X, Y)\rangle = 0, \quad i = 1, 2, \ldots, n, \tag{8.49}$$

i.e., the residual is made orthogonal to the weighting function. In Galerkin's method, the weighting function is chosen to be the same as basis functions $N_i(x, y)$.

As n basis functions are selected and as there are n undetermined coefficients $C_i(t)$ $(i = 1, 2, 3, \ldots, n)$, only n conditions of orthogonality can be satisfied. These are

$$\iint_{\Omega} L(\hat{U}) N_i(x, y)\, dx\, dy$$
$$= \iint_{\Omega} L\left[\sum_{j=1}^{n} C_j(t) N_j(x, y)\right] N_i(x, y)\, dx\, dy = 0, \quad i = 1, 2, \ldots, n. \tag{8.50}$$

Performing appropriate integration of the basis functions, and solving the resulting set of simultaneous equations, the coefficients $C_i(t)$ could be determined. Substituting these values in Equation (8.47), the solution to Equation (8.46) could be obtained. Forray (1968) gives the proof for convergence

$$\lim_{n \to \infty} \hat{U} = U \tag{8.51}$$

8.8.6.2 Steps in the finite element method solution

The following basic steps are involved in obtaining the solution to the differential equation using the finite element method.

(1) The domain of interest is divided into a number of finite elements described by node points.

(2) Suitable basis (shape) functions are selected to approximate the dependent variable at each nodal point. Values between these points are calculated using piecewise continuous interpolating functions defined over each element.

(3) The weighted integrals of the residuals are set to zero to obtain a set of simultaneous equations.

(4) The resulting set of equations are solved to obtain the values of the dependent variable at all nodal points.

(5) Partial integration can be carried out to reduce the continuity required and also satisfy only the kinematic boundary conditions.

8.8.6.3 Choice of the Element Shape and Basis Functions

Ergatoudis et al. (1968) introduced the curved isoparametric quadrilateral element and used it mainly in structural engineering applications. The basic shape of the element is quadrilateral, but the sides may be distorted in a prescribed way. Eragatoudis developed basis functions using the same order of approximations. Pinder et al. (1972) developed this deformed, mixed isoparametric quadrilateral element. The advantage of this element is that all four sides of the quadrilateral need not be of the same order. For each side of the element linear, quadratic, or cubic approximation could be used depending upon the shape of the boundaries or the expected shape of the unknown function.

The basis functions $N_i(x, y)$ in Equation (8.47) are selected such that $N_i(x, y)$ is unity at node i and zero at all other nodes. Because of the fact that the basis functions are chosen with this constraint, the undetermined coefficients $C_i(t)$ $(i = 1, 2, \ldots, n)$ in Equation (8.47) are equal to the required function $\hat{U}$ at the n node points in the domain.

To facilitate integration of the basis functions over each element a dimensionless and curvilinear local coordinate system (ζ, η) is introduced. All elements appear as squares in the local coordinate system with no side nodes (linear side), one side at the mid point (Quadratic node), or two side nodes at one-third points (cubic side). The functions $N_i(x, y)$ referred to as basis functions are also sometimes known as shape functions.

The mapping from local coordinates to cartesian coordinates x, y is by the following shape function mapping

$$X = N_1 X_1 + N_2 X_2 + \cdots = \sum_{j=1}^{n} N_j X_j, \tag{8.52}$$

$$Y = N_1 Y_1 + N_2 Y_2 + \cdots = \sum_{j=1}^{n} N_j Y_j. \tag{8.53}$$

The summation is over each element and N_j are functions of ζ and η. The unknown functions $\hat{U}_i$ are prescribed by shape functions as

$$\hat{U}_i = \sum_{i=1}^{n} N_i U_i. \tag{8.54}$$

As a point is moved around in the parent element its local coordinates in that element will generate the cartesian coordinates through Equations (8.52) and (8.53) and the corresponding U_i values through Equation (8.54). The functions $N_i(x, y)$ written in terms of ζ and η are selected such that in addition to satisfying the basic requirements of a basis function they also relate the global and local co-ordinate systems.

8.8.7 Application of Galerkin's method

The objective is to obtain the values of the dependent variable H (hydraulic he and C (concentration) at the n nodal points in the system. The trial solutions H and C are now assumed in the form

$$H \approx \hat{H} = \sum_{i=1}^{n} H_i(t) N_i(x, y) \tag{8.5}$$

and

$$C \approx \hat{C} = \sum_{i=1}^{n} C_i(t) N_i(x, y), \tag{8.56}$$

where $N_i(x, y)$ $(i = 1, 2, 3, \ldots, n)$ are the basis functions (N_i need not necessarily be the same for approximating H and C). The basis functions satisfy the essential boundary conditions of the original partial differential equations. $H_i(t)$ and $C_i(t)$ are a set of undetermined coefficients and at the specified points (referred to as nodes) in the domain Ω.

Define

$$L_h(H) \equiv \frac{\partial}{\partial x}\left(K_x \frac{\partial H}{\partial x}\right) + \frac{\partial}{\partial y}\left(K_y \frac{\partial H}{\partial y}\right) + q = 0 \tag{8.57}$$

and

$$\begin{aligned} L_c(C) \equiv & \frac{\partial}{\partial x}\left(D_{xx} \frac{\partial C}{\partial x}\right) + \frac{\partial}{\partial y}\left(D_{yy} \frac{\partial C}{\partial y}\right) + \frac{\partial}{\partial x}\left(D_{xy} \frac{\partial C}{\partial y}\right) + \\ & + \frac{\partial}{\partial y}\left(D_{xy} \frac{\partial C}{\partial x}\right) - \frac{\partial}{\partial X}(V_{xx} C) - \frac{\partial}{\partial y}(V_{yy} C) + S(x, y) \\ & - \theta \frac{\partial C}{\partial t} = 0. \end{aligned} \tag{8.58}$$

The approximating functions $\hat{H}$ and $\hat{C}$ when substituted in $L_h(H)$ and $L_c(C)$ will yield the residual. The residual is now minimized by using the orthogonality condition with respect to the basis functions.

$$\iint_{\Omega} L_h(\hat{H}) N_i \, \mathrm{d}\Omega = 0, \quad i = 1, 2, \ldots, n \tag{8.59}$$

and

$$\iint_{\Omega} L_c(\hat{C}) N_i \, \mathrm{d}\Omega = 0, \quad i = 1, 2, \ldots, n \tag{8.60}$$

The n orthogonality conditions can be satisfied as n basis functions have been selected. These are

$$\iint_{\Omega} L_h\left[\sum_{j=1}^{n} H_j N_j(x, y)\right] N_i(x, y) \, \mathrm{d}\Omega = 0, \quad i = 1, 2, \ldots, n \tag{8.61}$$

and

$$\iint_{\Omega} L_c[\sum_{j=1}^{n} C_j N_j(x,y)]\, N_i(x,y)\, \mathrm{d}\Omega = 0, \quad i = 1,2,\ldots,n. \tag{8.62}$$

Substituting $L_h(H)$ and $L_c(C)$ in Equation (8.61) and (8.62) the following equations are obtained:

$$\iint_{\Omega} \left\{ \left[\frac{\partial}{\partial x}\left(K_x \frac{\partial}{\partial x}\right) + \frac{\partial}{\partial y}\left(K_y \frac{\partial}{\partial y}\right) \right] \sum_{j=1}^{n} H_j N_j - q \right\} N_i \, \mathrm{d}\Omega = 0,$$
$$i = 1,2,\ldots,n. \tag{8.63}$$

and

$$\iint_{\Omega} \left\{ \left[\frac{\partial}{\partial x}\left(D_{xx} \frac{\partial}{\partial x}\right) + \frac{\partial}{\partial y}\left(D_{yy} \frac{\partial}{\partial y}\right) + \frac{\partial}{\partial x}\left(D_{xy} \frac{\partial}{\partial y}\right) + \frac{\partial}{\partial y}\left(D_{yx} \frac{\partial}{\partial x}\right) \right] \right.$$
$$\sum_{j=1}^{n} C_j n_j - \frac{\partial}{\partial x}\left(\sum_{j=1}^{n} C_j N_j V_x\right) - \frac{\partial}{\partial y}\left(\sum_{j=1}^{n} C_j N_j V_y\right)$$
$$\left. - \frac{\partial}{\partial t}\left(\theta \sum_{j=1}^{n} C_j N_j\right) + S \right\} N_i \, \mathrm{d}\Omega = 0, \quad i = 1,2,3,\ldots,n. \tag{8.64}$$

8.8.8 Application of Green's theorem

The second derivative terms appearing in Equations (8.63) and (8.64) would require continuity of the first derivative at all element interfaces. To avoid this restriction the second derivative terms are transformed to first derivatives by using Green's theorem in the following form

$$\iint_{\Omega} \Psi \frac{\partial^2 H}{\partial x_\alpha \partial x_\beta} \, \mathrm{d}\Omega = - \iint_{\Omega} \frac{\mathrm{d}H}{\mathrm{d}x_\alpha} \cdot \frac{\mathrm{d}\Psi}{\mathrm{d}x_\beta} \, \mathrm{d}\Omega + \int_s \frac{\partial H}{\partial n} \, \mathrm{d}s, \quad \alpha, \beta = 1,2, \tag{8.65}$$

where

$$\frac{\partial H}{\partial n} = \frac{\partial}{\partial x_\alpha} l_\alpha; \; \frac{\partial H}{\partial n}$$

being the outward normal derivatives of the function $H(x_\alpha)$ and l_α's are the direction cosines.

Application of Green's theorem is made element by element. Using Green's theorem Equations (8.63) and (8.64) are now transformed as follows:

$$\iint_{\Omega} \sum_{j=1}^{n} \left(K_x \frac{\partial N_i}{\partial x} \frac{\partial N_j}{\partial x} + K_y \frac{\partial N_i}{\partial y} \frac{\partial N_j}{\partial y} \right) H_j \, \mathrm{d}x \, \mathrm{d}y +$$

$$+\iint_{\Omega} qN_i \,\mathrm{d}x\,\mathrm{d}y - \int_s N_i \sum_{j=1}^{n} \Big(K_x \frac{\partial N_j}{\partial x} l_x + K_y \frac{\partial N_j}{\partial y} l_y\Big) H_j \,\mathrm{d}s,$$
$$i = 1, 2, 3, \ldots, n \tag{8.66}$$

and

$$\iint_{\Omega} \sum_{j=1}^{n} \Big(D_{xx} \frac{\partial N_i}{\partial x}\frac{\partial N_j}{\partial x} + D_{yy} \frac{\partial N_i}{\partial y}\frac{\partial N_j}{\partial y} + D_{xy} \frac{\partial N_i}{\partial x}\frac{\partial N_j}{\partial y} + D_{yx} \frac{\partial N_i}{\partial x}\frac{\partial N_j}{\partial y} +$$
$$+ N_j V_x \frac{\partial N_i}{\partial x} + N_j V_y \frac{\partial N_i}{\partial y} + N_j N_j \frac{\partial v_x}{\partial x} + N_j N_j \frac{\partial v_y}{\partial y}\Big) C_j \,\mathrm{d}x\,\mathrm{d}y +$$
$$+ \iint_{\Omega} \theta N_i N_j \,\mathrm{d}x\,\mathrm{d}y \frac{\partial C}{\partial t} - \iint_{\Omega} S N_i \,\mathrm{d}x\,\mathrm{d}y$$
$$- \int_s N_i \Big(D_{xx} \frac{\partial C}{\partial x} l_x + D_{xy} \frac{\partial C}{\partial y} l_x + D_{yx} \frac{\partial C}{\partial x} l_y + D_{yy} \frac{\partial C}{\partial y} l_y\Big) \,\mathrm{d}s = 0,$$
$$i = 1, 2, 3, \ldots, n. \tag{8.67}$$

The above equations in matrix form are :

$$[A]\{H\} + \{B\} = 0 \tag{8.68}$$

for groundwater flow and

$$([K] + [L]\{C\} + [M]\Big\{\frac{\partial C}{\partial t}\Big\} + \{F\} = 0 \tag{8.69}$$

for solute transport.

Typical elements of different matrices are

$$A_{ij} = \iint_{\Omega} \Big(K_x \frac{\partial N_i}{\partial x}\frac{\partial N_j}{\partial x} + K_y \frac{\partial N_i}{\partial y}\frac{\partial N_j}{\partial y} \,\mathrm{d}x\,\mathrm{d}y, \tag{8.70}$$

$$B_i = \iint_{\Omega} qN_i \,\mathrm{d}x\,\mathrm{d}y - \int_s N_i \sum_{j=1}^{n} \Big(K_x \frac{\partial N_i}{\partial x} l_x + K_y \frac{\partial N_j}{\partial y} l_y\Big) N_j \,\mathrm{d}s, \tag{8.71}$$

$$K_{ij} = \iint_{\Omega} D_{xx} \frac{\partial N_i}{\partial x}\frac{\partial N_j}{\partial x} + D_{yy} \frac{\partial N_i}{\partial y}\frac{\partial N_j}{\partial y} + D_{xy} \frac{\partial N_i}{\partial x}\frac{\partial N_j}{\partial y} +$$
$$+ D_{yx} \frac{\partial N_i}{\partial y}\frac{\partial N_j}{\partial x}, \tag{8.72}$$

$$L_{ij} = \iint_{\Omega} N_j \Big(V_x \frac{\partial N_i}{\partial x} + V_y \frac{\partial N_i}{\partial y}\Big) + N_i N_j \Big(\frac{\partial V_x}{\partial x} + \frac{\partial V_y}{\partial y}\Big), \tag{8.73}$$

$$M_{ij} = \iint_{\Omega} \theta N_i N_j \,\mathrm{d}x\,\mathrm{d}y, \tag{8.74}$$

$$F_i = -\iint_\Omega SN_i \,\mathrm{d}x\,\mathrm{d}y - \int_\Omega N_i\left(D_{xx}\frac{\partial}{\partial x}l_x + D_{yy}\frac{\partial}{\partial y}l_y + D_{xy}\frac{\partial}{\partial x}l_x + D_{yx}\frac{\partial}{\partial y}l_y\right) \sum_{j=1}^{n} C_j N_j \,\mathrm{d}s, \quad i = 1,2,3,\ldots,n;\ j = 1,2,3,\ldots,n. \tag{8.75}$$

8.8.8.1 Integration of the Approximating Equations

An analysis of Equations (8.68) to (8.73) indicates that integration of the basis functions and their derivatives has to be done. As the basis functions are selected in such a way that they are non-zero only over the elements over which they are located, the domain Ω at any time consists of one or more elements depending upon the location of a particular node.

As stated earlier, each element appears as a square in the local coordinate system. The integration is done on the square element in the local coordinate system and the limits of integration are $+1$ and -1. The basis function derivatives $\partial N_i/\partial x$ and $\partial N_i/\partial y$ are obtained using the relationship.

$$\begin{Bmatrix} \dfrac{\partial N_i}{\partial \zeta} \\ \dfrac{\partial N_i}{\partial \eta} \end{Bmatrix} = \begin{bmatrix} \dfrac{\partial X}{\partial \zeta} & \dfrac{\partial Y}{\partial \zeta} \\ \dfrac{\partial X}{\partial \eta} & \dfrac{\partial Y}{\partial \eta} \end{bmatrix} \begin{Bmatrix} \dfrac{\partial N_i}{\partial x} \\ \dfrac{\partial N_i}{\partial y} \end{Bmatrix} = [J] \begin{Bmatrix} \dfrac{\partial N_i}{\partial x} \\ \dfrac{\partial N_i}{\partial y} \end{Bmatrix} \tag{8.76}$$

$[J]$ is the Jacobian matrix evaluated numerically (Ergatoudis et al., 1968) noting that

$$[J] = \begin{bmatrix} \dfrac{\partial N_1}{\partial \zeta}, & \dfrac{\partial N_2}{\partial \zeta} & \cdots & \dfrac{\partial N_n}{\partial \zeta} \\ \vdots & & & \vdots \\ \dfrac{\partial N_1}{\partial \zeta}, & \dfrac{\partial N_2}{\partial \zeta} & \cdots & \dfrac{\partial N_n}{\partial \zeta} \end{bmatrix} \begin{bmatrix} X_1 & Y_1 \\ X_2 & Y_2 \\ \vdots & \vdots \end{bmatrix}. \tag{8.77}$$

The element area in the global coordinate system is also changed to local-coordinate system using the relationship:

$$\mathrm{d}x\,\mathrm{d}y = \det[J]\,\mathrm{d}\eta\,\mathrm{d}\zeta. \tag{8.78}$$

Numerical integration is done over each element using a Gaussian quadrature. In this technique a polynomial $f(x)$ of a degree $2N-1$ could be integrated exactly as a weighted mean of its N partial values at specified points referred to as Gauss points. If cubic elements are used a sixth degree polynomial in ζ and η is generated and 16 Gauss points (4×4 rule) are required. For all elements, generally 9 Gauss points (3×3) are used.

8.8.8.2 Boundary Conditions

The original Equations (8.4) and (8.14) are subject to certain boundary conditions and they should be incorporated in the calculations. For both the flow and transport equations two types of boundary conditions could be described. These are: (a) Dirichlet type wherein the hydraulic head or concentration at certain nodes is specified, and (b) Neumann type wherein the flux of water or contaminant entering the aquifer is specified.

The second term in the $\{B\}$ matrix (Equation (8.71)) for flow equation and the second term in the $\{F\}$ matrix (Equation (8.75)) for the transport equation are generated only if Neumann type boundary conditions are prescribed. For nodes where Dirichlet type boundary conditions are prescribed the elements of the matrices $[A]$ in the flow equation and $([K] + [L])$ and $[M]$ in the transport equation are not generated in the simultaneous set of equations resulting from Equations (8.68) and (8.69). The nodes for which the Dirichlet conditions are prescribed are known values. In the computational procedure the $\{B\}$ or $\{F\}$ matrix depending on whether the flow equation or the transport equation is being considered has to be modified as follows. Let the elements of $\{B\}$ and $\{F\}$ be denoted by b_i^* $(i = 1, 2, 3, \ldots, n)$ and f_i^* $(1, 2, 3, \ldots, n)$, respectively. Then modify $\{B\}$ or $\{F\}$ as per rule

$$b_i = b_i^* + \sum_{\text{all } j} A_{ij} H_j \tag{8.79}$$

and

$$f_i = f_i^* + \sum_{\text{all } j} ([K] + [L])_{ij} C_j, \tag{8.80}$$

where j is the number of nodes for which the Dirichlet type boundary conditions are prescribed in each case. The computational steps are shown as follows.

Consider Equation (8.69) in the form

$$[N]\{C\} + [M]\{\frac{\partial C}{\partial t}\} + \{F\} = 0. \tag{8.81}$$

The partitioning of the matrices $[N]$, $[M]$ and $\{F\}$ for a four-node element using the Dirichlet boundary conditions is as follows. Write the coefficient matrices in the following form.

$$\begin{bmatrix} N_{11} & N_{12} & N_{13} & N_{14} \\ N_{21} & N_{22} & N_{23} & N_{24} \\ N_{31} & N_{32} & N_{33} & N_{34} \\ N_{41} & N_{42} & N_{43} & N_{44} \end{bmatrix} \{C\} +$$

$$+ \begin{bmatrix} M_{11} & M_{12} & M_{13} & M_{14} \\ M_{21} & M_{22} & M_{23} & M_{24} \\ M_{31} & M_{32} & M_{33} & M_{34} \\ M_{41} & M_{42} & M_{43} & M_{44} \end{bmatrix} \{\frac{\partial C}{\partial t}\} + \{F\} = 0 \tag{8.82}$$

and assume the Dirichlet boundary conditions are prescribed for node 1, first the rows and columns of $[N]$ and $[M]$ matrices for node 1, are made zero.

Next, applying the backward finite difference approximation to $\mathrm{d}C/\mathrm{d}t$, Equation (8.82) will be as follows:

$$\begin{bmatrix} N_{22}+\frac{M_{22}}{\Delta t} & N_{23}+\frac{M_{23}}{\Delta t} & N_{24}+\frac{M_{24}}{\Delta t} \\ N_{32}+\frac{M_{32}}{\Delta t} & N_{33}+\frac{M_{33}}{\Delta t} & N_{34}+\frac{M_{34}}{\Delta t} \\ N_{42}+\frac{M_{42}}{\Delta t} & N_{43}+\frac{M_{43}}{\Delta t} & N_{44}+\frac{M_{44}}{\Delta t} \end{bmatrix} \{C\}_{t_0+\Delta t} =$$

$$= \begin{bmatrix} \frac{M_{22}}{\Delta t} & \frac{M_{23}}{\Delta t} & \frac{M_{24}}{\Delta t} \\ \frac{M_{32}}{\Delta t} & \frac{M_{33}}{\Delta t} & \frac{M_{34}}{\Delta t} \\ \frac{M_{42}}{\Delta t} & \frac{M_{43}}{\Delta t} & \frac{M_{44}}{\Delta t} \end{bmatrix} \{C\}_{t_0} - \begin{Bmatrix} N_{21}\cdot BC_1 \\ N_{31}\cdot BC_1 \\ N_{41}\cdot BC_1 \end{Bmatrix}. \tag{8.83}$$

The right side matrices are now combined into one matrix as given by

$$\begin{bmatrix} \frac{M_{22}}{\Delta t}C_2+\frac{M_{23}}{\Delta t}C_3+\frac{M_{24}}{\Delta t}C_4-N_{21}\cdot BC_1 \\ \frac{M_{32}}{\Delta t}C_2+\frac{M_{33}}{\Delta t}C_3+\frac{M_{34}}{\Delta t}C_4-N_{21}\cdot BC_1 \\ \frac{M_{42}}{\Delta t}C_2+\frac{M_{43}}{\Delta t}C_3+\frac{M_{44}}{\Delta t}C_4-N_{41}\cdot BC_1 \end{bmatrix} \tag{8.84}$$

where C_2, C_3, C_4 are concentrations at nodes 2, 3, and 4 respectively at time t_0, and BC_1 is the value of the concentration at node 1.

8.8.8.3 Calculation of Velocity Vectors

Apart from the dispersion coefficients, the main input to the transport equation is the seepage velocity coefficients. The flow Equation (8.4) is solved to obtain the hydraulic head values at the nodes. To obtain the seepage velocity components at the nodes of each element, the basis functions are differentiated at each Gauss point and the functional coefficient approach is used (Pinder, 1971).

$$V_x = K_x\frac{\partial H}{\partial x} = K_x\sum_{i=1}^{n}\frac{\partial N_i}{\partial x}H_i, \tag{8.85}$$

$$V_y = K_y\frac{\partial H}{\partial y} = K_y\sum_{i=1}^{n}\frac{\partial N_i}{\partial y}H_i. \tag{8.86}$$

Again, for calculating the gradient of the seepage velocity components a similar approach is used.

$$\frac{\partial V_x}{\partial x} = \sum_{i=1}^{n} \frac{\partial N_i}{\partial x} V_{x_i}, \tag{8.87}$$

$$\frac{\partial V_y}{\partial x} = \sum_{i=1}^{n} \frac{\partial N_i}{\partial y} V_{y_i}. \tag{8.88}$$

The gradients of the seepage velocity components are calculated at the center of the element ($\zeta = 0$; $\eta = 0$). For calculating the dispersion coefficients using the velocity relationships, the velocity components calculated at the center of the element are used. For each element the components of the hydrodynamic dispersion coefficient are assumed to be the same.

8.8.8.4 Calculation of the Time Derivative

The set of Equations (8.69) are a set of first order differential equations. To solve these equations a finite difference approximation to the time derivative dC/dt is adopted. The time derivative could be approximated either by central, forward or backward differences. The backward difference scheme, which is fully implicit, is found to be stable and accurate (Pinder, 1973). Using the backward difference method, Equation (8.69) can be written as:

$$([K] + [L])\{C\}_{t+\Delta t} + \frac{[M](\{C\}_{t+\Delta t} - \{C\}_t)}{\Delta t} + \{F\} = 0. \tag{8.89}$$

By rearranging, the following recurrence relationship is obtained:

$$([K] + [L]) + \frac{[M]}{\Delta t}\{C\}_{t+\Delta t} = \left(\frac{[M]}{\Delta t}\right)\{C\}_t - \{F\}. \tag{8.90}$$

Starting with the initial, known concentration distribution for all the nodes, this recurrence relationship calculates the concentration for the subsequent time steps.

8.8.8.5 Nature of the Matrices

Equations (8.68) and (8.69) as stated earlier are a set of simultaneous equations. The matrices $[A]$, $([K] + [L])$ and $[M]$ are $(n - m) \times (n - m)$ size where m is the number of passive nodes due to Dirichlet boundary conditions. While the matrix $[A]$ is symmetric, the matrix $([K] + [L])$ is non symmetric, the nonsymmetry being due to the nonequality of the components of the seepage velocity. This lack of symmetry requires an increased computational effort as the entire matrix is to be considered in computations, whereas in case of symmetric matrices only the upper triangular part need to be considered. Even though the matrices are nonsymmetric they are sparse and banded. The band width can be kept to a minimum by proper numbering of the nodes in the system.

Considering any two of the sixteen nodes which describe an element, let the differences in their node numbers be denoted by N_d. For a given element let the

maximum value of N_d be denoted by NE_d. Considering all elements in the system let the maximum value of NE_d be denoted as $NE_{\max}$. For minimization of the bandwidth and consequently the computational cost for a given analysis, the node points should be numbered so as to minimize the value of $NE_{\max}$.

8.8.8.6 Comments on Finite Element Method

The finite element method, in its standard form, can be considered to be somewhat more flexible than the standard form of the finite difference method. This may explain its popularity. It must be noted, however, that there exist variants of the finite difference method that are very similar to the finite element method. Thus, the two methods may be considered as more or less equivalent. The method to be used for a given class of problems may be determined to a large part by such considerations as the convenience of a program, its availability and cost, and the level of experience of the user.

A disadvantage of the finite element method, as compared to the finite difference method, is that the memory requirement is often considerably larger. The formulation of a finite element program leads to a system of linear equations, just as the finite difference method, but the general procedure is such that these equations are constructed by generating a system matrix, the coefficients of which have to be stored in the computer's memory. In the most elementary form this means that for a system with N unknowns a matrix of $N * N$ coefficients must be stored. Fortunately, the matrix is sparse, due to the fact that connections between nodes happen only when they are in one element. Thus, by careful programming, it is possible to reduce the memory requirement to about $5 * N$, taking into account the symmetry of the system matrix also. When these properties of the system matrix are used, the finite element method becomes a real competitor for the finite difference method. Of course, the advantage, that each element, at least in principle, may have its own values for all the physical parameters, also involves a certain penalty, in that all these values have to be stored in the computer's memory. In general one may conclude that, as a rule, the memory requirements of a finite element program are larger than those of a finite difference program. It is worth while to pay this price only when the flexibility of the finite element method is really needed.

Finite element models are widely used in engineering practice for the analysis of groundwater flow and contaminant transport problems (Sherif et al., 1988). In contrast with finite difference methods, however, the program codes are often proprietary, as some of them have been developed by private companies. For the analysis of transport problems the finite element method can also be used, and several models are currently available.

8.8.9 Boundary Element Models

A completely different numerical technique is the boundary element method or boundary integral equation method. The details of Boundary Element Method are given by Liggett and Liu (1983). The most common way of presenting this method is through integral theorems such as Green's second identity, or by starting from certain analytical solutions for typical singularities, such as a point source, or a

dipole, in an infinite field. The system of equations is obtained by an addition of the contribution of singularity distributions along the entire boundary of the domain, and then requiring that along each part of the boundary, the correct boundary conditions are satisfied, at least on the average. This leads to a system of N equations with N unknowns. Usually these equations are linear. The number N here is the number of points on the boundary, and not, as in the finite difference and finite element methods, the number of points in the interior of the domain and its boundary. Thus the number of unknowns and the number of equations are usually considerably lower. This fact is often considered as an important advantage of the method. It must be noted, however, that the matrix of the system of equations is completely filled with non-zero coefficients, whereas in the finite difference and finite element methods, the matrix is sparse, which means that only a small percentage of the coefficients of the matrix is actually different from zero. Thus a direct and general comparison of memory requirements, and computation time requirements, is not possible, and may depend upon the type of the problem considered.

A disadvantage of the boundary element method is that it is basically restricted to regions with homogeneous soil and fluid properties, as it uses a fundamental solution for a singularity in a homogeneous field. With some additional effort, the method can be generalized to a body consisting of a number of smaller subregions, each of them having constant physical properties. It then seems more natural, however, to use the finite element method.

8.8.10 Velocity Based Modelling

In finite difference or finite element models the velocity field generally has to be derived from the distribution of the groundwater head, as calculated by the model. This involves numerical differentiation, as an approximation of Darcy's law, or a similar numerical process, during which an order of magnitude of accuracy is lost. This is sometimes accepted as an unavoidable consequence of the numerical approximations, but in many cases the description of the flow process is so inaccurate, for instance in the vicinity of wells, that it becomes unacceptable. In order to overcome these difficulties an interesting alternative has been developed by Ziji (1984). In this method the basic equations are formulated in terms of velocity components. This leads to a higher order system of simultaneous differential equations, for each of the three components, together with appropriate boundary conditions. This system is then solved numerically by using a finite difference or a finite element approximation. Continuity and smoothness of the velocity field are assured in this way. The price to be paid is that of an increase in the number of equations and some difficulties in the handling of the boundary conditions.

8.9 Accuracy of Numerical Models

Numerical models can, of course, never give more than approximate representation of reality. In general these models use a mesh of discrete points, in which the physical values are calculated, and it can be expected that this creates a lower

bound for the fluctuations that the model can represent. In addition to this there are some more difficulties inherent to numerical modelling of transport processes. These are discussed in what follows.

8.9.1 NUMERICAL DISPERSION

The first difficulty is that there appears to be a contradiction in the basics of numerical modelling of transport by dispersion. As mentioned before, the dispersivities are of the order of magnitude of the pore size, or, more generally, of the size of the random irregularities in the porous medium. The size will be denoted by d. The transition to a continuum model involves an averaging process over some representative elementary volume (Bear, 1972), which must be an order of magnitude larger than the size of the irregularities, d. This means that the smallest physically relevant dimension in the continuum formulation is about $10d$,

$$x \gg d. \tag{8.91}$$

On the other hand, in a numerical solution a spatial discretization is used. The characteristic mesh size is denoted by x. As a result of the discretization various numerical errors are introduced. One of these is due to the approximation of the first order derivative which describes the advective transport. It can be shown that the resulting error is an additional dispersion-like term (numerical dispersion), having a dispersion coefficient $v\Delta x$. In order to suppress this error the numerical dispersion effect should be small compared to the physical dispersion. This leads to the requirement that the mesh size x should be small compared to the dispersivities, hence

$$x \ll d. \tag{8.92}$$

This is in contradiction with the earlier conclusion that the smallest dimension in the problem should be large compared to d.

For the solution of this problem of numerical dispersion various procedures have been suggested, and are currently being investigated. Most of these are based upon the analysis of the transport problem by using the method of characteristics (Konikow and Bredehoeft, 1974). This can also be viewed as a process of tracing particles, as they are being transported by the pore fluid. In this way the error in modelling the advective transport is eliminated. The dispersive transport now can be modeled using discrete particles also, and a convenient method to do so is to use a random walk simulation. A disadvantage of such a simulation is that it does not immediately give quantitative results in the form of the concentration at a point as a function of time. The results give a good insight into the real behavior, however.

8.10 Available Numerical Models

Numerous models are available to investigate groundwater contamination. These models are a result of different concepts and/or combinations of processes involved. It is difficult to give a summary of all the numerical models. Bachmat et al.

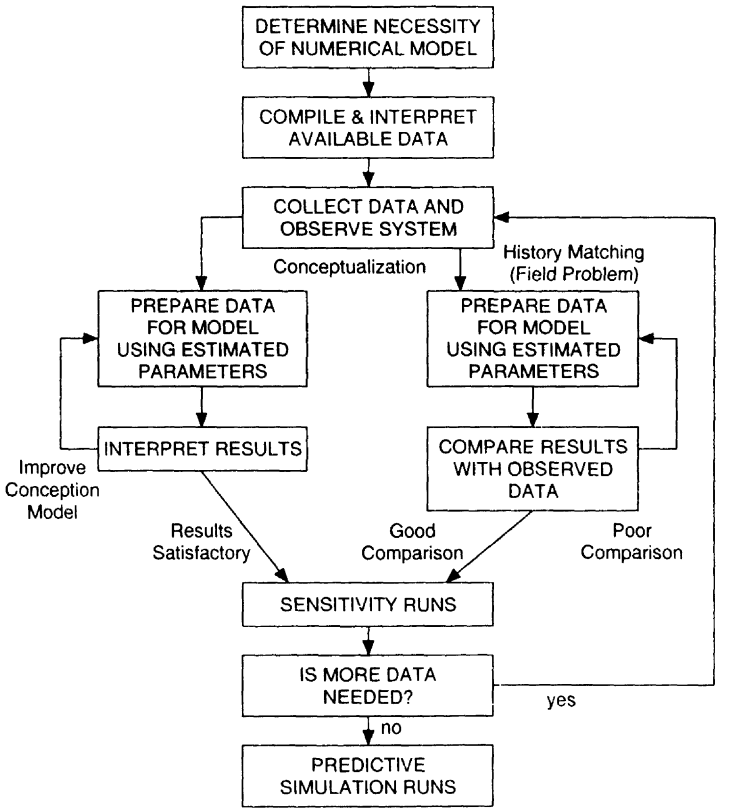


Fig. 8.8. The process of model application.

(1980) conducted a survey of numerical models mainly related to groundwater management. That report contained a list of 138 flow models and 39 mass transport models. Naymik (1987) presented a systematic review of 44 technically advanced articles on mathematical modelling of contaminant transport in the subsurface system. Another comprehensive review paper on the modelling of contaminant transport in groundwater was presented by Abriola (1987).

8.10.1 General Guidance for Groundwater Contamination Model Use

The application of a groundwater contamination model to an aquifer contamination problem is involved and requires several areas of effort. These are shown in Figure 8.8 and include (1) data collection, (2) data preparation for the model, (3) history of matching, and (4) predictive simulation. These tasks should not be considered separate steps of a chronological procedure, rather they should be viewed in a feedback mode. It is best to use a model not only as a predictive tool, but also as an aid in conceptualizing the aquifer behavior.

Data preparation for the groundwater model first involves determining boundaries of the region to be modeled. The boundaries may be physical (impermeable or no flow, recharge or no flow, recharge or specified flux, and constant head) or merely convenient (small subregion of a large aquifer). Once the boundaries of the aquifer are determined it is necessary to discretize the region, that is, subdivide it into a grid. Depending on the numerical procedure used, the grid may have rectangular or irregular polygonal subdivisions.

Once the grid is designed, it is necessary to specify aquifer parameters and initial data for the grid. Required program data includes aquifer properties for each grid block, such as storage coefficients and transmissivities. For nonconservative contaminant transport simulations, additional data such as chemical reaction properties are required.

Computed results generally include hydraulic heads at each of the grid blocks throughout aquifer. These spatial distributions of hydraulic heads are determined at each of a sequence of time levels covering the period of interest. For contaminant transport problems, computed results also include concentrations at each of the grid blocks.

Initial estimates of aquifer parameters constitute the first step in a trial-and-error procedure known as history matching. The matching procedure is used to refine initial estimates of aquifer properties and to determine boundaries, (i.e. the areal and vertical extent of the aquifer) and the flow conditions at the boundaries (boundary conditions). Aquifer tests generally provide the first estimates for storage coefficients and transmissivities. For certain groundwater problems, steady-state (or equilibrium) heads must also be determined and used as initial or beginning conditions. When called for, simulated wells in the aquifer grid system are allowed to pump at the observation rates, and computed (simulated) drawdowns are compared with observed drawdowns.

Assuming that the model is correct, a comparison between observed and predicted values indicates the accuracy of the initial estimates of input data. It may be necessary to modify some of the input data until all observed and calculated data compare sufficiently well. In the past, this has been done by trial and error. More recently the amount of work in matching has been reduced somewhat by using parameter estimation methods that modify initial estimates of input data in a more objective fashion. No hard and fast rules exist to indicate when a satisfactory match is obtained. The number of runs required to produce a satisfactory match depends on the objectives of the analysis, the complexity of the flow system, and length of observed history, as well as the patience of the analyst. Once completed, the model can be used to predict the future behavior of the aquifer. Of course, confidence in any predictive results must be based on (1) a thorough understanding of model limitations, (2) the accuracy of the match with observed historical behavior, and (3) the knowledge of data reliability and aquifer characteristics.

At the beginning of this chapter a number of specific applications for use of contaminant transport predictive techniques were discussed. Possible applications are wide ranging from environmental impact of new and existing land sited facilities to integration of surface and subsurface flow. However, the underlying purpose for applying predictive models is estimation of the rate of movement and concentration changes of contaminants that reach groundwater systems. In addition to

the cited uses, contaminant transport models also can be used to determine the effects of parameter changes on the overall contaminant process. This operation is referred to as sensitivity analysis and is useful to indicate when additional data is needed to address a particular problem at a particular site.

8.10.2 Data Requirements

There are a number of data requirements to be considered depending on the type of model to be used to predict contaminant transport (flow models for simpler but conservative prediction or more complex contaminant transport models) and the specific problem to be analyzed. Obtaining good field data is an important factor. Moore (1979) developed a list of data requirements which has been tailored to contaminant transport predictive techniques. The listed data requirements are as follows:

8.10.3 Physical Framework

8.10.3.1 Groundwater Flow

(a) Topographic map showing surface water bodies.

(b) Hydrogeologic map showing areal extent, boundaries, and boundary conditions of all aquifers.

(c) Water table, bedrock configuration, and saturated thickness maps.

(d) Transmissivity or hydraulic conductivity map showing aquifer and boundaries.

(e) Transmissivity and specific storage map of confining bed.

(f) Map showing variation in storage coefficient of aquifer.

(g) Relation of saturated thickness to transmissivity.

(h) Relation of streams and aquifer (hydraulic conductivity)

8.10.3.2 Contaminant Transport

(a) Estimates of the parameters that comprise hydrodynamic dispersion.

(b) Effective porosity distribution.

(c) Background information on natural concentration distribution (water quality) in aquifer.

(d) Estimates of fluid density variations and relationship of density to concentration.

(e) Hydraulic head distributions (used to determine groundwater velocities).

(f) Boundary conditions for concentrations.

(g) Chemical and biological reaction coefficients.

8.11 Stresses on System

8.11.1 Groundwater Flow

(a) Type and extent of recharge areas (irrigated areas, recharge basins, recharge wells, etc.).

(b) Surface water diversions.

(c) Groundwater pumpage (distributed in time and space).

(d) Precipitation.

8.11.2 Contaminant Transport (in addition to above)

(a) Areal and temporal distribution of water quality in aquifer.

(b) Stream flow quality (distribution in time and space).

(c) Sources and strengths of pollution.

8.12 Physical Interpretation

Just as the physical groundwater system is idealized as continuous in deriving the differential equations, it is also expedient to idealize the conditions on the boundaries of the system in order that they also can be given a mathematical expression. The boundary conditions of groundwater systems in nature are of several kinds, perhaps the most common being those describing the conditions at a well. Because the porous medium stops at the well face, the aquifer not only has a boundary around its perimeter but the outline of each well is also considered a boundary to the aquifer. The boundary conditions at wells are treated as constant or variable, specified flux, or constant head, depending on which best describes the actual physical conditions. If the well is discharging or recharging at a given contaminant concentration, then these may also be specified if the transport equations are being applied.

Impermeable or nearly impermeable boundaries are formed by (1) underlying or overlying beds of rock, (2) continuous rock masses (as long a fault or along the wall of a buried rock valley), or (3) dikes or similar structures. Permeable boundaries are formed by the bottoms of (a) rivers, (b) canals, (c) lakes, (d) pits, (e) ponds, and (f) other surface water bodies. These permeable boundaries may be treated as surfaces of equal head (specified) if the body of surface water is large in volume so that its level is uniform and independent of changes in groundwater flow. The uniform head on a boundary of this type, however, may change with time due to seasonal or other variations which affect the surface water level. Other

bodies of surface water, such as streams, may form boundaries with nonuniform distribution of head which may be either constant or variable with time. A small stream, for example, might be affected by a nearby withdrawal of groundwater if that withdrawal occurred at a rate of the same order of magnitude as the flow in the stream. In this case the boundary conditions would not be independent of the groundwater flow; that is, it would be a head-dependent flux condition.

In the case of a pond containing wastes, the contaminant transport equation boundary conditions at the holding pond must be specified. In theory, this may be simply accomplished by specifying some concentration at the bottom of the pond which may vary with time as the concentration of the waste in the pond changes. In practice, however, historical records rarely contain the necessary data to accurately define this boundary condition. This condition is perhaps the most critical portion of the model study because it describes the source of pollution. Often, it can only be estimated, and depends on the judgement of the individual performing the study. A more accurate definition of this boundary condition is of great importance to those who dispose of solid or liquid wastes in disposal sited facilities. The following results have been calculated by a finite difference model for a hypothetical groundwater aquifer (Bobba and Bukata, 1980).

8.12.1 Effect of Velocity Ratios on Contaminant Transport Model

Figure 8.9 displays the times to achieve the steady state concentration at different locations from the origin in areal plane. The velocity ratio is taken to be 20 and the adsorption coefficient K is taken to be 0.05 during the build-up phase. Steady state concentration conditions are then seen to be reached after 12, 15, and 17 days at different locations from the waste disposal site. In a similar manner, steady state concentration times may be determined for groundwater motion in the X-Y plane. The effect of anisotropic groundwater flow on the rise time is shown in Figure 8.10 for a variety of velocity ratios. The contaminant concentration observed at a distance of $X = 4L$ is plotted as a function of time for each considered velocity ratio V_x/V_z. It is clearly seen that as the velocity ratio increases, the steady-state condition is reached subsequent to considerably more rapid rise-times and the apparent steady-state concentration is dramatically increased. Figure 8.11 shows the steady state concentration profiles for the horizontal X-Y domain. The salient features of the iso-concentration curves of Figure 8.11 are: (a) Clearly, contaminant concentration decreases in a quasi-exponential manner as the distance from the source increases. (b) The contaminant is transported to shorter distances in the lateral (Y-axis) direction than in the downstream direction. (c) Distinct, albeit limited, contaminant transport is also apparent in the upstream direction, the cause of which may be the consequence of molecular diffusion acting opposite to the forcing velocity. Similarly, Figure 8.12 displays the steady state iso-concentration profile for the Z-X domain. The salient features of the groundwater motion in the vertical domain are obvious counterparts of the features characterizing the groundwater motion in the horizontal domain, including the apparent molecular diffusion in the "back flow" negative X-direction. The result of contaminant motion in the X-Y plain is shown in Figure 8.13. Herein are plotted the calculated concentrations as a function of distance along the X-axis for four distances along the Y-axis. The

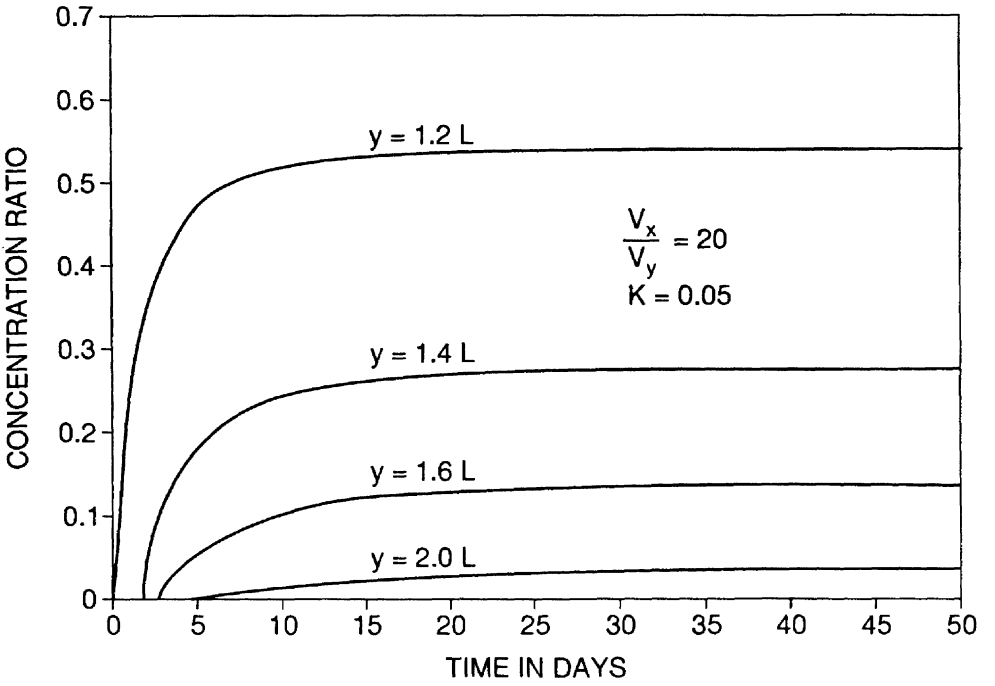


Fig. 8.9. Contaminant build-up stage at various stations along the water table (positive y-axis) (after Bobba and Bukata, 1980).

peak of contaminant concentration is seen to have moved both laterally and downstream from the source, asymptotically reaching zero concentration at distances far removed from the waste disposal site (recharge area).

Steady state concentration profiles for vertical sections at successively increasing downstream distances from the source are shown in Figure 8.14. Clearly evident is the reduction in peak concentration levels combined with penetration to shallower depths beneath the water table. Once again the contaminant concentration peak is seen to move along the resultant direction of groundwater flow.

8.12.2 Effect of Adsorption on Contaminant Concentration

The effect of contaminant reaction which results in the removal of contaminant from the fluid to the solid phase is considered to be the retarding effect due to the adsorption of contaminant on the porous medium. Mathematically, therefore, the porous medium acts as a sink. Figure 8.15 illustrates the steady state concentration profiles along the X-axis for different values of the adsorption coefficient K. It is evident that (a) with increasing values of K, the peak contaminant concentration moves closer to the source; (b) with increasing values of K, less contamination is

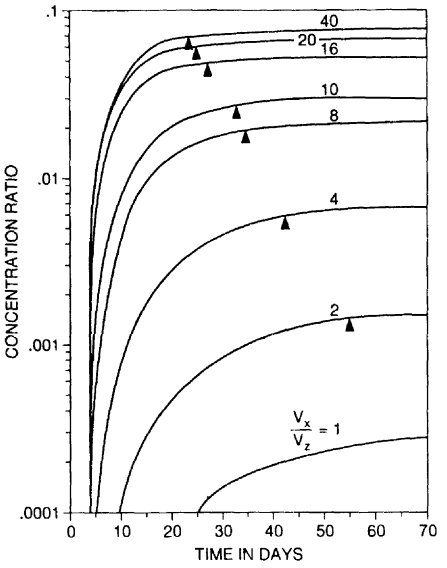


Fig. 8.10. The effect of principal velocity ratios on the build-up stage (after Bobba and Bukata, 1980).

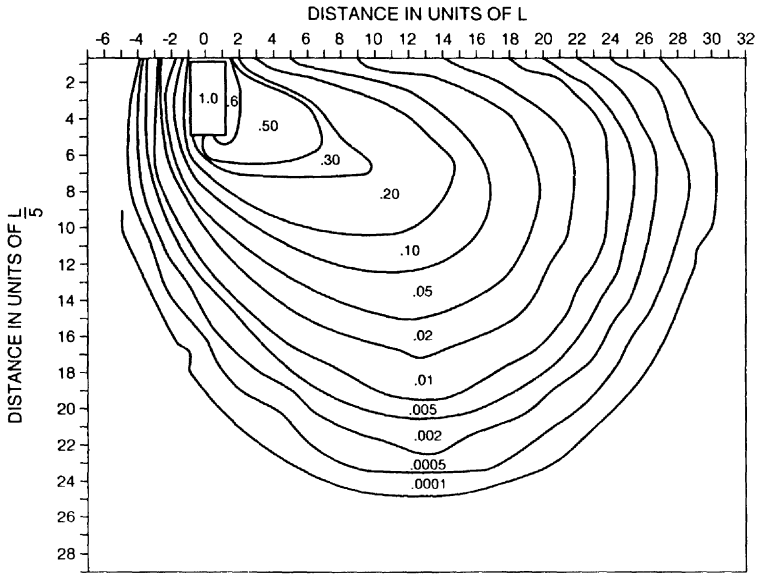


Fig. 8.11. Steady state contaminant isoconcentration lines in the horizontal (x-y) domain (after Bobba and Bukata, 1980).

apparent at any point along the water table, including backflow concentrations upstream from the source; and (c) contaminant concentration becomes an increasingly

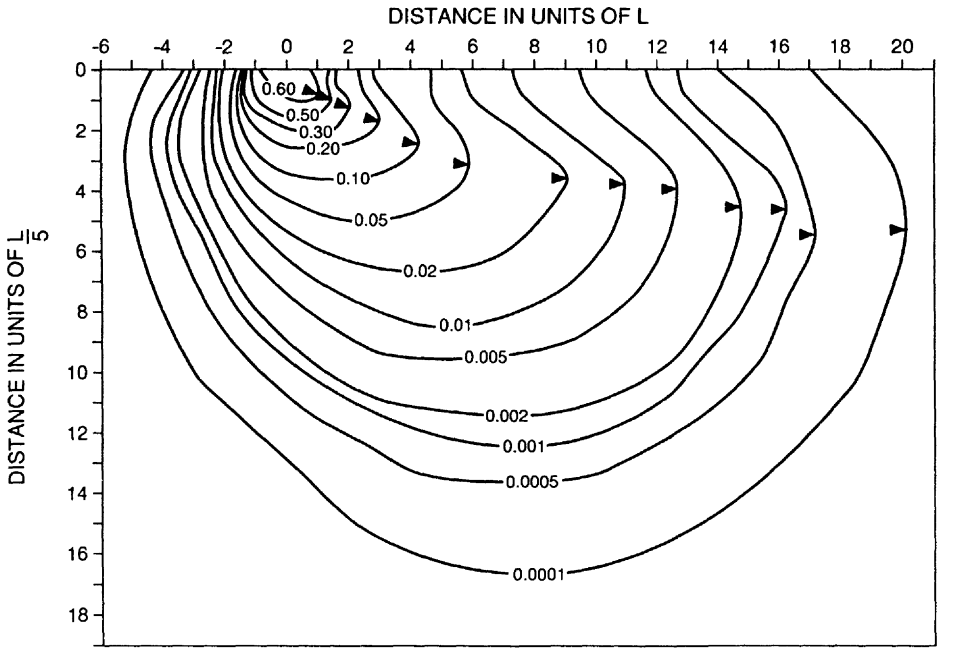


Fig. 8.12. Steady state contaminant isoconcentration lines in the vertical (z-x) domain (after Bobba and Bukata, 1980).

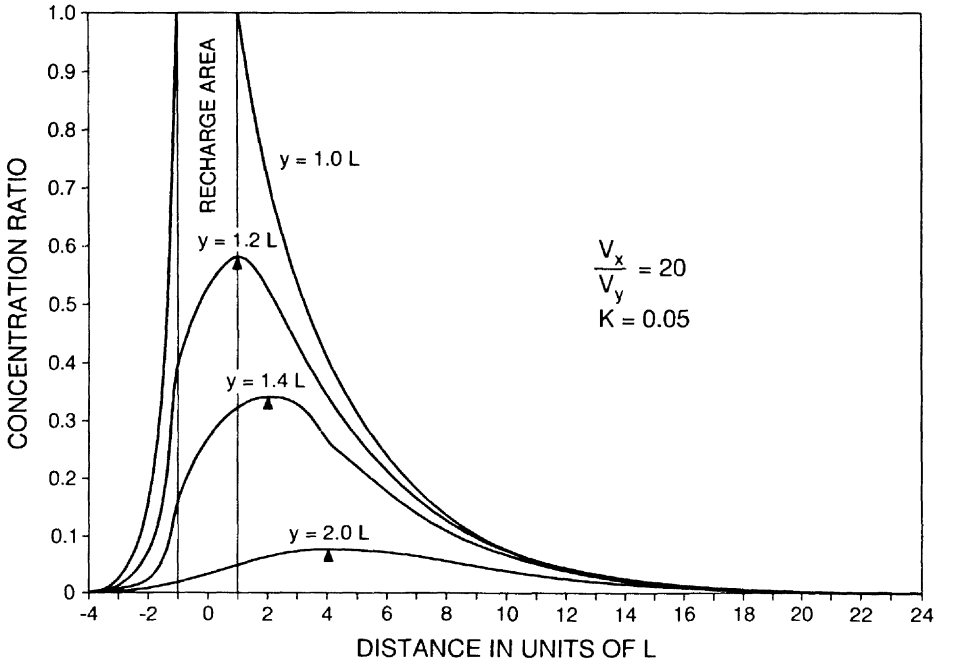


Fig. 8.13. Contaminant concentration profiles along the x-axis for various positions along the y-axis (after Bobba and Bukata, 1980).

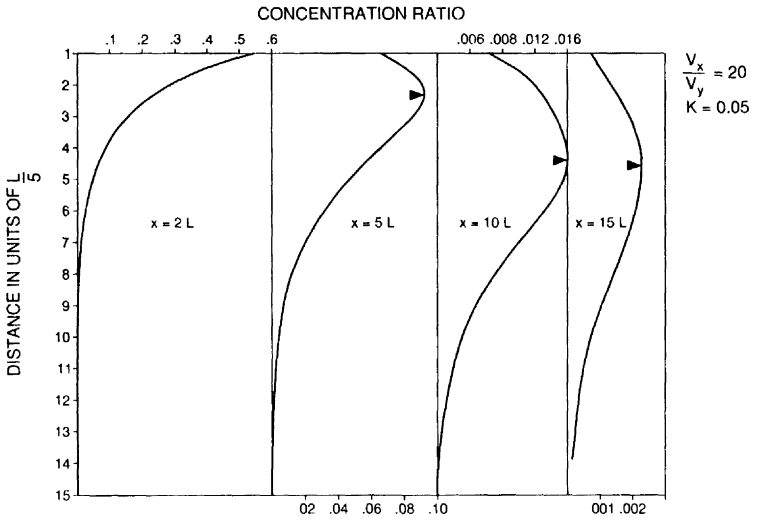


Fig. 8.14. Contaminant concentration profiles along the z-axis for various positions along the x-axis (after Bobba and Bukata, 1980).

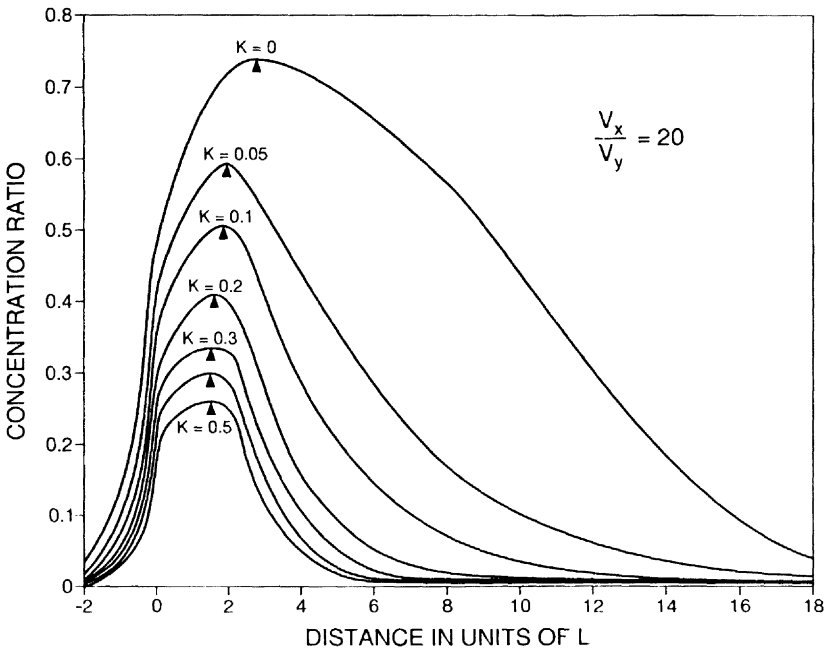


Fig. 8.15. Contaminant concentration profiles along the x-axis for various values of the adsorption coefficient K (after Bobba and Bukata, 1980).

more localized phenomenon with increasing K. Similarly, Figure 8.16 illustrates the effect of adsorption on the vertical propagation of injected contaminant. Once

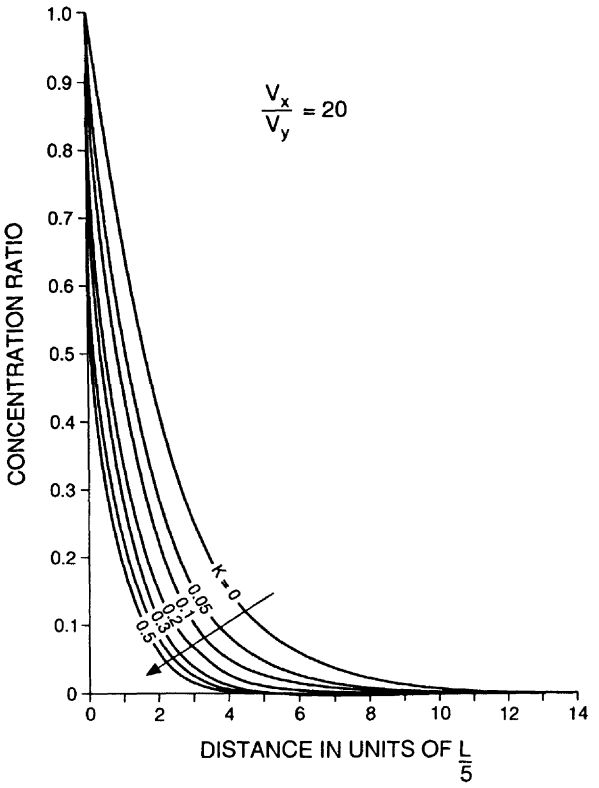


Fig. 8.16. Contaminant concentration profiles along the z-axis for various values of the adsorption coefficient K (after Bobba and Bukata, 1980).

again, it is seen that an increase in K results in decreasing concentration becoming more localized at shallower depths.

8.12.3 Effect of Velocity Ratios on Concentration

The influence of velocity ratios on contaminant transport along the principal flow direction is shown in Figures 8.17 and 8.18. In the vertical section, increasing the velocity ratio results in an obvious, more localized concentration of contaminant becoming apparent at shallower depths. In the horizontal domain, it is seen that an isotropic aquifer displays a Poisson type concentration distribution which becomes increasingly more skewed in the principal axis direction as the aquifer displays a higher order anisotropy resulting in a very gradual migration of reduced peak concentration to greater distances downstream from the source. It is evident, therefore, that an essentially isotropic aquifer will transport contaminant concentration to deeper depths, while a sharply anisotropic aquifer will transport contaminant concentrations to much greater distances downstream. The duo two dimensional contaminant transport dependence upon aquifer anisotropy is shown in Figure 8.19. Herein are illustrated the 0.5 isoconcentration curves determined

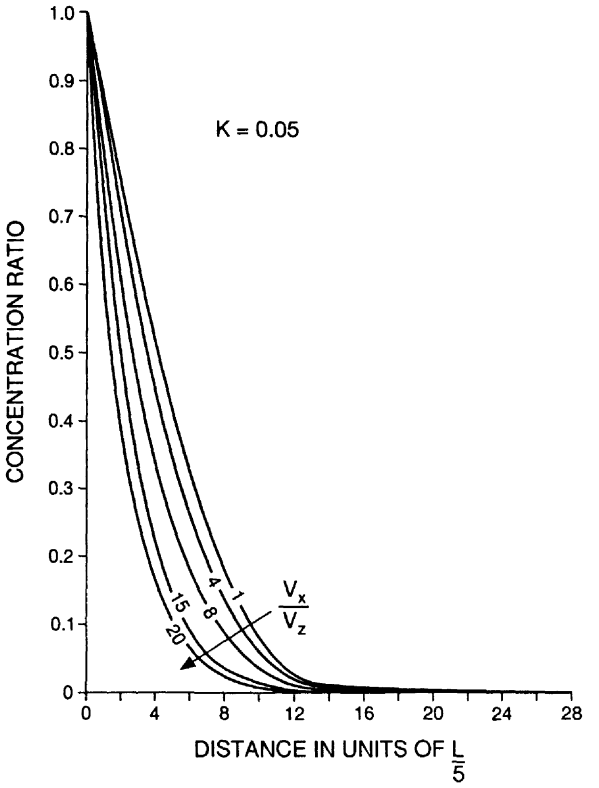


Fig. 8.17. Contaminant concentration profiles along the x-axis for various principal velocity ratios (after Bobba and Bukata, 1980).

for several different velocity ratios. The extent of contaminant influence is seen to move progressively shallower depths and larger downstream distances as the velocity ratio increases.

8.12.4 Effect of Dispersivity on Contaminant Transport

The hydrodynamic dispersion coefficient D may be expressed through Figures 8.20 and 8.21 which show the build up of contaminant concentration in groundwater system for two dissimilar velocity ratios. It is seen that the higher velocity ratio (Figure 8.20) displays an independence of rise times with varying diffusivity ratios, while a distinct dependence upon diffusivity is displayed by the steady state concentration values achieved by near isotropic subsurface systems. The steady contaminant concentrations of highly anisotropic porous media are considerably less dependent upon dispersivity ratios. This clearly indicates that groundwater flow through highly anisotropic subsurface systems are largely under control of advective transport processes, while groundwater flow through subsurface systems approaching isotropy are largely dominated by hydrodynamic dispersive transport processes.

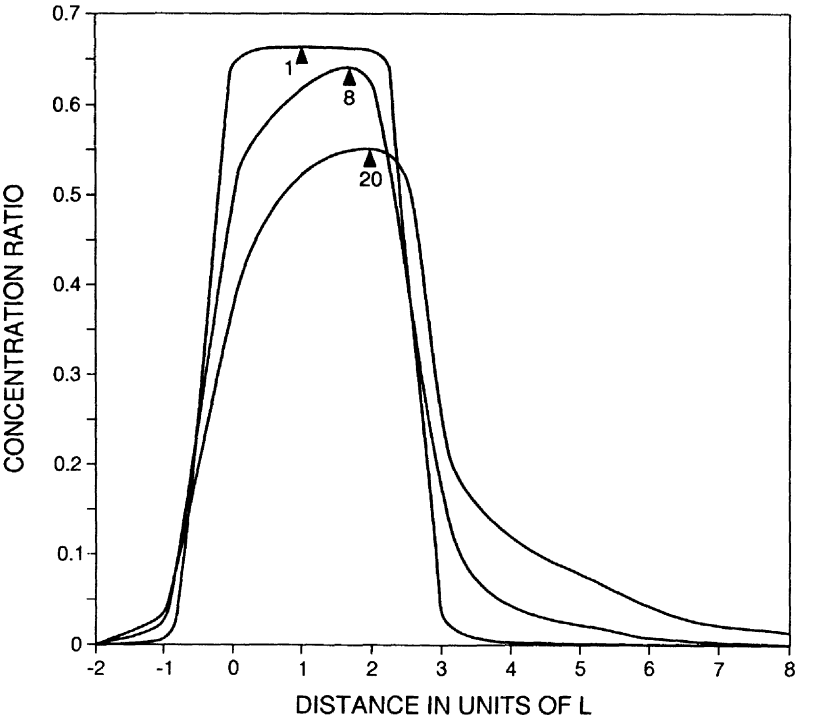


Fig. 8.18. Contaminant concentration profiles along the x-axis for various principal velocity ratios (after Bobba and Bukata, 1980).

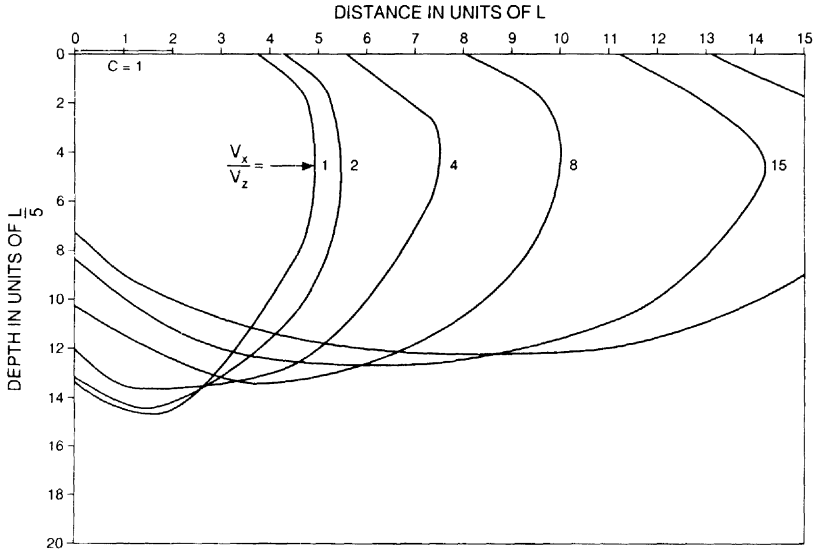


Fig. 8.19. A representative contaminant isoconcentration contours in the z-x domain for various velocity ratios (after Bobba and Bukata, 1980).

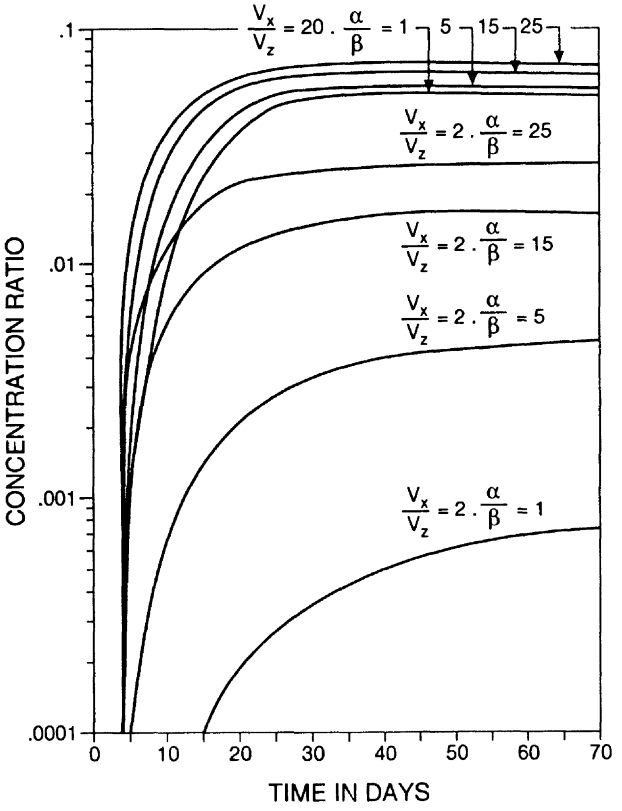


Fig. 8.20. Contaminant build-up stages of a near isotropic ($V_x/V_z = 2$) and a highly anisotropic ($V_x/V_z = 20$) aquifer for various dispersivity ratios $\alpha/beta$ (after Bobba and Bukata, 1980).

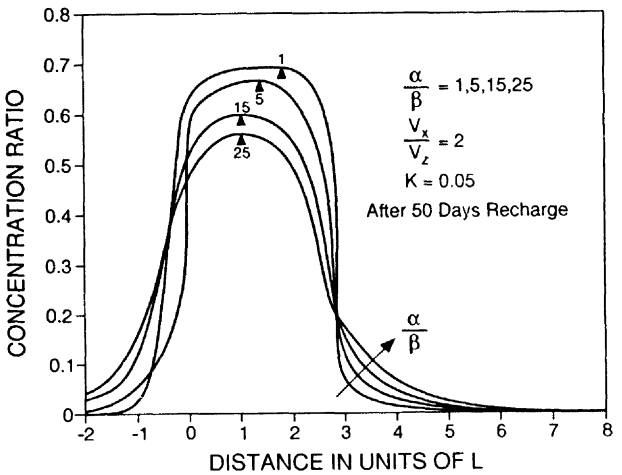


Fig. 8.21. Contaminant concentration profiles along the x-axis for various dispersivity ratios, near isotropic aquifer (after Bobba and Bukata, 1980).

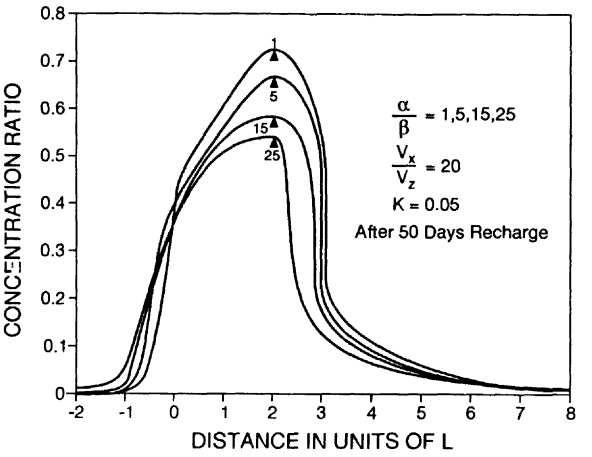


Fig. 8.22. Contaminant concentration profiles along the x-axis for various dispersivity ratios, highly anisotropic aquifer (after Bobba and Bukata, 1980).

The spatial distribution of steady state contaminant along the water table is illustrated in Figures 8.21 and 8.22. Figure 8.21 considers a near isotropic aquifer defined by a low velocity ratio, and Figure 8.22 considers an anisotropic aquifer defined by a high velocity ratio. For high velocity ratios it is seen that (a) the peak contaminant concentration decreases with increasing dispersivity ratio, and (b) the physical location of the peak contaminant concentration along the water table is essentially invariant. For low velocity ratios it is seen that (a) the peak contaminant concentration also decreases with increasing dispersivity ratio, but (b) the physical location of the peak contaminant concentration along the water table migrates farther downstream from the source as the dispersivity ratio decreases. As perhaps expected, a considerably more pronounced skewness is apparent in the X-axis contaminant distribution associated with anisotropic transport.

The spatial distribution of steady state contaminant along the Z-axis is illustrated in Figures 8.23 and 8.24. Again, the anisotropic condition is depicted in Figure 8.24 and near isotropic condition is depicted in Figure 8.23. Very slight differences are noticed between the anisotropic and isotropic situations. In both cases increasing the dispersivity ratios results in a more localized contaminant concentration being transported to shallower depths, although the effect of dispersion is slightly more significant for lower velocity ratios.

8.12.5 Contaminant Recovery Phase

Figures 8.25 and 8.26 display the calculated recovery of the aquifer zone in the horizontal and vertical directions, respectively, subsequent to the termination of 10 days of continuous source injection. From the concentration profiles depicted in Figure 8.25, it is seen that the recovery pattern along the downstream direction is one of an exponential decrease in peak concentration value coupled with an exponential migration of this peak concentration downstream.

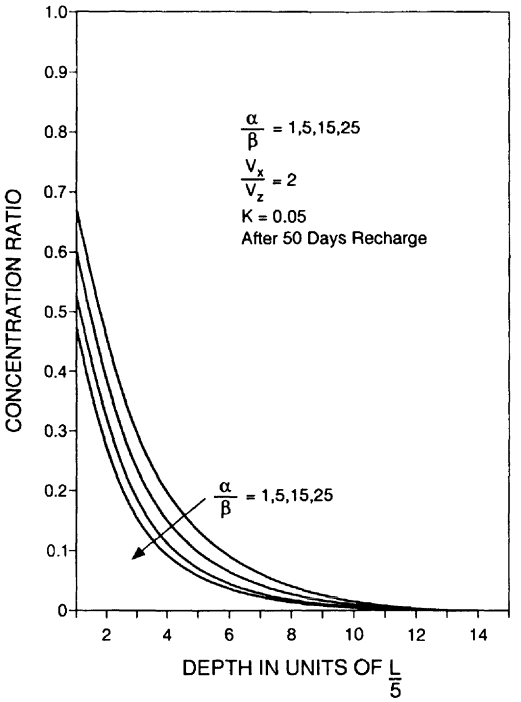


Fig. 8.23. Contaminant concentration profiles along the z-axis for various dispersivity ratios, near-isotropic aquifer (after Bobba and Bukata, 1980).

The recovery pattern is considerably more complex in the vertical domain; Figure 8.26 depicts the vertical concentration for stations located at distances $3L$, $6L$ and $10L$ downstream from the source. The profiles have been determined for several days following termination of source injection at day 10. The dotted lines in Figure 8.26 define the loci of the location in depth of the peak contaminant concentration. Close to the source a fairly rapid exponential collapse to pre-injection conditions is observed as the contaminant is transported downwards. For locations further downstream from the source the peak concentration continues to increase in magnitude and remain essentially at the steady state depth before a downward propagation of contaminant accompanies an exponential evaluation of contaminant from the area. Such decay phase phenomena are even more evident at locations substantially removed from the source where clearly the presence of contaminant is readily apparent for more than two months subsequent to the termination of continuous injection.

8.13 Application of Groundwater flow and Contaminant Transport Models

8.13.1 Prot Granby Radioactive Disposal Site, Ontario, Canada

Canada has only one uranium refinery at Port Hope, Ontario (Figure 8.27), where uranium is separated from the yellow cake to leave a residue with a radium to

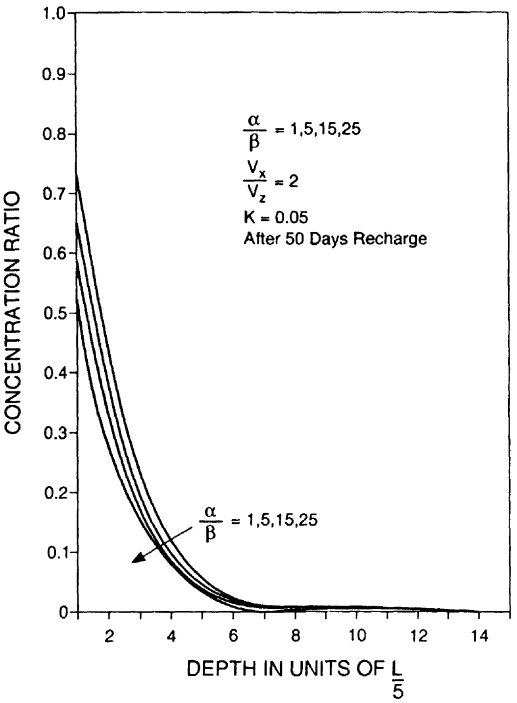


Fig. 8.24. Contaminant concentration profiles along the z-axis for various dispersivity ratios, highly anisotropic aquifer (after Bobba and Bukata, 1980).

uranium ratio near that of the original ore. The waste was disposed by Eldorado Resource Limited at its Port Granby waste management site located about 13 km west of Port Hope, on the north shore of Lake Ontario. The objective of the study was to predict the radioactive contamination plume to Lake Ontario by a finite element model. The details of this study were presented earlier by Bobba and Joshi (1988, 1989). The site covers a 10-ha area sloping down from the north toward Lake Ontario. The land surface along the shore rises abruptly as bluffs above the lake level which stands at an elevation of approximately 350 feet (105 m) above the sea level. The dominant topographic features on the site are two deep valleys which are formed due to erosion from surface runoff and groundwater (Figure 8.28). The 30-year (1956–86) average annual precipitation in the area is 868 mm and is made up of 704 mm of rainfall and 164 mm of snow. The average annual temperature is about 7 °C, and the mean annual potential evapotranspiration is about 504 mm.

The general geologic cross section of the site is similar to that shown in Figure 8.29 and is summarized in Table 8.5. The four major stratigraphic units overlay the limestone bedrock which occurs at a depth of about 140–145 feet (42–43.5 m) below the surface.

TABLE 8.5
Hydrogeologic properties of the Port Granby radioactive disposal site

Stratigraphic unit	Thickness in feet (m)	Lithographic description	Simplified geologic description	Water bearing characteristics	Hydraulic conductivity (cm/s)
Fill		Waste material		Contains radio active waste	10^{-3}–10^{-4}
Upper glacial till	50 (15)	Very dense brown to grey brown sand, silt	Upper sand and clayey silt strata	Low conductivity and little water	10^{-2}–10^{-4}
Middle glacial till	6-20 (2-6)	Mixture of boulders and pebbles in a matrix of sand, silt and clay.	Clayey to sandy till	Water occurs primarily in thin sand lenses in till	8×10^{-8}
Clarke sand deposits	30-35 (9-11)	Very dense, grey sandy silt, silty sand or sand	Lower sand and sandy silts	Most permeable deposits at the site	3×10^{-3}
Lower glacial till	26-30 (8-9)	Mixture of boulders and pebbles in a matrix of sand, silt and clay	Clayey silts and lower clayey till	Water occurs principally in the sand lenses in till	10^{-7}
Bedrock	Limestone	Bedrock	Water occurs along the fractures on the top portion		

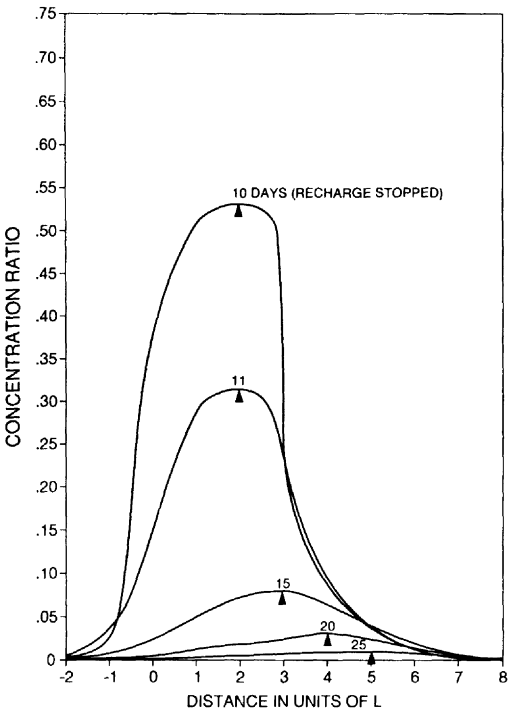


Fig. 8.25. Aquifer recovery stage contaminant concentration profile along the x-direction (after Bobba and Bukata, 1980).

8.13.1.1 Geology and Hydrogeology of Site

The groundwater flow at the site is generally influenced by the geology and topography of the site. Two permeable layers exist in the overburden deposits, separated by relatively semipermeable deposits that contain sand and gravel layers with microfractures. The hydraulic gradients are predominantly downward and springs are formed as small streams along the bluffs and the water flows to the lower deposits before reaching the beach. The water table of the Clarke sandy aquifer and sandy silty aquifer is shown in Figure 8.29. The bedrock hydraulic heads are two or three feet higher than the Lake Ontario beach deposits so that the water from piezometers continuously flow above the lake level. The hydraulic heads were not observed to fluctuate with wave action. In general, the groundwater flows downward and outward from the waste disposal site. The hydrogeological parameters used in the numerical computation to delineate the groundwater flow from the site to Lake Ontario are given in Table 8.5.

8.13.1.2 Numerical Modelling Results

A finite element groundwater contaminant model was applied to a waste disposal site. Figure 8.30 shows the computed hydraulic head distribution of the north south geological cross section. Both the computed and observed data showed that hy-

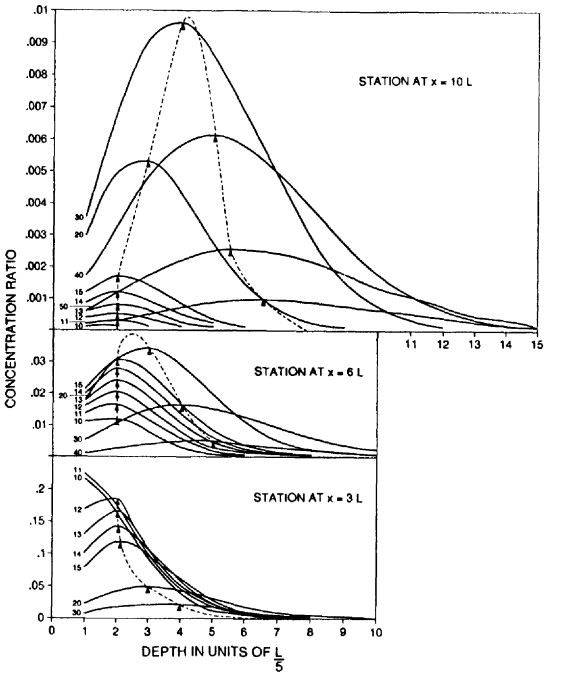


Fig. 8.26. Aquifer recovery stage contaminant concentration profiles along the x-direction for three locations on the x-axis (after Bobba and Bukata, 1980).

draulic gradients were predominantly downward. Very high hydraulic heads were observed above the middle till but due to semipermeable layer, the groundwater flows from the upper sandy silt deposits to lower sand aquifer. The variation in head difference within the upper till unit is probably due to the heterogeneous nature of the deposit. The actual pattern and velocity of groundwater flow in the heterogeneous till deposit beneath the silt will be dominated by the hydraulic head difference across the flow path and the hydraulic conductivity of the deposits within the upper till unit. Since the three-dimensional head distribution is predominantly downward, the overall groundwater flow path in the till beneath the disposal site will be primarily downward. Due to regional groundwater flow, the contaminated water flows towards Lake Ontario.

On the beach, the computed hydraulic heads of bedrock nodes are about three feet (1 meter) higher than Lake Ontario water level as is observed in the field. The groundwater discharge from the waste disposal site between the east and west creeks is computed to be 2.5×10^5 m^3 year^{-1} (Platford et al., 1984). The Darcy velocities were computed from the computed hydraulic heads and then used in the contaminant transport model to predict the ^{226}Ra concentrations at these locations before reaching Lake Ontario. The transport parameters α and β, which were computed by deterministic and optimization methods were used in the contaminant transport model (Bobba and Joshi, 1989). The computations show

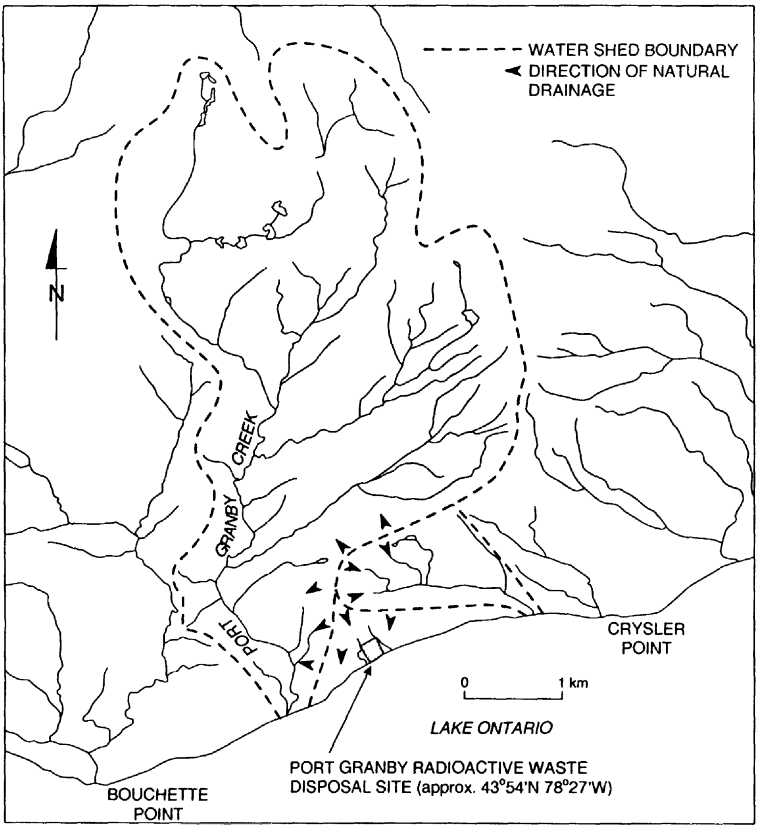


Fig. 8.27. General location of the Port Granby radioactive waste management site (after Bobba and Joshi, 1988).

that after 15 years the contaminant will reach the sandy aquifer and move towards Lake Ontario under the influence of regional groundwater flow system.

The computed ^{226}Ra concentrations were found to be in good agreement with those actually observed in the field. Assuming that the applicable site parameters remain unchanged, the ^{226}Ra concentrations were computed for the several future years. Figure 8.31 shows the computed ^{226}Ra concentrations for the north-south cross-section at 30 years. The observed pattern suggests consistent migration of ^{226}Ra to Lake Ontario over the projected time period. The previous migration of the predicted plume to Lake Ontario waters is discernible from the 1977 surface water ^{226}Ra measurements (Durham and Joshi, 1980) where the plume (Figure 8.32) is observed to transverse along the coastal zone (Platford et al., 1984). If the ^{226}Ra activity in the beach piezometers averages about 2.7 mCi m^{-3}, about 0.7 Mci $year^{-1}$ is transported into Lake Ontario. Thus, it appears that, barring unforeseeable circumstances, the long-lived radionuclide ^{226}Ra will continue its slow migration toward Lake Ontario in small amounts.

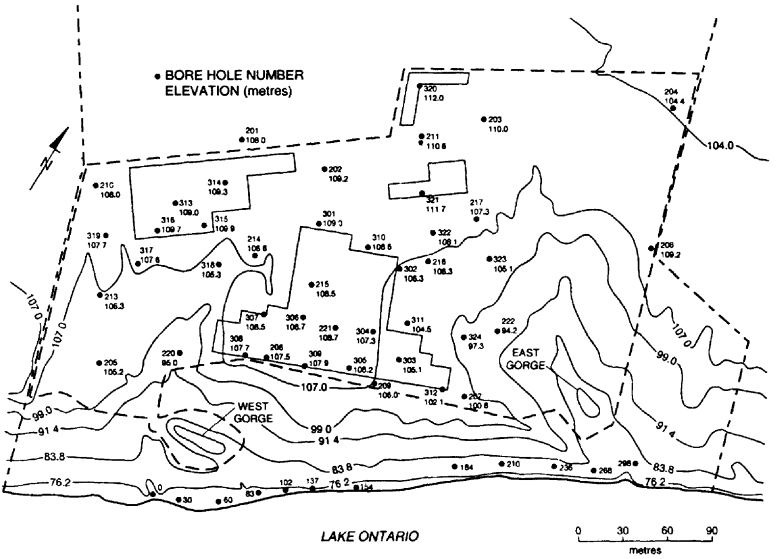


Fig. 8.28. Detailed borehole locations and corresponding elevations at the study site (after Bobba and Joshi, 1988).

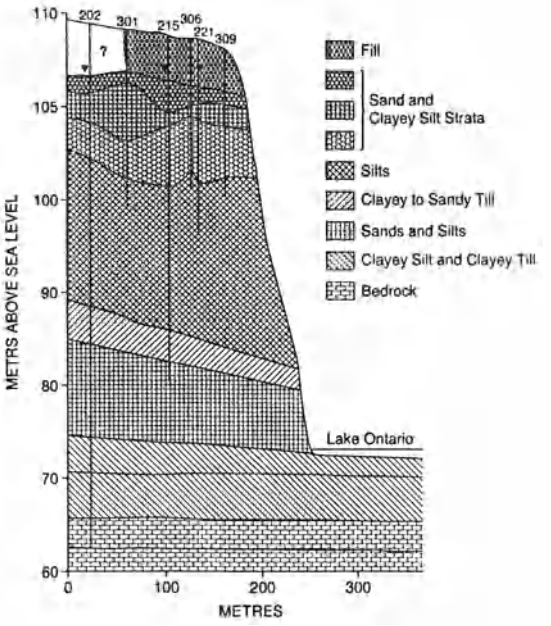


Fig. 8.29. General geological cross-section of the Port Granby site (after Bobba and Joshi, 1988).

Relocation of the waste from the Port Granby site is one of the remedial actions being considered by Eldorado Resources Limited. To assess the effectiveness of this remedial measure, the model was used to predict the recovery of ^{226}Ra

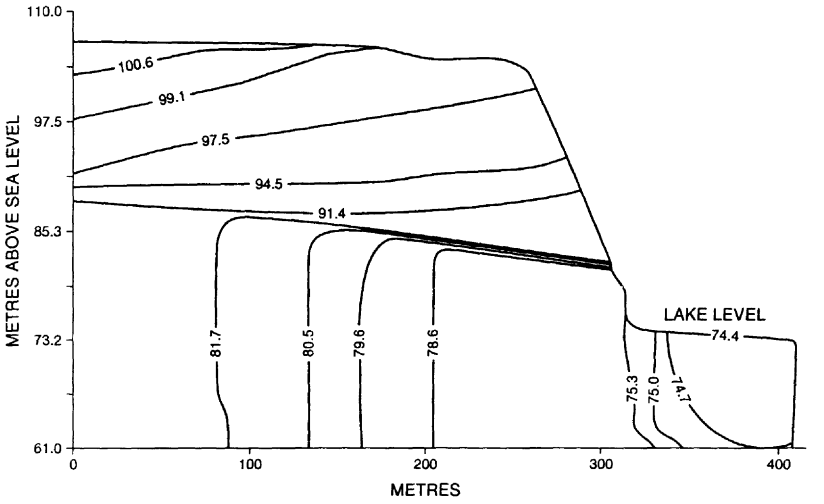


Fig. 8.30. Computed hydraulic head distribution of the North-South geological cross section (after Bobba and Joshi, 1988).

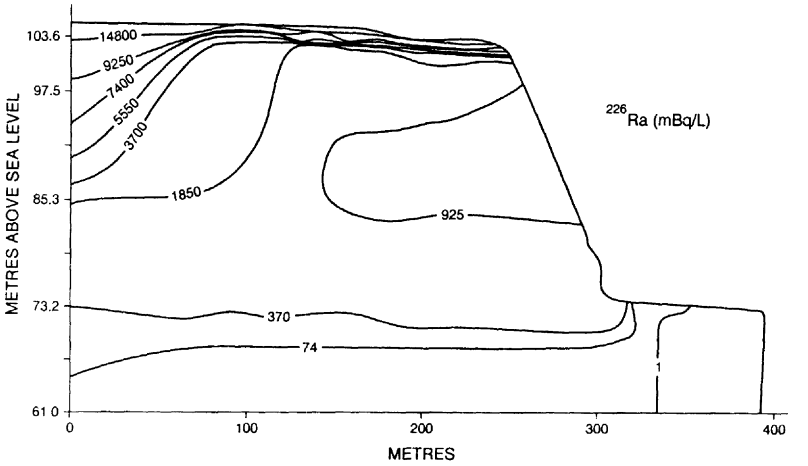


Fig. 8.31. Computed concentrations of ^{226}Ra and Uranium (after Bobba and Joshi, 1988).

concentrations at the site and in beach deposits. As expected, the contaminant concentrations decrease more rapidly near the waste disposal site when compared with the beach locations (Figure 8.33). Lower till and bedrock aquifer show higher contaminant build up presumably due to molecular diffusion.

8.13.2 Application of SUTRA model to Lambton County, Ontario, Canada

SUTRA groundwater contamination model was developed by Voss (1984). It is a finite element model. The objective of this work is to describe the application

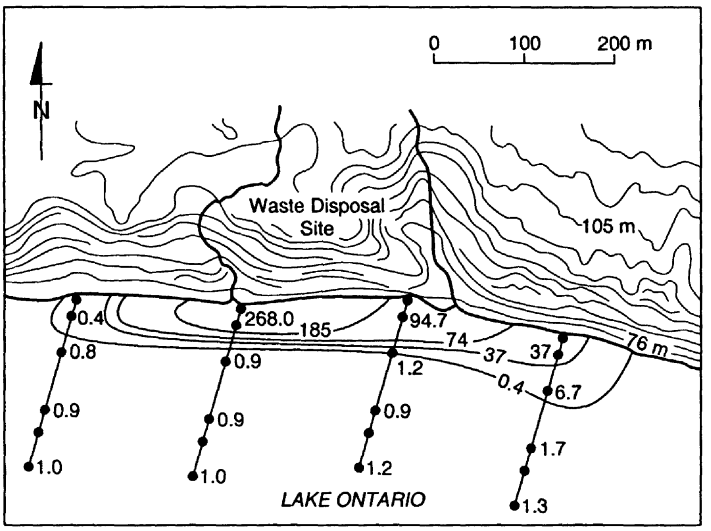


Fig. 8.32. Observed migration of ^{226}Ra (mBq/L) in Lake ontario coastal waters. (after Bobba and Joshi, 1988)

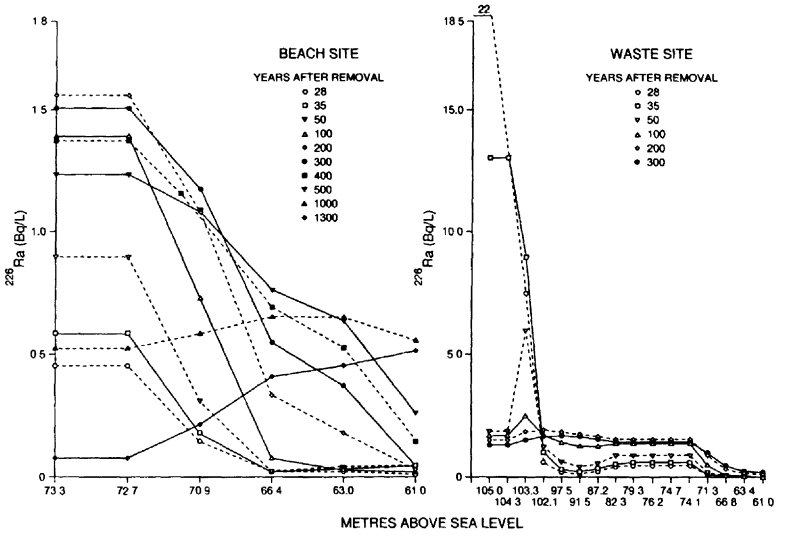


Fig. 8.33. Computed vertical domain concentrations of ^{226}Ra below the site and in beach deposits following removal of waste (after Bobba and Joshi, 1988).

of the SUTRA groundwater flow model to help reduce uncertainty in the conceptual model of groundwater flow patterns for Lambton County, Ontario, Canada. The simulation is based on the hydrogeologic framework and regional geology described in the literature (INTERA, 1989). Site specific data on geology, groundwater levels and hydrogeologic properties of the overburden at Sarnia area were

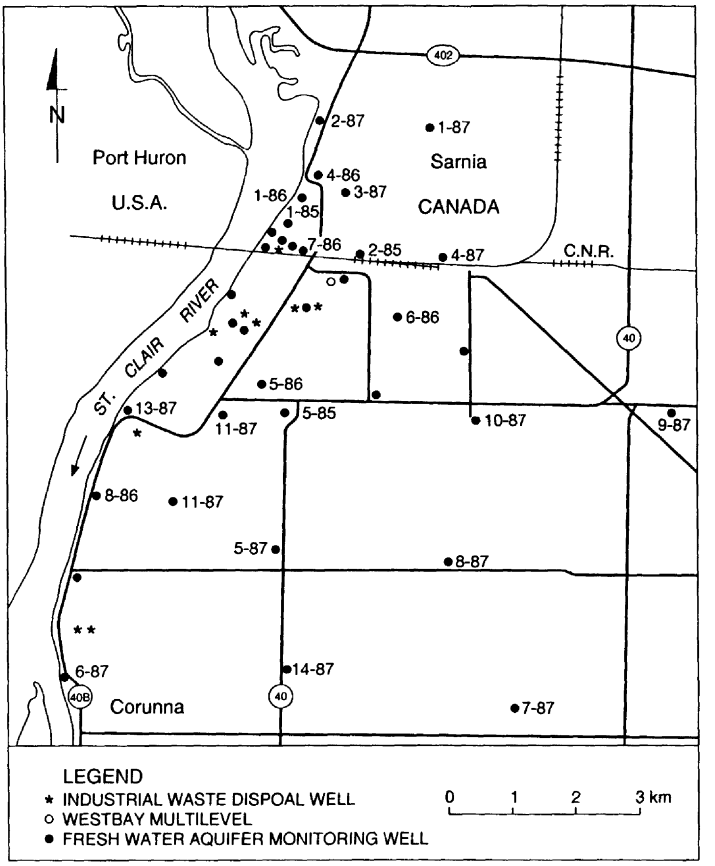


Fig. 8.34. Location map of Lambton County (after Bobba, 1993).

obtained from a drilling program and monitoring well observations carried out by INTERA during 1985–1988.

The capability to model the area consisted of the construction of a numerical model that would accurately reproduce field observations as embodied in a conceptual model of the area. A statistical approach was used to identify areas of uncertainty and incongruity within the conceptual model of the area. Once the conceptual model was understood well enough and could be modeled numerically, then the model could be used to derive conclusions regarding the hydrogeologic suitability of the area. The groundwater flow model can also be used to integrate and interpret information from stratigraphy, hydrogeology, and geochemical investigations, and to develop a groundwater flow balance.

8.13.2.1 Geology and Hydrogeology

The study area (Figure 8.34) is located over the Algonquin Arch which separates the Applanchian and Michigan basins. The sedimentary rocks at this location are about 1370 m thick and include Cambrian, Ordovician, Silurian and Devonian

PERIOD	GROUP	FORMATIONS	LITHOLOGY
Quaternary			clay, sand, gravel
Devonian		Port Lambton	sasndstone -shale
		Kettle Point	shale
	Hamilton	Ipperwash	alternating limestone
		Widder	and shale
		Hungary Hollow	
		Arkona	
		Backport	
		Bell	
		Marcellus	shale
		Dundee	dolostone
	Detroit River	Lucas	dolostone
		Amherstburg	limestone
		Oriskany	sandstone
Silurian		Bertie	dolostone
		Salina	shale-anhyd.-carbonate
	Lockport	Guelph	dolostone
		Eramosa	
		Goat Island	
		Gasport	
	Clinton	Rochester	shale-sandstone
		Irondequoit	and limestone
		Reynales	
		Thorold	
	Cataract	Grimsby	shale-dolostone
		Cabot Head	dolostone
		Manitoulin	dolostone
		Whirlpool	sandstone
Ordovician	Nottawasaga	Queenston	silty shale
		Whitby (Collingwood & Dundas-Meaford)	shale-limestone and siltstone
	Black River and Trenton		carbonate
Cambrian	Basal		sandstone-dolostone
Precambrian Basement			

Fig. 8.35. Regional geological section of southern Ontario (after Bobba,1993).

strata which dip gently toward the Michigan Basin at a gradient of 5 to 8 m/km. The regional geology of Southwestern Ontario is shown in Figure 8.35.

The area is overlain by glacial drift which is generally about 30 to 35 m thick in the Sarnia area and consists of silty and clayey tills, possible intermediate glaciolacustrine sediments and a thin discontinuous silty sand layer of variable texture and thickness. The drift is up to 68 m thick above a bedrock valley.

The Kettle Point formation forms the bedrock beneath most of Lambton county where it is 15 to 30 m thick and consists of black bituminous shales. Its upper surface is relatively flat and slopes gently to the north and northwest. The Kettle Point formation is absent where eroded to form the bedrock valley.

The Hamilton Group of formations consists of interbedded shales and limestones. It is about 90 m thick in the vicinity of the study area and forms the bedrock where the Kettle Point formation is absent.

The Dundee Formation is a fine-grained limestone with occasional chert. Oil and gas are produced from the Dundee Formation at many locations in Lambton County and ranges in thickness between 30 and 38 m in the vicinity of the study area.

The Lucas Formation comprises the upper portion of the Detroit River Group of formations and is the interval in which liquid industrial wastes were injected. It

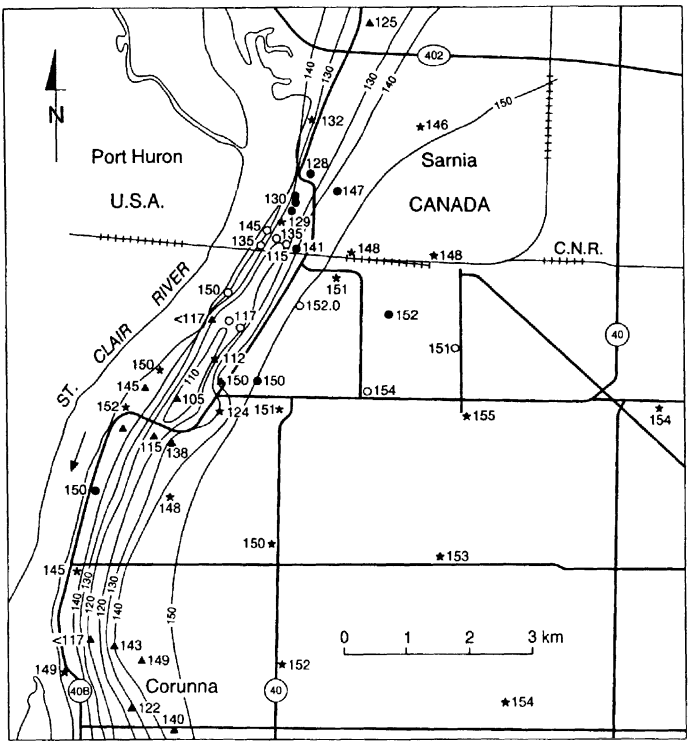


Fig. 8.36. Buried valley of Sarnia in Lambton County (after Bobba, 1993).

is fine-textured, crystalline limestone and dolomite with locally occurring, highly permeable zones farther from the St. Clair River. Thickness ranges between 60 and 75 m and structural contours on its upper surface indicate a southwesterly dip.

The Amherstburg Formation comprises the lower portion of the Detroit River Group. It is about 90 m thick and consists of dolomites and limestones.

The Bois Blanc Formation consists of interbedded limestone and dolomites with some chert. It is about 45 m thick in Lambton County. The Bass Islands Formation consists of dense dolomite and is about 110 m thick.

The Salina formation is about 400 m thick and is comprised of halite, anhydride, limestone, dolomite and shale. It has been subjected locally to dissolution over geologic time with resulting collapse and faulting in overlying formations.

In Lambton County, potable groundwater is usually obtained at depths of about 30 m from thin discontinuous deposits which constitute the Fresh Water Aquifer. Dug wells are used where surficial sand deposits are present. The groundwater is obtained from sand and gravel deposits at depths below 60 m. These occurrences are located along the probable trend of the bedrock valley and are an indication that the Fresh Water Aquifer in the valley is a source of water supply.

The Fresh Water Aquifer occurs at or near the bedrock surface beneath most of the area. Groundwater in the Fresh Water Aquifer is confined by overlying clayey till deposits. The basal unit occurs directly on the bedrock. The bedrock valley

was observed to have a deep channel cut into the bedrock and trends roughly north-south parallel to the St.Clair River (Figure 8.36). The valley appears to have relatively steep sides especially on the east.

The Sarnia area is covered by approximately 30 to 68 m of glacial drift. The drift at the ESSO site consists of one or more tills which range in texture from clayey silt to silty clay with occasional pebbles, cobbles, and a basal silty sandy deposit which contains varying amounts of pebbles, cobbles and boulders. The basal deposit is present as a thin, discontinuous veneer on the bedrock surface at most locations but is thicker in the bedrock valley. The drift thickness is about 33 m beyond the valley and about 68 m in the central part of the valley. It may be an outwash deposit with interlayers of sand and clay. In the bedrock valley it consists mainly of silt and clayey silt with sandy zones. Its upper contact is not clearly defined and it appears gradational with the overlying till. The aquifer portion of the deposit occurs in the coarser texture which is usually located near the base of the unit close to the bedrock surface.

For Kettle Point Formation, the Hamilton Group of formations and the Dundee formation, little information is available respecting the hydrogeology of formations between the Fresh Water Aquifer and the Lucas Formation. Permeable zones exist in some limestone/dolostone zones within the Dundee and Hamilton formations. Lateral gradients within these permeable zones are expected to be westward and vertical gradients are expected to be downward. Vertical gradients were reversed locally by pressurization resulting in deep well injection, but pressures have declined and downward gradients have been reestablished.

McLean (1968) notes that water is commonly encountered in the Lucas Formation. Vandenburg et al. (1977) analyzed water level data and produced a map showing persistent potentiometric features of several permeable zones within the Lucas Formation. West and north-west gradients exist but the authors conclude that the data are not adequate to infer groundwater flow directions on a local scale. Permeability conditions vary widely in the Lucas Formation.

Natural water levels and flow systems in the Detroit River Group and other deep formations in Lambton County have been disturbed by injection of waters, oil and gas production, re-injection of oil field waste waters, and the interconnection of previously isolated formations by improperly completed and abandoned wells. Overall, flow rates are slow because natural gradients are low. Flow is focused along zones of higher permeability.

8.13.2.2 Meteorology

The average temperatures found in Southwestern Ontario are regulated, to a certain extent, by the proximity of Lake Huron, Lake St. Clair and Lake Erie. The presence of Lake Huron and Lake Erie thus tend to warm the study area by 1–2 °C. The average daily maximum and daily minimum temperature of the area is 15 °C and −2 °C, respectively. The mean precipitation across Southern Ontario ranges from about 66 to 100 cm. This variation is closely related to the slope and elevation of the land location relative to the Great Lakes and prevailing wind direction.

8.13.2.3 Groundwater Flow Systems in Lambton County

The local scale groundwater flow system consists of the water table associated flow in the shallow subsurface. It is contained in the surficial sand and gravels and the weathered and fractured zone of the lacustrine clays and clay tills. Flow in this system is likely be strongly influenced by precipitation events and results in either discharge to shallow ditches, creeks and rivers or recharge to the deeper formations.

The thickness of the weathered zone is typically less than 6 m. Because of the fractured and weathered nature of the shallow clays, this zone is likely to be more active than the deeper intact clays. The presence of fractures in the weathered zone produce an increased permeability and relatively open conduits for the transmission of water. The hydraulic conductivities for the local groundwater flow system range from about 10^{-11} to 10^{-6} m/s with a mean of about 10^{-9} m/s (Lesage et al., 1991; Ruland et al., 1991). This wide range of values represents the weathered and fractured nature, as well as the geologic variability, of the local systems. Movement of groundwater and pollutants, if present, would be more rapid through the fissure networks than through similar unweathered clays.

The direction of groundwater flow in the local system will be dependent on the specific location but, in general, it is expected that this system will have both a horizontal and a vertical component of flow. The horizontal component will be the result of increased permeability of the weathered zone overlying the clay causing increased lateral groundwater flow. In the shallow system, groundwater will flow primarily towards the surface water bodies (i.e., rivers, creeks, etc.).

The intermediate scale of groundwater flow system in this area includes the thick sequences of intact lacustrine clay and clay till that overlie the sand and gravel aquifer at the bedrock contact. The thickness varies but is typically less than 40 m. The groundwater flow in this system is likely downward, toward the aquifer at the bedrock contact, at a very slow rate although locally there may be discharge to relatively deep rivers. The hydraulic conductivities for the intermediate zone are in the range of 10^{-11} to 10^{-8} m/s with mean of about 10^{-10} m/s. The low hydraulic conductivity results from the dense compact nature of the intact clays. This flow system is obviously of lower permeability than the overlying local system. Significant vertical gradients exist across these units but with low hydraulic conductivity, advective transport of pollutants can be expected to be very slow. With these low flow rates, diffusion is the most dominant pollutant transport mechanism.

The regional scale groundwater flow systems in this area are contained in the bedrock units. Of primary importance to this study is the "Fresh Water Aquifer" located at the overburden bedrock contact. Recharge to the Fresh Water Aquifer occurs primarily through the surficial sand and gravel deposits located to the east of Sarnia. A small recharge contribution is likely also obtained from the overlying clay units by downward hydraulic movement.

The Fresh Water Aquifer, which is the most important regional system, has a much higher hydraulic conductivity in the range of 10^{-10} to 10^{-6} m/s with a mean of about 10^{-7} m/s (INTERA, 1989). The yields in this aquifer are typically less than 7.5×10^{-4} m^3/s. If the aquifer is assumed to be about 1.5 to 3.0 meters thick,

then the aquifer transmissivities in the range of 1.5×10^{-10} to 3×10^{-6} m^2/s are likely.

Besides the Fresh Water Aquifer, several other deeper groundwater flow systems can be included in the regional flow system. One of these, important to this study, is that contained in the Detroit River Formation which is a permeable dolomitic and cherty limestone unit. This unit has been used for waste disposal by deep injection wells. Groundwater flow directions in this aquifer are predominantly westward and hydraulic heads for this aquifer in its undisturbed state are typically 75–100 meters below the surface. This aquifer is confined by the overlying shales and shaley limestones of the Kettle Point and Hamilton formations.

8.13.2.4 Groundwater Modelling

The details of the application of SUTRA model are presented earlier by Bobba (1993). Simulation of groundwater flow by mathematical models represents the essence of the physical flow system. The use of models helps improve the understanding of the hydrogeologic system as they incorporate virtually all available information about the behavior of the system. In addition, once these models are calibrated, they can be used to predict the response of the system to groundwater management practices. New plans for withdrawal or proposals for recharge can be evaluated by changing the locations or rates of withdrawal or recharge. Thus, these models can be used as management tools to evaluate the effects of alternative methods of controlling or developing the groundwater resources in the area of interest.

8.13.2.5 Calibration of Groundwater Flow Models

The main method of testing the accuracy of a groundwater model is to simulate historical conditions (i.e., groundwater levels in a number of observation wells) and compare the response of the model to that measured in the field. This process is known as calibration. A great deal of caution should be exercised in the process of calibrating a groundwater model because it is possible to obtain an excellent fit between the response of the model and the observed heads in a number of observation wells, simply by arbitrarily changing the values of the hydrogeologic parameters in the vicinity of these wells. An infinite number of combinations could lead to an "excellent fit" and the process of "calibration" would become meaningless.

A close examination of Equation (8.4) reveals the pitfalls that are associated with the process of calibrating a groundwater model. This equation describes the unsteady, groundwater flow in an aquifer. Its solutions is the spatially and temporarily continuous distribution of hydraulic heads determined by the combined effects of the initial and boundary conditions, spatial distributions of transmissivity and storage and space-time distributions of the flux. This implies that simulated hydraulic variations at a point (observation well) within the modelled system are dependent on the transmissivity and storage coefficient values assigned to those points, the initial hydraulic head distribution, the time variation of recharge and the boundary conditions. Assume now that the simulated hydraulic head at the

point of interest is completely different from its measured historical record and the objective is to close the gap between the two. What hydrogeologic parameter should be changed and how much? Unfortunately, the answer to this question is not simple because changes in the values of various parameters could produce the same result. For example, to reduce the amplitude of water level fluctuations at a point, one might increase the value of transmissivity, increase the value of storage coefficient or decrease the amount of recharge. The question now is: What course has to be followed? The only answer to this question in applied hydrogeology is to accept parameter values which are deemed to lie within a reasonable range. There is nothing wrong with this limitation if, indeed, this estimated range of "reasonable values" is based on sound judgement. Calibration procedures are not always applied in an entirely rational way and are often subjective.

The calibration procedure may be simplified and rationalized when groundwater models are applied to gauged basins that are under natural conditions, where the groundwater flow is from the divide towards the streams. For such basins, it is reasonable to assume that the long term groundwater recharge is equal to the long term groundwater discharge (baseflow) and that the groundwater flow is under steady state. A groundwater model, which is based on Equation (8.4) could be forced to simulate the steady state groundwater flow in a basin, by assuming that the net flux to and from the basin is equal to zero (i.e., recharge = discharge). The assumption of a steady state groundwater flow implies that the hydraulic head at any point within the system is not a function of time. This means that the time derivative of the hydraulic head, $\partial H/\partial t$, in Equation (8.4) is equal to zero and consequently the term involving the storage coefficient is equal to zero. Hence, the solution of the equation becomes independent of time and the storage coefficient. Furthermore, because the solution is independent of time, it is also independent of the initial conditions. It follows that the steady state solution depends only on the transmissivity distribution, the recharge distribution (constant in time and variable in space) and the boundary conditions. Thus, the groundwater flow model will reproduce a steady state water level distribution map which will be similar to the water level map for the basin derived from actual observations of the transmissivity and boundary conditions.

The boundary conditions in basins under natural conditions, where the groundwater flow is from the divide towards the streams, can be adequately approximated. This reduces the calibration process to the verification of the transmissivity and the recharge distributions. Even at this stage, the calibration of a groundwater model is not an easy task and should be based on sound hydrologic judgement and, to the greatest extent possible, on additional hydrologic data.

Once calibration is completed, the boundary conditions and verified transmissivity distribution can be used along with a storage coefficient distribution and variable recharge in time and space to simulate the unsteady groundwater flow. The calibration of the unsteady state model then follows. Again, calibration has to be performed with prudence and be based on sound hydrologic judgement.

Experience in calibrating groundwater models shows that there is no substitute for sound hydrologic and hydrogeologic data. The "calibrated" set of hydrogeologic parameters is not necessarily a true image of the parameters of the real system, rather it is a set which yields the best fit with observed records. The useful-

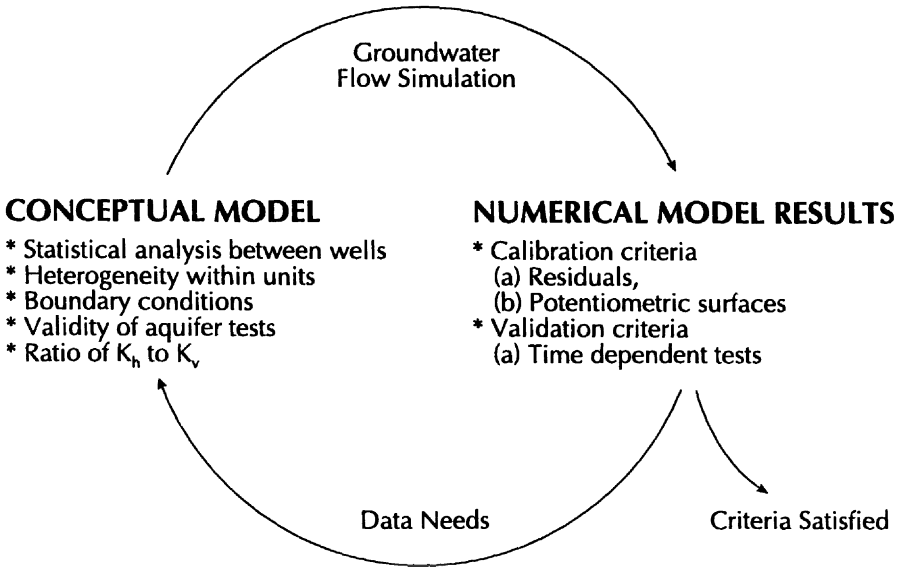


Fig. 8.37. Groundwater flow modelling approach (after Bobba, 1993)

ness of models in the study of the regional groundwater flow problems, however, is tremendous because they incorporate all the available information pertinent to the modelled system, and they verify the perception of the system and point out those areas where additional data are required.

Figure 8.37 depicts the iterative approach followed to reduce uncertainty within the groundwater flow conceptual model. The initial conceptual model was derived from the observed field data. It is important to note that due to the uncertainties listed in Figure 8.37 and others, the initial conceptual model was incomplete and uncertain in some respects. Once a coherent conceptual model was postulated, the construction of the numerical model began and simulations started.

At present, a large number of groundwater flow models (van der Heijde et al., 1986) are available in the literature. The main difference between these models is related to the type of numerical solution (finite element or finite difference) of the basic equations of groundwater flow. Minor differences are the type of groundwater problems being solved, the number of options available and the input-output formats. In this study, SUTRA (Saturated, Unsaturated, TRAnsport) a finite-element groundwater flow and energy (heat) and contaminant (contaminant) transport code, developed by U.S. Geological Survey (Voss, 1984), was used. The SUTRA model employs a two dimensional, finite-element method to simulate variable density groundwater flow, and transport of the total dissolved solids (TDS) in the groundwater. This model has many advantages associated with finite element models. The nature of finite element models provides substantial flexibility in designing the mesh to conform to the boundary conditions as well as in making the placement of material property boundaries such as hydraulic conductivity boundaries more realistic. Although the code does not allow triangular elements, it compensates for this by providing pinch nodes. Pinch nodes are nodes that end

at an element side. The input also does not allow direct variation of the storage coefficient but allows variability of this coefficient through changes in porosity, saturated thickness, and compressibility. This model has been successfully applied to real field data (Voss, 1984; Voss and Souza, 1987; Souza and Voss, 1987; Bush, 1988; Bobba, 1991)

8.13.2.6 Model Design

The constructed model was two dimensional in nature and employed 414 nodes to represent the approximate 10 km by 15 km region. Most of observation wells are located in Lambton County. The data for the groundwater model are the hydrogeologic parameters, the boundary conditions, the transmissivity, and storage coefficient, and mesh hydraulic head must be specified at each node. In addition, boundary nodes and stream nodes must be specified.

Prior to development of the model it was necessary to define the hydrologic boundary conditions which would be incorporated into the model. Since the primary objective was to simulate the Fresh Water Aquifer behavior under natural conditions, the natural water table configuration was used to determine the hydrologic boundaries of the Fresh Water Aquifer system. The hydrologic boundaries are described below.

(a) Seepage inflow boundary: It consisted of two separate segments, an equipotential line and a somewhat arbitrary boundary terminating to the west. Under natural conditions water enters the aquifer by natural flow across this boundary.

(b) Seepage outflow boundary: It corresponds to an equipotential line boundary to the west. The equipotential line was chosen so that it coincided approximately with the transition zone between the coarse-grained outwash deposits of the aquifer and the fine grained sediments to the St. Clair River. Under natural conditions water flows from the aquifer across the boundary.

(c) Groundwater divide boundary: It corresponded to a major flow line. This boundary defined the southern periphery of the aquifer flow system. Coarse-grained outwash sediments extend south of this flow line; however, beyond this boundary groundwater is part of a different flow system.

(d) Lake-aquifer boundary: The surface pressure boundary conditions were employed to represent the water levels of both St. Clair River and Lake Huron.

(e) The surface flux was distributed evenly over the basin. The flux was estimated as 4×10^{-12} to 6×10^{-12} m/s by using a gradient through the overburden of 0.02 to 0.03 m/m and hydraulic conductivity of 2.5×10^{-10} m/s.

(f) The losses to the underlying bedrock were considered negligible because the permeability of the bedrock was almost two orders of magnitude lower.

Calibration began by attempting to minimize the differences between the computed hydraulic heads and the averaged field data for each point. Calibration was

approached in two ways. One approach was to reduce the residual difference between the simulated potentiometric surfaces and the point data available for each well to within a 0.25 m mean and a 1.0 m standard deviation. In the second approach, the potentiometric surfaces were visually compared with observed data to determine if the groundwater flow direction and gradients matched. These two approaches defined the acceptance criteria that would determine if the calibration was close enough for the site to be considered modelled according to the requirements.

If the statistical analysis of the residuals or visual comparison of the simulated and actual potentiometric surfaces failed the calibration acceptance criteria, with input data being within observed data ranges, then more data are necessary. This is where the numerical model became very valuable for hypothesis testing. Areas where the model deviated the most from observed data could be evaluated to determine what additional data were necessary. Once this was determined, a focused phase of field investigations could be performed in those areas, and the resultant information incorporated into a revised conceptual model resulting in less uncertainty. Additional simulations could then be performed by varying parameters within known ranges until a closer calibration was achieved.

This iterative approach continued until the residuals were reduced to less than 0.25 m mean and 1.0 m standard deviation. This produced a quantitative measure of goodness of calibration. More importantly, the calibration continued until the actual and simulated potentiometric surfaces showed very similar flow direction patterns and gradients.

Once calibration was achieved in the steady-state model, the model was validated in a transient or time dependent test. This is an additional method to determine if the model accurately represents the groundwater flow system at the site and was accomplished by simulating a large-scale aquifer test. If the simulated drawdown did not match the observed measurements within the residual tolerances, or if the general shape and size of the simulated cone of depression was not correct, then more data would be necessary.

8.13.2.7 Statistical Analysis

Several statistical tests were used to assess the performance of the model. These tests compare the computed and observed hydraulic heads. The tests are: (1) Root Mean Square Error (RMSE), (2) Coefficient of Determination (CD), (3) Coefficient of Efficiency (CE) and (4) Coefficient of Residual Mass (CRM). The above statistical methods were developed by Aitken (1973) and have been successfully applied in hydrological modelling (Aitken, 1973, Bobba and Lam, 1990; Loague et al., 1991). These are discussed in Chapter 2.

8.13.3 Numerical Results

The steady state calibration of the model of the Fresh Water Aquifer was made using the piezometric record during the field season. This involved a laborious procedure consisting of several simulations with different input data; by varying, in particular, the transmissivity distribution and the boundary conditions at the eastern side. Figure 8.38 shows the computed hydraulic head map which best approximates the

TABLE 8.6
Statistical analysis of observed and computed hydraulic heads (Lambton County)

Contents	Statistical analysis
Root mean square error	0.007
Coefficient of determination	0.997
Coefficient of efficiency	0.995
Coefficient of residual mass	0.002
Maximum error	4.00

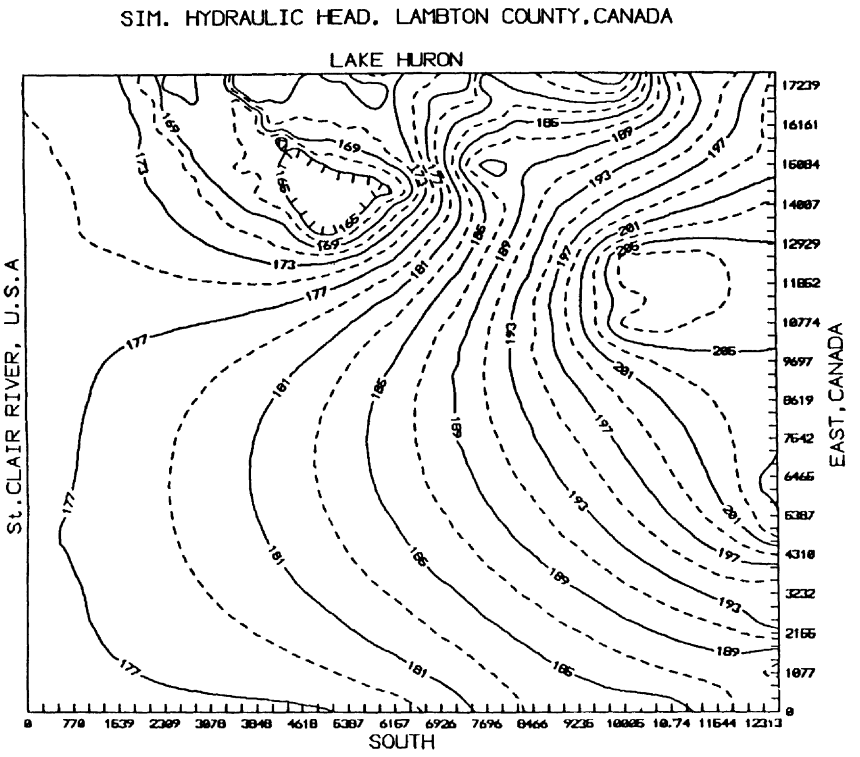


Fig. 8.38. Simulated hydraulic head map for Lambton County (after Bobba, 1993).

observed water table map (Figure 8.39). Figure 8.40 shows the contours of relative percentage errors between observed and computed data. Table 8.6 shows the results of statistical analysis of the computed and observed hydraulic heads. The various statistical parameters illustrate favourable agreement between the observed and calculated results. Maximum error magnitudes are observed on the eastern edge boundary. This may be due to an error recharge boundary condition. Large error magnitudes are also observed near the St. Clair river. This may be due to a poor

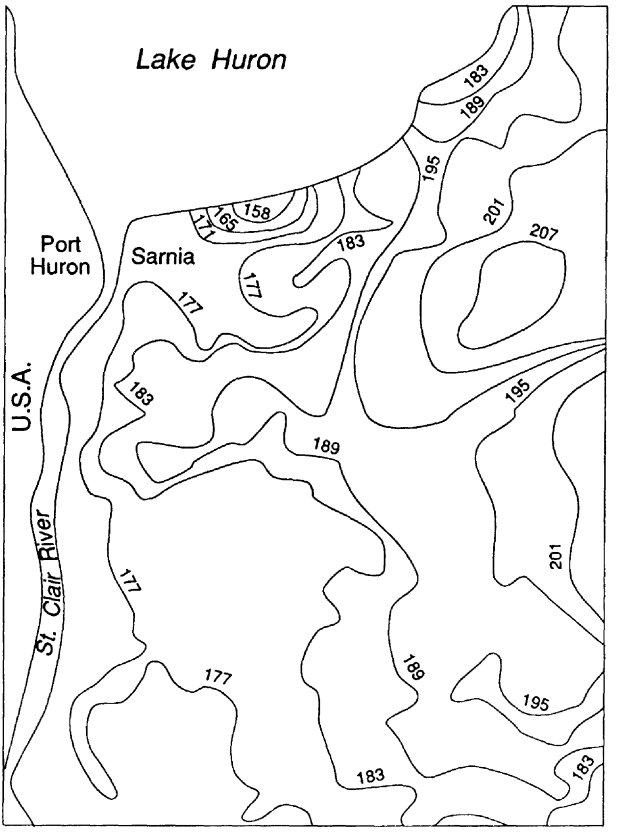


Fig. 8.39. Observed hydraulic head map for Lambton County (after Bobba, 1993).

description of the buried valley; that is, the thickness and hydraulic conductivity of the aquifer in the buried valley might have contributed to the error. The topography and hydraulic head of the freshwater aquifer apparently follow the westerly trend of the bedrock surface.

Similarly, the observed and computed hydraulic heads for the buried valley were compared. Figures 8.41 and 8.42 show the computed and observed hydraulic head data, respectively. Figure 8.43 shows the contours of relative percentage errors between observed and computed hydraulic heads. These figures have been drawn by using the SURFER computer program which uses the kriging method. By observing Figures 8.41 and 8.42, we can locate groundwater discharge areas to St. Clair river. These locations compare well with observed data (Sklash et al., 1986). Comparison of the observed hydraulic heads by error analysis shown in Table 8.7 also indicates a good agreement. The maximum error is observed at the center of area. This may be due to heterogeneity in this area. Flow rates in the drift are generally slow because the deposits which comprise the drift are fine textured and are characterized by low hydraulic conductivities. The hydraulic conductivity of the bedrock valley formation ranges from 4×10^{-8} to 2×10^{-6} m/s. The

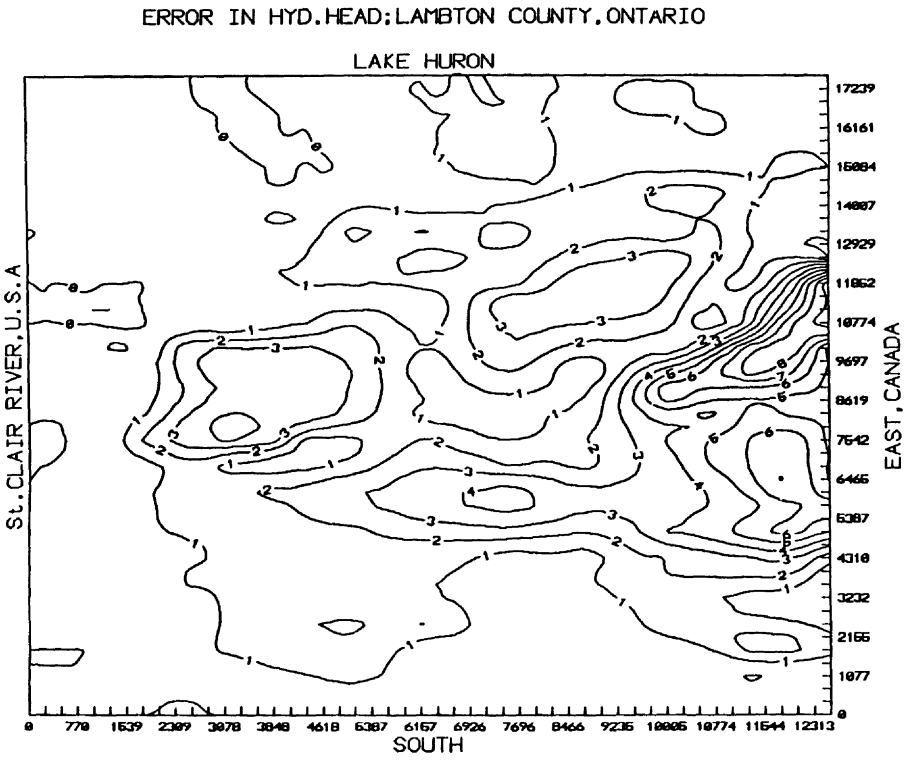


Fig. 8.40. Relative error map for Lambton County (after Bobba, 1993).

TABLE 8.7
Statistical analysis between observed and computed hydraulic heads (Buried Valley)

Contents	Statistical analysis
Nodes	442
Root mean square error	0.004
Coefficient of determination	0.941
Coefficient of efficiency	0.916
Coefficient of residual mass	0.001
Maximum error	3.10

Fresh Water Aquifer was encountered in the bedrock valley and is considered to be continuous along the bedrock valley. The water level contours in the bedrock valley indicate a generally westward groundwater flow system. Since hydraulic gradient varies across the bedrock valley, the rates of groundwater flow also vary. Calculations indicate that groundwater flux into and out of the valley is about 0.9 and 0.5 m^3/sec, respectively, for a 100 m strip of aquifer. The difference of about

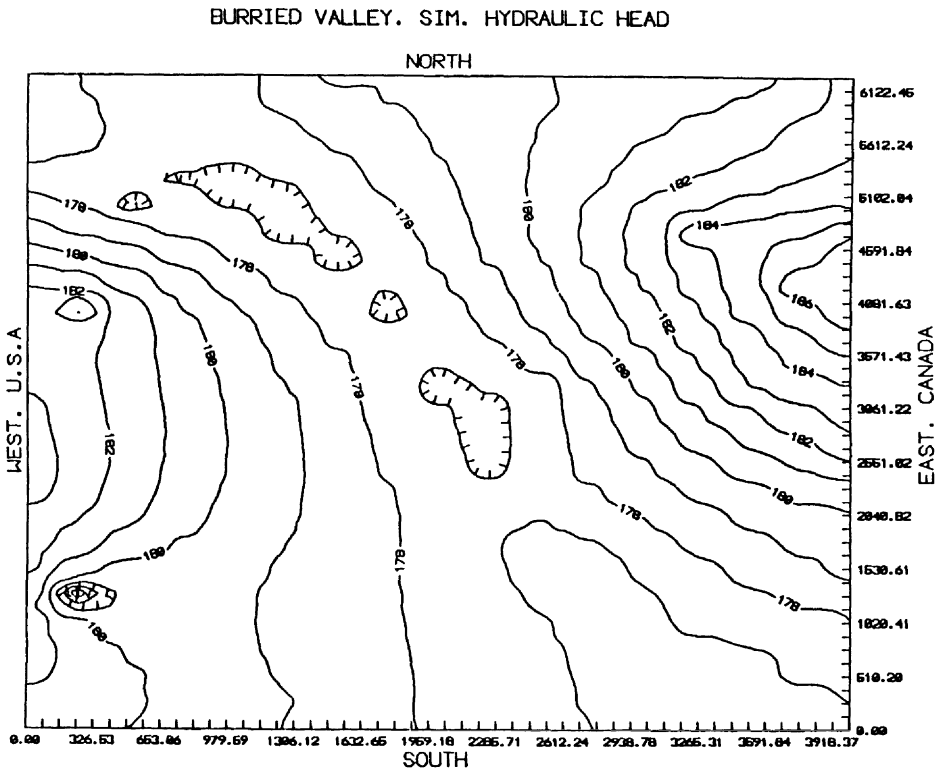


Fig. 8.41. Observed hydraulic head map for the buried valley (after Bobba, 1993).

0.4 m^3/sec is lost within the valley and may flow downward into the overlying Hamilton Formation. The amount which flows westward out of valley (0.5 m^3/sec) probably discharges to the St. Clair river. These values are confirmed with observed values (Chekaur and Taylor, 1990).

8.14 Limitations and Source of Error in Groundwater Contamination Models

In order to effectively apply numerical models, it is important to know and understand the limitations and possible sources of error in their use. All numerical models are based on a set of simplifying assumptions, which limit their use for solution of certain types of problems. To avoid applying an otherwise valid model to an inappropriate field situation, it is not only important to understand the field conditions but also to understand all of the assumptions that form the basis of the model. An areal (two dimensional) model, for example, should be applied with care to a three dimensional problem involving a series of aquifers, hydrogeologically connected by confining beds, because the model results may not be indicative of the field behavior. Errors of this type are referred to as conceptual errors.

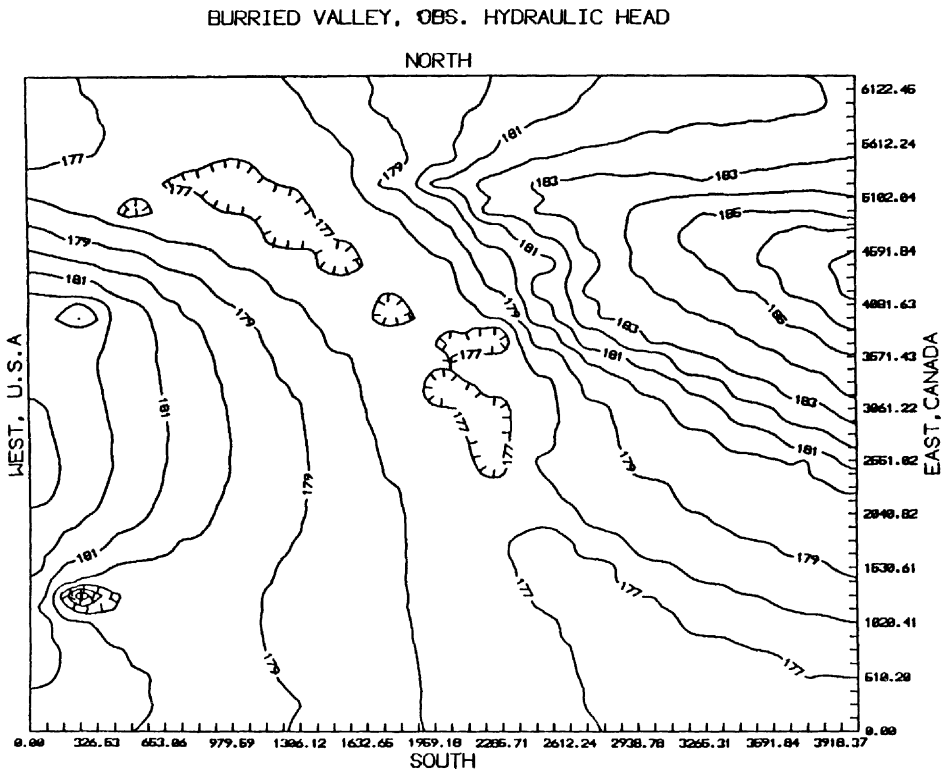


Fig. 8.42. Simulated hydraulic head map for the buried valley (after Bobba, 1993).

In addition to conceptual errors, there are several potential sources of error associated with the numerical model results. First, replacement of the model differential equations by a set of algebraic equations differs somewhat from the solution of the original differential equations. Second, the exact solution of the algebraic equations is not obtained due to round off error, as a result of the finite accuracy of computer calculations. Compared to the other error sources, roundoff error is generally negligible. Finally, and perhaps most importantly, aquifer data, for example, transmissivities, storage coefficients, and the distribution of heads within the aquifer, are seldom known accurately or completely, thus producing data errors in describing the system.

The level of truncation error in computed results may be estimated by repeating runs or portions of runs with smaller space and/or time increments. Significant sensitivity of computed results to changes in these increment sizes indicates a significant level of truncation error and the corresponding need for smaller spatial and/or time increments.

Models are collections of partial differential equations that contain a number of parameters which represent the physical properties of an aquifer. These properties must be measured in the field so that the simulated results can be verified to represent the real situation as closely as possible by statistical analysis.

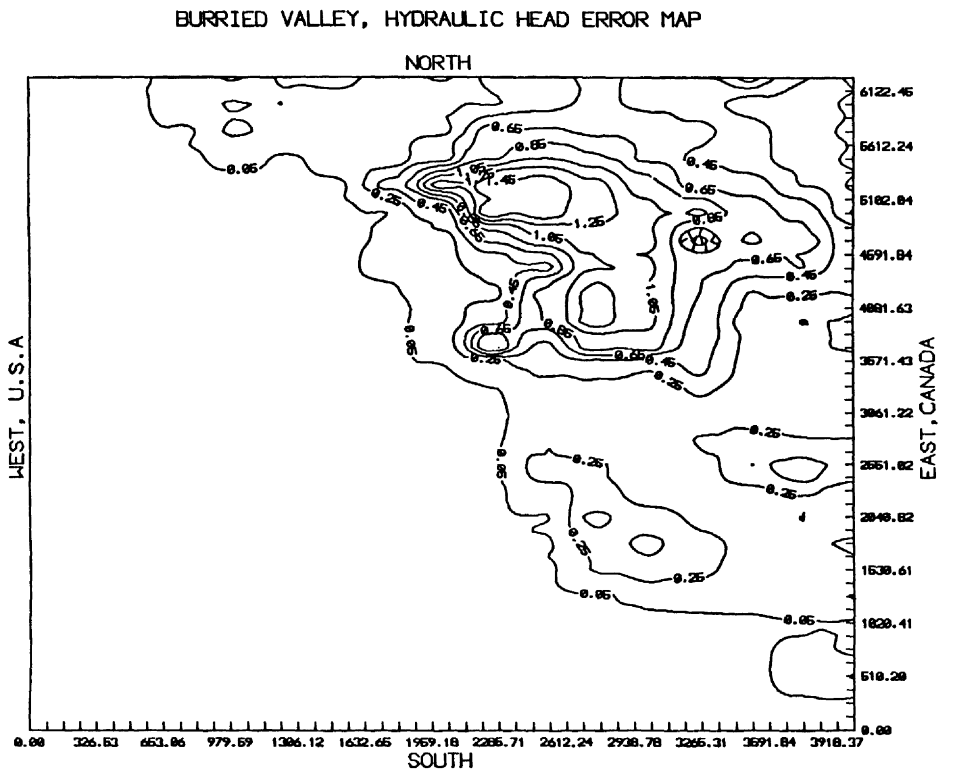


Fig. 8.43. Relative error map for the buried valley (after bobba, 1993).

Errors caused by erroneous aquifer description data are difficult to assess since the true aquifer description is never known. Although a combination of core analysis, aquifer tests, and geological studies often give valuable insight into the nature of transmissivity, storage coefficients, and aquifer geometry, much of this information may be very local in extent and should be regarded carefully when used in a model for a large area. As discussed, the final parameters that are used to characterize the aquifer are usually determined by obtaining the best agreement between model calculated and observed aquifer behavior during some historical period.

Of the various parameters involved, the hydraulic conductivity distribution is of major importance. Other parameters relating to sorption, dispersion, chemical and biological reactions are important as well, but hydraulic conductivity is more fundamental because, when combined with the hydraulic gradient and porosity, it relates to the fundamental mechanism of flow. Variations in hydraulic conductivity with depth result in significant variations in groundwater flow and transport velocities. It is the resulting velocity variations that give rise to the apparent scale dependency of dispersion. Theoretical and experimental studies indicate that the variation of horizontal hydraulic conductivity with respect to vertical position is the key aquifer property relating to the spreading of contaminants.

The contaminant transport is a natural phenomenon involving numerous environment dependent physical, chemical, and biological activities. It is impossible to simulate all properties by a model. Therefore, the validity of the assumptions made to reduce the complexities of the phenomenon is of paramount importance in modelling contaminant transport.

8.14.1 Model Misuse

There are a variety of ways to misuse models. Three common and related ones are (1) overkill, (2) inappropriate prediction, and (3) misinterpretation. The temptation to apply the most sophisticated computational tool to a problem is difficult to resist. A question that often arises is, under what circumstances should simulation be three-dimensional? Inclusion of flow in the third (nearly vertical) direction is usually recommended only if aquifer thickness is large in relation to areal extent or if pronounced heterogeneity exists in the vertical direction (for example, high stratification). Another type of overkill is that grid sizes are used which are finer (smaller) than necessary considering available information about aquifer properties, which results in additional work and expense.

In some applications, complex models are used too early in the study. For example, generally in the early stages of addressing a hazardous waste problem a contaminant transport model should not be used. Rather, the first step is to be sure that the groundwater hydrology (velocity in particular) can be characterized satisfactorily. Once this is done to satisfaction, then contaminant transport can be included in the evaluation. Basically, the analyst should assess the complexities of the problem, the amount of data that is available, and the objectives of the analysis, and then determine the best approach for the particular situation. A general rule might be to start with the simplest model and a coarse aquifer description and then refine the model and data until the desired estimation of the aquifer performance is obtained.

It is important to understand that the history matching portion of any simulation occurs under a given set of field conditions, and that these conditions may change during the prediction phase. For example, during the history match portion, the aquifer may be confined, but be on the verge of becoming desaturated. Using a confined model for prediction will give erroneous results because the saturated thickness and storage coefficient will be incorrect. Because groundwater models deal with the subsurface, there are always unknown factors such as those already cited that could affect results. In general, predictions should not be made for more than approximately twice the period used for matching, and only then under similar system stresses.

Perhaps the worst possible misuse of a model is blind faith in model results. Calculations that contradict normal hydrogeologic intuition almost always are the result of a (1) some data entry mistake, (2) a 'bug' in the computer program, or (3) misapplication of the model to a problem for which it was not designed. Proper application of a groundwater model requires an understanding of the specific aquifer. Without this conceptual understanding the whole exercise may become a meaningless waste of time and money.

8.15 Improvements in Modelling

8.15.1 Output Formats

A barrier to the more effective utilization of models, particularly by managers, is the nature of model output. Model results need to be presented in a way that is both meaningful and compatible. Constraints in the predictions for various alternatives are especially important in the decision making process and need to be adequately presented in the output. Careful use of graphics and indication of uncertainties contained in model predictions will greatly enhance the utility of model results.

8.15.2 Data Collection Methods

Inadequate or insufficient data can be an important limiting factor in the use of models for groundwater management. Data on groundwater systems are still collected largely by drilling wells, a costly and time consuming activity. Improved methods of data collection such as supplemental data generated by geophysical and other methods would be of great value to the management of groundwater and would probably significantly reduce the cost of applying models.

An improved process in data collection is the determination of which data are necessary to improve the reliability of the model output. Models of real systems should be tested through sensitivity analysis in an effort to discover which are critical data to a particular application, that is, to discover what data most critically affect the model output. The results of this sensitivity analysis, coupled with the value of the expected improvements in model results, will determine what data or additional data are necessary to be collected for improving the model and set the bounds for the cost involved to make these improvements.

8.15.3 Further Development of Flow and Transport Models

Although there have been significant advances in development of mathematical models in recent years, certain kinds of models are still not available or are unsatisfactory. There is a need to have models which fully integrate subsurface flow, surface flow, and flow through the unsaturated zone. Models which adequately simulate treatment in the saturated zone would also be a most useful addition. Although such models do exist, there is some dispute over their success in describing the treatment phenomena and their application. Models which describe flow and contaminant mass transport through fractured media would also be a valuable addition.

The unsaturated zone is a difficult zone to model. Models of unsaturated flow have been built, but they are not yet universally regarded as adequate. A complete understanding of the movement and removal of contaminants in the unsaturated zone is lacking and is, without question, of considerable importance. Development of satisfactory unsaturated zone models will likely have to await improved scientific understanding of the physical, chemical and biological phenomena involved. Contaminant transport models that include chemical and biological reactions are

needed. Such non-conservative transport models are currently being developed but have not been adequately tested.

Models describing the flow of two immiscible fluids would be another valuable addition. These models would be particularly useful for the analysis of the effects of waste disposal on groundwater systems. The petroleum industry has conducted extensive research in the area of immiscible fluid flow and has developed sophisticated models of the process for exploration purposes. However these models need to be modified for groundwater problems.

8.16 Further Research in Chemical Kinetics

The fact that certain types of models have not been sufficiently developed or are not considered reliable may be caused by factors not intrinsic to the models. Rather, the underlying theory concerning the system to be modeled may not yet have been adequately described. In other words, there are still major gaps in the scientific understanding of certain physical, chemical and biological phenomena. These gaps will have to be filled through further research geared to observation of the phenomena in question and the development of adequate theories to describe them. Once these phenomena have been described, the incorporation of the necessary equations is a natural next step.

There is no complete agreement concerning specific research gaps. However, four areas of particular concern are (a) an inadequate description of the kinetics of chemical and biological processes, (b) the movement of pollutants through the unsaturated zone, (c) better parameter identification methodologies, and (d) the effect of scale and heterogeneity on transport phenomena.

8.17 Summary

Recognition of existing and potential contamination of important groundwater resources will result in increased use of contaminant transport modelling for problem analysis and management purposes. To aid in understanding the state-of-the-art in groundwater modelling, this chapter provides a general introduction to the modelling of contaminant transport in groundwater systems. It discusses (a) groundwater model applications, (2) development of groundwater models and the principles and concepts used, (3) guidance for groundwater model use, and (3) areas of future research which would provide information that would lead to more useful simulation techniques.

In summary, the important findings include the following:

(A) Contaminant transport simulation techniques can be an important adjunct when dealing with (a) the environmental acceptability of new and existing land sited facilities, (b) design of groundwater monitoring systems, (c) the need for or extent of remedial actions required to remediate an aquifer, and (d) integration of subsurface and surface flow systems to provide a greater understanding of actual events that lead to environmental concerns.

(B) Further research and development is needed to improve contaminant transport models and increase confidence in their application. Areas for study include (a) understanding the kinetics of processes occurring in groundwater systems which is an important step to model improvement and utility, (b) improvements in data collection methods to decrease costs for obtaining information and to produce more representative model results, and (c) improvement in output formats to make modelling more useful for environmental management purposes.

Acknowledgment

The senior author (A. Ghosh Bobba) dedicates this chapter to Dr. K.V.N. Rao, who retired in 1993 as Research Scientist at Gould Batteries of Canada Ltd., Fort Erie, Ontario, Canada from 1966–1993. Dr. Rao helped the senior author to come to North America and created interest in research and further post graduate studies. The encouragement and support from J.H. Carey, D.S. Jeffries, and W. Strachatan are also appreciated. The authors acknowledge the publishers to permit to publish the figures in this chapter.

References

Abriola, L.M. 1987. Modeling contaminant transport in the subsurface: An interdisciplinary challange. U.S. Natl. Rep. Int. Union. Geod. Geophys., 1983-1986, Rev. Geophys. Vol: 25, 125-134.

Aitken, A.P., 1973. Assessing systematic errors in rainfall-runoff models. J. of Hydrology. Vol. 20: 121-136.

Bear, J. 1972. Dynamics of fluids in porous media. American Elsevier Publishing Company, Inc. 764pp.

Bear, J., 1979. Hydraulics of groundwater. McGraw-Hill, New York, 567p.

Blair, P.m., and Peaceman, D.W. 1963. An experimental verification of a two dimensional technique for computing performance of gas reservoirs. Society of Petroleum Engineers Journal, Vol. 3: 19-27.

Boast, C.W. 1973. Modeling the movement of chemicals in soils by water. Soil science, Vol. 115: 224-230.

Bobba, A.G., 1989. Numerical model of contaminant transport through conduit-porous matrix system. Mathematical Geology. Vol. 21, 861-890.

Bobba, A.G. 1991. Application of digital simulation model to fresh water aquifer of Lambton County, Ontario, Canada. Paper presented at International Conference on Hydrology and Hydrogeology in the '90s, Orlando, Florida, U.S.A.

Bobba, A.G., 1993. Field validation of "SUTRA" groundwater flow model to Lambton County, Ontario, Canada. Water Res. Management. Vol. 7, 289-310.

Bobba, A.G., and Bukata, R.P. 1980. Parameterization of contaminant transport through an anisotropic unconfined aquifer. Nordic Hydrology. Vol. 11: 187-208.

Bobba, A.G. and Joshi, S.R., 1988. Groundwater transport of radium-226 and uranium from Port Granby waste management site to Lake Ontario. Nucl. Chem. Waste Manage. Vol. 8: 199-209.

Bobba, A.G. and Joshi, S.R. 1989. Application of an inverse approach to a canadian radioactive waste disposal site. Ecological Modelling, Vol. 46, 195-211.

Bobba, A.G. and D.C.L. Lam, 1990. Hydrological modelling of Canadian watersheds. Ecological Modelling. Vol. 50: 5-32.

Bredehoeft, J.B., and Pinder, G.F. 1970. Digital analysis of areal flow in multiaquifer groundwater system. Water Resources Research. Vol. 6: 883-888.

Bredehoeft, J.D. and G.F. Pinder. 1973. Mass transport in flowing groundwater. Water Resources Research. Vol. 9: 194-210.

Bush, W.P., 1988. Simulation of saltwater movement in the Floridian aquifer system, Hilton Head Island, South Carolina. USGS. Water Supply Paper 2231. 19p.

Cherkauer, D. S. and Taylor R.W. 1990. The spatially continuous determination of groundwater flow to surface water bodies: Application to the connecting channels between lakes Huron and Erie. J. of Hydrology. Vol. 114: 349-369.

D'Astous, A.Y., W.W. Ruland, J.R.G. Bruce, J.A. Cherry and R.W. Gillham, 1988. Fracture effects in the shallow groundwater zone in weathered Sarnia-area clay. Can. Geotech. J. Vol. 26: 43-56.

Durham, R. W. and S.R. Joshi, 1980. Investigation of Lake Ontario water quality near Port Granby radioactive waste management site. Water Air Soil Pollution. Vol. 13: 17-26.

Douglas, J., Jr., Peaceman, D.W., Rachford, H.H. 1959. A method for calculating multidimensional immiscible displacement. Transactions AIME, Vol. 216: 297-308.

Ergatoudies, I., Irons, B.M., and Zienkiewicz, O.C. 1968. Curved isoparametric "Quadrilateral" elements for finite element analysis. International Journal of Solids and Structures. Vol. 4: 31-42.

Faust, C.R., and Mercer, J.W. 1980. Ground-Water modelling: Numerical models. Ground Water. Vol. 18: 115-125.

Freeze, R.A. and Cherry, J.A. 1979. Groundwater, Prentice-Hall, Englewood Cliffs, New Jersey.

Hoopes, J.A. and Harleman, D.R.F. 1967. Dispersion in radial flow from a recharge well. Journal of Geophysical Research. Vol. 72:3595-3607.

Hornberger, G.M., Ebert,J., and Remson, I. 1970. Numerical solution of the Boussineque equation for aquifer stream interaction. Water Resource Research, Vol. 6: 601-608.

Hubbert, M.K.L. 1940. The theory of ground water motion. J. Geology. Vol. 48: 765-944.

Isotok, I. 1989. Groundwater modelling by the finite element method. Water Resources Monograph # 13. American Geophysical Union. Washington, D.C. 495pp.

INTERA, 1989. Hydrogeologic study of the Fresh Water Aquifer and deep geologic formations Sarnia, Ontario. Intera Technologies Ltd., Ottawa, Ontario. 193 p.

Langmuir, I. 1918. The adsorption of gases on plane surfaces of glass, mica, and platinum. J. Am. Chem. Soc. Vol. 40: 1361-1402.

Larkin,B.K. 1964. Some stable explicit difference approximations to the diffusion equation. Journal of Mathematics of Computation Vol. 18:

Lesage, S, R.E. Jackson, M. Pridle, P. Beck and K.G. Raven, 1991. Investigation of possible contamination of shallow groundwater by deeply injected liquid industrial wastes. Groundwater Monitoring Review., Winter, 1991. pp 151-159.

Liggett, J.A. and Liu, P.L.F. 1983. The boundary integral equation method for porous media flow. Allen and Unwin, London, 1983.

Loague, K and R.E. Green, 1991. Statistical and graphical methods for evaluating contaminant transport models: Overview and application. J. of Contaminant Hydrology. Vol. 7: pp 51-73.

Longenbaugh, r.A. 1967. Mathematical simulation of a stream-aquifer system. Paper presented at the Third American Water Resources Conference, san Francisco, California.

McLean, D.D, 1968. Subsurface disposal of wastes in Ontario, Energy and Resources Management, Petroleum Resources Section, Paper 68-2, 91p.

Mercer, J.W., and Faust, C.R. 1980. Ground-Water modelling. An Overview. Ground Water Vol. 18: 108-114.

Mercer, J.W., and Faust, C.R. 1980. Ground-Water Modelling:Applications. Ground Water, Vol. 18: 486-497.

Moore, J.E. 1979. Contribution of ground-water modelling to planning. Journal of Hydrology. Vol. 43:121-128

Naymik, T.G. 1987. Mathematical modeling of solute transport in the subsurface. CRC Critical Reviews in Environmental Control. Vol: 17, 229-251.

Pinder, G.F., and Bredehoeft, J.D. 1968. Application of the digital computer for aquifer evaluation. Water Resources Research, Vol. 4: 1069-1093.

Pinder, G.F., Frind, E.O. and Papadopules, S.S. 1972. Fundamental coefficients in the analysis of groundwater flow using finite elements, Water Resources Research, Vol. 9: 222-226.

Pinder, G. F. and Frind, E.O. 1972. Application of Galerkin's procedure to aquifer analysis. Water Resources Research, Vol. 8: 108-120.

Pinder, G.F. 1973. A Galerkin finite element simulation of groundwater contamination on Long Island, New York, Water Resources Research. Vol. 9: 108-120.

Pinder, G.F, and W.G. Gray, 1977. Finite Element Simulation in Surface and Subsurface Hydrology. Academic Press; New York. N.Y., 295p.

Platford, R.F., Bobba, A.G., and Joshi, S.R. 1984. The Port Granby radioactive waste management. Water Pollut. Res. J. Can. Vol. 19: 90-96.

Pricket, T.A. 1979. Ground-Water computer models-State of the Art. Ground water, Vol. 17: 167-173.

Quon, d. 1965. A stable explicit computationally efficient method for solving two dimensional mathematical models of petroleum reservoirs. Journal of Canadian Petroleum Technology. Vol.4: 115-120.

Reddell, D.L. and Sunada, D.K. 1970. Numerical simulation of dispersion in groundwater aquifers. Hydrology Paper # 41 , Colorado State University, Fort Collins, CO., 95 pp.

Remson, I., Appel., C.A., and Webster, R.A. 1965. Groundwater models solved by digital computer. Proc. Am. Soc. of Civil Engineers (Hydraulics Division), Vol. 91: 133-149.

Remson, I., Hornberger, G.M., and Molz, F.J. 1971. Numerical methods in sub-surface hydrology. Wiley-Interscience, New York. 389pp.

Ruland, W.W, J.A. Cherry and S. Feenstra, 1991. The depth of fractures and active groundwater flow in a clayey till plain in Southwestern Ontario. Ground Water, Vol. 29, pp. 405-417.

Saulyev, V.K. 1964. Integration of equations of parabolic type by the method of nets. Macmillan Company, New York.

Shamir, U., and Dagan, G. 1971. Motion of seawater interface in coastal aquifers–A numerical solution. Water Resource Research. Vol. 7: 644-658.

Shaw, F.S. and Southwell, r.v. 1941. Relaxation methods applied to engineering problems. VII. Problems relating to the percolation of fluids through porous materials. Proc. Royal. Soc. , A178.

Sherif, M.M., Singh, V.P., and Amer, A.M. 1988. A two dimensional finite element model for dispersion (2D-FED) in coastal aquifer. Journal of Hydrology, Vol. 103: 11-36.

Sklash, M.S., S. Mason and S. Scott, 1986. An investigation of the quantity, quality and sources of groundwater seepage into the St. Clair River, near Sarnia, Ontario, Canada. Water Poll. Res. J. Canada. Vol. 21. pp 351-367.

Smith, I.M., Farraday, R.V. and O'Conner, B.A. 1973. Rayleigh-Ritz and Galerkin finite elements for diffusion convection problems, Water Resources Research. Vol. 9: 593-606.

Souza, W.R. and C.I. Voss, 1987. Analysis of an anisotropic coastal aquifer system using variable-density flow and contaminant transport simulation. J. of Hydrology. Vol. 92:pp 17-41.

Stone, H.L. and Brian, P.L.T. 1963. Numerical solution of convective transport problems. AIChE Journal, Vol. 9: 681-688.

Strack, O.D.L. 1989. Groundwater Mechanics, Prentice-Hall, Englewood Cliffs, N.J. 1989.
Taylor, G.S., and Luthin, J.N. 1969. Computer methods for transient analysis of water table aquifers. Water Resource Research, Vol. 5:144-152.
van der Heijde, P.V., Y. Bachmat., J. Bredehoeft., B. Andrews., D. Holtz., S. Sebastain. 1985. Groundwater Management: the use of numerical models. American Geophysical Union, Washington, D.C., 100 p.
Vandenberg, A., D.W. Lawson, J.E. Charron and B. Novakovic, 1977. Subsurface waste disposal in Lambton County, Ontario., Inland Waters Directorate, Water Resources Branch, Tech. Bull. 90., 64p.
Voss, C.I., 1984. SUTRA- A finite element simulation model for saturated-unsaturated fluid density dependent groundwater flow with energy transport or chemically reactive single species contaminant transport. USGS Water Resour. Investigate Report., 84-4269, 409p.
Voss, C.I. and W.R. Souza, W.R., 1987. Variable density flow and contaminant transport simulation of regional aquifers containing a narrow freshwater-saltwater transition zone. Water. Res. Res. Vol. 23. pp 1851-1866.
Wang, H.F., and M.P. Anderson. 1982. Introduction to Groundwater Modelling. Freeman and Co., San Francisco., 237p.
Zienkiewicz, O.C. (1971). The finite element method in Engineering Science, McGraw-Hill, New York.
Zijl, W. 1984. Finite element methods based on a transport velocity representation for groundwater motion. Water Resources Research. Vol. 20, 137-145.

CHAPTER 9

Modeling subsurface transport of microorganisms

Y. Tan and W.J. Bond

9.1 Introduction

The transport and fate of microorganisms in porous media was originally of interest because of concern for outbreaks of diseases caused by groundwater contaminated with pathogenic microorganisms (Craun, 1985). Transport of microorganisms is now also recognized to be important in fields as diverse as oil recovery (Jenneman et al., 1984) and biological control of plant root diseases (Parke et al., 1986; Tan et al., 1991). In-situ bioremediation of contaminated soil or aquifers using indigenous or exogenous bacteria has recently generated further interest in this topic, as well as in the simultaneous transport of microorganisms and biologically reactive solutes used by bacteria as substrates for growth (Baveye and Valocchi, 1989; Reynolds et al., 1989; Taylor and Jaffe, 1990c; Gannon et al., 1991).

Although there is growing interest in the topic, it is still in an early stage of development. Many important processes have been identified and mathematical formulations developed to enable them to be incorporated into microbial transport equations. These formulations have been drawn from various sources including microbiology and filtration literature. At present, the mathematical descriptions remain largely empirical and it is often not possible to determine independently the parameters contained in them.

Nevertheless, a number of conceptual and mathematical models have been developed to describe the transport of microorganisms and biologically reactive solutes in porous media. These models allow assessment of complex interrelationships among the physical, chemical and biological processes that affect the transport and fate of microorganisms in soils and groundwater. They also help characterise and quantify the environmental factors and processes that affect microbial transport. These models should also provide reliable tools for management

V. P. Singh (ed.), Environmental Hydrology, 321–355.
© 1995 *Kluwer Academic Publishers.*

and regulatory guidance, although thus far they have mainly been used for scientific research purposes.

This chapter reviews and summarises the progress made to date in this area.

9.1.1 Development of Mechanistic Models of Transport

Experimental studies in soils and other subsurface materials (Goldshmid et al., 1973; Bitton et al, 1974; Smith et al., 1985; Wollum and Cassel, 1978; Parke et al., 1986; Bales et al., 1991; Fontes et al., 1991; Gannon et al., 1991; Tan et al., 1991) have shown qualitatively that many environmental factors affect the transport of bacteria and viruses in porous media. These include the chemical composition of the interstitial solution, water flow velocity, and surface properties of the solid phase. Processes such as growth, decay, adsorption, straining, advection, motility, diffusion and dispersion, which have been studied in isolation by microbiologists, engineers, and hydrologists, are also known to be involved in the transport and fate of microorganisms.

Vilker and Burge (1980) and Matthess and Pekdeger (1981) appear to have been among the first to apply a transport equation to describe the distribution of bacteria and viruses in time and space. They each presented equations for microbial transport based on the advection-dispersion equation for solutes, but simplified in different ways by omitting various processes. Corapcioglu and Haridas (1984) presented a more complete transport equation, also based on the advection-dispersion equation, in which all processes known to affect microbial transport were included. Since then advances have been made in the mathematical description and parameterisation of the various processes, as well as in simultaneous transport of microbial growth substrates with bacteria.

Approaches other than those based on the conventional advection-dispersion equation have been used for microbial transport, adopting the alternative approaches for solute transport modeling suggested in recent years. These followed the apparent failure of the advection-dispersion equation to describe solute transport adequately under some field conditions, particularly highly variable fields and where flow was dominated by macropores. Jury (1982) introduced the transfer function approach, which replaces the advection and dispersion terms by a probability distribution of water flow velocities. Rather than determining the physical properties of the soil or aquifer, the probability distribution function is used to characterise the porous medium. Thus, the transfer function approach requires an initial calibration phase. White et al. (1986) applied the transfer function model to the transport of bacteria. Another approach that has been introduced for transport through subsurface materials with macropores is the kinematic wave approach, which was applied to bacteria by Germann et al. (1987).

Despite the introduction of alternative approaches, the advection-dispersion equation is widely accepted for most circumstances (Corapcioglu and Haridas, 1984; 1985; Taylor and Jaffe, 1990c; Dickinson, 1991; Harvey and Garabedian, 1991; Udoyara and Mostaghimi, 1991; Hornberger et al., 1992; McCaulou et al., 1994). In this chapter we concentrate on this approach and use it to derive a general transport equation for microorganisms in porous media. Various approaches to the mathematical description of the processes that affect microbial transport are

described and incorporated into the general transport equation. We also derive some special simplified cases of the general transport equation and present examples of experimental studies to illustrate their applicability.

9.1.2 Approaches to the Pore-Scale Distribution of Bacteria

Before describing processes affecting the transport of microorganisms, it is useful to introduce the different conceptual approaches to the pore-scale geometrical configuration of attached bacteria in porous media. Two major conceptual frameworks have emerged relating microbial activity to the pore scale spatial distribution of attached bacteria in porous media. The first approach assumes that the distribution of bacteria on the solid surfaces is unimportant and has no effect on their access to substrates for growth (Bazin et al., 1976; Sykes et al., 1982; Corapcioglu and Haridas, 1984). This will be referred to as the strictly macroscopic approach (Baveye and Valocchi, 1989).

The second approach assumes that collections of bacteria attached to solid surfaces have specific geometries, usually either biofilms or microcolonies (Rittmann et al., 1980; McCarty et al., 1984; Molz et al., 1986; Chen et al., 1992). The effect of these geometries on the access of bacteria to substrates in the solution phase is then considered explicitly. Although we consider the two approaches together for the purposes of this discussion, biofilms and microcolonies are based on slightly different conceptualisations. Biofilms are proposed to be thin, continuous layers composed of a homogeneous matrix of bacteria and extracellular polymers that bind them together and to the solid surfaces. Depending on the assumptions made about movement of substrate along concentration gradients through the biofilm, mathematical descriptions of different levels of complexity result (Baveye and Valocchi, 1989). The microcolony approach assumes that microorganisms exist in the form of discrete groups rather than in continuous films. Growth is accounted for by either allowing the colonies to increase in size (Watson and Gardner, 1986) or by allowing the number of colonies having fixed dimensions to increase as a result of the consumption of substrate (Molz et al., 1986; Chen et al., 1992). Although biofilm and microcolony approaches have traditionally neglected mass transport processes of bacteria such as advection and dispersion of bacteria, it will be shown later in this chapter this is not necessary.

In the remainder of this chapter we do not distinguish between the biofilm and the microcolony approaches and refer to both as the "biofilm" approach. We do, however, distinguish between the strictly macroscopic and "biofilm" pore-scale distribution approaches and present examples of both. Furthermore, we discuss the inter-relationship between the two and show that under some conditions the biofilm approach can be easily simplified to the macroscopic approach.

It should be noted that these concepts have been applied only to bacteria and are not relevant to viruses.

9.2 Processes Affecting Transport and Fate of Microorganisms

As pointed out in Section 9.1.1, the transport of microorganisms is determined by a variety of processes. Movement of microorganisms in and with the liquid phase is controlled primarily by the flow of water, which results in advective transport and hydrodynamic dispersion, as well as by Brownian diffusion and by their own active movement in the case of motile bacteria. When microorganisms move through a porous medium, they are also subject to physical and physico-chemical forces, which usually cause the microorganisms to be retarded relative to the water. Many investigators (e.g., Butler, 1954; Bitton et al., 1974; Corapcioglu and Haridas, 1984; Gerba and Bitton, 1984; McDowell-Boyer et al., 1986) have discussed the mechanisms for retention of microorganisms. These include processes such as straining (or filtering), adsorption, interception and sedimentation, the relative importance of which depends on the type of microorganism and environmental factors such as solution ionic strength and properties of the porous medium. In addition, microorganisms are subject to decay or death, and may grow or multiply at rates controlled by availability of substrates. The consumption and transport of growth-limiting substrates occur simultaneously with the transport of bacteria and therefore may also affect bacterial transport. All these processes need to be considered when constructing a comprehensive mathematical description for the transport of microorganisms in porous media.

Some processes that affect bacteria in porous media, particularly attachment processes which cause a gradual decrease in the water filled porosity, affect permeability and therefore the movement of water. This, in turn, affects the transport of bacteria. Although this feedback mechanism is not discussed here, Taylor and Jaffe (1990c) did explicitly consider changes in water-filled porosity, permeability, and the dispersion coefficient, caused by bacterial attachment, and their affect on microbial transport.

The mathematical formulations of the processes listed above vary in their completeness and in their biological, chemical and physical basis. Some processes are relatively well understood, while others rely on empirical descriptions. Often it is difficult to determine independently the parameters contained in these descriptions.

The discussion below examines a number of the descriptions that are currently in use and their inter-relationships.

9.2.1 Movement of Microorganisms

9.2.1.1 Advection

Moving soil water or groundwater often carries with it dissolved gases and solutes, as well as suspended colloidal material including clay particles, bacteria, viruses and organic matter. This form of transport is commonly referred to as advection, and is the prime mechanism for movement of bacteria in soils and groundwater. Hamdi (1974) found that the vertical movement of *Rhizobium trifolii* cells increased with the amount of water entering the soil. Similarly, Parke et al. (1986) showed that water flow caused pseudomonads to move rapidly and over a relatively long

distance in soils. The movement of viruses is also known to be influenced by water flow (Gerba and Bitton, 1984).

The mass flux of bacteria or viruses carried with the flow of soil water is proportional to the concentration of bacteria or viruses in solution (C), i.e.,

$$J_v = J_w C = v_w \theta C, \tag{9.1}$$

where J_v is the advective flux of microorganisms (the mass of microorganisms per unit cross sectional area of soil per unit time), J_w is the flux of water (the volume of water flowing through a unit cross-sectional area of soil per unit time), v_w is the average water flow velocity, and θ is the water filled porosity.

9.2.1.2 Chemotaxis

Some bacteria are motile (capable of active movement), using their flagella to propel them, and this movement can be in random or systematic directions. The random active movement of motile bacteria will be treated below in Section 9.2.1.3. The systematic active movement of bacteria caused by the response of bacteria to chemical gradients is known as chemotaxis and has been described by, e.g., Adler (1969), Berg and Brown (1972) and Dahlquist et al. (1972). Chemotactic bacteria move towards attractant chemicals (such as growth substrates), or away from toxins. The phenomenon of chemotaxis can be observed by inserting a capillary tube containing an attractant chemical into a suspension of motile bacteria. Bacteria will accumulate near the end of the tube where the concentration of the attractant is relatively high. It is thought that chemotactic bacteria actively avoid regions of lower concentration of an attractant by reversing direction or by choosing new directions at random.

The average chemotactic velocity, v_x, of the bacteria may be expressed as (Dahlquist et al., 1972; Corapcioglu and Haridas, 1984)

$$v_x = (k_x/S)\nabla S, \tag{9.2}$$

where S is the concentration of an attractant or substrate, and k_x is the chemotactic coefficient.

Keller and Segel (1971) stated that even though a cell may not be capable of making an accurate assessment of the gradient to which it is exposed, its average behavior can nevertheless reflect the gradient. Although individual cells may respond to different local chemical concentrations giving rise to fluctuations in their paths, an average cell flux proportional to the macroscopic substrate gradient can be determined. The resulting macroscopic flux due to chemotactic movement is expressed as

$$J_x = \theta v_x C, \tag{9.3}$$

where J_x is the chemotactic flux (Corapcioglu and Haridas, 1984). Combination of Equations (9.2) and (9.3) yields an equation for calculating the chemotactic flux.

Chemotaxis has been observed for some bacteria but not for viruses. Although the contribution of chemotaxis to the overall bacterial transport is likely to be

small it has been included in bacterial transport models. Experimental studies of the transport of bacteria in columns and in the natural environment have generally been described adequately without invoking chemotactic movement. There is a need for more information to quantify the contribution of chemotactic movement to the overall transport of bacteria in porous media.

9.2.1.3 Diffusion, Dispersion and Motility

The random motion of microorganisms in the water phase results in Brownian diffusion. Non-uniform distributions and concentration gradients of microorganisms often exist in the natural environment and the tendency to equalize the spatial distribution results in the diffusion of microorganisms from higher to lower concentration. The flux of microorganisms (J_B) in porous media due to Brownian diffusion is described by Fick's first law, i.e.,

$$J_B = -\tau\theta D_B \nabla C, \tag{9.4}$$

where D_B is the diffusion coefficient of bacteria and viruses in free solution. The volumetric water filled porosity (θ) is included because diffusion of microorganisms occurs only in water, and this is modified by the pore space tortuosity (τ) to account for the reduced effect of diffusion resulting from the tortuous flow paths in the pore space compared with free solution. According to Corapcioglu and Haridas (1984), D_B can be estimated by the Stokes-Einstein equation. Characklis (1990) stated that larger bacteria of size 1–4 μm^3 or greater undergo little Brownian motion. He also reported that Brownian diffusion coefficients for various bacteria in free solution were of the order of 5×10^{-11} $m^2\,s^{-1}$, and hence two orders of magnitude smaller than diffusion coefficients for common cations and anions in free water.

The flow of soil water affects the transport of microorganisms not only by carrying them in flowing water (advection), but also by mixing and equalizing the concentration differences between different portions of the flowing water. This process of mixing and spreading is known as hydrodynamic dispersion (Bear, 1972; Kirkham and Powers, 1972). This term was introduced to describe the transport of solutes in porous media and has also been accepted to describe mixing and spreading of microorganisms during transport (Bouwer, 1984; Matthess and Pekdeger, 1981; Corapcioglu and Haridas, 1984). Hydrodynamic dispersion differs from the diffusion process in both magnitude and mechanism. In the liquid phase, diffusion occurs regardless of whether the liquid solution is in motion or not while dispersion occurs only in moving water. Although the two processes differ, it is difficult to separate them in a flowing system.

Besides the systematic movement due to the response of bacteria to macroscopic gradients of stimuli (chemotaxis), motile bacteria also move in a tumbling fashion in random directions (Berg and Brown, 1972). This chaotic and random aspect of bacteria motility can be likened to Brownian motion and described by an effective diffusion coefficient or a motility coefficient, D_t. The value of D_t has been calculated to be of the order of 10^{-9} $m^2\,s^{-1}$ (Characklis, 1990), which is comparable with chemical diffusion coefficients. Viruses do not exhibit motility.

In a flowing system, diffusion, hydrodynamic dispersion and the random active movement of motile bacteria occur simultaneously. Because these processes are not practically distinguishable from each other and because they all contribute to a diffusion or smearing of the distribution of microorganisms along a concentration gradient, the effects are considered to be additive. The mathematical formula for diffusion [Equation (9.4)] may be adopted to describe the combined process of random active movement, diffusion and hydrodynamic dispersion with the diffusion coefficient replaced by the dispersion coefficient, i.e.,

$$J_d = -\tau\theta D\nabla C, \tag{9.5}$$

where J_d is the combined diffusive-dispersive flux of microorganisms and D is the combined diffusion-dispersion coefficient. The dispersion coefficient is the sum of the Brownian diffusion coefficient (D_B), the hydrodynamic dispersion coefficient (D_h) and, for bacteria, the motility coefficient (D_t). The combined flux therefore has both a diffusional and a hydrodynamic character, and it also takes into account the random active movement of motile bacteria. The dispersion coefficient of solutes has been found to be a function of the average pore water velocity (Bond and Smiles, 1983), and the dispersion coefficient of microorganisms is also dependent on the water flow velocity (Matthess and Pekdeger, 1981). Various formulations for relating D and v_w have been used (Biggar and Nielsen, 1980), but are not discussed here.

9.2.2 Growth and Decay of Microorganisms

9.2.2.1 Growth

This section deals specifically with growth of bacteria. Viruses are not capable of reproduction outside host organisms, and therefore reproduction (growth) is not relevant to a discussion of viruses in soil and groundwater where their hosts are unlikely to be present.

Although bacteria can utilize several substrates (or nutrients) simultaneously, growth is often controlled by a single substrate and the specific growth rate is assumed to be a function of the concentration of that rate-limiting substrate. Numerous mathematical models have been proposed to describe the relationship between the specific growth rate and substrate concentration. These can be classified into two categories, "structured" and "unstructured" (Tan et al., 1994a). The structured approach accounts for aspects of the internal structure of bacteria and variations in their internal composition. This is important for the description of transient growth, but the large number of parameters required for structured models has limited their use to a few well-defined growth conditions. Unstructured models ignore the complexity of the living organism and assume that the amount of biomass present in the culture provides sufficient information about microbial activity (Roels, 1983). Unstructured models are more applicable to describe steady state cell growth than transient growth, although they have been used successfully to describe slow transients in batch growth for certain conditions (Moser, 1985). Only unstructured models are considered here.

The Monod equation (Monod, 1942) is the most widely used unstructured model for microbial growth limited by a single substrate. This equation takes the form:

$$\mu = \frac{\mu_m S}{K_s + S}, \tag{9.6}$$

where μ is the specific growth rate, μ_m is the maximum specific growth rate, and S is the concentration of the limiting substrate. The saturation constant, K_s, is defined as the concentration of the limiting substrate at which the specific growth rate is half the maximum value. The maximum specific growth rate is achieved when $S \gg K_s$.

This equation has been used to quantify the specific growth rate of bacteria in both the water and attached biomass phases. However, it has been shown to fail to fit some experimental data, including Monod's original data (Dabes et al., 1973). Shehata and Marr (1971) proposed the use of multiple Monod equations (the sum of two or more single Monod equations) for the description of microbial growth after finding that using a double Monod equation improved prediction of the growth of *Escherichia coli* over a large range of substrate concentrations. Neither the single Monod equation nor the multiple Monod equation is able, mathematically, to describe S-shaped growth curves which are sometimes observed.

Moser (1958) proposed a modified form of the Monod equation which, although not widely used, provides a good description of S-shaped growth curves (Moser, 1958; Dabes et al., 1973). There has been much debate about which model provides a better fit to experimental data, and the goodness-of-fit obtained by use of particular models may often depend on the organism and environmental conditions (Bader, 1978).

Tan et al. (1994a) developed a general mathematical approach which was able to describe all types of single substrate limited bacterial growth using a statistical mechanical method. Most of the commonly used empirical models such as the Monod, Moser, and multiple Monod equations were shown to be special cases of this unifying approach. Furthermore, this statistical mechanical method showed that the parameters appearing in these empirical equations were physically meaningful and well defined. The general equation was found to provide a better description of the experimental data and to be more widely applicable because it includes the Monod, Moser and multiple Monod equations as special cases.

Despite recent progress in modeling substrate-limited bacterial growth, the Monod equation is the predominant equation used to describe the relation between bacterial growth and limiting substrate concentration in the soil and groundwater literature. This is partly due to the low substrate concentrations and slow growth rates in the groundwater environment under which conditions the Monod equation performs satisfactorily, therefore not warranting the added complexity of the improved models.

Using the Monod equation [Equation (9.6)], the growth rate of bacteria in the water phase is written as,

$$R_g^w = \frac{\mu_m S}{K_s + S}\theta C, \tag{9.7}$$

where θ is the water filled porosity (which may decrease as biomass accumulates in the pore space), C is the concentration of bacteria in the water phase defined as the mass of microbial cells per unit volume of water, and R_g^w is microbial growth rate (mass of bacteria per unit volume of porous medium per unit time). This equation describes microbial growth under balanced growth conditions, i.e., the environmental conditions are relatively stable and the composition of the biomass remains relatively constant.

Bacterial growth also occurs in the attached biomass phase. Assuming that access of bacteria to the substrate is not restricted by geometry of the attached biomass (i.e., adopting the strictly macroscopic approach), and that the growth rate in the attached biomass phase, R_g^a, is therefore a function of the substrate concentration in the water phase, the growth rate of bacteria is the same in both phases. In this case, Equation (9.7) can be rewritten for the attached biomass phase as (Corapcioglu and Haridas, 1984),

$$R_g^a = \frac{\mu_m S}{K_s + S} \rho_b \sigma, \tag{9.8}$$

where ρ_b is the density of bacteria (mass per unit volume of bacteria), and σ is the volume fraction of the biomass associated with the solid matrix (volume of bacteria per unit volume of porous material). It is noted that $\rho_b \sigma$ can be replaced by $\rho_s C^a$ where ρ_s is the bulk density of the porous medium and C^a is the mass of microorganisms associated with the solid matrix per unit mass of solid matrix.

The biofilm approach to bacterial growth in the attached phase, which accounts for the effect of biomass geometry on substrate access, also makes use of the Monod equation. Using our notation, the rate of biomass production in the attached biomass phase (biofilm) usually adopted is written (e.g. Taylor and Jaffe, 1990c):

$$R_g^a = M \int_0^{L_b} \frac{\mu_m S_b}{K_s + S_b} \rho_b \, \mathrm{d}z, \tag{9.9}$$

where M is the surface area of the water-biofilm phase boundary per unit volume of porous material, L_b is the biofilm thickness (z-direction), and S_b is the concentration of substrate at any given point within the biofilm. Taylor and Jaffe (1990a) simplified the above equation to

$$R_g^a = M L_b \rho_b \frac{\eta \mu_m S}{K_s + S} \,. \tag{9.10}$$

This simplification was achieved by assuming a steady state substrate concentration profile within the biofilm, expressing the integral term in Equation (9.9) in terms of the substrate concentration at the water-biofilm phase boundary and an effectiveness factor η (which accounts for diffusional resistance), and further approximating the substrate concentration at the water-biofilm phase boundary by that in the water phase. The effectiveness factor was introduced to correct for the diffusional resistance at the biofilm-water phase boundary. Equation (9.10) reduces to the macroscopic approach [Equation (9.8)] if we assume that there is no diffusional resistance at the biofilm-water boundary (i.e., $\eta = 1$), and that ML_b is equal to σ.

TABLE 9.1
Summary of mathematical expressions used for the rate of microbial growth and decay in the combined water and attached phases

	Growth Rate R_g	Decay Rate R_d
Macroscopic approach (Corapcioglu and Haridas, 1984)	$\frac{\mu_m S}{K_s + S}(\theta C + \rho_b \sigma)$	$b(\theta C + \rho_b \sigma)$
Biofilm approach (Taylor and Jaffe, 1990a)	$\frac{\mu_m S}{K_s + S}(\theta C + \eta M L_b \rho_b)$	$b(\theta C + \rho_b \sigma)$

The equations above are examples of those available for predicting bacterial growth limited by a single substrate, and are summarised in Table 9.1. More versatile models such as that of Tan et al. (1994a) also exist. In the natural environment, bacterial growth may be limited simultaneously by several nutrients and/or inhibited by the presence of metabolic by-products or the overabundance of substrates. Mathematical models are available that describe these cases. For conditions where such models are more appropriate, they can be incorporated readily into the macroscopic transport equations describing the transport of bacteria and growth limiting substrates in porous media in place of the Monod-based equations above (Roels, 1983; Moser, 1985; Tan et al., 1994a).

9.2.2.2 Decay

The life-span or survival of microorganisms obviously affects their transport and fate. The survival and decay of microorganisms in the environment have been reviewed by, e.g., Gerba et al. (1975), Crane and Moore (1984) and Yates and Yates (1988). Published values of survival times range from several weeks to a few months, and sometimes even longer. Survival is probably a more important process, from a practical viewpoint, in coarse porous media since bacteria move readily through such media. Thus the distance over which they may move and remain active depends on their survival. Survival times vary with species and depend on many environmental variables such as water status, pH, nutrient availability, temperature, microbial competition and solid phase properties. Often one variable becomes limiting, and controls the microbial decay rate. Availability of nutrients is often the limiting variable for introduced bacteria, which are typically copiotrophs (bacteria requiring abundant nutrients). By contrast most indigenous bacteria are autochthonous with relatively stable populations (Peterson and Ward, 1989).

Many processes may result in the decay of bacteria such as cell death, lysis (rupture of cell membrane), or exceeding the maintenance requirements due to endogenous metabolism (Bryers, 1988). Under certain laboratory conditions, decay is considered to be associated specifically with maintenance requirements. Bacterial cells in a nutrient-limited environment subject to the stress of starvation shift their

metabolic activities away from reproduction toward existing biological functions such as endogenous metabolism or respiration (Kurath and Morita, 1983).

Viruses are subject to inactivation, referring to the loss of infectivity toward host cells and the ability to cause disease. Vilker (1980) and Yates and Yates (1988) reviewed the inactivation of viruses and discussed its mathematical formulation in some detail. Although many factors have been reported to affect the inactivation of viruses, Gerba and Bitton (1984) stated that desiccation and temperature are the leading factors controlling virus persistence in soil. The inactivation of viruses is associated with the degradation of viral genome and dissociation of viral components.

The decay of bacteria or inactivation of viruses is usually described by an irreversible first order process, expressed mathematically as

$$R_d^w = b\theta C \tag{9.11}$$

in the water phase, and

$$R_d^a = b\rho_b\sigma \tag{9.12}$$

in the attached biomass phase, where R_d^w and R_d^a are the rates of decay in the water phase and in the attached or immobile bacterial phase (mass of bacteria per unit volume of porous medium and time), and b is the decay or inactivation constant.

Predation by other organisms also contributes to the decay of microorganisms. Attempts have been made to include predation of microorganisms in models of microbial transport in porous media (e.g., Peterson and Ward, 1989). When the prey concentration is low, a Monod type equation [Equation (9.6)] is valid for the description of the predation-prey relationship.

9.2.3 Attachment and Detachment Processes

9.2.3.1 Adsorption

Attraction between soil particles and microorganisms was probably first reported to exist by Cutler (1919) after noticing that soil particles settling through water carried suspended protozoa with them. The term "adsorption" was later accepted to describe the attraction of bacteria and viruses to solid surfaces (Peele, 1936; Daniels, 1972; Gerba and Bitton, 1984). The adsorption of bacteria and viruses to surfaces has been described by Marshall et al. (1971), Daniels (1972, 1980), Marshall (1976, 1980), Hendricks et al. (1979), Vilker and Burge (1980), and Gerba and Bitton (1984). Adsorption causes retention of bacteria and viruses in porous media and therefore delays their transport.

Brownian motion, active movement of motile bacteria, and fluid flow bring bacteria and viruses to the vicinity of the surfaces of soil particles where several types of forces may become effective in attracting or repelling them. These include the hydrodynamic force (i.e., the tendency of squeezing out fluid as the suspended particles approach the surfaces of the solid matrix), physical adsorption (caused by van der Waals forces), chemical adsorption (resulting from chemical reactions), or physico-chemical adsorption (referring to electrostatic interactions). There are

also other mechanisms such as hydrophobicity and polymer bridging proposed for bacteria adsorption (e.g., Marshall, 1976). Hydrophobicity is the tendency of bacteria having partly or wholly hydrophobic outer surfaces to be rejected from the aqueous phase and attracted towards any nonaqueous phase such as the surfaces of soil particles. Polymer bridging refers to the interaction with solid surfaces of fine extracellular polymer fibrils produced by some bacteria (Marshall, 1976). All but chemical adsorption and polymer bridging are characterized by their reversibility; polymer bridging is irreversible (Marshall, 1980) whereas chemical adsorption may be either reversible or irreversible. Physical adsorption is characterized by low energy involvement and rapid establishment of the adsorption equilibrium (Marshall, 1976). Electrostatic forces can result in either attraction or repulsion depending on the sign of the charge. If the surface of a soil particle has a net positive charge, bacteria and viruses may be attracted since most bacteria have a net negative charge at normal physiological pH (Plummer et al., 1962; Marshall, 1967; Corpe, 1970; Marshall, 1975; Corapcioglu and Haridas, 1984). When the surfaces are negatively charged, as are most soil particle surfaces, it is unlikely that the negatively charged microorganisms will be attracted to them. Electrostatic phenomena may not therefore be very significant for the attraction of bacteria to soil particles (Marshall, 1975). However, it is possible that the two negatively charged particles may be linked by positively charged organic macromolecules or by multivalent cations.

In most cases, sorption between microorganisms and solid surfaces is probably by double-layer mechanisms as explained by Marshall et al. (1971), Hattori (1973), Marshall (1975, 1976) and Corapcioglu and Haridas (1984). Around each negatively charged particle, there exists a diffuse electrical double-layer of cations. The cations are subject to two forces: Brownian motion which tends to distribute them uniformly in the solution, and negative surface charge attraction tending to concentrate them near surfaces. The thickness of the electrical double-layer decreases with increasing concentration of the solution and valence of the ions in it. When a microorganism and a solid particle approach one another, the electrical double-layers overlap and there is a tendency for mutual repulsion between the two particles. The two particles are also attracted to each other by van der Waals forces. The balance between these two opposing forces depends on ionic strength and the thickness of the double layer. There is a minimum interparticle distance, below which there is a net repulsion energy so that microorganisms do not come in direct contact with the solid surface. Therefore the adsorption by this mechanism may be reversible.

Vilker (1980) discussed the adsorption of viruses on solid surfaces and stated that the sorptive interactions with sand, clays and organic matter are rapid and reversible and do not eliminate infectivity for host cells. This is supported by the work of, for example, Moore et al. (1975) and Schaub and Sagik (1975), although Vilker and Burge (1980) suggested that virus adsorption could be a kinetic process.

9.2.3.2 Straining

Straining occurs when a particle in suspension flowing through a porous medium is too large to pass through the smallest pores or pore necks. Due to their extremely

small sizes (0.1–0.5 μm), straining is not considered to affect the movement of viruses in porous media. Many studies have noted the importance of straining in the retention of bacteria in porous media (Krone et al., 1958; Gerba et al., 1975; Corapcioglu and Haridas, 1984). Vinten and Nye (1985) stated that whatever the dimensions of the suspended particles and pores, some straining will occur. They explained that a particle suspended in laminar flow in a small pore may experience a torque caused by differences in flow velocity across its diameter. This sets up a rotational motion causing movement perpendicular to the principal direction of flow (i.e., towards stagnant volumes in pits and crevices in the walls of transmission pores). As these dead pore volumes fill, the rate of deposition will decrease, since shear forces and diffusion will tend to bring particles back into flowing suspension.

Straining is sometimes taken to be the main mechanism causing the retardation of the transport of larger bacteria in porous media (Butler et al., 1954; Wollum and Cassel, 1978; Gerba and Bitton, 1984; Smith et al., 1985). Matthess and Pekdeger (1981) pointed out that in gravelly aquifers, straining may not be effective in retaining bacteria due to their small sizes compared with the diameters of pore necks.

On geometric considerations alone, bacteria need to be of similar sizes to the pore necks if filtering or straining is to occur. The criterion for particle straining frequently used is the ratio of the suspended particle diameter (d_m) to the solid particle diameter (d_s). In filtration literature, straining is generally acknowledged to be an important retention mechanism if d_m/d_s is larger than 1/15 or 1/20 (Sakthivadivel and Einstein, 1970; Corapcioglu and Haridas, 1984). Other investigators (Sherard et al., 1984; McDowell-Boyer et al., 1986) have also discussed the criterion for the occurrence of straining. In most cases the criterion established takes into account only the geometry of the porous medium and the suspended particles. In practice, the magnitude of straining depends on the extent of coagulation or clumping of bacteria, which has the effect of increasing the unit size of the moving microorganism. Similarly, attachment of microorganisms to other mobile colloidal material may increase their susceptibility to straining.

The experimental and theoretical criterion provides only a qualitative measure of the importance of straining. For quantitative descriptive purposes, straining is included as part of the overall attachment process. This is described by a first order reaction and will be discussed in Section 9.2.4 below. The effects of clumping and movement with other mobile colloidal size particles are very case specific and difficult to quantify.

9.2.3.3 Sedimentation and Interception

Sedimentation is the settling of suspended particles caused by gravity. It occurs when the density of a particle is greater than that of the fluid. In principle, sedimentation of microorganisms can result in both their retention and their vertical movement in porous media. In practice, sedimentation has been found to be unimportant for bacteria (Smith et al, 1984; Gannon et al., 1991) largely because their densities are similar to that of water, ranging from 1.0 to 1.1 g cm^{-3} (Matthess and Pekdeger, 1981; Corapcioglu and Haridas, 1984; Characklis, 1990). Viruses are not expected to be affected by sedimentation because of their small size. However,

the attachment of bacteria and viruses to other colloidal size material may increase their effective density and enhance sedimentation. Matthess and Pekdeger (1981) concluded that the importance of sedimentation has been overestimated in the past and that straining is usually more important.

Interception, which refers to the collision of suspended particles with the surface of the soil particles, has also been shown to contribute to retention of suspended particles in soils (e.g., O'Melia and Stumm, 1967). However, interception is probably important only for relatively large particles (> 5 μm) (Yao et al., 1971) and is not likely to be significant for the retention of most bacteria (0.2–7 μm) and viruses (0.01–0.5 μm). Furthermore, it is likely to be indistinguishable from straining or adsorption.

9.2.4 Mathematical Description of Attachment and Detachment

The attachment and detachment processes described above fall into two categories: those which are equilibrium processes, and those which are kinetic or non-equilibrium processes. If more than one process is operating, their individual attachment and detachment rates are assumed to be additive to produce overall attachment and detachment rates.

9.2.4.1 Equilibrium Attachment and Detachment

The only attachment process in this category is adsorption. While adsorption can be both an equilibrium and a non-equilibrium process, it is most often described as an equilibrium one.

The equilibrium adsorption of viruses and bacteria to solid surfaces can be quantified by evaluating the relationship between microbial concentration in suspension and the concentration of viruses or bacteria associated with the solid surfaces, in a way analogous to chemical adsorption. Equilibrium adsorption of microorganisms has been variously reported to follow linear, Freundlich or Langmuir isotherms (Cookson, 1970; Filmer et al., 1971; Vilker and Burge, 1980; Matthess and Pekdeger, 1981; Moore et al., 1981; Tan et al., 1992). Reviewing adsorption of viruses, Vilker and Burge (1980) observed that virus adsorption isotherm data could generally be described by linear isotherms.

Although other expressions have been used, the most common representations of measured isotherms for microorganisms are the linear and the Freundlich isotherms. In practice, the choice of adsorption models may simply be one of convenience in adequately fitting the batch experimental data. The Freundlich isotherm is written as

$$C^a = k_f C^N, \tag{9.13}$$

where C^a is the mass of microorganisms adsorbed per unit mass of solid particles; C is the concentration of microorganisms in the solution defined as mass of microorganisms per unit volume of suspension and k_f and N are empirical constants. When N is equal to 1, the Freundlich isotherm becomes a linear isotherm.

In the absence of other processes, the difference between the rate of attachment (R_a) and the rate of detachment (R_y) as a result of reversible adsorption is defined by

$$R_a - R_y = \frac{\partial(\rho_s C^a)}{\partial t} . \tag{9.14}$$

Using the Freundlich isotherm this expression becomes

$$R_a - R_y = \rho_s k_f N C^{N-1} \frac{\partial C}{\partial t} . \tag{9.15}$$

If adsorption is irreversible or hysteretic, separate expressions for attachment and detachment rates are required.

9.2.4.2 Non-Equilibrium Attachment and Detachment

The non-equilibrium mechanisms for attachment and detachment include straining, interception and sedimentation as described above as well as the detachment of microorganisms caused by, for example, fluid shearing and diffusion back into the flowing water. Non-equilibrium (kinetic) adsorption also falls into this category. All mechanisms may operate simultaneously but are operationally indistinguishable (Gannon et al., 1991). They are therefore described by models that lump these processes together.

Expressions from the filtration literature have been adopted to describe the *attachment* rate of microorganisms in porous media. The colloid filtration approach recognises that the removal (attachment) rate of suspended particles is of first order with respect to the concentration of suspended particles. The basic expression for this approach is (Corapcioglu and Haridas, 1984) :

$$R_a = k_a \theta C, \tag{9.16}$$

where k_a represents the attachment coefficient, and is a function of porosity, flow rate, clogged pore volume and capture efficiency (Yao et al., 1971).

The attachment coefficient can be treated either as an empirical parameter in its own right or calculated from more complex empirical expressions. Taylor and Jaffe (1990c) obtained k_a using the empirical relationship of Deb (1969), who found that k_a was a linear function of the volume fraction of the attached biomass (σ). Equation (9.16) then becomes:

$$R_a = (\sigma c_1 + c_2)\theta C, \tag{9.17}$$

where c_1 and c_2 are constants.

Harvey and Garabedian (1991) explicitly set k_a proportional to the fluid velocity v_w, and used the colloid filtration expression of Yao et al. (1971) to evaluate it. Using their formulation, Equation (9.16) becomes:

$$R_a = \frac{3}{2} \frac{(1-\theta)}{d_s} \alpha \gamma v_w \theta C, \tag{9.18}$$

TABLE 9.2
Summary of mathematical formulation for microbial attachment and detachment in porous media

	Attachment Rate R_a	Detachment Rate R_y
Equilibrium		
Freundlich adsorption	$R_a - R_y = \rho_s k_f N C^{N-1} \frac{\partial C}{\partial t}$	
Non-Equilibrium		
Corapcioglu and Haridas (1984)	$k_a \theta C$	$k_y \rho_b \sigma^h$
Taylor and Jaffe (1990a)	$(\sigma c_1 + c_2)\theta C$	$b_s \rho_b \sigma + b'_s \rho_b M L_b \frac{\mu_m \eta S}{K_s + S}$
Tan et al. (1994b)	$k_a \left(1 - \frac{C^a}{C^a_m}\right) \theta C$	$k_y \rho_b \sigma$

where α and γ are empirical parameters referred to by Yao et al. as the collision efficiency factor and the single-collector efficiency, respectively.

Tan et al. (1994b) found experimentally that once the concentration of attached bacteria (C^a) reached a threshold value, the porous matrix did not retain bacteria as effectively. They therefore proposed the following equation to describe the retention of biomass in the porous medium:

$$R_a = k_a \left(1 - \frac{C^a}{C^a_m}\right) \theta C, \tag{9.19}$$

where C^a_m is the maximum concentration of retained bacteria, i.e., the maximum mass of bacterial cells that can be retained by unit mass of solid material. They reported that C^a_m was related not only to the intrinsic properties of the porous material but also to the bacterial cell concentration in solution.

Similarly, a number of different expressions have been proposed for the *detachment* rate. Rittmann (1982) attributed the detachment of biomass to loss due to shear and used the following expression,

$$R_y = b_s \rho_b \sigma, \tag{9.20}$$

where b_s is the specific shear loss coefficient, which is a function of the fluid shear stress and the thickness of the layer of attached microorganisms. Fluid shear stress is related to the permeability of the porous medium and the Darcy velocity. Equation (9.20) is analogous to an expression existing in the filtration literature (e.g., Rajagopalan and Chu, 1982). Corapcioglu and Haridas (1984) expressed detachment rate as a power of the attached bacteria concentration, i.e.,

$$R_y = k_y \rho_b \sigma^h, \tag{9.21}$$

where k_y is the detachment coefficient and h is an empirical constant, often assumed to be unity (Tan et al., 1994b). The detachment coefficient is thought to be a function of C and flow rate. Speitel and DiGiano (1987) proposed a two term equation with the first term similar to that proposed by Rittmann (1982) and the second term describing the dependence of biomass shearing on microbial growth rate. This model is based on the biofilm approach and is written as,

$$R_y = b_s \rho_b \sigma + b'_s M \int_0^{L_b} \frac{\mu_m S^b}{K_s + S_b} \rho_b \, \mathrm{d}z, \tag{9.22}$$

where b'_s is a dimensionless parameter describing the biological aspects of shearing. Taylor and Jaffe (1990c), applying the work of Atkinson and Davies (1974), simplified this equation to

$$R_y = b_s \rho_b \sigma + b'_s M L_b \frac{\mu_m \eta S}{K_s + S} \rho_b. \tag{9.23}$$

As in section 9.2.2.1, if we assume that ML_b is equal to σ and $\eta = 1$, then an equivalent expression can be written for the macroscopic approach.

It is thus evident that there are a variety of mathematical expressions describing attachment and detachment of bacteria to and from solid surfaces while moving in porous media. None of these expressions, which are summarised in Table 9.2, are universally accepted.

9.3 Transport of Microorganisms in Porous Media

The flow of water and solutes through porous media are described mathematically at a "macroscopic" or "Darcy" scale, ignoring microscopic variations of the flow velocity vector that occur within a representative elementary volume. The same approach is applied to transport of bacteria and viruses. Thus, the transport equation is written in terms of variables that can be quantified at this scale i.e., averaged over a representative elementary volume of a porous medium. This has been the approach followed in the preceding section in which the individual processes affecting microbial transport were enunciated.

As described in Section 9.2 above, many of the processes affecting bacteria in the soil and groundwater environment are not relevant to viruses. We therefore treat bacterial transport first and will show in a later section how this simplifies for the description of virus transport.

9.3.1 General Transport Equation for Microorganisms

As for solutes, the general transport equation for microorganisms in porous media is based on the continuity or mass balance equation. This states that the rate of change of the total biomass in an elementary volume of porous material is equal to the divergence of the flux through that volume plus the rate of contribution to the biomass of any sources in that volume (growth) less the rate of removal of the biomass in that volume by any sinks or by decay. At any point in time, the

total biomass in the elementary volume is the sum of that in solution (θC) plus that in the attached phase ($\rho_b \sigma$ or $\rho_s C^a$). In three dimensional space the continuity equation for a particular microorganism is therefore written as

$$\frac{\partial(\theta C)}{\partial t} + \frac{\partial(\rho_s C^a)}{\partial t} = -\nabla \cdot J + R_g - R_d, \tag{9.24}$$

where J is the combined microbial flux resulting from all mechanisms, and R_g and R_d are respectively the growth and decay rates of microorganisms combined for the water and attached biomass phases (mass of microorganisms per unit volume and time).

The combined flux of bacteria is the sum of the individual fluxes identified in Section 9.2.1, namely advection (J_v), chemotaxis (J_x), and the sum of diffusion, dispersion and random motility (J_d). Therefore, using Equations (9.1), (9.3) and (9.4), J is given by

$$J = v_w \theta C + v_x \theta C - \tau \theta D \nabla C \tag{9.25}$$

or

$$J = v \theta C - \tau \theta D \nabla C, \tag{9.26}$$

where v is the sum of the velocity of water flow (v_w), and active chemotactic velocity (v_x), and D is the sum of the diffusion coefficient (D_B), the hydrodynamic dispersion coefficient (D_h) and the random motility coefficient (D_t). For practical purposes, it has been concluded that the tortuosity, τ, can be set to unity for solute transport (e.g., Bond, 1987). This assumption has also been adopted for descriptions of microbial transport (Corapcioglu and Haridas, 1984; Taylor and Jaffe, 1990c; Harvey and Garabedian, 1991), and tortuosity will be omitted from all subsequent equations.

The combined growth and decay rates may be obtained from the expressions presented in Section 9.2.2. These have been summarised in Table 9.1 for the macroscopic biomass approach (following Corapcioglu and Haridas, 1984) and the biofilm approach (based on Taylor and Jaffe, 1990c).

The rate of change of the immobile or attached biomass phase can be written as

$$\frac{\partial(\rho_s C^a)}{\partial t} = R_a - R_y + R_g^a - R_d^a, \tag{9.27}$$

where R_a and R_y are the rates of microbial attachment and detachment (mass per unit volume and time), and R_g^a and R_d^a are the rates of growth and decay of microorganisms in that phase. As described in Section 9.2.4 and summarised in Table 9.2, there are various formulations for attachment and detachment rates depending on the assumptions that are most appropriate to the particular case. While most definitions of attachment rate are dependent on concentration of microorganisms in solution alone, most detachment rates are expressed in terms of microbial concentration in the attached phase. It is therefore impossible to obtain a single equation in terms of solution concentration alone as may be done for solutes (Bond and Phillips, 1990). Equations (9.24) and (9.27) must therefore be solved simultaneously for C and C^a, except in special cases.

9.3.2 Special Cases of the Microbial Transport Equation

As discussed in Section 9.2, not all of the processes included in Equations (9.24) to (9.27) are applicable to viruses. Similarly, for particular situations, simplifying assumptions can be made allowing certain terms to be omitted from these general equations for bacterial transport. In the section below, we describe some special cases for both bacteria and viruses.

9.3.2.1 Transport of bacteria

Tan et al. (1992) described a series of transport experiments in sand columns which were of short duration so that growth and decay could be ignored. Furthermore, the relative size of the bacteria and the sand grains was such that the only form of attachment and detachment that was relevant was adsorption, described by the Freundlich equation [Equation (9.13)]. The high water flow velocities also meant that the chemotactic velocity could be ignored. For this case, Equations (9.24) through (9.27) can be combined and simplified to produce a single equation in terms of solution bacteria concentration alone, namely:

$$\frac{\partial(\theta C)}{\partial t} + \rho_s k_f N C^{N-1} \frac{\partial C}{\partial t} = \frac{\partial}{\partial x}(\theta D \frac{\partial C}{\partial x} - v_w \theta C). \tag{9.28}$$

Harvey and Garabedian (1991) also assumed growth and decay to be unimportant for the description of bacterial transport through groundwater. They further assumed that adsorption was equilibrium and linear, that non-equilibrium attachment could be described by a first order rate equation with the attachment rate, k_a, given by Equation (9.18), and that detachment was unimportant. With regard to movement mechanisms, they assumed that the chemotactic velocity was negligible, along with the Brownian diffusion and motility diffusion coefficients. This also resulted in an equation in terms of C alone, which for one dimensional steady flow is:

$$\frac{\partial(\theta C)}{\partial t} + \rho_s k_f \frac{\partial C}{\partial t} = \theta D \frac{\partial^2 C}{\partial x^2} - v_w \theta \frac{\partial C}{\partial x} - k_a \theta C. \tag{9.29}$$

Tan et al. (1994b) examined the case of one-dimensional steady transport of bacteria in which no substrates were applied and the temperature was kept low so that growth and decay of bacteria could be ignored. They assumed that chemotactic movement was negligible and that attachment and detachment of bacteria were described respectively by Equation (9.19) and by Equation (9.21) with $h = 1$. The resulting simplifications of Equations (9.24) and (9.27) are:

$$\theta \frac{\partial C}{\partial t} = \theta D \frac{\partial^2 C}{\partial x^2} - v_w \theta \frac{\partial C}{\partial x} - k_a \frac{(C_m^a - C^a)}{C_m^a} \theta C + k_y \rho_s C^a, \tag{9.30}$$

$$\rho_s \frac{\partial C^a}{\partial t} = k_a \frac{(C_m^a - C^a)}{C_m^a} \theta C - k_y \rho_s C^a. \tag{9.31}$$

Taylor and Jaffe (1990c) assumed that advection and hydrodynamic dispersion were the dominant movement mechanisms in their one-dimensional model of bacteria transport during steady water flow. They related growth in the water phase to

the substrate concentration, S, with the simple Monod equation [Equation (9.7)], and adopted the simplified biofilm approach [Equation (9.10)] to describe growth in the attached phase. Decay was also included, using Equations (9.11) and (9.12). Attachment was described by Equation (9.17), in which the attachment rate coefficient was expressed in terms of the constants c_1 and c_2, while detachment was described by Equation (9.23), which is based on the biofilm approach. For these assumptions, Equations (9.24) and (9.27) become

$$\frac{\partial(\theta C)}{\partial t} + \frac{\partial(\rho_b \sigma)}{\partial t} = \frac{\partial}{\partial x}\left(\theta D \frac{\partial C}{\partial x}\right) - \theta v_w \frac{\partial C}{\partial x} + \\ + \frac{\mu_m S}{K_s + S}(\eta M L_b \rho_b + \theta C) - b(\rho_b \sigma + \theta C), \tag{9.32}$$

$$\frac{\partial(\rho_b \sigma)}{\partial t} = (1 - b'_s)\left(\frac{\eta \mu_m S}{K_s + S}\right) M L_b \rho_b - (b + b_s)\rho_b \sigma + (c_1 \sigma + c_2)\theta C. \tag{9.33}$$

9.3.2.2 *Transport of viruses*

In Section 9.2 it was pointed out that because of their size or because of the absence of hosts in the soil or groundwater environment viruses are not affected by all the processes which all affect the transport of bacteria. Growth, chemotaxis, motile diffusion, straining, sedimentation and interception are all unimportant for the transport of viruses. Omitting a description of these processes from the general transport equations [Equations (9.24) to (9.27)] results in a considerable simplification. Assuming that the inactivation (decay) of viruses is described by Equations (9.11) and (9.12), and that adsorption can be described by a linear equilibrium adsorption isotherm (Vilker and Burge, 1980), and ignoring the processes mentioned above, Equations (9.24) to (9.27) can be combined and written as:

$$(\theta + \rho_s k_f)\frac{\partial C}{\partial t} = \nabla \cdot [\theta(D_h + D_B)\nabla C - \theta v_w C] - b(\theta C + \rho_s C^a). \tag{9.34}$$

9.3.3 Simultaneous Transport of Substrates with Bacteria

The transport of growth-limiting substrates often occurs simultaneously with that of bacteria during soil and groundwater flow and therefore forms a natural part of a discussion of the transport of bacteria in porous media that includes growth. When transport of substrates occurs at the same time as transport of bacteria, the concentration of substrate is unlikely to remain constant. Thus, a substrate transport equation must be derived and solved simultaneously with that for bacteria transport in order to obtain the substrate concentration on which growth is based. As in Section 9.2.2.1, it is assumed here that growth is controlled by a single rate-limiting substrate. More complex models (Roels, 1983) dealing with multiple substrate limitation exist and can be easily applied.

The subsurface transport of solutes is well understood (Bear, 1972). The main processes affecting transport include advection with the flowing water, diffusion

and hydrodynamic dispersion relative to the water in response to concentration gradients, and sorption by the solid phase by specific adsorption reactions or ion exchange. Transport of a solute that is a substrate for microbial growth is also affected by its consumption. Taking these processes into account, the transport equation for biologically reactive solutes can be written as:

$$\frac{\partial(\rho_s S^a)}{\partial t} + \frac{\partial(\theta S)}{\partial t} = -\nabla \cdot [-\theta D_s \nabla S + \theta v_w S] - R_b, \tag{9.35}$$

where S_a is the mass of substrate adsorbed per unit mass of solid matrix, S the mass of substrate per unit volume of water, D_s the diffusion-dispersion coefficient of the substrate, and R_b the rate of substrate consumption by bacteria combined for the water and attached biomass phases.

Different approaches for describing the spatial distribution of bacteria in the pore space result in different mathematical expressions for the substrate consumption rate. This is because the specific growth rate in the water phase may differ from that in the attached biomass phase as seen in Section 9.2.2.1. The specific growth rate of bacteria in the attached biomass phase has been variously assumed to be a function of the substrate concentration in the water phase, that in the attached biomass phase, or that in the "immobile water phase" that lies between the free water phase and "attached biomass phase" (Baveye and Valocchi, 1989). The assumption that the specific growth rate is related to substrate concentration in the attached biomass phase, or in the immobile water phase, is equivalent to assuming that the response of biomass is determined directly by the environment that it "sees" in its immediate vicinity. This approach introduces considerable complexity and extra equations need to be included to ensure mathematical closure. We assume here that any immobile water phase present is indistinguishable from the attached biomass phase.

The simplest approach is to assume that the substrate concentration throughout the attached biomass phase is equal to that in the external soil solution, i.e., that there is no impedance of access of the bacteria to the substrate as a result of their attachment to the surface. This is the strictly macroscopic approach introduced in Section 9.1.2. The specific growth rates for both phases are therefore dependent on the substrate concentration of the water phase only. This following expression for R_b results [c.f. Equations (9.7) and (9.8)]:

$$R_b = \frac{1}{Y}\frac{\mu_m S}{K_s + S}(\theta C + \rho_b \sigma), \tag{9.36}$$

where Y is the maximum yield coefficient expressed as the amount of biomass produced per unit mass of substrate consumed.

The biofilm approach, on the other hand, assumes that the rate of substrate utilization by biomass in the attached phase is controlled by the transfer of substrate from the water phase to the biofilm phase as the biofilm grows, and by the transfer of substrate from the water phase to the biofilm phase by diffusion across the water-biofilm phase boundary. Taylor and Jaffe (1990c) expressed this as follows [c.f. Equations (9.7) and (9.10)]:

$$R_b = \frac{1}{Y}\left(\frac{\mu_m S}{K_s + S}\right)\theta C + \frac{1}{Y}\left(\frac{\eta\mu_m S}{K_s + S}\right) M L_b \rho_b. \tag{9.37}$$

Depending on the choice of model, either Equations (9.36) or (9.37) is substituted into Equation (9.35), which must then be solved simultaneously with the appropriate bacteria transport equation.

9.4 Comparison of Models with Experiments

Few experimental studies have been carried out to investigate transport of microorganisms in porous media and to test both specific mathematical models and the conceptual approaches. Of those studies reported, most have found that specific models provide a good description of the experimental results, but very few models have been validated using independently measured parameters. Similarly, few models have been tested beyond their initial development or applied to other scenarios. Much experimental work remains to be done before models of microbial transport can be used with confidence to make *a priori* predictions. Nevertheless, existing experimental studies have advanced the modeling efforts of microbial transport during the last decade or so and have provided valuable information for the study and evaluation of microbial transport and fate in soil and ground water.

We include a few examples of recent experimental studies here as illustrations of the capability of the mathematical models presented in earlier sections.

9.4.1 Transport of Bacteria

9.4.1.1 Tan et al. (1992)

Tan et al. (1992) measured the distribution *of Pseudomonas fluorescens* introduced during unsteady, unsaturated flow through columns of fine and coarse sand. In these experiments, d_m/d_s was very small, thus eliminating straining. In addition the experiments were of short duration ($<$ 1 hour) and no substrate was added. Therefore, adsorption was the only important attachment mechanism, and there was no need to consider growth or decay during the experiments. Thus the appropriate equation to describe these experiments is Equation (9.28). In the absence of independent measurements of the dispersion coefficient, Tan et al. (1992) solved Equation (9.28) without the dispersion term to obtain an expression for the average position of the bacteria front. Using independently measured adsorption isotherms for the bacteria, they accurately predicted the concentration fronts of bacteria in four out of five cases tested.

9.4.1.2 Harvey and Garabedian (1991)

A small scale field study was conducted by Harvey and Garabedian (1991) to evaluate a model incorporating attachment/detachment mechanisms as well as dispersion and advection [Equation (9.29)]. The assumptions in this model have been discussed in Section 9.3.2.1. The study followed the movement of bacteria with the natural flow of groundwater in sandy aquifer sediments. Labelled bacteria were injected into the aquifer at 8.5 m and 9.1 m below the land surface. Their

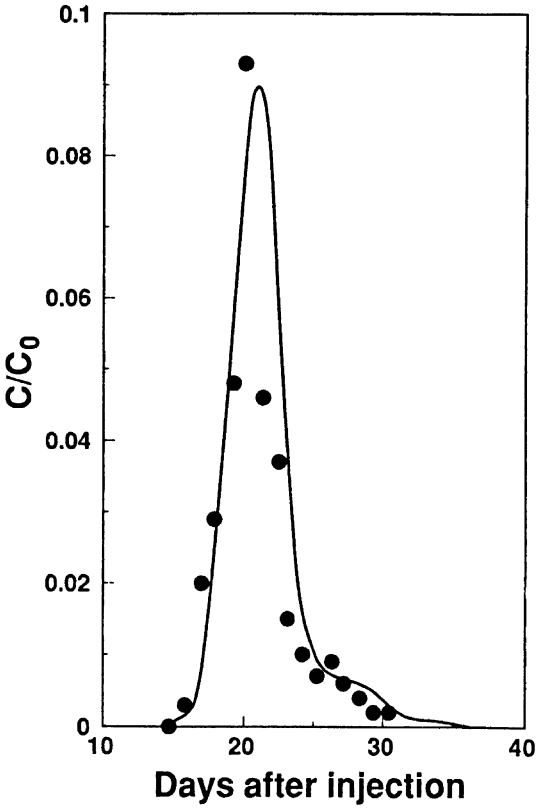


Fig. 9.1. Comparison of measured (points) and simulated (solid line) breakthrough of bacteria at a depth of 9.1 m and a distance of 6.8 m from injection (after Harvey and Garabedian, 1991).

concentration was monitored using a row of multilevel sampling devices at a point 6.8 m down gradient from the injection point as they moved with the natural flow of groundwater. Harvey and Garabedian (1991) obtained an analytical solution to Equation (9.29) for a pulse input, and calibrated it to the observed bacteria breakthrough curves.

Figure 9.1 shows their comparison of the measured and simulated [using Equation (9.29)] bacteria breakthrough at the depth of 9.1 m below the land surface. Although growth and decay were not incorporated into the transport model, the observed behavior appeared to be described quite well. The assumptions in Equation (9.29) include linear equilibrium adsorption. Harvey and Garabedian (1991) also tested a slightly more complex model that included kinetic adsorption and desorption, but found little difference between the ability of the two models to describe the observed concentration distributions of bacteria.

9.4.1.3 Tan et al. (1994b)

Breakthrough curves of bacteria were measured during steady-state water flow at three water flow velocities (0.05, 0.1 and 0.2 mm s^{-1}) in columns packed with a

coarse aquifer sand. Three different initial cell concentrations (10^7, 10^8, 10^9 cells mL^{-1}) were used. In order to limit growth and decay of bacteria, all experiments were conducted at a temperature of $3 \pm 1.5°C$ and no substrates were added. For these conditions, the appropriate transport equations [Equations (9.30) and (9.31)] were solved using the finite difference Crank-Nicolson scheme combined with the Newton-Raphson iterations. Water flow velocity was calculated from the measured inflow and average water content, and the dispersion coefficient was estimated from chloride breakthrough curves. The remaining three parameters (i.e., k_a, k_y and C_m^a) were adjusted using the downhill simplex optimization technique to minimize the sum of the squares of deviations between model simulations and experimental data to obtain the best fit. The model described the experimental data very well under all the experimental conditions tested. Figure 9.2a and 9.2b illustrate examples of these simulations together with observed bacterial breakthrough curves.

Tan et al. (1994b) also used their experimental data to test the alternative formulations for attachment and detachment proposed by Corapcioglu and Haridas (1984) [Equations (9.16) and (9.21)]. Simulations obtained when these formulations replaced those of Tan et al. (1994b) in Equations (9.30) and (9.31) are also presented in Figure 9.2. In this case the three parameters k_a, k_y and h were adjusted to obtain the best fit of simulations to experimental data. The model using the attachment/detachment expressions of Corapcioglu and Haridas (1984) did not represent the experimental data as well as Equations (9.30) and (9.31). The better representation of the experimental data using the model of Tan et al. (1994b) was achieved by the inclusion of a maximum retention capacity, C_m^a, i.e., by assuming that once a finite number of attachment sites in a given volume of porous medium are saturated, bacteria can move through that volume without significant retention. It was found that C_m^a increased with cell concentration and ionic concentration of microbial suspension, but was unaffected by water flow velocity.

9.4.1.4 Taylor and Jaffe (1990c)

Taylor and Jaffe (1990c) tested their model for microbial transport [Equations (9.32) and (9.33)] and simultaneous substrate transport [Equations (9.35) and (9.37)] against experimental data obtained by Taylor and Jaffe (1990a). Steady flow experiments were conducted using columns packed with sand (0.59–0.84 mm in diameter) to study the growth and transport of bacteria cultured from primary sewage and activated sludge. Methanol, added with the inflowing water together with excess of other nutrients as mineral salts, was assumed to be the growth-limiting substrate. The biomass distribution in the column was determined at various intervals up to 85 days and this was used to calculate biofilm thickness as a function of distance as described by Taylor and Jaffe (1990b).

The models for the transport of microorganisms and growth-limiting substrate derived by Taylor and Jaffe (1990c) were based on the biofilm concept for the distribution of microorganisms in the pore space. Other assumptions used in deriving their models are discussed in Sections 9.3.2.1 and 9.3.3. The system of equations describing the transport of bacteria and substrate were solved numerically.

A large number of parameters were required to use this model to simulate biomass growth and transport. Some of these parameters were determined from

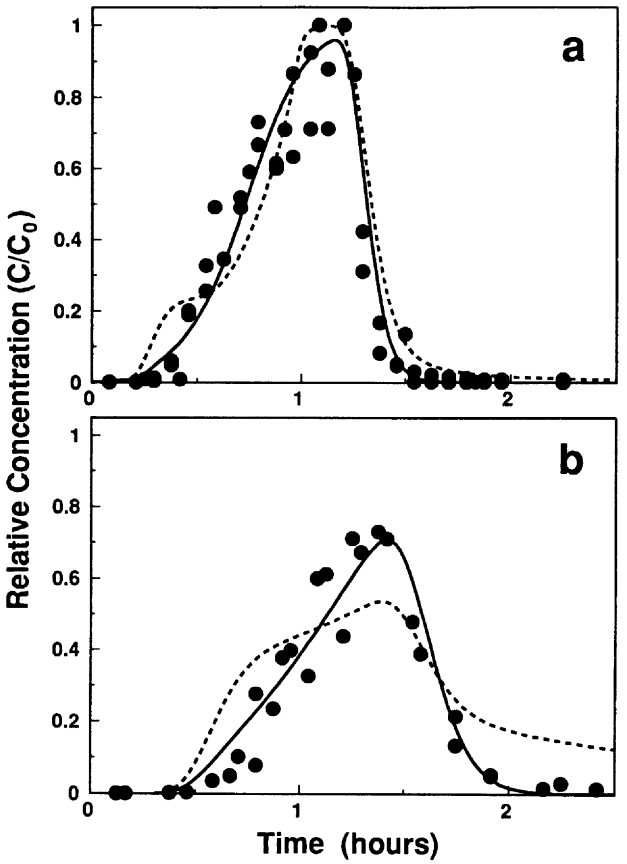


Fig. 9.2. Comparison of measured bacteria breakthrough curves (points) with simulations using the model of Tan et al. (1994b) [Equations (9.30) and (9.31)] (—) and Corapcioglu and Haridas (1984) (- - -) following injection of a pulse with a concentration of 10^9 cells mL^{-1} (after Tan et al., 1994b). (a) $v_w = 0.1$ mm s^{-1} (b) $v_w = 0.05$ mm s^{-1}.

independent measurements or observations of tracer transport, while others were calibrated by fitting the simulated to the measured biofilm thickness data for one column. The parameters thus obtained were then used to "predict" biofilm thickness in a second column in which a different water flow velocity and a different methanol input concentration were imposed. The results of the simulation are compared with the measured biofilm thickness distribution in the second column in Figure 9.3. The model predicted the observed advance of biofilm thickness through the column reasonably well at some times (early and late) but not at others. A similar feature was observed in the calibration phase using data from the first column. The modeled and observed steady state biofilm thickness were in good agreement.

Taylor and Jaffe (1990c) included the diffusional resistance of substrate at the biofilm and water boundary in their model formulation through the introduction of an effectiveness factor, η. This parameter was, however, set to equal to unity. It was also assumed that ML_b (i.e. the product of the specific surface of the biofilm and the

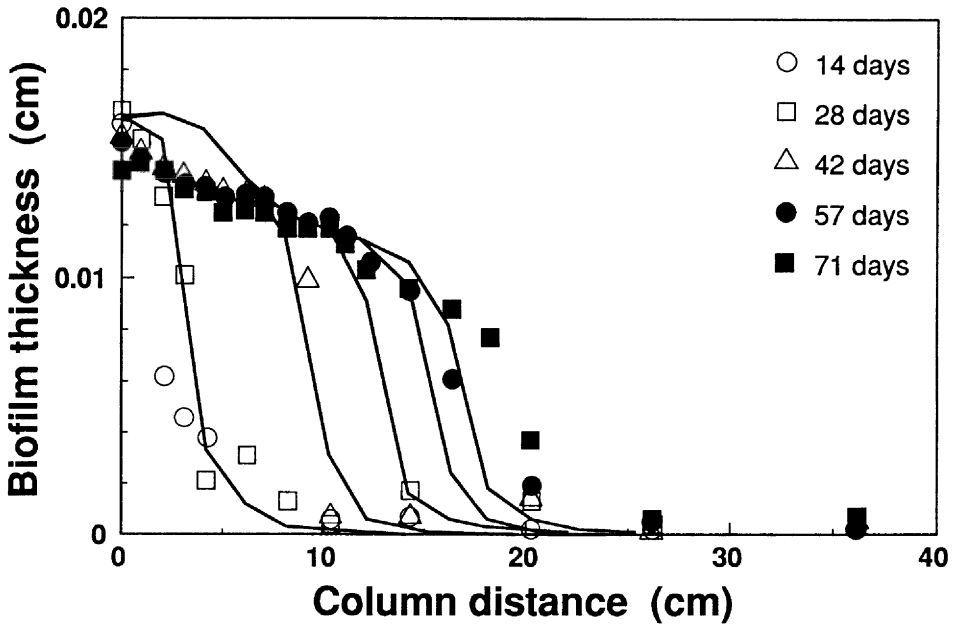


Fig. 9.3. Measured (points) and simulated (lines) distributions of biofilm thickness with distance at different times after introduction of bacteria at the column surface (after Taylor and Jaffe, 1990c).

biofilm thickness) is equal to the volume fraction of bacteria in the attached phase, σ, and that the specific growth rate of microorganisms in the biomass phase is a function of the substrate concentration in the water phase. When these assumptions are made, this approach becomes identical to the strictly macroscopic approach.

9.4.2 Transport of Viruses

Yates and Ouyang (1992) tested a simple model for transport of viruses based on Equation (9.34) against data from two experimental studies. Only one of these tests is presented here, namely that using the data of Powelson et al. (1990), who determined the transport of MS-2 bacteriophage through columns of a loamy fine sand under both saturated and unsaturated conditions. Soil water samples were taken daily from sampling ports at 0.2 m, 0.4 m, 0.8 m, and 1.05 m from the column inlet for the determination of viral concentration.

Yates and Ouyang (1992) used a numerical solution of Equation (9.34) to simulate the experimental data for day 4 for the unsaturated soil column. It is not clear how all the parameters required for this model were obtained, but at least some of them were calculated from the measured data set, rather than independently measured. The model simulations are compared with the experimental data in Figure 9.4, which shows that there is reasonable agreement between the two.

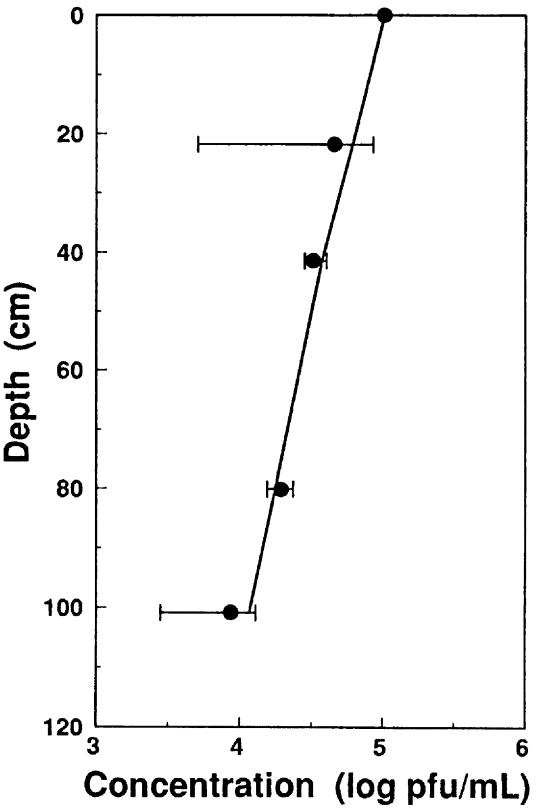


Fig. 9.4. Comparison of measured (points with 95% confidence limits) and simulated (lines) distributions of virus concentration in soil water 4 days after introduction of the virus into the soil column (after Yates and Ouyang, 1992).

9.4.3 Summary of Experimental Studies

The experimental studies described above for a variety of experimental conditions appear to have all been well represented by the various models applied to them. The models vary in conceptual approach and complexity and it is not possible to say that any one model is better than the others, particularly as different models have rarely been compared with each other. Furthermore, the ability of the models to match experimental data by no means proves that the models are based on correct descriptions of the chemical, physical, and biological processes occurring. Many processes are often lumped together and many parameters required are difficult to measure independently.

9.5 Summary and Conclusions

This chapter provides an overview of the current state of progress with the development of models for the subsurface transport of microorganisms. Recent years have seen the development of a number of mathematical models which vary in

conceptual approach and degree of complexity. Most models are based on the advection-dispersion equation, with terms added to account for processes such as growth, decay, attachment, and detachment. When transport of bacteria is considered, the simultaneous transport of growth-limiting substrate has often been included as well.

The principal differences between models depend on which of these processes are included and how they are described. Expressions used for virus transport are, in general, simpler than those used for bacteria transport because fewer processes affect viruses. One of the key differences between bacteria transport models results from the assumption made concerning the pore-scale distribution of bacteria, viz. biofilm versus macroscopic approach. It has been shown, however, that even when a biofilm approach is adopted in the theoretical development, simplifying assumptions are often made which reduce the model to the macroscopic approach. It is not expected that these two different approaches for pore-scale bacteria distribution will result in operationally significant differences in the capabilities of specific models to predict the transport of microorganisms and growth-limiting substrates.

Several of the different models are able to describe experimental data adequately. However, few models have been tested for their ability to provide *a priori* prediction of microbial transport using independently measured parameters.

We conclude that modeling the transport of microorganisms in porous media has not reached a stage where independently measured hydrological, physical and biological parameters can be used to predict transport and fate of microorganisms in the subsurface environment. This, together with simplifying the models for use as management tools, are the challenges for future work in this field.

Acknowledgement

We thank Dr David Smiles for interesting us in this topic several years ago, and for helpful comments on this manuscript.

References

Adler, J. 1969. 'Chemoreceptors in bacteria'. Science. 166:1588-1597.

Atkinson, B., and I. J. Davies. 1974. 'The overall rate of substrate uptake (reaction) by microbial films: 1. A biological rate equation'. Trans. Inst. Chem. Eng. 52: 248-259.

Bader, F. G. 1978. 'Analysis of double-substrate limited growth'. Biotechnol. Bioeng. 20:183-202.

Bales, R.C., S.R. Hinkle, T.W. Kroeger, K. Stocking and C.P. Gerba. 1991. 'Bacteriophage adsorption during transport through porous media: Chemical perturbations and reversibility'. Environ. Sci. Technol. 25:2088-2095.

Baveye, P., and A. J. Valocchi. 1989. 'An evaluation of mathematical models of the transport of biologically reacting solutes in saturated soils and aquifers'. Water Resour. Res. 25:1413-1421.

Bazin, M.J., P.T. Saunders, and J.I. Prosser. 1976, 'Models of microbial interactions in the soil'. CRC Critical Rev. Microbiol. 4:463-498.

Bear, J. 1972. Dynamics of fluids in porous media. American Elsevier, New York.

Berg, H. C., and D. A. Brown. 1972. 'Chemotaxis in Escherichia coli analyzed by three-dimensional tracking'. Nature. 239:500-504.

Biggar, J.W., and D.R. Nielsen. 1980. 'Mechanisms of chemical movement in soils'. p. 213-227. In A. Banin and U. Kafkafi (ed.) Agrochemicals in soils. International Irrigation Information Centre, Bet Dagan, Israel.

Bitton, G., N. Lahav, and Y. Henis. 1974. 'Movement and retention of Klebsiella aerogenes in soil columns'. Plant Soil. 40:373-380.

Bond, W.J. 1987. 'Solute transport during unsteady, unsaturated soil water flow: The pulse input'. Aust. J. Soil Res. 25:223-241.

Bond, W. J., and I. R. Phillips. 1990. 'Approximate solutions for cation transport during unsteady, unsaturated soil water flow'. Water Resour. Res. 26:2195-2205.

Bond, W. J., and D. E. Smiles. 1983. 'Influence of velocity on hydrodynamic dispersion during unsteady soil water flow'. Soil Sci. Soc. Am. J. 47:438-441.

Bouwer, H. 1984. 'Elements of soil science and groundwater hydrology'. p. 9-38. In G. Bitton and C. P. Gerba (ed.) Groundwater pollution microbiology. Wiley-Interscience, New York.

Bryers, J.D. 1988. 'Modeling biofilm accumulation'. p. 109-144. In M.J. Bazin and J.I. Prosser (eds.) Physiological Models in Microbiology (vol. 2). CRC Press, Boca Raton. Fla.

Butler, R. G., G. T. Orlob, and P. H. McGauhey. 1954. 'Underground movement of bacterial and chemical pollutants'. J. Am. Water Works Assoc. 46:97-111.

Characklis, W. G. 1990. 'Biofilm processes'. p. 195-231. In W. G. Characklis and K. C. Marshall (ed.) Biofilms. Wiley-Interscience, New York.

Chen, Y.M., L.M. Abriola, P.J.J. Alvarez, P.J. Anid, and T.M. Vogel. 1992. 'Modeling transport and biodegradation of benzene and toluene in sandy aquifer material: Comparisons with experimental measurements'. Water Resour. Res. 28:1833-1847.

Cookson. J. T. 1970. 'Removal of submicron particles in packed beds'. Environ. Sci. Technol. 4:128-134.

Corapcioglu, M. Y., and A. Haridas. 1984. 'Transport and fate of microorganisms in porous media: A theoretical investigation'. J. Hydrol. 72:149-169.

Corapcioglu, M. Y., and A. Haridas. 1985. 'Microbial transport in soils and groundwater: A numerical model'. Adv. Water Resour. 8:188-200.

Corpe, W. A. 1970. 'Attachment of marine bacteria to solid surfaces'. p.73-87. In R. S. Manly (ed.) Adhesion in biological systems. Academic Press, New York.

Crane, S. R., and J. A. Moore. 1984. 'Bacterial pollution of groundwater: A review'. Water, Air and Soil Pollut. 22:67-83.

Craun, G. F. 1985. 'A summary of waterborne illness transmitted through contaminated groundwater'. J. Environ. Health. 48:122-127.

Cutler, D. W. 1919. 'Observations on soil protozoa'. J. Agr. Sci. 9:430-444.

Dabes, J.N., R.K. Finn, and C.R. Wilke. 1973. 'Equations of substrate-limited growth: The case for Blackman kinetics'. Biotechnol. Bioeng. 15:1159-1177.

Dahlquist, F.W., P. Lovely, and D.E. Koshland. 1972. 'Quantitative analysis of bacterial migration in chemoxis'. Nature New Biol. 236:120-123.

Daniels, S. L. 1972. 'The adsorption of microorganisms onto surfaces: A review'. Dev. Indust. Microbiol. 13:211-253.

Daniels, S. L. 1980. 'Mechanisms involved in sorption of microorganisms to solid surfaces'. p. 7-58. In G. Bitton and K. C. Marshall (ed.) Adsorption of microorganisms to surfaces. Wiley-Interscience, New York.

Deb, A. K. 1969. 'Theory of sand filtration'. J. Sanit. Eng. Div. Am. Soc. Civ. Eng. 95:399-422.

Dickinson, R.A. 1991. 'Problems with using existing transport models to describe microbial transport in porous media'. p. 21-47. In C.J. Hurst (ed.) Modeling the environmental fate of microorganismes. Am. Soc. Microbiol., Washington, D.C.

Filmer, R. W., M. Felton, and T. Yamamoto. 1971. 'Virus sized particle adsorption on soil: 1. Rate of adsorption'. p. 75-101. In V. Snoeyink (ed.) Proc. 13th Water Qual. Conf. Virus and Water Qual. Occurrence Control. University of Illinois, Urbana-Champaign, Illinois.

Fontes, D. E., A. L. Mills, G. M.Hornberger, and J. S. Herman. 1991. 'Physical and chemical factors influencing transport of microorganisms through porous media'. Appl. Environ. Microbiol. 57:2473-2481.

Gannon, J. T., Y. Tan, P. Baveye, and M. Alexander. 1991. 'Effect of sodium chloride on transport of bacteria in a saturated aquifer material'. Appl. Environ. Microbiol. 57:2497-2501.

Gerba, C. P., C. Wallis, and J. L. Melnick. 1975. 'Fate of wastewater bacteria and viruses in soil'. J. Irrig. Drain. Div. Proc. Am. Soc. Civ. Eng. 101:157-174.

Gerba, C. P., and G. Bitton. 1984. 'Microbial pollutants: Their survival and transport pattern to groundwater'. p. 65-88. In G. Bitton and C. P. Gerba (ed.) Groundwater pollution microbiology. Wiley-Interscience, New York.

Germann, P. F., M. S. Smith, and G. W. Thomas. 1987. 'Kinematic wave approximation to the transport of Escherichia coli in the vadose zone'. Water Resour. Res. 23:1281-1287.

Goldshmid, J., D. Zohar, Y. Argaman, and Y. Kott. 1973. 'Effect of dissolved salts on the filtration of coliform bacteria in sand dunes'. p. 147-157. In S.H. Jenkins (ed.) Advances in water pollution research. Pergamon Press, New York.

Hamdi, Y. A. 1974. 'Vertical Movement of Rhizobia in soil'. Zbl. Bakt. Abt. II, Bd. 129:373-377.

Harvey, R. W., and S. P. Garabedian. 1991. 'Use of colloid filtration theory in modeling movement of bacteria through a contaminated sandy aquifer'. Environ. Sci. Technol. 25:178-185.

Hattori, T. 1973. Microbial life in the soil: An introduction. Marcel Dekker, New York.

Hendricks, D. W., F. J. Post and D. R. Khairnar. 1979. 'Adsorption of bacteria on soils: Experiments, thermodynamic rationale, and application'. Water, Air and Soil Pollu. 12:219-232.

Hornberger, G.M., A.L. Mills, and J.S. Herman. 1992. 'Bacterial transport in porous media: Evaluation of a model using laboratory observations'. Water Resour. Res. 28:915-938.

Jenneman, G. E., R. M. Knapp. M. J. McInerney, D. E. Menzie, and D. E. Revus. 1984. 'Experimental studies of in-situ microbial enhanced oil recovery'. Soc. Pet. Eng. J. 24:33-37.

Jury, W. A. 1982. 'Simulation of solute transport using a transfer function model'. Water Resour. Res. 18:363-368.

Keller, E. F., and L. A. Segel. 1971. 'Model for chemotaxis'. J. Theor. Biol. 30:225-234.

Kirkham, D., and W. L. Powers. 1972. Advanced soil physics. Wiley-Interscience, New York.

Krone, R.B., G.T. Orlob, and C. Hodgkinson. 1958. 'Movement of coliform bacteria through porous media'. Sewage and Industrial Wastes. 30:1-13.

Kurath, G., and R. Y. Morita. 1983. 'Starvation-survival physiological studies of a marine Pseudomonas sp.'. Appl. Environ. Microbiol. 45:1206-1211.

Marshall, K. C. 1967. 'Electrophoretic properties of fast and slow growing species of Rhizobium'. Aust. J. Biol. Sci. 20:429-438

Marshall, K. C. 1975. 'Clay mineralogy in relation to survival of soil bacteria'. Ann. Rev. Phytopathol. 13:357-373.

Marshall, K. C. 1976. Interfaces in microbial ecology. Harvard University Press, Cambridge.

Marshall, K. C. 1980. 'Adsorption of microorganisms to soils and sediments'. p. 317-329. In G. Bitton and K. C. Marshall (ed.) Adsorption of microorganisms to surfaces. Wiley-Interscience, New York.

Marshall, K. C., R. Stout, and R. Mitchell. 1971. 'Mechanism of the initial events in the sorption of marine bacteria to surfaces'. J. Gen. Microbiol. 68:337-348.

Matthess, G., and A. Pekdeger. 1981. 'Concepts of a survival and transport model of pathogenic bacteria and viruses in groundwater'. Sci. Total Environ. 21:149-159.

McCarty, P.L., B.E. Rittmann, and E.J. Bouwer. 1984. 'Microbiological processes affecting chemical transformations in groundwater'. p. 89-115. In G. Bitton and C. P. Gerba (ed.) Groundwater pollution microbiology. Wiley-Interscience, New York.

McCaulou, D.R., R.C. Bales and J.F. McCarthy. 1994. 'Use of short-pulse experiments to study bacteria transport through porous media'. J. Contam. Hydrol. 15:1-14.

McDowell-Boyer, L. M., J. R. Hunt, and N. Sitar. 1986. 'Particle transport in porous media'. Water Resour. Res. 22:1901-1921.

Molz, F.J., M.A. Widdowson, and L.D. Benefield. 1986. 'Simulation of microbial growth dynamics coupled to nutrient and oxygen transport in porous media'. Water Resour. Res. 22:1207-1216.

Monod, J. 1942. Recherches sur la Croissance des Cultures Bacteriennes. Hermann et Cie, Paris.

Moore, B.E., B.P. Sagik, and J.F. Malina. 1975. 'Viral association with suspended solids'. Water Res. 9:197-203.

Moore, R.S., D.H. Taylor, L.S. Sturman, M.M. Reddy, and G.W. Fuhs. 1981. 'Poliovirus adsorption by 34 minerals and soils'. Appl. Environ. Microbiol. 42:963-975.

Moser, A. (1958). The Dynamics of Bacterial Populations Maintained in The Chemostat. Ph.D. thesis, The Carnegie Institute, Washington DC.

Moser, A. 1985. 'Kinetics of batch fermentations'. p. 243-283. In H.-J. Rehm and G. Reed (ed.) Biotechnology (Vol. 2). VCH Verlagsgesellschaft mbH, Weinheim.

O'Melia, C. R., and W. Stumn. 1967. 'Theory of water filtration'. J. Am. Water Works Assoc. 59:1393-1412.

Parke, J. L., R. Moen, A. D. Rovira, and G. D. Bowen. 1986. 'Soil water flow affects the rhizosphere distribution of a seed-borne biological control agent, Pseudomonas fluorescens'. Soil Biol. Biochem. 18:583-588.

Peele, T. C. 1936. Adsorption of bacteria by soils. Cornell Agric. Exp. Stn. Memoir. 197:3-18.

Peterson, T.C., and R.C. Ward. 1989. 'Development of a bacterial transport model for coarse soils'. Water Resour. Bulletin. 25:349-357.

Plummer, D.T., A.M. James., H. Gooder, and W.R. Maxted. 1962. 'Some physical investigations of the behaviour of bacterial surfaces: 5. The variation of the surface structure of Streptococcus pyogenes during growth'. Biochim. Biophys. Acta. 60:595-603.

Powelson, D. K., J. R. Simpson, and C. P. Gerba. 1990. 'Virus transport and survival in saturated and unsaturated flow through soil columns'. J. Environ. Qual. 19:396-401.

Rajagopalan, R. and R. Q. Chu. 1982. 'Dynamics of adsorption of colloidal particles in packed beds'. J. Colloidal and Interface Sci. 86:299-317.

Reynolds, P. J., P. Sharma, G. E. Jenneman, and M. J. McInerney. 1989. 'Mechanisms of microbial movement in subsurface materials'. Appl. Environ. Microbiol. 55:2280-2286.

Rittmann, B. E. 1982. 'The effect of shear stress on biofilm loss rate'. Biotechnol. Bioeng. 24:501-506.

Rittmann, B. E., P. L. McCarty, and P. V. Roberts. 1980. 'Trace-organics biodegradation in aquifer recharge'. Ground Water. 18:236-243.

Roels, J. A. 1983. Energetics and Kinetics in Biotechnology. Elsevier Biomedical Press, Amsterdam.

Sakthivadivel, R. and H. A. Einstein. 1970. 'Clogging of porous column of spheres by sediment'. J. Hydraul. Div. Proc. Am. Soc. Civ. Eng. HY2:461-471.

Schaub, S. A., and B. P. Sagik. 1975. 'Association of enteroviruses with natural and artificially introduced colloidal solids in water and infectivity of solids-associated virions'. Appl. Microbiol. 30:212-222.

Shehata, T. E., and A. G. Marr. 1971. 'Effect of nutrient concentration on the growth of Escherichia coli'. J. Bacteriol. 107:210-216.

Sherard, J. L., L. P. Dunnigan, and J. R. Talbot. 1984. 'Basic properties of sand and gravel filters'. J. Geotech. Eng. 110:684-700.

Smith, M. S., G. W. Thomas, R. E. White, and D. Ritonga. 1985. 'Transport of Escherichia coli through intact and disturbed soil columns'. J. Environ. Qual. 14:87-91.

Speitel, G. E., and F. A. DiGiano. 1987. 'Biofilm shearing under dynamic conditions'. J. Environ. Eng., Am. Soc.Civ. Eng. 113:464-475.

Sykes, J.F., S. Soyupak, and G.J. Farquhar. 1982. 'Modeling of leachate organic migration and attenuation in groundwaters below sanitary landfills'. Water Resour. Res. 18:135-145.

Tan, Y., W. J. Bond, A. D. Rovira, P. G. Brisbane, and D. M. Griffin. 1991. 'Movement through soil of a biological control agent, Pseudomonas fluorescens'. Soil Biol. Biochem. 23:821-825.

Tan, Y., W. J. Bond, and D. M. Griffin. 1992. 'Transport of bacteria during unsteady unsaturated soil water flow'. Soil Sci. Soc. Am. J. 56:1331-1340.

Tan, Y., Z. X. Wang, R. P. Schuneider, and K. C. Marshall. 1994a. 'Modeling bacterial growth kinetics: A statistical thermodynamic approach'. J. Biotechnol. 34:97-106.

Tan, Y., J.G. Gannon, P. Baveye, and M. Alexander. 1994b. 'Transport of bacteria in a saturated aquifer sand'. Water Resour. Res. 30:3243-3252.

Taylor, S. W., and P. R. Jaffe. 1990a. 'Biofilm growth and the related changes in the physical properties of a porous medium. 1. Experimental investigation'. Water Resour. Res. 26:2153-2159.

Taylor, S. W., and P. R. Jaffe. 1990b. 'Biofilm growth and the related changes in the physical properties of a porous medium. 3. Dispersivity and model verification'. Water Resour. Res. 26:2171-2180.

Taylor, S. W., and P. R. Jaffe. 1990c. 'Substrate and biomass transport in a porous medium'. Water Resour. Res. 26:2181-2194.

Udoyara, S.T. and S. Mostaghimi. 1991. 'Model for predicting virus movement through soils'. Ground Water. 29:251-259.

Vilker, V. L. 1980. 'Simulating virus movement in soils'. p. 223-253. In I. K. Iskandar (ed.) Modeling wastewater renovation: Land treatment. Wiley-Interscience, New York.

Vilker, V. L., and W. D. Burge. 1980. 'Adsorption mass transfer model for virus transport in soils'. Water Res. 14:783-790.

Vinten, J. A., and P. H. Nye. 1985. 'Transport and deposition of dilute colloidal suspensions in soils'. J. Soil Sci. 36:531-541.

Watson, J. E., and W. R. Gardner. 1986. 'A mechanistic model of bacterial colony growth response to substrate supply'. Paper presented at the Chapman Conference on Microbial Processes in the Transport, Fate and in situ Treatment of Subsurface Contaminants. Snowbird, Utah. Oct., 1986.

White, R. E., J. S. Dyson, R. A. Haigh, W. A. Jury, and G. Sposito. 1986. 'A transfer function model of solute transport through soil: 2. Illustrative applications'. Water Resour. Res. 22:248-254.

Wollum, A. G., and D. K. Cassel. 1978. 'Transport of microorganisms in sand columns'. Soil Sci. Soc. Am. J. 42:72-76.

Yates, M. V., and Y. Ouyang. 1992. 'VIRTUS, a model of virus transport in unsaturated soils'. Appl. Environ. Microbiol. 58:1069-1616.

Yates, M. V., and S. R. Yates. 1988. 'Modeling microbial fate in the subsurface environment'. CRC Crit. Rev. Environ. Control. 17:307-344.

Yao, K., M. T. Habibian, and C. R. O'Melia. 1971. 'Water and wastewater filtration: Concepts and applications'. Environ. Sci. Technol. 5:1105-1112.

Notation

b	decay or inactivation constant T^{-1}
b_s	specific shear loss coefficient T^{-1}
b'_s	dimensionless parameter for biological aspects of shearing
c_1, c_2	constants in Deb (1969) expression for attachment rate
C	concentration of microorganisms in the water phase M L^{-3}
C^a	mass of microorganisms per unit mass of porous material M M^{-1}
C^a_m	maximum concentration of microorganisms in attached phase M M^{-1}
C_0	inflowing concentration of microorganisms
d_m	diameter of microorganisms L
d_s	diameter of solid matrix particles L
D	combined dispersion coefficient L^2 T^{-1}
D_h	hydrodynamic dispersion coefficient L^2 T^{-1}
D_B	diffusion coefficient of microorganisms L^2 T^{-1},
D_s	diffusion-dispersion coefficient of substrate L^2 T^{-1}
D_t	bacterial diffusion coefficient due to motility L^2 T^{-1}
h	dimensionless constant
J_v	advective flux of microorganisms M L^{-2} T^{-1}
J_B	flux of microorganisms due to Brownian diffusion M L^{-2} T^{-1}
J_d	flux of microorganisms due to motility, diffusion and hydrodynamic dispersion M L^{-2} T^{-1}
J_w	water flux M L^{-2} T^{-1}
J_x	chemotactic flux of motile bacteria M L^{-2} T^{-1}
k_a	attachment coefficient T^{-1}
k_f	Freundlich adsorption isotherm slope parameter
k_x	chemotactic coefficient
k_y	detachment coefficient T^{-1}
K_s	growth saturation constant M L^{-3}
L_b	biofilm thickness L
M	specific surface area of water-biofilm phase boundary L^2 L^{-3}
N	Freundlich adsorption isotherm exponent
R_a	rate of microbial attachment M L^{-3} T^{-1}
R_b	total rate of substrate consumption of bacteria M L^{-3} T^{-1}
R_d	rate of decay or inactivation of microorganisms in combined water and attached biomass phases M L^{-3} T^{-1}
R^w_d, R^a_d	decay or inactivation rate of microorganisms in water phase and attached phase, respectively M L^{-3} T^{-1}
R_g	rate of growth of bacteria in combined water and attached biomass phases M L^{-3} T^{-1}

R_g^w, R_g^a	bacterial growth rate in water phase and attached phase M L^{-3} T^{-1}
R_y	rate of microbial detachment M L^{-3} T^{-1}
S	substrate concentration in the water phase M L^{-3}
S^a	substrate concentration associated with solid matrix M M^{-1}
S_b	microscopic concentration of substrate in biofilm M L^{-3}
v	total velocity of microorganisms L T^{-1}
v_t	motility velocity of motile bacteria L T^{-1}
v_w	average pore water flow velocity L T^{-1}
v_x	chemotactic velocity of motile bacteria L T^{-1}
α	collision efficiency factor for filtration
γ	single-collector efficiency for filtration
η	effectiveness factor used to correct for diffusional resistance at the biofilm-water phase boundary
θ	water filled porosity L^3 L^{-3}
μ	specific growth rate T^{-1}
μ_m	maximum specific growth rate T^{-1}
ρ_b	biomass density of microorganisms (mass of microorganisms per unit volume of biomass) M L^{-3}
ρ_s	bulk density of the porous medium M L^{-3}
σ	volume fraction of biomass associated with solid matrix (volume of microorganisms per unit volume of porous material) L^3 L^{-3}
τ	tortuosity of the pore space

CHAPTER 10

Assessment and control of loading uncertainty for managing eutrophication and toxic chemical fate in lakes

T.C. Young and D.M. Dolan

Abstract. Loadings and other forcing functions are key elements of numerical models that are used for various tasks that arise in the management and planning of water quality for present and future uses. Of particular concern are loadings of nutrients and toxic chemicals. First order analysis is presented as an approach to evaluating uncertainty that has moderate computational requirements, but that permits ranking of sources of uncertainty and guides future monitoring for reducing uncertainty. Major sources of uncertainty associated with tributary loads, other non-point sources (rural and urban runoff, in-place pollutants, groundwater, and atmospheric), and point source loads (municipal and industrial) are reviewed, with particular emphasis on fluvial loadings. Alternative monitoring and sampling designs are discussed that have potential to maximize certainty for fixed resource investments, and discussion also is presented on methods for minimizing uncertainties associated with censored data.

10.1 Introduction

Eutrophication and toxic chemical exposure have been and continue to be among the most significant challenges to continued beneficial use of our aquatic environments. Eutrophication, or lake enrichment, was the object of intense study during the late 1960's and much of the 1970's, and research on lingering questions about eutrophic systems (nutrient bioavailability, process representation for numerical simulation, and uncertainty analysis, to name a few) continues to the present. On balance, however, the causes and consequences of eutrophication are relatively

V. P. Singh (ed.), Environmental Hydrology, 357–375.
© 1995 *Kluwer Academic Publishers.*

well understood and have been for several years (National Academy of Sciences, 1970; Likens, 1972). More importantly, perhaps, corrective measures that have been demonstrated to be effective against the effects of excessive, culturally-induced eutrophication have been identified, tested, and described (Cooke et al., 1986; Olem and Flock, 1990; Wedepohl et al., 1990).

Toxic chemical contamination of lakes and rivers has become increasingly recognized as a significant world-wide problem during the past two decades. In contrast with the technology available for eutrophication control, however, the alternatives for remediation of toxic chemical contamination are considerably less well developed. In North America, the Laurentian Great Lakes provide a prototypic example of the toxic chemical contamination problem that has become common in western nations. In the Great Lakes, active sources of new pollutants are comparatively rare. This may be due in part to increased enforcement activities, improvements in recycle and recovery systems, better or more efficient manufacturing or industrial processes, and other developments. Rather than active pollution, however, the primary toxic contamination problem in the Great Lakes is caused by "in-place" pollutants; these consist of those contaminants that are associated with sediments or terrestrial disposal sites located adjacent to water courses and that release their contaminant burden to the watercourse over time (IJC, 1987, 1992).

It has been widely recognized that a key element for managing the effects of cultural activities on aquatic resources is control of the amounts of contaminants that enter such systems of interest (Vollenweider, 1968; Schnoor et al., 1987). The effectiveness of such controls often is examined through the use of a numerical representation of the system of concern (Thomann and Mueller, 1987). These numerical representations, usually termed mathematical simulation models, provide a mechanistic approach to forecasting aquatic system response to implementation of possible water quality management alternatives. Contaminant inputs or loadings, a necessary component of these mathematical models, constitute one example of forcing functions, a group of phenomena required for modeling aquatic system. In this paper, we discuss such forcing functions, especially loadings, from various points of view relevant to water quality modeling. Emphasis, however, is given to steps that may be taken that could lead to possible increases in the confidence that may be attributed to loadings and other forcing functions when they are used in aquatic system models.

At this point, two definitions would be useful:

* *Forcing Function*: An independent variable in a model formulation. By independent it is meant that the variable is not affected by changes in the model. Thus, forcing functions can appear on the right hand sides of model equations, but not the left. Some forcing functions, such as loadings, can often be controlled and thus are of management interest, others, such as weather, are basically uncontrollable.

* *Uncertainty*: An indication of the value or reliability of information. In this paper, the indicator for uncertainty used will be the standard deviation of the model estimate of an output variable. It should be recognized that reductions

in uncertainty occur in a variety of ways. It is assumed that, whenever possible, state-of-the-art methods for data handling will be used, so that this will not be a source of uncertainty not related to the forcing function uncertainty. Further, it is assumed that all data will be collected in a manner that reduces uncertainty whenever possible, and that it is desirable to optimize this effort in some way, possibly because of budget constraints.

10.2 Loadings and Sources of Uncertainty

Contaminant loads frequently are handled in models of chemical fate and transport as if they were independent variables, or forcing functions, that drive the response of the modeled system. However, loadings are not truly independent variables; rather, they respond to numerous external forces, true forcing functions, in achieving their observed states. The coupling between loads, forcing functions, and chemical fate is partly deterministic and partly stochastic. Here, the emphasis will be on identifying the uncertain element in the linkages between loads and forcing functions. Chemical loads of concern to this discussion will be identified as belonging to one of three categories. These include:

- tributary inflows,
- direct point sources (industrial and municipal inputs), and
- direct non-point sources (rural and urban runoff, in-place pollutants, groundwater, and atmospheric inputs).

Loads in each of these categories possess attributes that vary in a manner that makes load quantitation difficult without encountering substantial uncertainty. On the other hand, several issues related to uncertainty are common to most protocols for estimating loads and forcing functions. An important issue common to any load or forcing function estimate is selection of appropriate time and space scales for aggregating information. This consideration is important because it can determine the accuracy with which uncertainty in the estimates may be known for given actual variability in the quantities of interest.

A second issue of common concern is that of observation uncertainty. This is uncertainty that may be attributed to the physical and chemical methods used for data collection (e.g. sampling schedule, field and laboratory methods for sampling, handling, preservation, storage, and analytical quantitation). Such methodological uncertainties may be systematic or random and provide the target for quality assurance and control programs (QA/QC). In the design of a monitoring program, however, the need to randomize is not always recognized. Nevertheless, the need to randomize at some stage of monitoring is always present. Otherwise, the analysis of results by statistical models cannot be accomplished without an unknowable amount of bias. Uncertainties due to sampling and analysis methodologies will be addressed to a limited degree in this paper, but a detailed discussion is outside of the scope of this paper. One special area of data uncertainty that frequently arises in estimating loads for mass balance models is the issue of non-detect observations.

Dealing with non-detect data will be examined in this paper in the context of estimating contaminant loads and doing so to minimize uncertainty.

A third common issue involving uncertainty arises over the approach to be used for data analysis, or the statistical estimation method to be employed. To a large extent, such concerns parallel those that emanate from uncertainties over model specification, including process parameterization. This is appropriate, because the statistical analysis applied to yield a loading estimate or time series must be faithful to an underlying mathematical (i.e. statistical) model. This aspect of uncertainty also will be given detailed treatment in this discussion.

10.2.1 TRADEOFFS

As has been alluded above, the feasibility of obtaining perfect load estimates is reduced by a number of difficult issues: scientific, logistic, economic and natural. It is important not only to divide these issues into ones that can be controlled and those that can't, but also into those that represent tradeoffs or compromises and those that don't.

Tradeoffs occur when a task must be performed within a fixed budget (either dollars or manpower or both). In the context of load estimation, tradeoffs involve questions such as: how many tributaries and sources can be sampled? how many samples per source can be taken?, etc. These are mainly economic and logistic considerations. Scientific issues in load estimation tend not to be as subject to tradeoffs because the usual approach is to apply state-of-the-art science without advancing to unproven areas.

10.3 Assessing Loading Uncertainty

As is true in the larger context of model uncertainty analysis, the evaluation of load estimation uncertainty, itself, is a procedure that may incur considerable error. Except in simple models, it is presently not possible to objectively and uniquely assess uncertainty due to model (statistical) formulation, dependent and independent variables, and parameters. Most effort in uncertainty analysis has been put toward evaluating uncertainty due to errors in parameters and independent variables or forcing functions.

Two fundamentally different approaches to the assessment of parameter uncertainty have been developed: Monte Carlo simulation and first order analysis (variance propagation). The approach taken with the former method involves assigning probability density functions to each uncertain model parameter or variable, and then "sampling" the distributions under some randomization scheme for calculation of a model result. This procedure is repeated a large number of times, and the distribution of model results may be analyzed statistically to ascertain the combined effects of all the uncertain parameters on the overall uncertainty exhibited by the model (Scavia et al., 1981).

Instead of Monte Carlo simulation, it may be possible to substitute the less computationally intensive first order analysis. The term, first order analysis, is apt because the approach is based on using the first order terms from a Taylor

series expansion of the model about the mean of each uncertain parameter or input variable (Cornell, 1972). For a single equation, multi-parameter model (e.g. $Y = f(X)$, where X is the vector of uncertain parameters), the expression for the estimated uncertainty of the model result ($V(Y)$, variance of outcome, Y), due to parameter uncertainty, is given by

$$V(Y) = \sum_{i=1}^{n} \left(\frac{\partial f}{\partial X_i}\right)^2 V(X_i) + 2\sum_{i=1}^{n}\sum_{j=1}^{i-1} \left(\frac{\partial f}{\partial X_i}\right)\left(\frac{\partial f}{\partial X_j}\right) \mathrm{COV}(X_i, X_j). \tag{10.1}$$

If the uncertain parameters are uncorrelated, then Equation (10.1) simplifies to the first term on the right and the total prediction uncertainty is due to the sum of the individual contributions by each uncertain parameter.

As noted by Reckhow and Chapra (1983), Monte Carlo simulation has two main advantages over first order analysis. First, Monte Carlo simulation does not lose exactness of model representation, as first order analysis does due to truncation of the Taylor series expansion. Thus, within the limits of the deterministic model and parameter distributions, Monte Carlo simulation is complete. Second, the results of Monte Carlo simulation provide a complete probability density function estimate to describe parameter error, rather than only an estimate of the mean and variance. The Monte Carlo procedure, as noted above, is computationally intensive. Consequently, first order analysis may be favored for simple models due to its convenience. On the other hand, an uncertainty analysis for complex models (dynamic, multi-state, multi-parameter, detailed spatial resolution) may require more computation for Monte Carlo simulation than can be justified. Statistically efficient methods of directing the sampling required for the simulations (e.g. the latin hypercube design, Iman and Shortencarier, 1984) can greatly reduce the computational requirements.

First order analysis provides a reasonably uncomplicated approach to assessing uncertainty that may be adapted for use with models of high complexity. Because of its wide utility, it will be used here to demonstrate a procedure for ranking the uncertainty in lake water concentration for a contaminant of concern that is associated with uncertainty in each of the loading categories discussed previously. To do this, of course, a model must be specified that relates concentration and loads. Given in Equation (10.2) for the purposes of illustration is a mass balance equation for a lake of volume, V, an outflow rate of Q, a contaminant of concern with water column concentration, C, and which receives loads of the contaminant from tributaries ($\sum W_{\mathrm{trib,i}}$), direct point sources ($\sum W_{\mathrm{PTS},j}$), external direct non-point sources ($\sum W_{\mathrm{NPS},k}$), atmospheric sources (W_{ATM}), and which loses mass due to the net of sedimentation, resuspension, diffusion, and bioturbation over the sediment with surface area A_s, ($J_{\mathrm{SED}}A_s$), by evaporation/volatilization over the water surface with area A, ($J_{\mathrm{EVAP}}A$), and which reacts within the water column by a first order process with rate law $-kC$:

$$V\frac{\mathrm{d}C}{\mathrm{d}t} = \sum_i W_{\mathrm{TRIB},i} + \sum_j W_{\mathrm{PTS},j} + \sum_k W_{\mathrm{NPS},k} + \\ + W_{\mathrm{ATM}} - J_{\mathrm{SED}}A_s - J_{\mathrm{EVAP}}A - QC - kCV. \tag{10.2}$$

Equation (10.2) has the steady-state solution:

$$C = \frac{\sum_i W_{\text{TRIB},i} + \sum_j W_{\text{PTS},j} + \sum_k W_{\text{NPS},k} + W_{\text{ATM}} - J_{\text{SED}} A_s - J_{\text{EVAP}} A}{Q + kV}. \tag{10.3}$$

The partial derivatives in Equation (10.1) ($\partial f/\partial X_i, \partial f/\partial X_j$) are sensitivity coefficients; they indicate the rate of change in the dependent or predicted variable with respect to the uncertain parameters. By analogy, the steady-state concentration of contaminant in Equation (10.3) may be taken to be the dependent variable and the loads may be taken to represent uncertain parameters, giving

$$\frac{\partial C}{\partial W_m} = \frac{1}{Q + kV}. \tag{10.4}$$

In this case, the objective is projecting the uncertainty in loads onto predictions made about the steady-state contaminant concentration. Because of the simplicity of this model, the sensitivity of the lake water concentration at steady-state to each of the loads is identical, as shown in Equation (10.4) (W_m is an arbitrary loading term from Equations (10.2) or (10.3)).

The contributions of the loads to the uncertainty of the concentration prediction, however, are not identical. Rather, they depend on their individual variances. If the loads are assumed to be independent of each other to permit simple extension of the illustration (seasonality alone suggests that this is not a particularly good assumption), then individual load contributions to the uncertainty of concentration is given by Equation (10.5), which is an estimate of concentration variance due to the variance of a given contaminant loading source, W_m.

$$S^2_{C,W_m} = \frac{S^2_{W_m}}{(Q + kV)^2}. \tag{10.5}$$

In addition, Equation (10.6) gives the total uncertainty due to the uncertainty in loading from all sources.

$$S^2_{C_w} = \left(\frac{1}{(Q + kV)^2}\right) \sum_{m=1}^{n} S^2_{C,W_m}. \tag{10.6}$$

10.3.1 Demonstration

The approach of using first order analysis to assess lake concentration uncertainty due to uncertain loadings may be illustrated using data on contaminant loadings from the steady-state model of Mackay et al. (1993) for Lake Ontario (PROGRAM NOTL); also other attributes of Lake Ontario required for Equation (10.3) were taken from the Mackay model ($V = 1.67 \times 10^{12}$ m^3, $V/Q = 6.5$ y, $k = 6.07 \times 10^{-7}$ y^{-1}). The loads of contaminant range from 100 to 200 kg y^{-1}. The analysis

is shown in Table 10.1. To perform the analysis, coefficients of variation were required for the load estimates. The values used are given in Table 10.1 and are hypothetical; they are suggestive of the relative ranking of the uncertainty with which each load may be known.

To put the concentration uncertainty estimates into context, Equation (10.3) may be calculated for the volume, discharge, and chemical decay attributes given in parentheses above. To do this with the minimal information given, it was assumed that the net interaction with sediment and atmosphere is zero. As indicated in Table 10.1, the given loadings total 700 $\mathrm{kg\,y^{-1}}$. This would yield a steady-state concentration estimate of 2.7 $\mathrm{ng\,L^{-1}}$ by Equation (10.3). Because the partial uncertainties hypothesized for the loadings total 2.0 $\mathrm{ng\,L^{-1}}$, the steady-state concentration estimate has an overall coefficient of variation of 74 percent. The relative ranking of the partial uncertainties closely tracks the assumed coefficients of variation. This underscores the importance to uncertainty analysis of knowing the properties of the underlying error structure as well as possible.

In the example represented in Table 10.1, the direct point sources have been split into municipal and industrial. In general, this split is desirable because these sources represent different management options (pretreatment programs vs. direct effluent controls). Sufficient information currently exists to make this division of the point source load.

In the case of the tributary load, a split has been made into the Niagara River and "all others". This was done because the Niagara load is comparatively well known (Data Interpretation Group, 1986, 1988, 1989, 1990, 1991, 1992, 1993); the differences in coefficients of variation reflect this situation.

For non-point sources, no such division is possible given existing knowledge, although it would be desirable because of the implications for control options.

After agreement on a particular model formulation, the above analysis should be repeated using as much current data as possible. The partial uncertainties should be estimated and ranked, after which sampling efforts may be concentrated on the highest sources of uncertainty. The importance of historical data is clear from this analysis. Another point suggested from this analysis is that methods of load estimation for a system should continue development. These include methods based on extrapolation, the use of typical concentrations, or other procedures involving estimates in lieu of real data. This is important since improving the effectiveness of these methods will improve the utility of the uncertainty analysis.

10.4 Managing Loading Uncertainty

The average load, L [$\mathrm{M\,T^{-1}}$], delivered to a receiving water during some specific time interval, t [T], through an interface of area, A [$\mathrm{L^2}$], for a contaminant that is carried by water flowing at an instantaneous velocity, ν [$\mathrm{L\,T^{-1}}$], and concentration, C [$\mathrm{M\,L^{-3}}$], may be represented by Equation (10.7) when contaminant transport occurs by purely advective processes. If contaminant transport occurs by purely diffusive or dispersive processes along the instantaneous concentration gradient normal to the interface, $\mathrm{d}C/\mathrm{d}z$ [$\mathrm{M\,L^{-4}}$], with a transport coefficient, D_z [$\mathrm{L^2\,T^{-1}}$], then Equation (10.8) may be used to represent the average load to the receiving

TABLE 10.1
First-order analysis of lake water concentration uncertainty due to loading variability; hypothetical illustrative data for Lake Ontario

Source of Uncertainty	Load Magnitude (kg/y)	Assumed Load Coefficient of Variation (%)	Load Standard Deviation ($ng\,y^{-1}$)	Partial Uncertainty in Concentration Estimate Due to Load Uncertainty ($ng\,L^{-1}$)
Municipal Point Sources	100	25	25×10^{12}	0.1
Industrial Point Sources	120	25	30×10^{12}	0.1
Direct Non-Point Sources	130	100	130×10^{12}	0.5
All Other Influent Tributaries	150	200	300×10^{12}	1.2
Niagara River	200	20	40×10^{12}	0.2
TOTAL	700	–	–	2.0

water.

$$S = \frac{1}{\Delta t}\int_{\Delta t}\int_{A} \nu(A(\tau),\tau)C(A(\tau),\tau)\,\mathrm{d}A\,\mathrm{d}\tau, \tag{10.7}$$

$$S = -\frac{1}{\Delta t}\int_{\Delta t}\int_{A} D_z(A(\tau),\tau)\frac{\mathrm{d}C(A(\tau),\tau)}{\mathrm{d}z}\,\mathrm{d}A\,\mathrm{d}\tau, \tag{10.8}$$

Equations (10.7) and (10.8) are one-dimensional, idealized, deterministic simplifications of the complex, stochastic situation that exists and determines loads during contaminant transport. If it were possible to monitor continuously in time and space and with absolute accuracy all boundaries of a receiving water body, then one could compute loads without uncertainty, at least in theory. Even if monitoring with that intensity were technologically feasible, however, it would be impractical. Nevertheless, these equations can serve to illustrate the major sources of uncertainty that affect estimates of chemical loading and other forcing functions of interest for aquatic system modeling. In particular, any condition or change in conditions that alters any of the variables in Equations (10.7) and (10.8) will affect, in turn, the resulting estimate of load. By corollary, uncertainty in the variables of Equations (10.7) and (10.8) will propagate through the relationship to affect the load estimate. Any quantity, therefore, that produces uncertainty in flow velocities, concentrations, dispersion or diffusion coefficients, or concentration gradients can lead to uncertainty in contaminant loading estimates. The role of sampling design in reducing uncertainty is examined below after a survey of the major sources of uncertainty associated with the loads in each of the categories identified earlier.

10.4.1 Tributary Mouth Loadings

Contaminant loading estimates for rivers derive uncertainty from a combination of several sources. Because tributary inputs constitute the largest single load for many contaminants, considerable attention has been given to tributary load estimation and uncertainty (Young et al., 1988; Cohen et al. 1989; Preston et al., 1989). Exhaustive coverage of tributary load uncertainty will not be given here, however, the main contributory elements will be examined. The discussion will focus on uncertainty associated with estimating loads of contaminants that partition strongly to the solid phase. This will be done for two reasons. First, phosphorus, which is of paramount importance in eutrophication, and many anthropogenic organic contaminants of concern are surface active due to hydrophobicity or polyvalency. As a consequence, these chemicals may be predominantly associated with transported solids. This phase partitioning is a complex phenomenon involving thermodynamic and kinetic constraints that result from chemical and physical characteristics of the solid and liquid phases, and the contaminant(s) of concern. Some of the more important factors include pH, particulate and dissolved organic carbon levels, solid phase levels of hydrous metal oxides, sorbent and sorbate concentrations and kinetics of adsorption and desorption (Benjamin and Leckie, 1981; Wu and Gschwend, 1986). The second reason for focusing this discussion on particulate matter is the fact that particulate matter, and associated contaminants, generally are subject to greater spatial and temporal heterogeneity than contaminants in the dissolved phase. Consequently, the uncertainty of total tributary loads will be heavily

influenced by the uncertainty of particulate loads of such contaminants. Moreover, reductions in particle-associated tributary loading uncertainty by improved sample design will yield more certain soluble phase loadings, too.

Sediment and sediment-associated pollutants that reach the mouth of a tributary originate largely from non-point sources in the watershed and result from a complex interplay of numerous forces. Of special importance are (1) climate, in general, and the nature, amount, and intensity of precipitation, in particular; (2) orientation, degree, and length of slopes; (3) geology and soil types; (4) land use; (5) condition and density of the channel system; (6) particle setting velocity; and (7) streamflow regime. In addition, the finest resolution of sediment load variability, and that of associated contaminants, for a receiving water generally will depend on the coarsest resolution of either meteorologic events, receiving water events, or channel geometry and geology. Fluvial loading uncertainty, therefore, will depend in large measure on the variability of these physical factors. This dependence, of course, is in addition to the influence of physicochemical factors specific to the contaminant of interest, such as, solubility, volatility, partitioning, and temperature dependencies, plus the biochemical influences of uptake and decay.

The interplay of gravitational and viscous forces on particulate matter in water results in fluvial systems having a selective transport capacity for particle-bound pollutants. Thus, as fluid forces vary in time and space, so too does the movement of particulate contaminants. In general, therefore, the total loading of particle-bound contaminants is related to discharge in a non-linear manner that consists of watershed-determined movement of fine particles superimposed on a discharge-dependent and spatially varying flux of coarse sediments. The consequence of selective transport for loading uncertainty arises immediately as a requirement to design sampling protocols that address concentration and velocity variability over a sample transect.

Cross-sectional variability in particulate contaminant concentration also is a problem consisting of two components, because fine and coarse particulate matter behave differently. Fine matter (coarse silt and smaller) may be considered to be uniformly distributed vertically and horizontally at an appropriate distance downstream from a source of sediment, so long as flows are sufficiently turbulent. This criterion is not well defined but will depend on characteristics of the particles, the fluid, and the site of concern (discharge, cross-sectional area, channel width and depth). Coarse sediment (fine sand and greater), on the other hand, shows a strong dependence on the hydraulic character of the system. In particular, the transport of sand-size particles depends on bedform (depth and dune forms)and flow regime (development of turbulence). The combined effect of these components is relatively low variability across the cross-section if the amount of fine sediment is high relative to sand.

Temporal variability, of course, adds an additional dimension of uncertainty to the fluvial load problem. Concentrations of suspended matter and associated pollutants at a stream cross-section during a hydrologic event often shows a time-dependence that reflects the character of the hydrograph, generally increasing and then decreasing during the event, due to erosion, scour and deposition. The peak concentration, however, often does not coincide with the peak of flow and may lead or lag the flow peak. Most frequently the concentration peak leads the hydrograph

somewhat; however, for suspended matter the actual pattern realized by a given drainage system depends on the size of the watershed, location of the causative event, and antecedent streamflow. Time of year exerts significant influence on the solids levels found in most systems. For fine suspended matter, this influence is a result of the dependence of concentration on rate of supply, which in turn depends on erodability, land use patterns, precipitation, intensity, and pattern, antecedent soil moisture, and spring freeze-thaw cycles. Seasonal variability in coarse suspended matter concentrations depends mainly on the streamflow conditions that are realized in the systems of interest.

Three additional sources of uncertainty in tributary mouth loading estimates will be discussed: seiching, near upstream point sources, and gauge to confluence distance. Seiching is a condition that gives rise to periodic reversals of flow at the tributary-lake confluence and leads to an especially difficult load estimation problem. The problem can be so severe that loads of conservative as well as non-conservative contaminants cannot be estimated by projection from a near upstream monitoring point without unreasonably high uncertainty. The main factors prohibiting extrapolation of the upstream estimate are transport time phase error and physical phenomena dependent on the particle size distribution of the solid phase matter, sedimentation, resuspension, and delivery to the receiving water (Bedford and Mark, 1988).

Consideration must be given to uncertainty attributable to upstream point sources that are proximate to the confluence of concern. The uncertainty such sources contribute to the load estimate is that due to inadequate mixing over the cross-section and the resulting transverse gradients. Combined sewer overflows, which often occur in near-mouth reaches, also give rise to uncharacteristic flow-concentration responses during wet weather, and their influence may contribute uncertainty in the fluvial load estimate.

Commonly, an existing gauging site may not be the most desirable location for sample collection, or the chemical observations of interest may be pre-existing and from a site that is a substantial distance upstream from the confluence. Uncertainty in "extrapolating" the upstream measurements over the unmonitored portion of the tributary give rise to uncertainty in the tributary load estimate.

10.4.1.1 Data Collection Improvements

In the illustrative example given in Table 10.1, the largest source of loading uncertainty was "all other influent tributaries". Accordingly, the following discussion will concentrate on ways to reduce the uncertainty for tributary loads. Some of the recommendations discussed below will be applicable to other sources of uncertainty, which will be noted when appropriate.

Economic and logistical considerations can be optimized for given fixed dollars and resources (sampling crews, lab capacity, etc.). However, it is recommended that certain sampling *not* be optimized. Rather, some sampling should be designed for redundancy when possible. In general, physical and chemical variables that are relatively inexpensive to obtain but critical to load estimation and subsequent modeling efforts (such as flow, temperature, pH, chloride, suspended solids, nutrients and other conventional contaminants) should not be subject to optimization.

Ideally, existing jurisdictional programs for these parameters should continue or even be strengthened during the study. In addition, such optimal sampling that is undertaken for heavy metal or organic contaminants should also be designed to permit overlap sampling of the less costly parameters; this provides both a redundant measure and a way of relating the more costly measurements to less expensive ones in the historical record.

The sampling that should be optimized, of course, is that established for the more expensive, toxic contaminants. Tradeoffs actually occur at this level of sampling, due to the cost per sample for many of these chemicals, plus factors related to resources, such as laboratory capacity, field equipment, and expertise. The compromises will come down to accuracy of load estimates vs. cost of sampling. As a hypothetical example, suppose one wished to sample 95% of the tributary flow to Lake Ontario on a monthly basis for a suite of organic and metal contaminants. The list of tributaries that would provide this coverage includes:

Black River	Moir River	Salmon River
Buffalo River	Napanee River	Twelve Mile Creek
Credit River	Niagara River	Trent River
Genesee River	Oswego River	Welland River

In reality the largest of these tributaries, the Niagara River, is well sampled by an existing program, as referenced earlier (e.g. Data Interpretation Group, 1986, ff.). However, new efforts may be required to cover the remaining rivers. For monthly sampling, an estimate of the increased annual data acquisition cost would be:

$$\text{Annual Cost} = 11 \text{ Tributaries} \times 12 \text{ Samples/year} \times \text{Cost per Sample}$$

If this total is within the budget, then no optimization is necessary. However, if it is too expensive and no additional funds are available, the tradeoff occurs in number of samples (fewer tributaries or fewer months). Alternatively, if a specific maximum uncertainty is the constraint rather than costs, then a program could be designed to maintain the achieve this at minimum cost. Lesht (1991) provides a detailed treatment of this approach.

An additional alternative approach to this type of optimization is the Balanced Incomplete Block Design proposed by El-Shaarawi and Williams (1989) for Niagara River point sources. Briefly, it allows the estimation of the total load of a contaminant, the individual loads of larger tributaries and standard errors of the loads with substantially less samples needed than conventional designs. Again, assuming a fixed number of samples allocated to the effort, it should be possible to estimate the total tributary load, the relative contribution of at least four of the major tributaries and provide the equivalent of at least three replications at better than half the cost of a conventional design. This is a type of tradeoff also, but it occurs without giving up the primary objective, which is an accurate total load. This approach may not be appropriate for all scales of modeling but it should be considered because of the substantial savings it offers.

10.4.1.2 Data Handling Improvements

Improvements to reduce uncertainty in load and forcing function estimation in the area of data handling should not be subject to tradeoffs. The best available methods should be used so that no information that is collected is unused or partially used. In past projects this has occurred in two areas: load estimation methods and censored data handling.

Considerable effort has been devoted to demonstrating that load estimation methods that ignore the correlation between measured variables are unacceptable. Besides the fact that they do not use all available information, they introduce significant biases into load estimates. Ratio and regression estimates are two methods that do consider the correlation between variables and these should be used whenever possible.

One problem that inhibits the use of these methods is that of data censoring. This practice should be avoided whenever possible in the laboratory, but in many cases it is unavoidable. Similarly, if data are to be censored, they should have one detection limit reported. Again, sometimes this is impossible. Therefore, a generalized load estimation procedure is being developed for use with censored data that will maximize the information that can be obtained from partially censored loading data sets (up to 80% censoring). This involves a maximum likelihood estimate for statistical parameters that describe loading. This procedure will solve the problem of "minimum and maximum" load estimates that was caused by partial censoring of loadings data in the Green Bay study. The procedure is an extension of work by Dolan and El-Shaarawi (1989, 1991) and El-Shaarawi and Dolan (1989).

10.4.2 Point Sources

Quantitation of contaminant point source loadings to the Great Lakes may be done with greater certainty than non-point source loadings, because the technical and logistical aspects of sampling are less complicated for point sources. The simplicity of point source load estimation results from ignoring the spatial component of variability during monitoring, a convenience permitted by the point character of the source (the interfacial area, A, in Equations (10.7) or (10.8) is negligible). The degree of simplification this affords is such that a continuous record of an instantaneous load is possible for some contaminants (e.g. oxygen deficit). Of course, the extent to which the point source model is inappropriate for the monitoring data will lead to uncertainty in estimates based on it.

Most important to the overall uncertainty of most point source load estimation questions is the uncertainty associated with temporal variability of the load, whether due to variability in concentration or discharge or both. Temporal variability becomes a problem whenever a continuous record of the contaminant load is not available for the point source of interest. This is, of course, generally the case for Great Lakes contaminants, even though permitted dischargers are required to monitor their effluents routinely. As will be discussed later, proper sampling design can significantly reduce the influence of temporal variability on the uncertainty of point source load estimates.

The temporal variability in point source loads may possess a considerable level of determinism, or structure, that can be exploited to improve load estimate accuracy. Such structure that may be present will depend on the nature of the point source. Municipal wastewater effluents, for example, may show periodicity over 8 to 12 and 24 hour cycles for some contaminants, weekly cycles for others, and seasonal or annual cycles for still others. Industrial point sources, on the other hand, tend to show less temporal structure, though they do reflect manufacturing or other production cycles. In general, variability in point sources is associated with community size, meteorologic and geographic factors, mixtures of influent sources (industrial/sanitary), and design flow rate of the facility of concern.

10.4.2.1 Data Collection Improvements

The location, design flow and industrial category of all point sources in the basin are known. This information should be used in optimizing the sampling design for point sources. A GIS would be most useful in identifying accurately the direct point sources.

The same optimization considerations discussed under tributaries are relevant with point sources. However, the practicality of mixing effluents to achieve the *Balanced Incomplete Block Design* is dubious, especially from a legal and analytical chemistry viewpoint. The question of utilizing effluent self-monitoring data is unresolved. Some self-monitoring data will probably be used, but toxic contaminant data obtained this way will usually be subject to less stringent quality assurance procedures.

10.4.2.2 Data Handling Improvements

In general, load estimation with point sources is easier, given adequate data. Ratio and/or regression estimations are not normally necessary, and a simple summation of the products of flow and concentration should be adequate. The same type of considerations for censored data apply to point sources.

10.4.3 Non-Point Sources

The major phenomena that produce uncertainty in non-point source load estimates for use in toxic chemical mass balance models consist of "true" forcing functions in the sense that they are truly independent variables that govern both the deterministic elements of the quantities represented in Equations (10.7) and (10.8) and the stochastic elements that are only implied by the previous discussion.

Direct non-point loads from rural runoff derive uncertainty from numerous sources, but it is possible to generalize these by stating that rural runoff is mainly associated with (1) random hydrologic events that have their roots in meteorologic or climatic phenomena, (2) sediment movement that accompanies runoff, and (3) partitioning of the chemical of interest between water and sediment. The forcing functions that govern uncertainty in rural runoff load estimates, therefore, relate mainly to these three phenomena. Physical forcing functions that affect rural runoff include meteorologic and climatic variables such as temperature, wind speed and

direction, precipitation quantity, intensity, and form; terrestrial attributes such as distribution and size of source areas, soil type and erodability, land use and management practices, slope, aspect, depression storage, and antecedent moisture conditions. Chemical forcing functions that lead to uncertainties in rural runoff load estimation include the phase partitioning behavior of the contaminants of interest and the rate, timing, and mode of environmental release of the contaminant in the source area. Biological variables that influence the degree of certainty with which rural runoff loads may be determined include seasonal and ecological succession, evapotranspiration, migration, and distribution of ground cover.

In a general way, urban runoff loading of contaminants depends on the same factors as those cited above for rural runoff loads. Urban runoff loads normally differ substantially, however, in that they generally reflect much higher rates of runoff compared to rural ones. This is due to the generally greater imperviousness of surfaces that intercept precipitation. Further, urban runoff normally is collected for routing to ultimate disposal. Under circumstances that are favorable to surface receiving waters, such routing is through separate storm sewers and setting basins followed by application to infiltration beds to recharge groundwater. Under less favorable circumstances, the routing is through combined sewers (runoff and sanitary wastewater) that may be allowed to overflow during runoff events that exceed the volumetric design capacity of wastewater treatment facilities. Urban runoff loads also differ from rural ones in that urban loads depend on population density and street cleaning effort. Empirical models (e.g. STORMS - US Army Corps of Engineers) are available to aid in estimating urban runoff loads, however, monitoring is required for any quantitative assessment of estimate uncertainty, which may be large.

Dissolved phase contaminants predominate in *groundwater* and are transported into (and out of) the receiving waterbody by advection, dispersion, and diffusion at the sediment-water interface.

Contaminant loads from infiltration of contaminated groundwater are difficult to quantify directly and frequently are estimated by mass balance closure; consequently, they are often notably uncertain. Among the numerous determinants of groundwater contaminant transport are included aquifer thickness, hydraulic conductivity or transmissivity and degree of anisotropy, hydraulic gradient, proximity of recharge and contaminant source areas, and such contaminant attributes as partitioning behavior, susceptibility to decay, and concentration gradients. When groundwater loading of contaminants is estimated directly, then each of the parameters named above can contribute uncertainty to the result.

Contaminated sediments, or in-place pollutants, may enter the overlying water in the dissolved or particulate phase. Mechanisms that lead to the introduction of in-place pollutants into the water column include bioturbation, the net result of scour, resuspension and sedimentation; diffusion, and advection (via groundwater). At least in principle, each of these mechanisms may be quantified, but each is determined by parameters that are difficult to measure directly, that vary spatially and temporally, and that generally possess high uncertainty. The determinants of in-place contaminant loading rates are numerous. Some of the more significant ones include temperature, and contaminant-specific diffusion and dispersion coefficients, hydraulic gradient and conductivity, viscosity (water and surficial sed-

iment slurry), porosity and bulk density of sediments, thickness of active sediment layer and sediment cohesiveness, velocity of overlying water and boundary shear stress; state of water column stratification (thermal) and internal current structure; burrowing invertebrate population type, density, activity level; migration of biota and seasonal succession; chemical-specific characteristics, such as partitioning behavior, volatility, solubility; sediment particle fraction of organic carbon; concentrations of contaminants in water, pore water, and sediment particles; and, antecedent conditions.

Direct atmospheric deposition of contaminants consists of two spatially and temporally variable processes which are subject to significant uncertainty: dry and wet deposition (scavenging). Dry deposition fluxes consists of the air-to-water interfacial transport of contaminants bound to particulate matter, in the gas phase, or dissolved in water vapor. The physical processes that lead to dry deposition of particulate matter at the water surface are size dependent and include gravitational settling for large particles (> 5 μm), Brownian motion for fine particles (< 0.3 μm), plus interception and impaction for intermediate size particles. In addition to the transport rate dependence on particle size, the phase distribution of atmospheric contaminants also depends on particle size, and smaller particles normally contain higher amounts of contaminants per unit of weight or volume. Gas phase contaminants also may be deposited after transport to the water surface. Wet deposition of contaminants occurs as a result of the tendency of precipitation to cleanse the atmosphere by scavenging particulate matter and absorbing gases. Transport of wet deposition occurs largely by gravity, though the transport processes that permit scavenging and absorption processes are more diverse. The principal factors causing uncertainty in estimating atmospheric loads of contaminants include the meteorologic factors involved in atmospheric transport (i.e. wind speed and direction, temperature, precipitation kind and amount), attributes of the contaminants of interest (concentration gradients, partitioning behavior), and the receptor area.

10.4.3.1 Data Collection Improvements

Of the non-point sources discussed above, the most important is probably atmospheric deposition. Fortunately, this source also has the best prospects for data collection improvements because of current research and monitoring programs on air toxics. Urban and rural non-point sources are less important because only direct sources need to be estimated; indirect sources are covered by tributary monitoring. At the scale of Lake Ontario, it is doubtful that new samples of urban and rural runoff would be taken, unless these sources are much larger than anticipated. The importance of groundwater on a lakewide basis is unclear. Sampling should occur in the most susceptible areas and a judgement made whether further sampling is warranted. In-place pollutants may not be considered as external forcing functions, depending on the model formulation. It is probable that the level of modeling used on Lake Ontario would include contaminated sediments.

10.4.4 Data Handling Improvements

Once again, censored data are an issue for non-point sources, especially for atmospheric deposition monitoring. Load estimation methods will have to be worked out that allow use of partially censored data. Also, the estimation procedures for runoff and groundwater need to be researched thoroughly.

10.5 Conclusions

Several opportunities exist for reducing uncertainty in forcing functions or their estimates. Many of these involve "up front" planning and analysis and require historical or typical data to use in this effort. Other methods involve using the best available methods for processing and analyzing data once it is collected.

It is recommended that the first order uncertainty analysis, as outlined and illustrated in this paper, be conducted with the model formulation to be used in the specific application or management problem analysis. Some type of optimization should be done, subsequently, to reduce the sources of uncertainty indicated as most significant by the first order analysis. This effort should use real data rather than synthetic or "literature values" wherever possible. For data collection not subject to optimization, it is recommended that redundancy be built into the sampling effort. Additionally, a thorough understanding of the sources of uncertainty that affect loadings and other forcing functions is important to the success of efforts at improving the usefulness of monitoring programs.

References

Bedford, K.W. and D.J. Mark. 1988. Effect of storms and gage location on tributary load estimate. J. Environ. Engr. Div. ASCE 114(2):352-367.

Benjamin, M.M. and J.O. Leckie. 1981. conceptual model for metal-ligand interactions during adsorption. Environ. Sci. Technol. 15:1050-1057.

Cohen, T.A., L.L. DeLong, E.J. Gilroy, R.M. Hirsch, and D.K. Wells. 1989. Estimating constituent loads. Water Resources Research 25(5):937-942.

Cooke, G.D., E.B. Welch, S.A. Peterson, and P.R. Newroth. 1986. Lake and Reservoir Restoration. Butterworths/Ann Arbor Science, Boston MA. 392 pp.

Cornell, C.A. 1972. First order analysis of model and parameter uncertainty. pp. 1245-1274. in: Proceedings international Symposium on Uncertainties in Hydrologic and Water Resource Systems. Vol. III. C. Kisiel and L. Duckstein, eds. University of Arizona, Tucson AZ.

Data Interpretation Group. 1986. Joint Evaluation of Upstream/Downstream Niagara River Monitoring Data 1984-1986. Joint publication of New York State Department of Environmental Conservation, Environment Canada, U.S. Environmental Protection Agency, and Ontario Ministry of the Environment.

Data Interpretation Group. 1988. Joint Evaluation of Upstream/Downstream Niagara River Monitoring Data 1986-1987. Joint publication of New York State Department of Environmental Conservation, Environment Canada, U.S. Environmental Protection Agency, and Ontario Ministry of the Environment.

Data Interpretation Group. 1989. Joint Evaluation of Upstream/Downstream Niagara River Monitoring Data 1987-1988. Joint publication of New York State Department of Envi-

ronmental Conservation, Environment Canada, U.S. Environmental Protection Agency, and Ontario Ministry of the Environment.

Data Interpretation Group. 1990. Joint Evaluation of Upstream/Downstream Niagara River Monitoring Data 1988-1989. Joint publication of New York State Department of Environmental Conservation, Environment Canada, U.S. Environmental Protection Agency, and Ontario Ministry of the Environment.

Data Interpretation Group. 1991. Joint Evaluation of Upstream/Downstream Niagara River Monitoring Data 1989-1990. Joint publication of New York State Department of Environmental Conservation, Environment Canada, U.S. Environmental Protection Agency, and Ontario Ministry of the Environment.

Data Interpretation Group. 1992. Joint Evaluation of Upstream/Downstream Niagara River Monitoring Data 1990-1991. Joint publication of New York State Department of Environmental Conservation, Environment Canada, U.S. Environmental Protection Agency, and Ontario Ministry of the Environment.

Data Interpretation Group. 1993. Joint Evaluation of Upstream/Downstream Niagara River Monitoring Data 1991-1992. Joint publication of New York State Department of Environmental Conservation, Environment Canada, U.S. Environmental Protection Agency, and Ontario Ministry of the Environment.

Dolan, D.M. and A.H. El-Shaarawi. 1991. Applications of mass balance approach with censored data. J. Great Lakes Res. 17(2):220-228.

Dolan, D.M. and A.H. El-Shaarawi. 1989. Inferences about point source loadings from upstream/downstream river monitoring data. Environmental Monitoring and Assessment 12:343-357.

El-Shaarawi, A.H. and D.M. Dolan. 1989. Maximum likelihood estimation of water quality concentrations from censored data. Can. J. Fish. Aquat. Sci. 46(6):1033-1039.

El-Shaarawi, A.H. and D.J. Williams. 1989. Status of the Niagara River point source discharge information: Sampling design and estimation of loading. Environmental Monitoring and Assessment 12:331-341.

Goldman, C.R. and A.J. Horne. 1983. Limnology. McGraw-Hill, New York. 464 pp.

IJC. 1987. Guidance on Characterization of Toxic Substances Problems in Areas of Concern in the Great Lakes Basin. International Joint Commission. Ottawa ON, Washington DC. 179 pp.

IJC. 1992. Sixth Biennial Report on Great Lakes Water Quality. International Joint Commission. Ottawa ON, Washington DC. 60 pp.

Iman, R.L. and M.J. Shortencarier. 1984. A Fortran 77 Program and User's Guide for the Generation of Latin Hypercube and Random Samples for Use with Computer Models. NUREG/CR-3624, SAND83-2365. Sandia National Laboratories. Albuquerque NM. 50 pp.

Lesht, B.M. 1991. Design of an optimal tributary monitoring strategy for estimating contaminant loads. International Journal of Environmental Studies 37:225-234.

Likens, G.E. 1972. Nutrients and Eutrophication: The Limiting Nutrient Controversy. Special Symposia Volume I. American Society of Limnology and Oceanography. G.E. Likens [ed.] Allen Press, Lawrence KS. 328 pp.

Mackay, D., M. Diamond, S. Sang, and P. Ulahos. 1993. A simple rate constant model of chemical dynamics in a lake ecosystem. Journal of Great Lakes Research. (submitted May 1993).

National Academy of Sciences. 1970. Eutrophication: Causes, Consequences, Correctives. National Academy of Sciences, Washington DC. 661 pp.

Ohlem, H. and G. Flock. 1990. Lake and Reservoir Restoration Guidance Manual. 2nd Edition. EPA 440/4-90-006. U.S. Environmental Protection Agency, Washington DC.

Preston, S.D., V.J. Bierman, Jr., and S.E. Silliman. 1989. An evaluation of methods for the estimation of tributary mass loads. Water Resources Research 25(6):1379-1389.

Reckhow, K.H. and S.C. Chapra. 1983. Engineering Approaches for Lake management, Volume I: Data Analysis and Empirical Modeling. Butterworth Publ. Boston MA.

Scavia, D.W., F. Powers, R.P. Canale, and J.L. Moody. 1981. Comparison of first-order error analysis and Monte Carlo simulation in time dependent lake eutrophication models. Water. Resour. Res. 17(4):1051-1059.

Schnoor, J.L., C. Sato, D. McKechnie, and D. Sahoo. 1987. Processes, Coefficients, and Models for Simulating Toxic Organics and Heavy Metals in Surface Waters. EPA/600/3-87/015. U.S. Environmental Protection Agency, Washington DC. 303 pp.

Thomann, R.V. and J.A. Mueller. 1987. Principles of Surface Water Quality Modeling and Control. Harper and Row, New York. 644 pp.

Vollenweider, R.A. 1968. Scientific Fundamentals of the Eutrophication of Lakes and Flowing Waters,with Particular Reference to Nitrogen and Phosphorus as Factors in Eutrophication. Organization for Economic Cooperation and Development. DAS/CSI/68.27. Paris. 274 pp.

Wedepohl, R.E., D.R. Knauer, G.B. Wolbert, H. Olem, P.J. Garrison, and K. Kepford. 1990. Monitoring Lake and Reservoir Restoration (Technical Supplement). EPA 440/4-90-007. U.S. Environmental Protection Agency, Washington DC.

Wetzel, R.G. 1983. Limnology. 2nd Edition. Saunders, Philadelphia PA. 767+91 pp.

Wu, S.C. and P.M. Gschwend. 1986. Sorption kinetics of hydrophobic organic compounds to natural sediments and soils. Environ. Sci. Technol. 20:717-725.

Young, T.C., J.V. DePinto, and T.M. Heidtke. 1988. Some factors affecting fluvial load estimation efficiency. Water Resources Research 24(9):1535-1540.

CHAPTER 11

Modeling watershed water quality

Anthony S. Donigian, Jr., John C. Imhoff and Robert B. Ambrose, Jr.

11.1 Introduction

In recent years, a risk-based management approach to water quality protection is gaining increased popularity. This approach relies upon the identification and targeting of worst case problems as those needing initial emphasis. One of the currently available engineering tools that is believed to satisfy the greatest needs and provide the best foundation for future problem solving is the development and application of watershed[1] models to understand and predict the relationship of watershed conditions and management activities to sediment and chemical loadings to rivers and lakes. Usable, scientifically-sound tools that predict combined point and nonpoint source pollution loadings must be coupled to riverine and lake models so that the results of various management and control scenarios and options can be simulated and predicted. These dynamic models must be able to evaluate changes in loadings over temporal and spatial scales, and must consider both the amounts and speciation of contaminants.

The concept and practice of watershed water quality modeling is not new. This chapter begins with a brief synopsis (Section 11.2) of the evolution of watershed-scale water quality modeling. In Section 11.3, the role of modeling as a tool for environmental analysis is explored, and the types of analyses that benefit from modeling are identified. In Section 11.4 a brief introduction to watershed models and modeling is provided. Modeling fundamentals are introduced and relevant model features are discussed. Sections 11.5 and 11.6, respectively, describe methods and models for nonpoint source and for surface receiving water quality modeling. General guidance on model selection is provided for each category of models. We

[1] A formal definition of 'watershed modeling' does not exist; environmental professionals tend to use the terms 'watershed modeling', 'catchment modeling' and 'basin modeling' interchangeably to connote a modeling exercise that represents a closed system of water and its associated water quality constituents. Generally, basin-scale modeling is considered to be larger in scale than catchment- and/or watershed-scale modeling.

V. P. Singh (ed.), Environmental Hydrology, 377–426.
© 1995 *Kluwer Academic Publishers.*

present our projections concerning the future of watershed water quality modeling in Section 11.7. Needed improvements in both algorithmic/computational capabilities and in user interaction features are addressed. Closing remarks are provided in Section 11.8.

11.2 Evolution of Watershed-Scale Water Quality Modeling

For decades, the need has existed to address certain *water quantity* problems such as flood hazard and control and water supply on a watershed scale. Hydrologic and hydraulic models predicated on watershed-scale analysis have been in popular use in the United States and throughout many parts of the world since the early 1970's. Of particular note are the "HEC" series of models created and maintained by the U.S. Army Corps of Engineers Hydrologic Engineering Center in Davis, CA.

The concept and common practice of *water quality* modeling[2] at a watershed scale have lagged behind water quantity modeling. While certain pockets of professionals, such as the EPA Environmental Research Laboratory in Athens, Georgia, have been involved in watershed-scale process and modeling research for nearly two decades, practical application of such models has been somewhat infrequent and scattered. However, watershed modeling has been gaining in usage over the past few years and is likely to continue for the foreseeable future.

The development and application of watershed-scale water quality models have been largely driven by the need for tools that can be used to evaluate two issues: nonpoint source pollution effects (i.e., exposure and/or loadings assessment), and chemical management plans such as agricultural best management plans for fertilizers and pesticides or, more recently, total maximum daily loads (TMDLs), including both point and nonpoint contaminant sources.

In the late 1970's the regulatory requirements resulting from Section 208 of the Federal Water Pollution Control Act Amendments of 1972 supported the development and use of watershed-scale to basin-scale computer models to serve as the mechanism for evaluating long term, basin-scale water quality management plans for conventional pollutants. Unfortunately, at that time, nonpoint source and watershed water quality modeling tools were not available to fully satisfy the technical needs required by this regulatory mandate. Beginning in the early 1980's concern with the fate and transport of pesticides and other nonpoint source toxicants, again, pointed out the need for models that could evaluate water quality at a watershed scale. By the mid-1980's, in geographic areas where receiving waters continued to deteriorate in spite of the implementation of strong controls on point sources, the necessity of using watershed analysis tools grew to the point where certain progressive governmental bodies began to take the first steps to implement long term watershed/basin water quality modeling efforts. Of particular note is the watershed and bay modeling effort that was initiated by the Chesapeake Bay Program in the early 1980's, and which continues today. Further impetus was given to the watershed-scale modeling movement by Section 303(d) of the Clean Water Act by EPA's Water Quality Planning and Management Regulations (40 CFR Part

[2]For the purposes of this chapter 'water quality modeling' will include consideration of water temperature, sediment, dissolved oxygen, nutrients, biota, and toxic substances.

130) which require the consideration of tradeoffs between point and nonpoint source pollutant controls to develop what are essentially total watershed chemical management strategies.

In recent years, the full community of environmental professionals, including scientists, engineers, planners and regulators, have come to recognize that understanding the full impact of chemical releases to the environment and developing management strategies to protect environmental health are best accomplished by performing analyses at a watershed scale. In 1992, the EPA ORD Nonpoint Sources Research Plan for the following five years was prepared at the Athens Laboratory (EPA, 1992a). The plan outlined a broad array of research activities to be conducted in a watershed context, noting that "the logical assessment, management, and restoration unit is the watershed, the surrounding land and associated land uses that contribute to the aquatic system. Problems are identified in a watershed context; (therefore) long-term solutions require watershed or geographic area management plans that address the multiple sources of, and solutions to, the nonpoint source problem".

In addition to regulatory and management needs, the development of the current generation of watershed models has been shaped by the state of scientific understanding and the availability of data and computational/programming resources. With increased scientific understanding over the past two decades, models' ability to represent the important processes that determine chemical fate and transport has commensurately improved. More and better data have allowed modelers to more effectively calibrate and validate their models, which in turn has resulted in increasing confidence in both model results and in the use of models to evaluate alternative management strategies. Faster, more powerful, and less expensive computers have allowed use of models by a broader base of modelers, and for increasingly complex applications. Improved computer software has facilitated interaction between data, the model, and the modeler. All these factors will continue to contribute to the future directions that watershed model development and application will venture.

11.3 Why Model?

In modern life, modeling has become common practice in almost every field of endeavor, and water quality analysis and planning is no exception. The ultimate value of models of any kind is twofold: first, to understand and reproduce how the modeled system works, and then to predict how the system will respond to future alternative conditions and actions. Within a watershed, the water quality at any location is determined by the net contributions from all point and nonpoint sources. Evaluating the net effect that can be expected from implementing one or more actions within the watershed is not at all straightforward. For example, contour plowing of agricultural lands may reduce the surface runoff of nitrates from fertilizer, but at the same time may result in increased nitrate percolation to subsurface flows, and a subsequent threat to drinking water supplies derived from wells. Hence, reliance on intuition, engineering experience and/or simplified analysis techniques is not an appropriate means for making management decisions

that can have extreme economic and environmental consequences. The need to consider the full breadth of processes that determine the fate and transport of water-driven materials has given rise to the current popularity of the term "comprehensive water quality analysis". As a rule, truly comprehensive analysis cannot be achieved without the use of complex mathematical models.

11.3.1 Role of Modeling in Water Quality Analyses

The fate of chemicals in the environment is determined by the complex interaction of numerous factors including: physical and chemical properties of the chemical; the natural occurrence or application practices associated with the chemical; the characteristics of the surface and subsurface media through which transport occurs; and climatological conditions. In a similar manner, sediment fate and transport is affected soils properties, topography, weather and man-induced disturbances and controls.

Simulation models provide a mechanism to evaluate these interactions and estimate the extent that chemicals may move to non-target environments, such as surface receiving waters or groundwater aquifers. Further, modeling allows such evaluations to be performed in a time- and cost-effective manner when compared to conducting resource-intensive field monitoring studies. Generally, a high degree of confidence is associated with the results of a field study, because field studies reflect measured concentrations. However, field studies are most often not practical when dealing with watershed-scale issues. The results of a field study represent only one specific location under one set of environmental conditions. In contrast, simulation models permit the relative impacts of numerous environmental parameters to be evaluated. Since simulation models are mathematical representations of physical, chemical and biological processes, confidence in model results reflects the degree and accuracy that input parameters can be defined and the proven ability of the model to reproduce observed phenomena.

11.3.2 Environmental Analyses That Benefit from Modeling

A considerable breadth of water quality issues can be addressed by performing watershed-scale modeling. Modern, computer-based tools and approaches to watershed evaluation are appropriate to support activities including the following:

- Performing detailed exposure assessments as part of special chemical reviews;
- Providing total maximum daily loads (TMDLs) for combined point and non-point sources by comparing, in an integrated framework, the potential effects of implementing point source reductions, urban runoff control strategies and/or non-urban best management practices;
- Evaluating the effects of agricultural and silvicultural management practices on sediment erosion and potential exposure to pesticides and nutrients in groundwater, surface water and air;

- Evaluating the impacts of contaminant migration and exposure from uncontrolled hazardous waste sites;
- Evaluating the effectiveness of proposed remediation activities for nonpoint source Superfund sites, such as heavily mined areas;
- Evaluating multimedia exposure associated with disposal strategies for municipal wastewater sludges and pesticides;
- Evaluating the effectiveness of both structural and nonstructural nonpoint source control strategies;
- Evaluating both upland and receiving water habitat alteration and ecorisk problems;
- Designing water quality sampling networks.

11.4 Introduction to Watershed Models and Modeling

The focus of this chapter is to describe methods and models that are commonly used for watershed-scale analyses. However, before we do so, it is useful to provide a brief introduction to watershed models and modeling. In this section three topics are addressed. First we highlight the fundamentals involved in any modeling exercise. Next, we describe two critical model features (i.e., level of analysis, breadth of state variables/process representation) that most often determine the appropriateness of selecting a particular model to evaluate a given problem. Finally, the three major components of traditional watershed water quality analysis (urban runoff, nonurban runoff, surface receiving water quality) are introduced to set the stage for subsequent , more detailed discussions in Sections 11.5 and 11.6. The concept of integrated watershed water quality modeling is introduced, and current capabilities are briefly discussed.

11.4.1 Modeling Fundamentals

Introductions to modeling and modeling caveats have been presented by several authors including James and Burges (1982), Kibler (1982), Huber (1985, 1986), Donigian and Huber (1991), and EMWG (Exposure Modeling Work Group of the National Agricultural Chemicals Association) (1994); additional summaries have been provided in at least one manual of practice (WPCF, 1989), as well as the Federal Register (EPA, 1992b). While space does not permit a full discussion of the necessary activities for successful model applications, a few fundamentals are highlighted sequentially below.

1. Have a clear statement of project objectives. Verify the need for quality modeling. (Perhaps the objectives can be satisfied without modeling).
2. Identify the processes that govern chemical behavior.
3. Formulate questions to be answered and scenarios to address them.

4. Establish constraints (time frame, costs, technology, availability and accuracy of data).

5. Assess the required level of sophistication.

6. Identify and evaluate candidate models.

7. Select appropriate model. Use the simplest model that will satisfy the project objectives. Often a screening model, e.g., regression or statistical, can determine whether more complex simulation models are needed. To the extent possible, utilize a quality prediction method consistent with available data.

8. Only predict the quality parameters of interest and only over a suitable time scale. Seasonal or annual loads will sometimes be all that is required.

9. Perform a sensitivity analysis on the selected model and familiarize yourself with the model characteristics.

10. If possible, calibrate and verify the model results. Use one set of data for calibration and another independent set for verification. If no such data exist for the application site, perhaps they exist for a similar catchment nearby.

11. Apply model to answer questions.

12. Reduce results; perform visualization and statistical analysis.

13. Assure quality assurance/quality control.

14. Implement management strategies based on model results; educate managers or regulators on appropriate modeling caveats.

Proper model applications account for relevant governing processes at an appropriate detail (time step and spatial scale) and accuracy relative to achieving the objectives of the study. The level of sophistication required in a modeling study reflects constraints such as: (1) accuracy required, (2) allotted time frames, (3) current technology in describing environmental fate behavior and (4) availability and reliability of input data. These constraints dictate model selection, whether modeling is an appropriate tool for achieving the objectives, and the limitations in interpreting the model results. Model selection should also consider the availability of pre- and post-processors which are designed to facilitate model setup and interpretation of results, quality and organization of model documentation, and availability of user support. Significantly, managers and regulatory officials must be educated on assumptions and limitations associated with interpreting the results of a modeling study.

Confidence in model results is based on the quality of data used to construct the model, the capabilities of the modeler, and the proven ability of the model to simulate observed phenomena (EMWG, 1994).

11.4.2 Relevant Model Features

In this section we examine two critical model features, and how they relate to potential model usage: (1) level of analysis and (2) breadth of "watershed processes" considered.

11.4.2.1 Level of Analysis

The level of analysis is measured by the amount and resolution of the data used, and the sophistication of the analysis employed. It is determined by the purpose of the assessment and the resources available to perform the assessment. Typically, individual water quality models are classified by using one of three designations that depend on the level of analysis that can be achieved by their use: screening, primary or secondary models. Since models featuring each of these three levels of analysis are useful for performing specific tasks related to watershed analysis, a brief description of each level is provided; the descriptions below are adapted from the work of the EMWG (1994). Note that subsequent discussions of specific models in Sections 11.5 and 11.6 are limited to "primary" or "secondary" models to better reflect the inherent complexity of most watershed-scale modeling exercises.

Screening Models – Screening models are used to provide a rapid examination of the environmental fate of a chemical and typically involve either simple algorithms employing "default" values or assumptions for environmental variables. Screening models focus on a limited number of state variables and do not attempt to represent complex interactions. A screening model is not intended to give a comprehensive evaluation of a chemical's behavior in a site-specific context. The accuracy of screening model results can often not be ascertained, and the results are more appropriately interpreted as qualitative as opposed to quantitative. Example applications of screening models for watershed water quality analysis include: (1) evaluating the relative impact that different soil conditions have on chemical behavior to determine sensitive areas or areas that require more detailed analysis, (2) comparing the environmental fate of a chemical relative to a group of other chemicals, (3) comparing the relative chemical migration in variable climatological regimes, (4) evaluating the relative erosion potential of the predominant soil types within a study area and (5) evaluating the relative effects of rainfall characteristics on chemical or sediment washoff using simplified rainfall/runoff relationships.

Primary Models – Primary models provide a more detailed representation of predominant environmental fate processes. Primary models can provide an intermediate level of quantitative answers; that is to say the numbers are produced by a sufficient level of process representation that they have some meaning. Example applications of primary models for watershed water quality analysis include: (1) evaluating, at a field scale, the surface and subsurface transport of applied pesticides taking into account the effects of agricultural management practices, (2) evaluating the risk of exposure to contaminants at, or above, maximum concentration levels (MCLs), and (3) estimating, on a unit area basis, the sediment or chemical loadings contributed to downfield or downstream locations by conditions/actions at an upstream location.

The most appropriate use of screening models is to identify dominant processes and variables. As a rule, primary models do not feature capabilities to model systems that have spatially heterogenous conditions (e.g. variable soil types, weather conditions, etc.), or to evaluate the net effects of fate and transport though linked, heterogenous flow paths. However, by providing a means for identifying the relative importance of conditions/actions, they can be useful in determining the spatial resolution necessary to evaluate systems or issues that require more detailed modeling via secondary models. By using primary models and standardized approaches, a consistent basis for making regulatory and management decisions can often be achieved in a timely and cost-effective manner.

Secondary Models – Secondary models are appropriate for chemical- and site-specific concerns that cannot be adequately addressed through primary models. This category of models are the only ones that can be used to perform "comprehensive" analyses of water quality phenomena. The distinction between primary and secondary models is admittedly somewhat arbitrary. A model tends to be classified as secondary instead of primary as its capability to represent complex interrelated processes increases; as its capability to represent spatially heterogenous conditions increases; and/or as its capability to evaluate the net effects of fate and transport though linked, heterogenous flow paths increases. Example applications of secondary models for watershed water quality analysis include: (1) evaluating the ultimate fate and transport of fertilizer nutrients via a model that utilizes detailed algorithms for surface, subsurface and receiving water processes; (2) modeling both the point and nonpoint sources of a specific contaminant to determine the relative risk to the studied area associated with each of the sources; and (3) evaluating, quantitatively, the effects of agricultural best management practices throughout a watershed on the resulting pesticide loadings to endpoint receiving waters.

As a rule, secondary models consider more than one environmental medium, and the interaction of multiple water quality constituents with each other and with the physical characteristics of each medium.

To conserve resources, most assessments are done in an iterative fashion, with a screening done first; successive iterations add more detail and sophistication (EPA, 1992b). Typically, the process of moving from one level of assessment to another has required the use of different modeling tools in different frameworks. This transition has often resulted in a major investment in time and resources. Since different models have been designed to be fueled by different database structures and to require different input parameters, much of the preliminary work (i.e., database development, model configuration, model parameterization) required prior to running the model and generating the results needed to perform a more detailed analysis must be repeated. However, a limited number of software products enable the use of analysis techniques directed at the same problem, or a similar problem, but representing different levels of analysis, to be used within a single modeling system. The analysis techniques share a common database and interaction framework, thus minimizing the effort involved in moving from a simpler to a more sophisticated level of analysis.

11.4.2.2 State Variable/Process Representation

Comprehensive watershed water quality modeling must represent the dominant processes that determine the interactions between the following state variables:

- Water
- Temperature
- Sediment
- Dissolved oxygen
- Nutrients (nitrogen and phosphorus species) and bacteria (coliforms)
- Pesticides and other toxic organics
- Metals
- Selected biological variables (phytoplankton, benthic algae, zooplankton, herbivorous fish, predatory fish))

With the exception of biological variables, all those listed above can be critical to modeling water quality in all three of the watershed components (i.e., urban nonpoint sources, non-urban nonpoint sources, and surface receiving waters) described in detail in this chapter. The most important processes that must be modeled for each of the state variables, or category of state variables, are described briefly below.

Water itself is, of course, the quintessential state variable for a watershed-scale modeling system. Critical processes that must be represented are all those contained in the hydrologic cycle: precipitation, evaporation, interception, frozen ground infiltration, snowmelt, infiltration, evapotranspiration, runoff, subsurface flow, streamflow, and lake/reservoir hydrodynamic processes (wave action, stratification, turnover, etc.).

Modeling temperature is fundamental to virtually all hydrodynamic, water quality, and soil process modeling activities (the same is true of aerodynamic and air quality models). Natural and man-made processes that result in sources or sinks of temperature to each medium that is modeled must be considered. Important processes include: heat exchange due to the incident shortwave radiation, longwave back radiation, conduction-convection, evaporation and precipitation; thermal enrichment from heat absorbing surfaces; and shading.

The ability to model sediment fractions (sand, silt, clay) for eroded sediments, from both pervious and impervious land surfaces, and within receiving waters is needed. Processes that must be represented include: influx from air due to wind action or atmospheric deposition; attachment (compaction); detachment by rainfall; washoff of detached sediment from pervious areas; rill erosion; accumulation and washoff from impervious land surface area; receiving water bank erosion; and deposition, scour, and resuspension within receiving waters. In addition to natural processes, man-driven processes such as tillage operations, construction and mining disturbances, silvicultural practices, and dredging need to be represented.

Dissolved oxygen modeling within aquatic media (i.e., surface, subsurface, receiving waters) is needed to support analysis of nutrients and biological parameters. Important processes include transfer across media boundaries; depletion by sediment oxygen demand, nitrification, carbonaceous oxygen demand, and biotic respiration; and influx due to photosynthesis.

Nitrogen species that must be represented are ammonium, ammonia, nitrite, nitrate, and organic nitrogen. Important processes in the nitrogen cycle include volatilization, ammonification (or mineralization), nitrification, denitrification, immobilization, sorption/desorption, plant uptake, nitrogen fixation, algal uptake and death of organisms. Phosphorus species that must be represented are orthophosphorus and organic phosphorus. Important processes in the phosphorus cycle include sorption/desorption, organic decomposition, plant uptake, algal uptake and death of organisms. Interactions between the water column and the benthos are also critical processes for nutrient cycling in many waterbodies. In rural and agricultural regions, coliforms and bacteria are associated with many nutrient forms from septic tanks and animal waste disposal/handling procedures, leading to major health and water quality concerns. Quantifying contaminant sources, in addition to the transport and fate (die-off) of bacteria are the key modeling issues to be addressed.

Additional chemicals that must be modeled by a comprehensive water quality model are (1) pesticides and other toxic organics and (2) metals. Given the large number of constituents in these two categories of abiotic state variables, a generalized approach to consideration of processes is necessary. The general physical and chemical processes important for fate and transport are adsorption/desorption, diffusion, atmospheric deposition, volatilization, and chemical reactions/degradation. Individual chemical reaction processes that need to be represented include hydrolysis, oxidation/reduction, ionization, photolysis, complexation, and dissolution/precipitation. Many of these chemical fate processes need to be considered in both the water column and the benthos compartments. In addition to the processes that comprise the transformation of these chemicals, the following natural chemical source processes must be represented: surface and subsurface release of metals from ore; generalized washoff of chemicals from the land surface; release of metals and organic toxics from disposal or storage sites; application of pesticides to pervious areas; and atmospheric deposition.

For biological state variables, the processes that need to be represented are different for the different trophic levels. For phytoplankton, important processes are growth, death, sinking and predation. Benthic algae share all these processes except sinking; instead, they are subject to "sloughing", or detachment from the benthic stratum. Important processes affecting zooplankton include growth, respiration, death and predation. Fish share all zooplankton processes; additional important processes for fish include exposure to environmental stresses such as high temperatures, low dissolved oxygen, toxic chemicals, and sedimentation. The effects on biological systems induced by acid rain and global climate change must also be represented.

11.4.3 Watershed Water Quality Modeling Components

Most of the water quality models that are commonly used today pre-date the current emphasis on comprehensive watershed water quality analysis, as described in Section 11.2. Typically, models were developed to provide a means to independently evaluate either urban runoff, non-urban runoff, or receiving water quality. Given this development history, it is most convenient in the following sections to describe the currently available models according to these three categories. Section 11.5 describes the methods and models that have been developed and most successfully used to evaluate watershed nonpoint source pollution; urban and non-urban methods and models are discussed separately. Section 11.6 describes the methods and models that we believe are the most useful for evaluating surface receiving water quality. The topics of air, groundwater and ecosystem modeling are addressed only in terms of their interface with land surface nonpoint source and receiving water models. Since the focus of this chapter is using models to evaluate watershed-scale issues, near-field receiving water models are not considered.

A limited number of modeling packages described in the above sections allow modelers to perform comprehensive watershed analysis (i.e., point and nonpoint source modeling plus surface water modeling) within a single framework; such models can be defined as "integrated watershed quality models". Included in this category are the HSPF, SWWM and SWWRB models. It has been common practice to use others of the component models described to perform integrated analyses via providing external linkages that utilize the output of nonpoint source models as input to receiving water models. This practice is both time- and cost-intensive, and is likely to give way to a greater focus on the development and use of integrated watershed modeling software in future years.

11.5 Methods and Models for Nonpoint Source Water Quality Modeling

Stormwater runoff from all categories of land use — urban, agricultural cropland, pasture, and forest — is involved in many water quality problems. Models can be used to help in setting total maximum daily loads (TMDLs) for basins or watersheds, in registering new pesticides or assessing the risk of current pesticide use, or in assessing the risk from hazardous waste sites. Alternate methods are available, and in common use, for performing water quality analyses for watershed components. In this section, we discuss methods for urban and non-urban nonpoint source analysis, and the models that incorporate these methods. The primary reference for the following discussion of methods and models for nonpoint source water quality modeling is Donigian and Huber (1991).

11.5.1 Urban Runoff Modeling Methods

Several modeling options exist for simulation of quality in urban storm and combined sewer systems. These have been reviewed by Huber (1985; 1986) and range from simple to involved, although some "simple" methods, e.g., the EPA statistical methods, can incorporate quite sophisticated concepts. The principal methods

available to the contemporary engineer, in a rough order of complexity, are as follows:

- Constant concentration
- Spreadsheet
- Statistical
- Rating curve or regression
- Buildup/washoff

A comprehensive description of each of these five methods is provided in two references: Donigian and Huber (1991) and Donigian et al. (1993). Of the five methods the latter three have practical applications to watershed water quality modeling, and are discussed below.

11.5.1.1 Statistical Method

The "Statistical Method" utilizes the fact that event mean concentrations, or EMCs (i.e., total storm event load divided by total storm event runoff volume), are not constant but tend to exhibit a lognormal frequency distribution. When coupled with an assumed distribution of runoff volumes (also lognormal), the distribution of runoff loads may be derived. When coupled again to the distribution of streamflow, an approximate (lognormal) probability distribution of in-stream concentrations may be derived (DiToro, 1984) – a very useful result, although assumptions and limitations of the method have been pointed out by Novotny (1985) and Roesner and Dendrou (1985). Further analytical methods have been developed to account for storage and treatment (DiToro and Small, 1979; Small and DiToro, 1979). The method was used as the primary screening tool in the EPA NURP studies (EPA, 1983) and has also been adapted to combined sewer overflows (Driscoll, 1981) and highway-related runoff (Driscoll et al., 1989). This latter publication is one of the best for a concise explanation of the procedure and assumptions and includes spreadsheet software for easy implementation of the method.

A primary assumption is that EMCs are distributed lognormally at a site and across a selection of sites. The concentrations may thus be characterized by their median value and by their coefficient of variation (CV = standard deviation divided by the mean). Although the validity of the lognormality assumption has been demonstrated (Driscoll, 1986), the method is then usually combined with weak hydrologic assumptions, e.g., prediction of runoff using a runoff coefficient. (The accuracy of a runoff coefficient increases as urbanization and imperviousness increase.) However, since many streams of concern in an urban area consist primarily of stormwater runoff during wet weather, the ability to predict the distribution of EMCs is very useful for assessment of levels of exceedance of water quality standards. The effect of BMPs can again be estimated crudely through constant removal fractions that lower the EMC median, but it is harder to determine the effect on the coefficient of variation. Overall, the method has been very successfully applied as a screening tool.

Input to the method includes statistical properties of rainfall (mean and coefficient of variation of storm event depth, duration, intensity and inter-event time), area, and runoff coefficient for the hydrologic component, plus EMC median and coefficient of variation for the pollutant. Generalized rainfall statistics have already been calculated for many locations in the U.S. Otherwise, the EPA SYNOP model (Hydroscience, 1979; EPA, 1983; Woodward-Clyde, 1989) must be run on long-term hourly rainfall records. If receiving water impact is to be evaluated the mean and CV of the streamflow are required plus the upstream concentration.

The choice of median concentration may be difficult, and the Statistical Method requires a coefficient of variation as well. Fortunately, from NURP and highway studies, CV values for most urban runoff pollutants are fairly consistent, and a value of 0.75 is typical. If local and/or NURP data are not available or inappropriate, local monitoring may be required, as in virtually every quality prediction method. The estimation of the whole EMC frequency distribution for a pollutant is a definite advantage of the Statistical Method over some applications of constant concentration and simple spreadsheet approaches. Frequency analyses of water quantity and quality parameters may also be performed on the output of continuous simulation models such as HSPF, SWMM and STORM. The derived distribution approach of the Statistical Method avoids the considerable effort required for continuous simulation at the expense of simplifying assumptions that may or may not reflect the prototype situation adequately.

11.5.1.2 Regression - Rating Curve Approaches

With the completion of the NURP studies in 1983, there are measurements of rainfall, runoff and water quality at well over 100 sites in over 30 cities. Some regression analysis has been performed to try to relate loads and EMCs to catchment, demographic and hydrologic characteristics (e.g., Miller et al., 1978; Brown, 1984), the best of which are recent results of the USGS (Tasker and Driver, 1988), to be described briefly below. Regression approaches have also been used to estimate dry-weather pollutant deposition in combined sewers (Pisano and Queiroz, 1977), a task at which no model is very successful. What are termed "rating curves" herein are just a special form of regression analysis, in which concentration and/or loads are related to flow rates and/or volumes.

A rating curve approach is most often performed using total storm event load and runoff volume although intra-storm variations can sometimes be simulated in this manner as well (e.g., Huber and Dickinson, 1988). It is usually observed (EPA, 1983; Driscoll et al., 1989) that concentration (EMC) is poorly or not correlated with runoff flow or volume, implying that a constant concentration assumption is adequate. Since the load is the product of concentration and flow, load is usually well correlated with flow regardless of whether or not concentration correlates well. Manifestation of spurious correlation (Benson, 1962) is often ignored in urban runoff studies. If load is proportional to flow to the first power (i.e., linear), then the constant concentration assumption holds; if not, some relationship of concentration with flow is implied. Rating curve results can be used by themselves for load and EMC estimates and can be incorporated into some models (e.g., SWMM, HSPF).

Rainfall, runoff and quality data were assembled for 98 urban stations in 30 cities (NURP and other) in the U.S. for multiple regression analysis by the USGS (Driver and Tasker, 1988). Thirty-four multiple regression models (mostly log-linear) of storm runoff constituent loads and storm runoff volumes were developed, and 31 models of storm runoff EMCs were developed. Regional and seasonal effects were also considered. The two most significant explanatory variables were total storm rainfall and total contributing drainage area. Impervious area, land use, and mean annual climatic characteristics also were significant explanatory variables in some of the models. Models for estimating loads of dissolved solids, total nitrogen, and total ammonia plus organic nitrogen (TKN) generally were the most accurate, whereas models for suspended solids were the least accurate. The most accurate models were those for the more arid Western U.S., and the least accurate models were those for areas that had large mean annual rainfall.

These USGS equations represent the best generalized regression equations currently available for urban runoff quality prediction. Note that such equations do not require preliminary estimates of EMCs or local quality monitoring data except for the very useful exercise of verification of the regression predictions. Regression equations only predict the mean and do not provide the frequency distribution of predicted variable, a disadvantage compared to the statistical approach. (The USGS documentation describes procedures for calculation of statistical error bounds, however). Finally, regression approaches, including rating curves, are notoriously difficult to apply beyond the original data set from which the relationships were derived. That is, they are subject to very large potential errors when used to extrapolate to different conditions. Thus, the usual caveats about use of regression relationships continue to hold when applied to prediction of urban runoff quality.

11.5.1.3 Buildup and Washoff

In the late 1960s, a Chicago study by the American Public Works Association (1969) demonstrated the (assumed linear) buildup of "dust and dirt" and associated pollutants on urban street surfaces. During a similar time frame, Sartor and Boyd (1972) also demonstrated buildup mechanisms on the surface as well as an exponential washoff of pollutants during rainfall events. These concepts were incorporated into the original SWMM model (Metcalf and Eddy et al., 1971) as well as into the STORM, USGS and HSPF models to a greater or lesser degree (Huber, 1985). "Buildup" is a term that represents all of the complex spectrum of dry-weather processes that occur between storms, including deposition, wind erosion, street cleaning, etc. The idea is simply that all such processes lead to an accumulation of solids and perhaps other pollutants that are then "washed off" during storm events.

Although ostensibly physically based, models that include buildup and washoff mechanisms really employ conceptual algorithms because the true physics is related to principles of sediment transport and erosion that are poorly understood in this framework. Furthermore, the inherent heterogeneity of urban surfaces leads to use of average buildup and washoff parameters that may vary significantly from what may occur in an isolated street gutter, for example. Thus, except in rare instances of measurements of accumulations of surface solids, the use of buildup and

washoff formulations inevitably results in a calibration exercise against measured end-of-pipe quality data. It then holds that in the absence of such data, inaccurate predictions can be expected.

Different models offer different options for conceptual buildup and washoff mechanisms, with SWMM having the greatest flexibility. In fact, with calibration, good agreement can be produced between predicted and measured concentrations and loads with such models, including intra-storm variations that cannot be duplicated with the methods discussed earlier. (When a rating curve is used in SWMM instead of buildup and washoff, it is also possible to simulate intra-storm variations in concentration and load.) A survey of linear buildup rates for many pollutants by Manning et al. (1977) is probably the best source of generalized buildup data, and some information is available in the literature to aid in selection of washoff coefficients (Huber 1985; Huber and Dickinson, 1988). However, such first estimates may not even get the user in the ball park (i.e., quality – not quantity – predictions may be off by more than an order of magnitude); the only way to be sure is to use local monitoring data for calibration and verification. Thus, as for most of the other quality prediction options discussed herein, the buildup-washoff model may provide adequate comparisons of control measures, ranking of loads, etc. but cannot be used for prediction of absolute values of concentrations and loads, e.g., to drive a receiving water quality model, without adequate calibration and verification data. Since buildup and washoff are somewhat appealing conceptually, it is somewhat easier to simulate potential control measures such as street cleaning and surface infiltration using these mechanisms than with, say, a constant concentration or rating curve method. In the relatively unusual instance in which intra-storm variations in concentration and load must be simulated, as opposed to total storm event EMC or load, buildup and washoff also offer the most flexibility. This is sometimes important for the design of storage facilities in which first-flush mechanisms may be influential.

As mentioned above, generalized data for buildup and washoff are sparse (Manning et al., 1977) and such measurements are almost never conducted as part of a routine monitoring program. For buildup, normalized loadings, e.g., mass/day-area or mass/day per curb-length, or just mass/day, are required, along with an assumed functional form for buildup vs. time, e.g., linear, exponential, Michaelis-Menton, etc. For washoff, the relationship of washoff (mass/time) vs. runoff rate must be assumed, usually in the form of a power equation. When end-of-pipe concentration and load data are all that are available, all buildup and washoff coefficients end up being calibration parameters.

In the discussion above, washoff is assumed proportional to the runoff rate, as for sediment transport. Erosion from pervious areas may instead be proportional to the rainfall rate. HSPF does the best job of including this mechanism in its algorithms for erosion of sediment from pervious areas. SWMM includes a weaker algorithm based on the Universal Soil Loss Equation.

Many pollutants, particularly metals and organics, are adsorbed onto solid particles and are transported in particulate form. The ability of a model to include "potency factors" (HSPF) or "pollutant fractions" (SWMM) enhances the ability to estimate the concentration or load of one constituent as a fraction of that of another constituent, e.g., solids (Zison, 1980).

The groundwater contribution to flow in urban areas can be important in areas with unlined and open channel drainage. Of the urban models discussed, HSPF far and away has the most complex mechanisms for simulation of subsurface water quality processes in both the saturated and unsaturated zones. Although SWMM includes subsurface flow routing, the quality of subsurface water can only be approximated using a constant concentration.

The precipitation load may be input in some models (SWMM, HSPF), usually as a constant concentration. Point source and dry-weather flow (baseflow) loads and concentrations can also be input to SWMM, STORM and HSPF to simulate background conditions. Other quality sources of potential importance include catchbasins (SWMM) and snowmelt (SWMM, STORM, HSPF).

Scour and deposition within the sewer system can be very important in combined sewer systems and some separate storm sewer systems. The state of the art in simulation of such processes is poor (Huber, 1985). SWMM offers a crude but calibratable attempt at simulation of such processes.

11.5.2 Urban Runoff Models

Four models (DR3M-QUAL, HSPF, STORM, SWMM) are considered the best choices as full-scale simulation models for urban areas. Other models have been adapted from SWMM (e.g., FHWA, RUNQUAL) and STORM (e.g., SEMSTORM) and given modified names, but the principles are fairly similar. Still other models, such as the Illinois State Water Survey ILLUDAS model (Terstriep and Stall, 1974) have sometimes been adapted for water quality simulation for a specific project (Noel and Terstriep, 1982), but such modifications and quality procedures remain undocumented, and the quality model cannot be considered operational. At least two European models are available that simulate water quality. These are described briefly at the end of this section. Finally, there are many models well known in the hydrologic literature, such as those developed by the HEC and SCS, that might be useful in the hydrologic aspect of water quality studies, but that do not simulate water quality directly. This review is limited to models that directly simulate water quality.

A general comparison of model attributes is given in Table 11.1. This table includes the EPA Statistical Method since, with the publication of the recent FHWA study, it can be considered a formalized procedure (Driscoll et al., 1989). The regression approach described earlier is more generic in nature and not included in Table 11.1.

DR3M-QUAL. A version of the USGS Distributed Routing Rainfall Runoff Model that includes quality simulation (DR3M-QUAL) is available from that agency for general use (Alley and Smith, 1982). Runoff generation and subsequent routing use the kinematic wave method, and parameter estimation assistance is included in the model. Quality is simulated using buildup and washoff functions, with settling of solids in storage units dependent on a particle size distribution. The model has been used in some of the NURP studies that were conducted by the USGS (Alley, 1986). No microcomputer version is available.

TABLE 11.1
Comparison of Urban model attributes (from Donigian and Huber, 1991)

Attribute	DR3M-QUAL	HSPF	Model Statistical[a]	STORM	SWMM
Sponsoring agency	USGS	EPA	EPA	HEC	EPA
Simulation type[b]	C,SE	C,SE	N/A	C	C,SE
No. pollutants	4	10	Any	6	10
Rainfall/runoff analysis	Y	Y	N[c]	Y	Y
Sewer system flow routing	Y	Y	N/A	N	Y
Full, dynamic flow routing equations	N	N	N/A	N	Y[d]
Surcharge	Y[e]	N	N/A	N	Y[d]
Regulators, overflow structures, e.g., weirs, orifices, etc.	N	N	N/A	Y	Y
Special solids routines	Y	Y	N	N	Y
Storage analysis	Y	Y	Y[f]	Y	Y
Treatment analysis	Y	Y	Y[f]	Y	Y
Suitable for screening (S), design (D)	S,D	S,D	S	S	S,D
Available on microcomputer	N	Y	Y[g]	N	Y
Data and personnel requirements[h]	Medium	High	Medium	Low	High
Overall model complexity[i]	Medium	High	Medium	Medium	High

[a] EPA procedure.

[b] C = continuous simulation, SE = single event simulation.

[c] Runoff coefficient used to obtain runoff volumes.

[d] Full dynamic equations and surcharge calculations only in Extran Block of SWMM.

[e] Surcharge simulated by storing excess inflow at upstream end of pipe. Pressure flow not simulated.

[f] Storage and treatment analyzed analytically.

[g] FHWA study, Driscoll et al. (1989).

[h] General requirements for model installation, familiarization, data requirements, etc. To be interpreted only very generally.

[i] Reflection of general size and overall model capabilities. Note that complex models may still be used to simulate very simple systems with attendant minimal data requirements.

HSPF. The Hydrological Simulation Program - Fortran (HSPF) is the culmination of hydrologic routines that originated with the Stanford Watershed Model in 1966 and eventually incorporated many nonpoint source modeling efforts of the EPA Athens Laboratory (Johanson et al., 1984). This model has been widely used for non-urban nonpoint source modeling and is described additionally in that section of this chapter. The user's manual includes information on all hydrologic and water quality routines, including the IMPLND (impervious land) segment for use in urban area. Additional guidelines for application are provided by Donigian et al. (1984). The model has special provisions for management of time series that result from continuous simulation. A microcomputer version is available. Ongoing support, development, and refinement of the model has been jointly sponsored by the U.S. EPA and the U.S. Geological Survey for the most recent version (Bicknell et al. 1993). A bibliography of HSPF documents, related references, and applications was recently developed (Donigian, 1994).

STORM. The first significant use of continuous simulation in urban hydrology came with the Storage, Treatment, Overflow, Runoff Model (STORM), developed by the Corps of Engineers, Hydrologic Engineering Center (HEC, 1977; Roesner et al., 1974) for application to the San Francisco master plan for combined sewer outflow (CSO) pollution abatement. The HEC also provides application guidelines (Abbott, 1977). The current (1977) version includes dry-weather flow input for combined sewer simulation. The support of the HEC has led to the wide use of STORM for planning purposes, especially for evaluation of the trade-off between treatment and storage as control options for CSOs (e.g., Heaney et al., 1977). Statistics of long-term runoff and quality time series permit optimization of control measures.

STORM utilizes simple runoff coefficient, SCS and unit hydrograph methods for generation of hourly runoff depths from hourly rainfall inputs. No flow routing is performed, but runoff may be routed through a constant-rate treatment device, with excess flow diverted to a storage device. Flows exceeding the treatment rate cause CSOs when storage is filled. The buildup and washoff formulations are used for simulation of six pre-specified pollutants. However, the model can be manipulated to provide loads for arbitrary conservative pollutants (e.g., Najarian et al., 1986). The model is hampered somewhat by lack of an operational microcomputer version. However, various individual consultants have adapted the nonproprietary code to their own project needs.

SWMM. The original version of the Storm Water Management Model (SWMM) was developed for EPA as single-event model specifically for the analysis of CSOs (Metcalf and Eddy et al., 1971), but its scope has vastly broadened since the original release. Version 4 (Huber and Dickinson, 1988; Roesner et al., 1988) of the model performs both continuous and single-event simulation throughout the whole model, can simulate backwater, surcharging, pressure flow and looped connections (by solving the complete dynamic wave equations) in its Extran Block, and has a variety of options for quality simulation, including traditional buildup and washoff formulations as well as rating curves and regression techniques. Subsurface flow routing (constant quality) may be performed in the Runoff Block in addition to surface quantity and quality routing, and treatment devices may be sim-

ulated in the Storage/Treatment Block using removal functions and sedimentation theory. A hydraulic design routine is included for sizing of pipes, and a variety of regulator devices may be simulated, including orifices (fixed and variable), weirs, pumps, and storage. A bibliography of SWMM usage is available (Huber et al., 1986) that contains many references to case studies.

SWMM is segmented into the Runoff, Transport, Extran, Storage/Treatment and Statistics blocks for rainfall-runoff, routing, and statistical computations. Water quality may be simulated in all blocks except Extran, and metric units are optional. Since the model is non-proprietary, portions have been adapted for various specific purposes and locales by individual consultants and other federal agencies, e.g., FHWA. A microcomputer version is available.

Two European Models. The four U.S. models discussed above do not take advantage of graphics and other "user-friendly" capabilities of microcomputers. Two well-known and commercially-available European models, MOUSE and Wallingford, are excellent examples of application of the full power of the microcomputer when used in conjunction with recent programming languages and graphics hardware and software. Both models feature menu-driven pre-processors for data input, graphical display and interactive editing of catchment boundaries and sewer networks, and post-processing of predicted hydrographs and pollutographs, including graphical displays and statistical analysis. Although the quality algorithms are relatively simple, the hydrologic and hydraulic components of both models are relatively sophisticated.

The Danish Hydraulic Institute, in cooperation with various other laboratories and private software firms has produced the MOUSE (Modeling of Urban Sewers) model. Included in the package are modules for generation of runoff from rainfall, sewer routing (the S11S model, comparable to the SWMM Extran Block), and a simple quality routine that uses the constant concentration approach (Jacobsen et al., 1984).

The Wallingford model is maintained by Hydraulics Research Ltd. in the United Kingdom. It also consists of a cluster of modules, including runoff generation from rainfall (WASSP), simple and fully-dynamic sewer routing (WALLRUS and SPIDA, respectively), and a quality routine (MOSQITO)featuring processes similar to those in SWMM (Henderson and Moys, 1987).

11.5.3 Non-Urban Runoff Modeling Methods

As with the options for modeling nonpoint source pollutants from urban areas, a wide range of techniques are available for modeling these contributions from non-urban land uses, from simple annual "loading functions" to detailed process simulation models. The key issue in estimating nonpoint pollution loads from a watershed, or parcel of land, is the type and extent of human activities occurring (or not occurring) on the land. The same hydrologic, physical, and chemical/biological processes that determine nonpoint pollutant loads occur on all land surfaces (and in the soil profile) whether they are urban, forest, agricultural cropland, pasture, mining, etc. The relative importance and magnitude of these processes, in determining nonpoint loads, will vary among land use categories and associated human activities. Even within an urban region, the parameters required for the various

modeling options will differ for commercial, industrial, transportation, and various densities of residential land. Many of these same urban modeling options have been used for non-urban land areas with parameters (e.g., constant concentrations) estimated for the specific non-urban land use.

The focus of the majority of non-urban nonpoint source estimation procedures and models has been on agricultural cropland, although the procedures and models have often been adapted and applied to many other land use categories. The agricultural research community, comprised of the U.S. Department of Agriculture (including the Soil Conservation Service and Agricultural Research Service) regional laboratories and state universities, have developed a significant body of knowledge of soil processes and procedures for estimating runoff and soil erosion that have formed the 'building blocks' for loading functions and nonpoint source models. This section discusses some of the loading function procedures available for agricultural and other non-urban areas and general concepts underlying the more detailed process simulation models.

11.5.3.1 Nonpoint Source Loading Functions

The term "loading function" has been used in the nonpoint pollution literature to describe simple calculational procedures usually for estimating the average annual load, and sometimes the storm event load, of a pollutant from an individual land use category. A number of different loading functions have been developed and proposed over the past two decades, the most widely used of which are the EPA Screening Procedures, also referred to as the EPA Water Quality Assessment Methodology. These procedures are described below, followed by a brief discussion of a few other loading functions in the literature.

EPA Screening Procedures. The EPA Screening Procedures (Mills et al., 1982) are a revision and expansion of water quality assessment procedures initially developed for nondesignated 208 areas (Zison et al., 1977). The Procedures have been expanded and revised to include consideration of the accumulation, transport, and fate of toxic chemicals, in addition to conventional pollutants included in the earlier versions. The manual includes a separate chapter describing calculation procedures for estimating nonpoint loads for urban and non-urban land areas in addition to chapters on procedures for rivers and streams, impoundments, and estuaries. The most recent update includes consideration of toxics loadings and fate/transport in groundwater systems.

The procedures for nonpoint load assessments described in the manual are essentially a compilation and integration of techniques developed earlier by Midwest Research Institute (MRI), Amy et al. (1974), Heaney et al. (1976) and Haith (1980). However, the presentation of the procedures is well-integrated, supplemented with additional parameter estimation guidance, and includes sample calculations. The procedures for non-urban areas are derived from the MRI loading functions for average annual estimates based on the Universal Soil Loss Equation (USLE) (Wischmeier and Smith, 1978), while the storm event procedures use the Modified USLE (Williams, 1975) and the SCS Runoff Curve Number procedure (Mockus, 1972) for storm runoff volume. Pollutant concentrations in runoff and soil, and enrichment ratios are required for estimating pollutant loads; precipita-

tion contributions of nutrients can be included. Specific information is included for estimation of salinity loads in irrigation return flows. Separate equations are provided for estimating loads for sorbed pollutants, dissolved pollutants, and partitioned pollutants (i.e., both sorbed and dissolved phases); this latter category is primarily for pesticides for which the procedures were developed by Haith (1980) for storm event loads. Guidance is provided for estimating *all* required parameters.

The primary strengths and advantages of the EPA Screening Procedures are as follows:

- Excellent user documentation and guidance, including occasional workshops sponsored by the EPA Center for Exposure Assessment Modeling, in Athens, GA.
- No computer requirements since the procedures can be performed on hand calculators; associated programs have been developed for river quality analyses (Mills et al., 1979).
- Loading calculations and procedures can be linked to water quality procedures in other chapters to assess water quality impacts of nonpoint source loads.
- Relatively simple procedures with minimal data requirements that can be satisfied from the user manual when site-specific data are lacking.

These screening procedures are well suited for general screening-level assessments; however, they suffer from the same disadvantages as all such gross estimation techniques. As with the urban options discussed above, the accuracy of the loads depends on the accuracy of the user-assumed pollutant concentration; the impacts of management options is usually represented by a simple, constant 'removal fraction'; snowmelt and associated loadings are not represented; and calculations can be tedious and time consuming for complex multi-land use basins. In spite of these disadvantages and limitations, the EPA Screening Procedures are appropriate for many types of nonpoint load assessments. They have enjoyed wide popularity, partly due to the availability of training workshops sponsored by EPA, and have been applied in a number of regions, including the Sandusky River in northern Ohio and the Patuxent, Ware, Chester, and Occoquan basins in the Chesapeake Bay region (Davis et al., 1981; Dean et al., 1981a; Dean et al., 1981b). Although the procedures are quite amenable to a computerized or spreadsheet implementation, to our knowledge no effort has been made to implement such a format.

Other Loading Functions. As noted earlier, other loading functions have been developed and proposed by various groups and authors, though none have the support, nor have they demonstrated the longevity, of the EPA Screening Procedures. The WRENS handbook (Water Resources Evaluation of Nonpoint Silvicultural Sources, U.S. Forest Service (1980)) is similar to the EPA procedures but its focus is directed to the effects of forestry activities on water quality. The handbook provides quantitative techniques for estimating potential changes in streamflow, surface erosion, soil mass movement, total potential sediment discharge, and water temperature for comparative analyses of alternative silvicultural management

practices. Runoff and erosion estimation techniques are similar to those used in the EPA Screening Procedures with parameters modified for forestry conditions.

Haith and Tubbs (1981) developed watershed loading functions as a screening tool to evaluate agricultural nonpoint source pollution in large watersheds. These functions also use the SCS Curve Number procedure for runoff estimation and the USLE for erosion; then, based on user-defined pollutant concentrations in runoff and attached to sediment, the procedures allow calculation of loadings to receiving waters. These functions have been added to the most recent update of the EPA Screening Procedures manual (i.e., 1985).

More recently, Li et al. (1989) have proposed loading functions for estimating the average annual pesticide loads in surface runoff. They developed regression equations derived from 100-year simulations of daily pesticide runoff using the Haith (1980) pesticide model. The regression equation coefficients are based on pesticide half life and soil partition coefficients, and are tabulated in the article for a wide range of values. Two different regressions are described: one based simply on mean annual soil erosion, and the other based on mean annual soil erosion and surface runoff volume during the month of pesticide application.

The loading functions discussed above, and other similar techniques in the literature, differ from the simulation models primarily in time scale definition and their simplified, mostly empirical techniques for estimating nonpoint loads. They are used primarily to estimate average annual or event loads, and potential changes in these loads with land use and management practice. These procedures and associated calculations can usually be performed with a hand calculator, and with proliferation of personal computers and advances in computer technology we can expect to see more and more of these techniques available on PCs. However, users should not interpret the aura of implementation on a PC as an improvement in the capabilities, accuracy, or validity of these techniques. There are significant limitations in these procedures, because of their simplified nature, especially for evaluation of the impacts of management practices. As described above, they often require user-specified concentrations of pollutants in runoff and/or attached to sediment; some assume the total pollutant load can be estimated as a function of sediment alone.

Unfortunately, a comprehensive data base, comparable to the NURP data base for urban areas, does not exist for estimating the needed input concentrations for the wide range of non-urban land use categories. Also, there appears to be much greater variability in runoff concentrations from non-urban land than from urban land areas; consequently, extrapolation of concentrations from other sites may be less appropriate for non-urban land categories. Agricultural cropland is especially difficult to represent by single-valued 'representative' concentrations due to differences in crops, fertilizer applications, tillage practices, agronomic practices, soils characteristics, etc.

In spite of these limitations, loading functions can be useful for general screening assessments to identify relative nonpoint contributions under different conditions as long as their assumptions and limitations are recognized. They are more popular than the detailed simulation models and have thus been applied more frequently, primarily because of their ease of use. Also, the loading functions are a very useful precursor to more detailed modeling studies for general prob-

lem assessment and identification and to determine if such detailed studies are warranted.

11.5.3.2 Simulation Models

The primary differences between nonpoint runoff simulation models and the loading functions described above relate to the temporal and spatial detail of the analysis, along with (usually) a more refined representation of the processes that determine nonpoint pollutant loadings. Whereas the loading functions can be used with only a hand calculator (or spreadsheet), the added detail of most simulation models requires a computer code, computer facilities, and significantly more input data, such as daily rainfall and possibly other meteorologic timeseries. These models are most often computerized procedures that perform hydrologic (runoff), sediment erosion, and pollutant (chemical/ biological) calculations on short time intervals, usually ranging from one hour to one day, for many years. The resulting values for each time interval, e.g., runoff, sediment, pollutant load or concentration, can be analyzed statistically and/or aggregated to daily, monthly, or annual values for estimates of nonpoint loadings under the conditions simulated.

As with the urban models, a wide range of nonpoint models appropriate for non-urban areas are available and have been used for many different types of land categories. The available models also cover a large range of complexity depending on the extent to which hydrologic, sediment erosion, and chemical/ biological processes are modeled in a mechanistic manner or based on empirical procedures. Similar to urban modeling, many of the same simple procedures and assumptions used in the loading functions are also incorporated into a number of simulation models, e.g., USLE, SCS Curve Number, constant pollutant concentration. Section 11.5.4 provides brief summaries of a number of the more widely used and 'operational' non-urban models.

11.5.4 Non-Urban Runoff Models

In this section we provide brief summaries of the most useful non-urban runoff quality models that are currently available; included are HSPF, CREAMS/GLEAMS, ANSWERS, AGNPS, PRZM, SWRRB, and UTM-TOX; HSPF was also briefly described earlier as it also includes urban runoff modeling capabilities. Table 11.2 shows a comparison of selected model attributes and capabilities.

HSPF. The Hydrological Simulation Program - FORTRAN (HSPF) (Johanson et al., 1984; Bicknell et al. 1993) is a comprehensive package for simulation of watershed hydrology and water quality for both conventional and toxic organic pollutants. HSPF incorporates the watershed scale ARM and NPS models into a basin-scale analysis framework that includes fate and transport in one-dimensional stream channels. It is the only comprehensive model of watershed hydrology and water quality that allows the integrated simulation of land and soil contaminant runoff processes with instream hydraulic, water temperature, sediment transport, nutrient, and sediment-chemical interactions. The runoff quality capabilities include both simple relationships (i.e. empirical buildup/washoff, constant concen-

TABLE 11.2
Comparison of non-urban model attributes (from Donigian and Huber, 1991)

Attribute	Model						
	AGNPS	ANSWERS	CREAMS	HSPF	PRZM	SWRRB	UTM-TOX
Sponsoring Agency	USDA	Purdue Univ.	USDA	EPA	EPA	USDA	ORNL and EPA
Simulation type	C,SE	SE	C,SE	C,SE	C	C	C,SE
Rainfall/Runoff analysis	Y	Y	Y	Y	Y	Y	Y
Erosion Modeling	Y	Y	Y	Y	Y	Y	Y
Pesticides	Y	N	Y	Y	Y	Y	N
Nutrients	Y	Y	Y	Y	N	Y	N
User-Defined Constituents	N	N	N	Y	N	N	Y
Soil Processes							
Pesticides	N	N	Y	Y	Y	Y	N
Nutrients	N	N	Y	Y	N	Y	N
Multiple Land Type Capability	Y	Y	N	Y	N	Y	Y
Instream Water Quality	N	N	N	Y	N	N	Y
Simulation Available on Micro-computer	Y	Y	Y	Y	Y	Y	N
Data and Personnel Requirements	M	M/H	H	H	M	M	H
Overall Model Complexity	M	M	H	H	M	M/H	H

Y = yes, N = no, M = Moderate, H = High C = Continuous, SE = Storm Event

trations) and detailed soil process options (i.e., leaching, sorption, soil attenuation and soil nutrient transformations).

The result of this simulation is a time history of the runoff flow rate, sediment load, nutrient, pesticide, and/or user-specified pollutant concentrations, along with a time history of water quantity and quality at any point in a watershed. HSPF simulates three sediment types (sand, silt, and clay) in addition to a single organic chemical and transformation products of that chemical. The instream processes are included that are needed to simulate DO, BOD, nitrogen and phosphorus, pH, phytoplankton, zooplankton, and benthic algae.

The organic chemical transfer and reaction processes included are hydrolysis, oxidation, photolysis, biodegradation, a volatilization, and sorption. Sorption is modeled as a first-order kinetic process in which the user must specify a desorption rate and an equilibrium partition coefficient for each of the three solid types. Resuspension and settling of silts and clays (cohesive solids) are defined in terms of shear stress at the sediment-water interface. For sands, the capacity of the system to transport sand at a particular flow is calculated and resuspension or settling is defined by the difference between the sand in suspension and the capacity. Calibration of the model requires data for each of the three solids types. Benthic exchange is modeled as sorption/desorption and desorption/scour with surficial benthic sediments. Underlying sediment and pore water are not modeled.

As noted earlier, a microcomputer version is available, and ongoing support, development, and refinement of the model has been jointly sponsored by the U.S. EPA and the U.S. Geological Survey for the most recent version (Bicknell et al. 1993). A bibliography of HSPF documents, related references, and applications was recently developed (Donigian, 1994).

CREAMS/GLEAMS. Chemicals, Runoff, and Erosion from Agricultural Management Systems (CREAMS) was developed by the U.S. Department of Agriculture - Agricultural Research Service (Knisel, 1980) for the analysis of agricultural best management practices for pollution control. CREAMS is a field scale model that uses separate hydrology, erosion, and chemistry submodels connected together by pass files.

Runoff volume, peak flow, infiltration, evapotranspiration, soil water content, and percolation are computed on a daily basis. If detailed precipitation data are available then infiltration is calculated at histogram breakpoints. Daily erosion and sediment yield, including particle size distribution, are estimated at the edge of the field. Plant nutrients and pesticides are simulated and storm load and average concentrations of sediment-associated and dissolved chemicals are determined in the runoff, sediment, and percolation through the root zone.

User defined management activities can be simulated by CREAMS. These activities include aerial spraying (foliar or soil directed) or soil incorporation of pesticides, animal waste management, and agricultural best management practices (minimum tillage, terracing, etc.).

Calibration is not specifically required for CREAMS simulation, but is usually desirable. The model provides accurate representation of the various soil processes. Most of the CREAMS parameter values are physically measurable. The model has the capability of simulating 20 pesticides at one time.

Groundwater Loading Effects of Agricultural Management Systems (GLEAMS) was developed by the United States Department of Agriculture - Agriculture Research Service (Leonard et al., 1987) to utilize the management oriented, physically based CREAMS model (Knisel, 1980) and incorporate a component for vertical flux of pesticides. GLEAMS is the vadose zone component of the CREAMS model.

GLEAMS consists of three major components: hydrology, erosion/sediment yield, and pesticides. Precipitation is partitioned between surface runoff and infiltration and water balance computations are done on a daily basis. Surface runoff is estimated using the Soil Conservation Service Curve Number Method as modified by Williams and Nicks (1982). The soil is divided into various layers, with a minimum of 3 and a maximum of 12 layers of variable thickness used for water and pesticide routing (Knisel et al., 1989)

ANSWERS. Areal Nonpoint Source Watershed Environment Response Simulation (ANSWERS) was developed at the Agricultural Engineering Department of Purdue University (Beasley and Huggins, 1981). It is an event based, distributed parameter model capable of predicting the hydrologic and erosion response of agricultural watersheds.

Application of ANSWERS requires that the watershed be subdivided into a grid of square elements. Each element must be small enough so that all important parameter values within its boundaries are uniform. For a practical application element sizes range from one to four hectares. Within each element the model simulates the processes of interception, infiltration, surface storage, surface flow, subsurface drainage, and sediment drainage, and sediment detachment, transport, and deposition. The output from one element then becomes a source of input to an adjacent element.

As the model is based on a modular program structure it allows easier modification of existing program code and/or addition of user supplied algorithms. Model parameter values are allowed to vary between elements; thus, any degree of spatial variability within the watershed is easily represented.

Nutrients (nitrogen and phosphorus) are simulated using correlation relationships between chemical concentrations, sediment yield and runoff volume. A research version of the model uses "clay enrichment" information and a very descriptive phosphorus fate model to predict total, particulate, and soluble phosphorus yields.

AGNPS. Agricultural Nonpoint Source Pollution Model (AGNPS) was developed by the U.S. Department of Agriculture - Agriculture Research Service (Young et al., 1986) to obtain uniform and accurate estimates of runoff quality with primary emphasis on nutrients and sediments and to compare the effects of various pollution control practices that could be incorporated into the management of watersheds.

The AGNPS model simulates sediments and nutrients from agricultural watersheds for a single storm event; a continuous simulation version of the model has been under development for the past few years. Watersheds examined by AGNPS must be divided into square working areas called cells. Grouping of cells results in the formation of subwatersheds, which can be individually examined. The output

from the model can be used to compare the watershed examined against other watersheds to identify sources of water quality problems, and to investigate possible solutions to these problems.

AGNPS is also capable of handling point source inputs from feedlots, waste water treatment plant discharges, and stream bank and gully erosion (user specified). In the model, pollutants are routed from the top of the watershed to the outlet in a series of steps so that flow and water quality at any point in the watershed may be examined. The Modified Universal Soil Loss Equation is used for predicting soil erosion, and a unit hydrograph approach used for the flow in the watershed. Erosion is predicted in five different particle sizes: sand, silt, clay, small aggregates, and large aggregates.

The pollutant transport portion is subdivided into one part handling soluble pollutants and another part handling sediment attached pollutants. The methods used to predict nitrogen and phosphorus yields from the watershed and individual cells have been developed by and are also used in CREAMS (Knisel, 1980). The nitrogen and phosphorus calculations are performed using relationships between chemical concentration, sediment yield and runoff volume.

Data needed for the model can be classified into two categories: watershed data and cell data. Watershed data includes information applying to the entire watershed which would include watershed size, number of cells in the watershed, and if running for a single storm event then the storm intensity. The cell data includes information on the parameters based on the land practices in the cell.

Additional model components that are under development are unsaturated/saturated zone routines, economic analysis, and linkage to a geographic information system.

PRZM. Pesticide Root Zone Model (PRZM) was developed at the U.S. EPA Environmental Research Laboratory in Athens, Georgia by Carsel et al. (1984). It is a one-dimensional, dynamic, compartmental model that can be used to simulate chemical movement in unsaturated zone within, and immediately below, the plant root zone. The model is divided into two major components: hydrology (and hydraulics) and chemical transport. The hydrology component calculates runoff and erosion based on the Soil Conservation Service curve number procedure, and the Universal Soil Loss Equation is used to compute erosion. Evapotranspiration is estimated directly from pan evaporation or by an empirical formula if pan evaporation data is not available. Soil-water capacity terms including field capacity, wilting point, and saturation water content are used for simulating water movement within the unsaturated zone. Irrigation application is also within model capabilities.

Pesticide application on soil or on the plant foliage are considered in the chemical transport simulation. Dissolved, adsorbed, and vapor-phase concentrations in the soil are estimated by simultaneously considering the processes of pesticide uptake by plants, surface runoff, erosion, decay, volatilization, foliar washoff, advection, dispersion, and retardation. The user has two options to solve the transport equations using the original backward difference implicit scheme or the method of characteristics. As the model is dynamic it allows considerations of pulse loads.

Recently, a new version of PRZM called PRZM-2 has been released. PRZM-2 (Mullins et al., 1993) links an improved version of the original PRZM model with

an unsaturated zone model called VADOFT. VADOFT is a one-dimensional finite element model which solves Richard's equation for water flow in the unsaturated zone. VADOFT can also simulate the fate and transport of two parent and two daughter products. In order to perform exposure assessments, the PRZM-2 code is equipped with a Monte Carlo pre- and post-processor.

SWRRB. Simulator for Water Resources in Rural Basins (SWRRB) was developed by Williams et al. (1985), and Arnold et al. (1989, 1991) for evaluating basin-scale water quality. SWRRB operates on a daily time step and simulates weather, hydrology, crop growth, sedimentation, and nitrogen, phosphorous, and pesticide movement. The model was developed by modifying the CREAMS (Knisel, 1980) daily rainfall hydrology model for application to large, complex, rural basins.

Surface runoff is calculated using the Soil Conservation Service Curve Number technique. Sediment yield is computed for each basin by using the Modified Universal Soil Loss Equation (Williams and Berndt, 1977). The channel and floodplain sediment routing model is composed of two components operating simultaneously (deposition and degradation). Degradation is based on Bagnold's stream power concept and deposition is based on the fall velocity of the sediment particles (Arnold et al., 1989).

Return flow is calculated as a function of soil water content and return flow travel time. The percolation component uses a storage routing model combined with a crack flow model to predict the flow through the root zone. The crop growth model (Arnold et al., 1989) computes total biomass each day during the growing season as a function of solar radiation and leaf area index.

The pollutant transport portion is subdivided into one part handling soluble pollutants and another part handling sediment attached pollutants. The methods used to predict nitrogen and phosphorus yields from the rural basins are adopted from CREAMS (Knisel, 1980). The nitrogen and phosphorus calculations are performed using relationships between chemical concentration, sediment yield and runoff volume. The nutrient capabilities were the most recent capabilities to be incorporated and have been undergoing testing and validation (Arnold et al. 1991).

The pesticide component is directly taken from Holst and Kutney (1989) and is a modification of the CREAMS (Smith and Williams, 1980) pesticide model. The amount of pesticide reaching the ground or plants is based on a pesticide application efficiency factor. Empirical equations are used for calculating pesticide washoff which are based on threshold rainfall amount. Pesticide decay from the plants and the soil are predicted using exponential functions based on the decay constant for pesticide in the soil, and half life of pesticide on foliar residue.

The Pesticide Runoff Simulator (PRS) was developed for the U.S. EPA Office of Pesticide and Toxic Substances to simulate pesticide runoff and adsorption into the soil on small agricultural watersheds. PRS is based on SWRRB. Thus, the PRS hydrology and sediment simulation is based on the USDA CREAMS model, and the SCS curve number technique is used to predict surface runoff. Sediment yield is simulated using a modified version of the Universal Soil Loss Equation and a sediment routing model.

The pesticide component of PRS is a modified version of the CREAMS pesticide model. Pesticide application (foliar and soil applied) can be removed by atmospheric loss, wash off by rainfall, and leaching into the soil. Pesticide yield is divided into a soluble fraction and an adsorbed phase based on an enrichment ratio.

The model includes a built-in weather generator based on temperature, solar radiation, and precipitation statistics. Calibration is not specifically required, but is usually desirable.

UTM-TOX. Unified Transport Model for Toxic Materials (UTM-TOX) was developed by Oak Ridge National Laboratory for the U.S. EPA Office of Pesticides and Toxic Substances, Washington, D.C. UTM-TOX is a multimedia model that combines hydrologic, atmospheric, and sediment transport into one computer code.

The model calculates rates of flux of a chemical from release to the atmosphere, through deposition on a watershed, infiltration and runoff from the soil, to flow in a stream channel and associated sediment transport. From these calculations mass balances can be established, chemical budgets made, and concentrations in the environment estimated.

The atmospheric transport model (ATM) portion of UTM-TOX is a Gaussian plume model that calculates dispersion of pollutants emitted from point (stack), area, or line sources. ATM operates on a monthly time step, which is longer than the hydrologic portion of the model and results in the use of an average chemical deposition falling on the watershed.

The Terrestrial Ecology and Hydrology Model (TEHM) describes soil-plant water fluxes, interception, infiltration, and storm and groundwater flow.

The hydrologic portion of the model is from the Wisconsin Hydrologic Transport Model (WHTM), which is a modified version of the Stanford Watershed Model (SWM). WHTM includes all of the hydrologic processes of the SWM and also simulates soluble chemical movement, litter and vegetation interception of the chemical, erosion of sorbed chemical, chemical degradation in soil and litter, and sorption in top layers of the soil. Stream transport includes transfer between three sediment components (suspended, bed, and resident bed).

11.5.5 Guidance on Nonpoint Source Model Selection and Use

The models discussed briefly here do not represent all of the modeling options available for runoff quality simulation, but they are certainly the most notable, widely used and most operational. Selection from among these models is often made on the basis of personal preference and familiarity, in addition to needed model capabilities. For example, for urban modeling various in-house versions of STORM are still used by consultants even though the "official" HEC version has not been updated since 1977, because these versions have been adapted to the needs of a particular organization and because STORM has proven to provide useful continuous simulation results. The USGS DR3M-QUAL model has perhaps been used the least by persons outside that agency, but has worked satisfactorily in several applications. Support for both STORM and DR3M-QUAL is minimal. CREAMS has been used most extensively for field-scale agricultural runoff

modeling because of its agricultural origins and ties to the agricultural research community.

HSPF and SWMM are probably the most versatile and most widely applicable of the models, with the nod to SWMM if the urban hydrology and hydraulics must be simulated in detail. On the other hand, the water quality routines in HSPF for sediment erosion, pollutant interaction and groundwater quality are superior, and the capability to efficiently handle all types of land uses and pollutant sources, (including urban and agriculture, point and nonpoint), is a definite advantage when needed for large complex basins. Both models appear somewhat overwhelming in terms of size to the novice user, but only the components of interest of either model need be used in a given study, and the catchment schematization can often be coarse for purposes of simulation of water quality at the outlet. Thus, although the installation of these models on a microcomputer may occupy several megabytes of a hard disk, they may be applied in simple ways (i.e., applied to a simplified schematization of the catchment) with a significant reduction in data requirements. Furthermore, the several quality modeling options within SWMM permit simple conceptual water quality simulation using constant concentration and rating curves as well as the more formidable buildup-washoff methods. Similarly for HSPF, the ability to use the simple SWMM-type formulations for urban and non-agricultural areas, and detailed soil/runoff process simulation for agricultural areas provides the user with great flexibility in representing the watershed system.

Continuing model development and testing within the agricultural research community will likely lead to further enhancements and development of many of the agricultural models, like CREAMS, SWRRB, and AGNPS. In fact, USDA has supported, and continues to support, a wide range of model development work in individual research facilities, many of which are (or at least appear to be) very similar in terms of using similar algorithms or model formulations (e.g. EPIC (Williams et al, 1984), Opus (V. Ferreira, 1989, personal communication), SWAM (DeCoursey, 1982). The SWRRB development effort appears to be focussing in on a middle ground (in terms of complexity) between HSPF and the detailed field-scale models which are limited to small areas; its use of daily rainfall, as opposed to smaller time interval measurements (usually hourly is needed for HSPF) is seen as a definite advantage by many users. However, most of these efforts still focus primarily on agricultural areas, with limited abilities to be used in large, complex multi-land use basins.

Regression, spreadsheet, statistical methods, and loading functions are most useful as screening tools. Indeed if the Statistical Method or EPA's Screening Procedures, indicate that there should be no water quality problem (as defined by exceedance of a specified concentration level with a specified frequency), then more detailed water quality simulation may not be required at all. If sensitivity analyses and 'worst-case' evaluations further support the conclusions, detailed water quality modeling will probably not be needed.

11.6 Methods and Models for Surface Receiving Water Quality Modeling

Surface water bodies are valuable resources, providing habitat for fish and wildlife and supporting human populations with drinking water, commercial fisheries, recreational opportunities, wastewater receptors, and navigational routes. Surface water quality can be impacted by many human activities, including land use such as farming and urban development, discharge of municipal and industrial wastewater, and direct use such as dredging for navigation or impoundment for water supply or flood control. Environmental assessments are often used to manage human activities to preserve or enhance water quality. Surface water quality models are a central component of many environmental assessments, including waste load allocations and watershed TMDL's, risk assessments for pesticides and hazardous waste sites, and operation of reservoirs.

A primary division for surface water quality models can be made between "far-field" models, which simulate water quality in entire sections of a water body, and mixing zone models, which simulate near-field dilution processes. Since the focus of this chapter is on large-scale water quality modeling, only far-field models are identified and described.

Far-field water quality models can be subdivided on the basis of water body type, i.e., streams and rivers, lakes and reservoirs, and estuaries (including tidal rivers). Because some models can be applied to more than one water body type, they are summarized together in the following section. The primary reference for the following discussion of methods and models for surface receiving water quality is Ambrose et al. (1994).

11.6.1 Methods

It is common practice to use *simulation models* to assess the water quality of surface receiving water bodies such as rivers, lakes and reservoirs. The approaches and process algorithms used to model different types of water bodies reflect the dominant processes and phenomena for each particular water body type. For example, one-dimensional modeling of free-flowing rivers is typically appropriate, while multi-dimensional modeling of large lakes and reservoirs is usually a necessity.

Procedures that have been developed and integrated into quality simulation models for surface receiving water bodies utilize many of the same assumptions and methods described above for urban and non-urban nonpoint source pollution assessment. Selective assumption of constant concentrations, use of derived frequency distributions to implement statistical approaches to analysis, and use of built-in or user-supplied rating curves or regression data are all techniques that have found applications within receiving water models and modeling practices. As a rule, the methods used to model particular state variables and processes move away from the simpler techniques and move towards a more deterministic representation as the degree of regulatory interest, scientific understanding, and data availability increase.

To a large extent, receiving water quality simulation models differ from one another via their representation of four key attributes: dimensionality, time, hydrodynamics, and/or pollutant loadings. Dimensionality determines how a model

can spatially represent real water bodies. Most river and stream models are one-dimensional and can be applied to branching systems. Two dimensional models can cover either longitudinal and lateral (X/Y) or longitudinal and vertical (X/Z) dimensions. "Box" models can be applied in 1, 2, or 3 dimensions.

Time determines how a model can represent water body dynamics. Steady-state models predict concentrations that do not vary in time. These are useful primarily for rivers under low-flow design conditions. Both dynamic and quasidynamic solutions predict concentrations that vary with time. Quasidynamic solutions allow some major forcing functions, such as flow, loading, or solar effects on photosynthesis, to vary with time. Typically, one or more forcing functions will be held constant or allowed to vary slowly. Dynamic solutions allow all important forcing functions to vary with time. This would include daily changes in inflow and loading, or hourly changes in a tidal boundary.

All receiving water quality models require information on the movement of water. Some provide this information in the input dataset. These models are steady or quasidynamic in nature. Others require input information on flow or velocity; still others can either accept input flows, or be linked to simulated hydrodynamics. Some water quality models provide for the hydrodynamic calculations internally. Most, however, require linkage to an external hydrodynamics file generated by a separate hydrodynamics program. This may provide additional flexibility, allowing linkage with several different hydrodynamics programs or generated files. Often, however, the linkage to external programs is not as smooth or "transparent" as internal linkage. Some linked water quality/hydrodynamic models include processes to simulate control structures, such as weirs or low-head dams in rivers, or selective withdrawal from reservoirs. These can be more easily applied to highly regulated or engineered bodies of water.

All models mentioned in Section 11.6.2 simulate advection, and most include dispersion. Most 1-D riverine simulations do not require dispersion because most rivers are not highly dispersive and the model network and solution technique introduce some degree of "numerical dispersion". Accurate dispersion coefficients are usually required for simulations of lakes, reservoirs, and estuaries. Transport to and from benthic compartments is an important component of toxic chemical models. More recently, benthic exchange has been included as part of conventional water quality simulations.

All models allow the user to input steady pollutant loads. Some allow the specification of variable loads from the input dataset. A few models provide internal or external linkage to nonpoint source loading simulations. In Section 11.6.2 we provide further details on which models incorporate which of the methods introduced above.

11.6.1.1 Models

In this section we summarize eleven water quality models that are documented and available to the public. These models cover a range of capabilities from conventional water quality to toxicants and from steady state, one-dimensional to dynamic, multidimensional solutions.

QUAL2E. The Enhanced Stream Water Quality Model QUAL2E and QUAL2E-UNCAS (Brown and Barnwell, 1987) permits simulation of several conventional water quality constituents in a branching stream system under steady conditions. The model uses a finite difference solution to the one-dimensional advective-dispersive mass transport and reaction equation. QUAL2E predicts DO, CBOD, temperature, and phytoplankton dynamics as affected by nutrients and organic material. It comes with a pre- an post-processor and software that facilitates sensitivity and uncertainty analysis.

SMPTOX3. The Simplified Method Program - Variable Complexity Stream Toxics Model (SMPTOX3) is an interactive computer program for performing waste load allocations for toxics (LimnoTech, 1990). This steady, one dimensional model calculates suspended solids and dissolved and particulate toxicant concentrations in the water column and stream bed resulting from point source discharges into streams and rivers. The model is imbedded in a user-friendly shell that handles data input, graphical output, sensitivity analysis, and uncertainty analysis.

HSPF. The Hydrological Simulation Program - FORTRAN (HSPF) (Bicknell et al., 1993) is a comprehensive package for simulation of watershed hydrology and water quality for both conventional and toxic organic pollutants. As noted earlier, HSPF incorporates field scale models for nonpoint loading into a watershed and basin-scale analysis framework that includes fate and transport in one-dimensional stream channels. Several flow routing and water quality options are available. The conventional water quality routines predict DO, CBOD, temperature, and phytoplankton dynamics as affected by nutrients and organic material. The toxicant routines combine organic chemical process kinetics with sediment balance algorithms to predict dissolved and sorbed chemical concentrations in the upper sediment bed and overlying water column.

WASP5. The Water Quality Analysis Simulation Program, WASP5 (Ambrose, et al., 1987) is a generalized compartment modeling framework for simulating contaminant fate and transport in surface waters. Two major subcomponent models are provided with WASP5. The toxics subcomponent, TOXI5, combines organic chemical process kinetics with simple sediment balance algorithms to predict dissolved and sorbed chemical concentrations in the sediment bed and overlying water column. The dissolved oxygen/eutrophication subcomponent EUTRO5 predicts DO, CBOD, and phytoplankton dynamics as affected by nutrients and organic material. Transport can be driven by specified flows or by coupling to external hydrodynamic models.

EXAMSII. The Exposure Analysis Modeling System EXAMSII (Burns, 1990) is an interactive modeling system that allows a user to specify and store the properties of chemicals and ecosystems, modify either via simple commands, and conduct rapid evaluations and error analyses of the probable aquatic fate of synthetic organic chemicals. EXAMS combines chemical loadings, transport, and transformation into a set of differentiai equations using the law of conservation of mass as an accounting principle. The model is imbedded in an interactive, command-driven shell that handles data input and graphical or tabular output.

CE-QUAL-RIV1. The U.S. Army Waterways Experiment Station's hydrodynamic and water quality model for streams, CE-QUAL-RIV1, accommodates one-dimensional, branching systems with multiple hydraulic control structures.

The model simulates DO, BOD, and phytoplankton dynamics as affected by nutrients and organic material in highly unsteady systems.

CE-QUAL-W2. The U.S. Army Waterways Experiment Station's hydrodynamic and water quality model CE-QUAL-W2 provides dynamic two-dimensional (x-z) simulations for stratified waterbodies such as reservoirs and narrow estuaries. The hydrodynamic and water quality modules are directly coupled, so that predicted flows are influenced by variable water density calculated from temperature, salinity, and solids concentrations. The model can simulate DO and phytoplankton dynamics as affected by nutrients and organic material.

CE-QUAL-ICM. The U.S. Army Waterways Experiment Station originally developed CE-QUAL-ICM as an integrated compartment model for application to Chesapeake Bay. Transport can either be specified by the user or imported from coupled multidimensional hydrodynamic models. The model features advanced phytoplankton-nutrient balance routines and a submodel for predicting sediment oxygen demand. This model is under active development, and the capability to simulate toxicants is planned.

HEC5Q. The U.S. Army Hydrologic Engineering Center developed HEC-5Q to simulate the operation of river-reservoir systems in basin wide studies. The stream component is a one-dimensional unsteady model driven by a selection of available hydrologic routing models. The reservoir component is a one-dimensional vertical model. The model simulates DO, BOD, a selection of conservative and non-conservative constituents, and contains a phytoplankton option.

SALMON-Q. HR Wallingford assembled SALMON-Q from a suite of programs to model unsteady flow, siltation, and water quality in one-dimensional channel networks, including rivers and estuaries. SALMON-Q incorporates the modules in a user-friendly graphics interface. Water quality parameters include DO, BOD, nutrients, and phytoplankton.

MIKE-11. The Danish Hydraulic Institute developed MIKE-11 for the simulation of unsteady flow, sediment transport, and water quality in one-dimensional channel networks, including rivers and estuaries. MIKE-11 is operated through an interactive menu system that supports modules for hydrology, hydrodynamics, cohesive and non-cohesive sediment transport, and water quality. Water quality parameters include DO, BOD, nutrients, phytoplankton, zooplankton, bottom vegetation, and metals.

Tables 11.3 and 11.4 summarize important model attributes. The first major category in Table 11.3 is the water body type. Some models specialize in rivers, including QUAL2E, SMPTOX3, HSPF, CE-QUAL-RIV1, SALMON-Q, and MIKE-11. Other models are more general, though perhaps less efficient for particular types of water bodies. EXAMS, WASP5, and CE-QUAL-ICM are "box" or "compartment" models that can be applied to all types of water bodies. CE-QUAL-W2 was developed as a reservoir model, but has been applied to deep, slow-moving rivers as well as estuaries. HEC5Q is used for river basin evaluation, including rivers and reservoirs.

The next major category, dimensionality, determines how a model can spatially represent real water bodies. Most river and stream models are one-dimensional. All that are listed here can be applied to branching systems. Two dimensional models

TABLE 11.3
Surface water quality models; a. basic information (from Ambrose et al., 1994)

Model		QUAL2E	SMPTOX3	HSPF	WASP5	EXAMS	CEQUALRIV1	CEQUALW2	CEQUALICM	HEC5Q	SALMONQ	MINTEQA1
Water Body Type	Stream, River	X	X	X	X	X	X	X	X	X	X	X
	Lake, Reservoir				X	X		X	X	X		
	Estuary	X		X		X	X	X				
Dimension	1-D, Branching	X	X	X	X	X	X	X	X	X	X	X
	2-D, X/Y				X	X			X			
	2-D, X/Z				X	X		X	X			
	3-D, Box				X	X			X			
Time	Steady	X	X			X						
	Quasidynamic	X		X	X	X					X	X
	Dynamic			X	X		X	X	X	X	X	X
Hydro-dynamics	Input	X	X	X	X	X						X
	Simulated			X	X		X	X	X	X	X	X
	Control Struct.			X			X	X	X	X	X	X
Transport	Advection	X	X	X	X	X	X	X	X	X	X	X
	Dispersion	X	X		X	X	X	X	X	X	X	X
	Benthic Exchange		X	X	X	X		X	X	X		
Loading	Input, steady	X	X	X	X	X	X	X	X	X	X	X
	Input, variable			X	X	X	X	X	X	X	X	X
	Simulated			X	X						X	X
Other	Preprocessor	X	X		X	X					X	X
	Postprocessor	X	X	X	X	X					X	X

can cover either longitudinal and lateral (X/Y) or longitudinal and vertical (X/Z) dimensions. CE-QUAL-W2 is a true X/Z model. The three "box" models listed here can be applied in 1, 2, or 3 dimensions. The user should be aware, however, that the underlying equations are not inherently multidimensional in form.

TABLE 11.4
Surface water quality models; b. Variables and Processes (from Ambrose et al., 1994)

	Model	QUAL2E	SMPTOX3	HSPF	WASP5	EXAMS	CEQUALRIV1	CEQUALW2	CEQUALICM	HEC5Q	SALMONQ	MIKE11
Chemical Processes	First-Order Decay	X	X	X	X	X	X	X	X	X	X	X
	Process Kinetics			X	X	X						
	Daughter Products			X	X	X						
	Sorption		X	X	X	X					X	
Sediment Processes	Input Rates		X	X	X		X		X	X	X	X
	Noncohesive Processes			X							X	?
	Cohesive Processes			X							X	?
Water Quality Processes	Temperature	X		X			X	X	X	X	X	X
	Salinity	X		X			X	X	X	X	X	X
	Bacteria			X			X	X	X	X	X	X
	DO-BOD	X		X	X		X	X	X	X		
	DO-Carbon Balance							X	X			
	Nitrogen Cycle	X		X	X		X	X	X	X	X	X
	Phosphorus Cycle	X		X	X		X	X	X	X	X	X
	Silicon Cycle							X	X		X	X
	Phytoplankton	X		X	X		X	X	3	X	X	X
	Zooplankton			X					X		X	
	Benthic Algae			X							X	X
	Simulate SOD								X			

The third major category, time, determines how a model can represent water body dynamics. Steady-state models, such as QUAL2E and SMPTOX3, predict concentrations that do not vary in time. These are useful primarily for rivers under low-flow design conditions. Both dynamic and quasidynamic solutions predict concentrations that vary with time. Quasidynamic solutions allow some major forcing functions, such as flow, loading, or solar effects on photosynthesis, to vary with time. Typically, one or more forcing function will be held constant or allowed to vary slowly. Dynamic solutions allow all important forcing functions to

vary with time. This would include daily changes in inflow and loading, or hourly changes in a tidal boundary.

The fourth major category is hydrodynamics. All water quality models require information on the movement of water. Some provide this information in the input dataset. These models are steady or quasidynamic in nature. QUAL2E, SMPTOX3, and EXAMS require input information on flow or velocity; HSPF, WASP5, and MIKE-11 can either accept input flows, or be linked to simulated hydrodynamics. Some water quality models, such as HSPF and CE-QUAL-W2 provide for the hydrodynamic calculations internally. Most, however, require linkage to an external hydrodynamics file generated by a separate hydrodynamics program. This may provide additional flexibility, allowing linkage with several different hydrodynamics programs or generated files. Often, however, the linkage to external programs is not as smooth or "transparent" as internal linkage. Some linked water quality/hydrodynamic models include processes to simulate control structures, such as weirs or low-head dams in rivers, or selective withdrawal from reservoirs. These can be more easily applied to highly regulated or engineered bodies of water.

The fifth major category deals with transport. All models mentioned in this section simulate advection, and all but HSPF include dispersion. Most 1-D riverine simulations do not require dispersion because most rivers are not highly dispersive and the model network and solution technique introduce some degree of "numerical dispersion". Accurate dispersion coefficients are usually required for simulations of lakes, reservoirs, and estuaries. Transport to and from benthic compartments is an important component of toxic chemical models, such as SMPTOX3, HSPF, WASP5, and EXAMS. More recently, benthic exchange has been included as part of conventional water quality simulations, as with CE-QUAL-ICM and, at the user's discretion, with WASP5.

The sixth major category is pollutant loading. All models allow the user to input steady pollutant loads. Some allow the specification of variable loads from the input dataset. Finally, a few models provide internal or external linkage to nonpoint source loading simulations. HSPF features internal linkage with pervious and impervious land modules. WASP5, on the other hand, provides for external linkage to many loading models through a formatted nonpoint source loading file. MIKE-11 and SALMON-Q come with linkage to urban wastewater modules.

The last category in Table 11.3 summarizes software features and availability. Many of the models come with preprocessors to aid the user in specifying input data, and graphical postprocessors to aid in the interpretation of simulations. The scope and style of support software vary from tightly integrated packages to loosely connected modules that may link to commercial software.

Table 11.4 summarizes variables and water quality processes employed by surface water quality models. The first major category, chemical processes, determines whether and how a model can be used to simulate toxicants. Some models deal explicitly with toxicants in a simple manner, including SMPTOX3, HSPF, WASP5, and MIKE-11. Other models have the capability to simulate toxicants in a more complex fashion. These include HSPF, WASP5, and EXAMS. Similar algorithms are being added to CE-QUAL-ICM.

Almost all models include a state variable that degrades at a user-specified first order rate. Some models that specialize in the simulation of organic chemicals will include process kinetics. These algorithms predict chemical degradation rates from chemical properties in combination with various environmental forcing functions such as pH, light, and temperature. Most of the chemical process models can track the daughter products of the transformation reactions.

In addition to degradation, toxicant models should contain algorithms to handle sorption to suspended and benthic solids. The simplest technique is equilibrium linear sorption, which is characterized by a simple partition coefficient. This coefficient may be specified directly or calculated internally for organic chemicals from such chemical and environmental properties as octanol-water partition coefficient and sediment organic carbon content. At high concentrations, the sorption process becomes nonlinear with respect to the chemical concentrations. None of the models listed here explicitly contains nonlinear sorption algorithms and sorption kinetics, although the latter can be handled indirectly in WASP5 or EXAMS using transformation reactions along with partition coefficients.

The second major category in Table 11.4, sediment processes, is most important for models that simulate toxicants. Most of the models that simulate sediment or solids as a state variable employ a simple mass balance approach with user input deposition, resuspension, and burial rates. Some models, such as HSPF and MIKE-11 contain algorithms to predict cohesive and/or noncohesive sediment deposition and resuspension based upon flow and sediment characteristics. There are many models that specialize in solids transport, but do not include water quality variables or processes. These sediment transport models are not listed here.

The final major category covers water quality processes. Most of the conventional pollutant water quality models listed here simulate temperature using simple heat balance techniques. In WASP5 and EXAMS, temperature is provided as user input and is not simulated. Salinity is another variable that is explicitly included in most models. It is used both as a tracer to calibrate transport, and as a component of the DO saturation calculations. Although salinity is not explicitly included in WASP5 or EXAMS, it can be simulated using a chemical state variable with no decay or sorption. Bacteria, such as total or fecal coliform, is included in many models. Those that do not simulate bacteria explicitly can nevertheless use a chemical state variable with first-order decay and no sorption to simulate bacterial dynamics. The user must then be careful to properly specify bacterial loadings and boundary conditions given the different analytical units reported for bacteria (i.e., counts/100 ml versus mg/l).

Most conventional pollutant models handle dissolved oxygen (DO) linked to carbonaceous biochemical oxygen demand (CBOD), as well as to nitrification, phytoplankton growth and respiration, and sediment oxygen demand (SOD). CE-QUAL-ICM calculates the DO budget using organic carbon rather than CBOD. The nitrogen and phosphorus cycles are treated explicitly by those models addressing eutrophication. Most models treat the nutrient cycles in a relatively simple manner. The phosphorus cycle, for example, contains inorganic phosphorus (or orthophosphate), phytoplankton phosphorus, and nonliving organic phosphorus. The nitrogen cycle contains ammonium nitrogen, nitrite nitrogen, nitrate nitrogen, phytoplankton nitrogen, and nonliving organic nitrogen. Some models, such as CE-

QUAL-ICM employ more complex nutrient cycles that simulate organic nitrogen and phosphorus in various forms, such as particulate and dissolved refractory, and particulate and dissolved non-refractory. Some also treat the silicon cycle in order to handle the growth of diatoms.

Most eutrophication models treat phytoplankton as a single variable. CE-QUAL-ICM handles three phytoplankton classes separately. Zooplankton are included as a variable in three models: CE-QUAL-ICM, HSPF, and MIKE-11. Benthic algal routines have been added to some models recently in experimental versions; HSPF, MIKE-11 and SALMON-Q include benthic algae in their standard release. Sediment oxygen demand (SOD), along with benthic nutrient fluxes, is almost always specified by the user at present. Although WASP5 has primitive capabilities to simulate benthic nutrients and SOD, only CE-QUAL-ICM offers this capability in a fully-developed form.

11.6.2 Guidance on Receiving Water Model Selection and Use

Simple analytical water quality models have been employed since the 1920's. More complex, computerized water quality models have now had more than 25 years of use. Many of the models listed here have evolved over a long period, and have stood the test of time. An excellent example is QUAL2E, which has a worldwide standard for conventional water quality analysis in streams. Other models are of recent origin and employ promising new scientific algorithms. A good example is the algorithms contained in CE-QUAL-ICM for conventional water quality interactions in water column and benthic sediment.

Toxicant pollutant models have been operational over the past decade. Some models, like EXAMS, have emphasized organic chemical reactions, and have been used primarily for pesticide analysis in simple aquatic systems. Models such as HSPF and WASP5 have imbedded EXAMS chemical kinetics in transport frameworks that simulate flow, sediment transport, and chemical dynamics in more complex aquatic systems. HSPF handles watersheds with internal coupling of land surface and river reach modules. WASP5 features external linkage to hydrodynamic and runoff models. SPMTOX3 has simplified the chemical kinetics and transport algorithms to give a more easily-used but less predictive modeling system for toxicants.

The series of models produced by the U.S. Army Corps of Engineers have emphasized dynamic flow and transport in regulated rivers and reservoirs. CE-QUAL-RIV1, CE-QUAL-W2, and HEC5Q contain a solid description of conventional water quality interactions. CE-QUAL-ICM, designed to link with external multidimensional hydrodynamic models, contains the most complex water quality dynamics.

Several emerging trends are apparent in surface water quality modeling. One key area is facilitating realistic simulations by linking models within graphical shells and decision support systems. These shells can not only link models with databases, but also with other models through standardized file transfers. In particular, we are now seeing more routine linkage of water quality models with 3-dimensional hydrodynamics programs. Linked watershed-waterbody modeling systems are also being produced. An important trend in conventional pollutant

modeling is the replacement of the benthic boundary with direct simulation of sediment oxygen demand and benthic nutrient fluxes.

11.7 The Future of Watershed Water Quality Modeling

Neither environmental water quality concerns nor the models that are used to evaluate them are static. Enhancement activities for popular models such as SWMM and HSPF have continued for over a decade, and the models will likely serve as a framework for near-term model development and application activities. However, the ultimate need in watershed water quality analysis is for comprehensive analysis tools that enable evaluation of the full impact of multiple chemical releases from both point and nonpoint sources to the environment, combined with habitat alteration/loss, to develop management strategies protective of human and ecological health. Satisfying this need necessitates the integration of flexible, interactive simulation models for air, surface water, estuarine and groundwater quality, and habitat and aggregate impacts with supportive databases and computer graphics for the visualization of information, problems and solutions. Projection of the directions that watershed modeling will venture, and the rate at which their evolution will progress, requires consideration of the following factors:

- *Objectives.* Future development efforts will recognize the need for improved man-data, man-model and data-model interaction capabilities; data storage and management capabilities integrated within the system; high quality visualization tools; and process algorithms capable of representing a range problems characterized by various levels of complexity and considering a broad range of interrelated state variables.

- *Institutional support.* Institutional support for development efforts is currently in a state of flux. Support within the certain governmental agencies, such as the U.S. Army Corps of Engineers (Davis, 1993) appears to be increasing, while future support in others, such as the EPA, is uncertain.

- *Development approach.* Two different approaches to meeting the needs for watershed analysis tools are likely to progress side-by-side: one places the focus on *short-term* needs, the other on long-term needs. The short-term approach will focus on rapid development of system capabilities that are much improved over those that are currently available. It will feature the use of "state-of-the-art", rather than "cutting edge", technologies. The *long-term* approach will require labor-intensive efforts and investigation and the use of technologies the are "cutting edge", at least as far as engineering applications are concerned. Both approaches will advance the state-of-the-art of watershed modeling, and will provide improved tools.

- *Development platform.* Development efforts will most likely be focused on a workstation environment, with efforts being made to allow portability to high-end PCs.

- *Modeling system architecture.* Future development efforts will feature "open systems" that accommodate the integration and use of multiple models, databases, tools, and development languages.

- *Interaction.* Institutional support is widespread for the use of graphical user interfaces (GUIs). In future modeling systems "direct manipulation" capabilities that allow model users to exert much more direct control over modeling activities will proliferate.

- *Process models.* Future development efforts will recognize the need to support multiple users and multiple applications of their systems to problems needing different levels of analysis detail and complexity. It is likely that current process algorithms will need to be enhanced, restructured and re-coded to enable a more flexible framework for exploring and solving engineering problems.

- *Databases and database support.* Modeling systems will recognize the ultimate need to accommodate and interact with multiple types of data including time-varying, space-varying and tabular data. Initial system development of the various efforts will feature different strengths and weaknesses in terms of accommodating these data types. Systems will be based on providing a persistent repository for a common, shared representation for domain-specific modeling information.

- *Graphics.* Future watershed models will feature the use of advanced visualization tools. Loosely coupled, industry-standard graphics packages that can be readily replaced as the need arises will likely provide the basis for implementing model graphics capabilities.

Perhaps the two areas that will experience the greatest growth as watershed models evolve over the upcoming years are (1) algorithm/computation capabilities and (2) user interaction capabilities. Consequently, these two categories are described in greater detail below. The primary reference for this discussion is Imhoff and Donigian (1993).

11.7.1 Algorithm/Computation Enhancements

While the current generation of models provide some capabilities to perform the general functions listed below, there is room for significant improvement. Possible enhancements to process algorithms and computational techniques that are potentially the most beneficial to improved watershed modeling include the following:

- *Capability to define physical setting of modeling effort.* Developing and verifying model input that defines the physical setting of the modeling effort and specifies spatially-referenced materials properties can be confusing and time-consuming. In the future, a graphical approach will likely be used to perform such tasks as inserting, deleting and moving rows and columns to define cells or segments; zooming, panning and window capabilities for detailed

design; overlaying a digitized map to aid in visualization or assign parameter values; establishing connectivity between cells or segments; assigning parameter values (global, spatially variable, time dependent) depicting natural and anthropogenic conditions; and assigning boundary and initial conditions.

- *Representing important environmental state variables and processes.* In order to provide the basis for multi-stressor analysis of whole-ecosystem effects, many chemical and biological state variables and processes must be represented. While the majority of these state variables are already considered in existing models, future models might be enhanced to include additional state variables including additional biological variables (herbivorous fish, predatory fish), selected habitat variables (% pools and riffles, streambank vegetation and shading, substrate character, turbidity), and selected ecosystem variables (elemental dynamics, energy dynamics, trophic dynamics, biodiversity, critical species (presence/abundance), genetic diversity, dispersal and migration, natural disturbance, ecosystem development).

- *Representing man-made effects on environmental state variables and processes.* In addition to representing natural processes, watershed modeling systems must provide process algorithms that represent the effects of man-induced sources or processes on environmental state variables. Models must include algorithms that can be used to represent any environmental disturbance that could influence the behavior of the natural watershed system. Examples of such phenomena include nutrient and pesticide application, tillage practices, crop harvest and residue practices, tile drainage, livestock grazing, feedlot runoff, highway drainage, urban development, stormwater detention structures, stream channelization, combined sewers, construction practices, mine drainage, silvicultural practices, municipal and industrial discharges, etc. Many of these effects can be represented by adjusting values for parameters contained in existing algorithms; others may require development of new algorithms. We envision that considerable work will be done to develop additional sets of parameter value changes (i.e., model scenarios) that reflect our best understanding of the physical/chemical changes resulting from a particular modification or activity. This may be the most critical area of model development activity as it directly affects our ability to use models for environmental management and decision-making.

- *Using process algorithms that utilize available data.* The most popular of the current models were developed prior to the proliferation of a new generation of data and data generation techniques that offer refined spatial detail for a number of parameters critical to watershed modeling. In some cases these new data are best used to support existing process algorithms that are solved for a higher resolution grid. However, the potential also exists to replace or enhance certain process algorithms to improve the simulation of natural processes by taking advantage of new data. For example, satellite data, GIS and digital elevation models (DEMs) have made it possible to compute the *aspect* (i.e., the direction toward which a slope faces) for watersheds or watershed segments at a high level of detail. The availability of techniques

to reliably compute aspect invites the incorporation of improved process algorithms for snowmelt, soil temperature, and water temperature in areas of significant topographical relief.

Interfacing with other environmental media. As the demand for truly "comprehensive" watershed analysis increases, so, too, will the need for interfacing nonpoint source and receiving water quality models with models that represent other traditional media such as air and groundwater. At the same time the need will be created to interface with additional media for which the first generation of models is currently being developed; examples of such newly modeled media are wetlands, and ecosystems dynamics.

11.7.2 User Interaction

User interaction, or more accurately the lack thereof, is a major weakness of most currently available watershed models and a significant stumbling block to many new model users. At the time development efforts for these models began, interactive water resources and environmental models were non-existent, and many of the popular menu-driven, business and word processing programs (e.g. Lotus, WordPerfect) did not emerge until the mid-1980's. Today, sophisticated users have developed a variety of software tools to automate specific functions needed to facilitate model application, data manipulation, and results analysis. Possible tools that could improve user interaction within watershed models can be categorized into five categories of functionality:

- Data exploration/visualization;
- Model support before processing;
- Model support during processing;
- Model support after processing;
- Decision support.

Potential enhancements for each of the categories are noted below.

11.7.2.1 Data Exploration/Visualization

- Basic display of data spatial and/or temporal variation (e.g., superimpose model results on location maps; time series plot of user-defined model variables; 2-dimensional and 3-dimensional plots, contour plots).
- Graphical comparison of different data types at common locations or times to make preliminary assessments of cause and effect.

The primary source of the list of capabilities to support modeling support activities (i.e., pre-processing, processing, post-processing) is the GEO-WAMS effort (DePinto, 1993).

11.7.2.2 Model Support Before Processing

- Site specification and problem definition. Provide an interactive graphical site schematic configuration and manipulation system with sophisticated capabilities (e.g., insert, delete and move rows and columns to define cells or segments; global and zonal modification; zoom, pan and window capabilities for detailed design; overlay of multiple information types).
- Graphical assignment of parameter values (global, spatially variable, time dependent)
- Hooks to retrieve raw data. Integrate custom programs to automate retrieval of data from commonly-used external databases (archival or commercial).
- Data transformation tools. Provide automated techniques for aggregating or disaggregating data. The modeling system should automatically transform data to meet process model requirements, and hence should be transparent to the user.
- Data derivation tools. When commonly accepted relationships of either an empirical or physical nature are available to estimate data needed by the component process models, the modeling system should include tools that enable automated value estimation based on available data.
- Computation of segment-dependent data volumes, interfacial areas.
- Graphical prescription of initial conditions and boundary conditions.

11.7.2.3 Model Support During Processing

- Spatial and temporal animation of results;
- Model interrupt/restart capability.

11.7.2.4 Model Support After Processing

- Basic display of primary variable spatial and/or temporal variation (e.g., superimpose model results on location maps; 3-dimensional plots, contour plots);
- Sensitivity analysis;
- Segment/system mass balance diagrams;
- Uncertainty analysis;
- Data export for use in other modeling systems. The modeling system needs a full library of tools designed to expedite data export from the application data base.

11.7.2.5 Decision Support

- Improve on-line model capabilities to assist users developing model input: estimating parameter values, selecting appropriate distributions for parameters used in uncertainty analysis, developing physically correct and mathematically compatible configurations and discretization (spatial and temporal) for modeled areas.
- Improve on-line model capabilities to assist users in interpreting model output: enhanced interpretive graphics, interpretation of time series data statistical analysis results as a potential calibration aid.
- Limited expert systems for specific processes, state variables.
- Additional scenario generators (urban and non-urban).

11.8 Closure

In this chapter, we have attempted to summarize the historical development of watershed-scale water quality modeling; explain the role of modeling in environmental analysis; introduce modeling fundamentals and relevant model features; describe the best of the current generation of modeling tools; offer guidance in model selection; and explore potential areas for future model development and research.

Clearly, the need for improved watershed analysis models will continue to exist and evolve over the foreseeable future. The future of watershed model development and application will be driven by regulatory and management needs, the state of scientific understanding and the availability of data and computational/programming resources. Creating the modeling systems that will be needed in the 21st century will require the following activities:

- Improving and integrating appropriate advances in natural process and control strategy representation;
- Extending watershed-scale process representation to include habitat and ecosystem parameters;
- Redesigning tools to make use of a continually expanding and evolving body of environmental data;
- Enabling multiple models to interact directly with each other and with a common database;
- Making tools more flexible by recoding process algorithms using new languages and technologies;
- Making tools more universally usable through on-line support and improved methods of user interaction with both data and evaluation techniques;

- Building a consistent user interface for dealing with all operations carried out in modeling systems; and
- Developing modeling systems that adhere to strict software quality assurance procedures.

By implementing these improvements, scientifically defensible evaluation techniques and tools can be adopted and used within the environmental community as a sound and common means for analyzing watershed problems.

References

Abbott, J. 1977. Guidelines for Calibration and Application of STORM. Training Document No. 8, Hydrologic Engineering Center, Corps of Engineers, Davis, CA.

Alley, W.M. and P.E. Smith. 1982. Distributed Routing Rainfall-Runoff Model Version II. USGS Open File Report 82-344. Gulf Coast Hydroscience Center, NSTL Station, MS.

Alley, W.M. 1986. Summary of Experience with the Distributed Routing Rainfall-Runoff Model (DR3M). In: Urban Drainage Modeling , C. Maksimovic and M. Radojkovic, (eds). Proc. of International Symposium on Comparison of Urban Drainage Models with Real Catchment Data, Dubrovnik, Yugoslavia. Pergamon Press, New York, NY. pp 403-415.

Ambrose, R.B. Jr., T.A. Wool, J.P. Connolly, and R.W. Schanz. 1987. WASP4, A Hydrodynamic and Water Quality Model. EPA/600/3-87/039. U.S. EPA, Athens, GA.

Ambrose, R.B., Jr., T.O. Barnwell, Jr., S.C. McCutcheon and J.R. Williams. 1994. "Computer Models for Water Quality Analysis". Chapter 14 In: Handbook of Water Resources. McGraw-Hill, New York, NY.

American Public Works Association. 1969. Water Pollution Aspects of Urban Runoff. Report 11030DNS01/69 (NTIS PB-215532), Federal Water Pollution Control Administration, Washington, DC.

Amy, G., R. Pitt, R. Singh, W.L. Bradford and M.B. LaGraff. 1974. Water Quality Management Planning for Urban Runoff. EPA 440/9-75-004 (NTIS PB 241 689/AS), U.S. Environmental Protection Agency, Washington, DC.

Arnold, J.G., J.R. Williams, A.D. Nicks and N.B. Sammons. 1989. SWRRB, A Basin Scale Simulation Model for Soil and Water Resources Management. Texas A&M Press.

Arnold, J.G., J.R. Williams, R.H. Griggs, and N.B. Sammons. 1991. SWRRBWQ, A Basin Scale Model for Assessing Management Impacts on Water Quality. USDA-ARS, West Lafayette, IN.

Beasley, D.B. and L.F. Huggins. 1981. ANSWERS Users Manual. EPA-905/9-82-001. U.S. EPA, Region V. Chicago, IL.

Bensen, M.A. 1962. Spurious Correlation in Hydraulics and Hydrology. J. Hydraulics Div., Proc. ASCE, 91(HY4):57-71.

Bicknell, B.R., J.C. Imhoff, J.L. Kittle, A.S. Donigian Jr., and R.C. Johanson. 1993. Hydrological Simulation Program FORTRAN (HSPF): User's Manual for Release 10. EPA-600/R-93/174. U. S. Environmental Protection Agency, Athens, GA.

Brown, R.G. 1984. Relationship between Quantity and Quality of Storm Runoff and Various Watershed Characteristics in Minnesota, USA. Proc. Third International Conference on Urban Storm Drainage, Chalmers University, Goteborg, Sweden, Vol.3, pp. 791-800.

Brown, L. C., and T. O. Barnwell. 1987. The Enhanced Stream Water Quality Model QUAL2E and QUAL2E-UNCAS: Documentation and User Manual. EPA-600/3-87/007. U.S. EPA, Athens, GA.

Burns, L. A. 1990. Exposure Analysis Modeling System: User's Guide for EXAMSII Version 2.94. EPA/600/3-89/084, U.S. EPA, Athens, GA.

Carsel, R.F., C.N. Smith, L.A. Mulkey, J.D. Dean and P. Jowise. 1984. Users Manual for the Pesticide Root Zone Model (PRZM) Release 1. EPA-600/3-84-109. U.S. EPA, Athens, GA.

Davis, M.J., M.K. Snyder and J.W. Nebgen. 1981. River Basin Validation of the Water Quality Assessment Methodology for Screening Nondesignated 208 Areas - Volume I: Nonpoint Source Load Estimation. U.S. Environmental Protection Agency, Athens, GA.

Dean, J.D., W.B. Mills and D.B. Porcella. 1981a. A Screening Methodology for Basin Wide Water Quality Management. Symposium on Unified River Basin Management. R.M. North, L.B. Dworsky and D.J. Allee (eds.) May 4-7, 1980, Gatlinburg, TN.

Dean, J.D., B. Hudson and W.B. Mills. 1981b. River Basin Validation of the MRI Nonpoint Calculator and Tetra Tech's Nondesignated 208 Screening Methodologies, Volume II. Chesapeake-Sandusky Nondesignated 208 Screening Methodology Demonstration. U.S. Environmental Protection Agency, Athens, GA.

DeCoursey, D.G. 1982. ARS Small Watershed Model. Paper No. 82-2094. ASAE Summer Meeting, University of Wisconsin, Madison, WI. June 27-30. pp. 33.

DePinto, J. 1993. "Use of ARC/INFO in Developing a Modeling Support System for Great Lakes Watersheds", In Developing Dynamic Watershed Loadings Modeling Capabilities for Great Lakes Tributaries. The Proceedings of an EPA Workshop at Heidelberg College in Tiffin, Ohio, August 4-6, 1993. Office of Research and Development, U.S. Environmental Protection Agency, Athens, GA, (J.C. Imhoff, Ed)

Di Toro, D.M. and M.J. Small. 1979. Stormwater Interception and Storage. J. Environmental Engineering Div., ASCE, 105(EE1):43-54.

Di Toro, D.M. 1984. Probability Model of Stream Quality Due to Runoff. J. Environmental Engineering Div., ASCE, 110(3):607-628.

Donigian, A.S., Jr., J.C. Imhoff, B.R. Bicknell and J.L. Kittle, Jr. 1984. Application Guide for Hydrological Simulation Program Fortran (HSPF). EPA-600/3-84-065, U.S. Environmental Protection Agency, Athens, GA.

Donigian, A.S. Jr. and W. C. Huber. 1991. Modeling of Nonpoint Source Water Quality in Urban and Non-urban Areas. EPA/600/3-91/039. U.S. Environmental Protection Agency, Athens, GA.

Donigian, A.S. Jr., W.C. Huber and T.O. Barnwell. 1993. Modeling of Nonpoint Source Water Quality in Urban and Nonurban Areas. In: Nonpoint Pollution and Urban Water Management. Technomic Pub. Co., V. Novotny, Ed.

Donigian, A.S. Jr. 1994. Bibliography of HSPF and Related References. AQUA TERRA Consultants, Mountain View, CA. 13 p.

Driscoll, E.D. 1981. Combined Sewer Overflow Analysis Handbook for Use in 201 Facility Planning. Two Volumes, Final Report for Contract 68-01-6148, U.S. Environmental Protection Agency, Facility Requirements Division, Policy and Guidance Branch, Washington, DC.

Driscoll, E.D. 1986. Lognormality of Point and Nonpoint Source Pollutant Concentrations. Proc. Stormwater and Water Quality Model Users Group Meeting, Orlando, FL, EPA/600/9-86/023, pp. 157-176, (NTIS PB87-117438/AS), U.S. Environmental Protection Agency, Athens, GA.

Driscoll, E.D., P.E. Shelley, and E.W. Strecker. 1989. Pollutant Loadings and Impacts from Highway Stormwater Runoff. Vol. I Design Procedure (FHWA-RD-88-006) and Vol. III Analytical Investigation and Research Report (FHWA-RD-88-008). Office of Engineering and Highway Operations R & D, Federal Highway Administration, McLean, VA.

EMWG. 1994. Primary, Secondary & Screening Models for Pesticide Registration. FIFRA Exposure Modeling Work Group, National Agricultural Chemicals Association, Washington, DC.

EPA. 1983. Results of the Nationwide Urban Runoff Program, Volume I, Final Report. NTIS PB84-185552, U.S. Environmental Protection Agency, Washington, DC.

EPA. 1992a. Guidelines for Exposure Assessment. Federal Register. Vol. 57, No. 104. Friday, May 29, 1992.

EPA. 1992b. Nonpoint Sources Research Plan, FY93-FY97. Office of Research and Development, U.S. Environmental Protection Agency, Athens, GA.

Heaney, J.P., W.C. Huber and S.J. Nix. 1976. Storm Water Management Model, Level I, Preliminary Screening Procedures. EPA-600/2-76-275 (NTIS PB-259916). U.S. Environmental Protection Agency, Cincinnati, OH.

Haith, D.A. 1980. A Mathematical Model for Estimating Pesticide Losses in Runoff. J. of Env. Qual 9(3):428-433.

Haith, D.A. and L.J. Tubbs. 1981. Watershed Loading Functions for Nonpoint Sources. ASCE. J. Env. Engr. 107(EE1):121-137.

Heaney, J.P., W.C. Huber, M.A. Medina, Jr., M.P. Murphy, S.J. Nix, and S.M. Hasan. 1977. Nationwide Evaluation of Combined Sewer Overflows and Urban Stormwater Discharges, Vol. II: Cost Assessment and Impacts. EPA-600/2-064b (NTIS PB-266005), U.S. Environmental Protection Agency, Cincinnati, OH.

Henderson, R.J. and G.D. Moys. 1987. "Development of a Sewer Flow Quality Model for the United Kingdom", Topics in Urban Storm Water Quality, Planning and Management, Proc. Fourth Int. Conference on Urban Storm Drainage, W. Gujer and V. Krejci, eds., Ecole Polytechnique Federale, Lausanne, Switzerland, pp. 201-207.

Holst, R.W. and L.L. Kutney. 1989. U.S. EPA Simulator for Water Resources in Rural Basins. Exposure Assessment Branch, Hazard Evaluation Division, Office of Pesticide Programs, U.S. EPA (draft).

Huber, W.C. 1985. Deterministic Modeling of Urban Runoff Quality. In: Urban Runoff Pollution, H.C. Torno, J. Marsalek and M. Desbordes, eds., NATO ASI Series, Series G: Ecological Sciences, 10:167-242, Springer-Verlag, New York, NY.

Huber, W.C. 1986. Modeling Urban Runoff Quality: State of the Art. Proceedings of Conference on Urban Runoff Quality, Impact and Quality Enhancement Technology, B. Urbonas and L.A. Roesner, eds., Engineering Foundation, ASCE, pp. 34-48, New York, NY.

Huber, W.C. and R.E. Dickinson. 1988. Storm Water Management Model User's Manual, Version 4. EPA/600/3-88/001a (NTIS PB88-236641/AS), U.S. Environmental Protection Agency, Athens, GA.

Hydrologic Engineering Center (HEC). 1977. Storage, Treatment, Overflow, Runoff Model, STORM, User's Manual, Generalized Computer Program 723-S8-L7520. U.S. Army Corps of Engineers, Davis, CA.

Hydroscience, Inc. 1979. A Statistical Method for Assessment of Urban Storm Water Loads - Impacts - Controls. EPA-440/3-79-023 (NTIS PB-299185/9), U.S. Environmental Protection Agency, Washington, DC.

Imhoff, J.C. and A.S. Donigian, Jr. 1993. Strategies for Developing a Watershed Scale Model Capable of Multi-Media, Multi-Stressor Sensitive, Total Ecosystem Risk-Based Analysis for the Great Lakes Basin. Office of Research and Development, U.S. Environmental Protection Agency, Athens, GA.

Jacobsen, P., P. Harremoes and C. Jakobsen. 1984. "The Danish Stormwater Modelling Package: The SVK-System", Proc. Third Int. Conference on Urban Storm Drainage, P. Balmer, P-A. Malmqvist and A. Sjoberg, eds., Chalmers University, Goteborg, Sweden, 2:453-562.

James, L.D. and S.J. Burges. 1982. Selection, Calibration, and Testing of Hydrologic Models. In: Hydrologic Modeling of Small Watersheds, C.T. Haan, H.P. Johnson and D.L Brakensiek, eds., Monograph No. 5, Chapter 11, pp. 435-472, American Society of Agricultural Engineers, St. Joseph, MI.

Johanson, R.C., J.C. Imhoff, J.L. Kittle and A.S. Donigian. 1984. Hydrological Simulation Program – Fortran (HSPF): User's Manual for Release 8. EPA-600/3-84-066. U.S. Environmental Protection Agency, Athens, GA.

Kibler, D.F., ed. 1982. Urban Stormwater Hydrology. American Geophysical Union, Water Resources Monograph 7, Washington, DC.

Knisel, W. (ed). 1980. CREAMS: A Field-Scale Model for Chemicals, Runoff, and Erosion-from Agricultural Management Systems. U.S. Department of Agriculture. Conservation Research Report No. 26, 640 pp.

Knisel, W.G., R.A. Leonard and F.M. Davis. 1989. Agricultural Management Alternatives: GLEAMS Model Simulations, Proceedings of the Computer Simulation Conference. Austin, Texas, July 24-27, pp. 701-706.

Leonard, R.A., W.G. Knisel and D.A. Still. 1987. GLEAMS: Groundwater Loading Effects of Agricultural Management Systems, Trans. of the ASAE, 30(5):1403-1418.

Li, W., D.E. Merrill and D.A. Haith. 1989. Loading Functions for Pesticide Runoff. J. WPCF 62(1):16-26.

Limno-Tech. 1990. Simplified Method Program - Two Phase Toxics with Bed Interactions (SMPTOX3) User's Manual. Prepared for U.S. EPA Region IV, Atlanta, GA., supported by U.S. EPA Assessment and Watershed Protection Division, Washington, DC.

Manning, M.J., R.H. Sullivan and T.M. Kipp. 1977. Nationwide Evaluation of Combined Sewer Overflows and Urban Stormwater Discharges - Vol. III: Characteristics of Discharges. EPA-600/2-77-064c (NTIS PB-272107), U.S. Environmental Protection Agency, Cincinnati, OH.

Metcalf and Eddy, Inc. 1971. University of Florida, and Water Resources Engineers, Inc., Storm Water Management Model, Volume I – Final Report. EPA Report 11024DOC07/71 (NTIS PB-203289), U.S. Environmental Protection Agency, Washington, DC.

Miller, R.A., H.C. Mattraw, Jr. and M.E. Jennings. 1978. Statistical Modeling of Urban Storm Water Processes, Broward County, Florida. Proc. International Symposium on Urban Storm Water Management, University of Kentucky, Lexington, KY. pp. 269-273.

Mills, W.B., V.H. Colber and J.D. Dean. 1979. Hand-Held Calculator Programs for Analysis of River Water Quality Interactions. Tetra Tech, Inc., Lafayette, CA.

Mills, W.B., J.D. Dean, D.B. Porcella, S.A. Gherini, R.J.M. Hudson, W.E. Frick, G.L. Rupp and G.L. Bowie. 1982. Water Quality Assessment: A Screening Procedure for Toxic and Conventional Pollutants. EPA-600/6-82-004a and b. Volumes I and II. U.S. Environmental Protection Agency.

Mockus, J. 1972. Estimation of Direct Runoff from Storm Rainfall. In: National Engineering Handbook, Sec. 4, Hydrology. U.S. Soil Conservation Service, Washington, DC.

Mullins, J.A., R.F. Carsel, J.E. Scarbrough, and A.M. Ivery. 1993. PRZM-2, A Model for Predicting Pesticide Fate in the Crop Root and Unsaturated Zones: Users Manual for Release 2.0. EPA/600/R-93/046. U.S. EPA, Athens, GA.

Najarian, T.O., T.T. Griffin and V.K. Gunawardana. 1986. Development Impacts on Water Quality: A Case Study. J. Water Resources Planning and Management, ASCE, 112(1):20-35.

Noel, D.D. and M.L. Terstriep. 1982. Q-ILLUDAS – A Continuous Urban Runoff/Washoff Model. Proc. Int. Symp. on Urban Hydrology, Hydraulics, and Sediment Control, UKY BU128, University of Kentucky, Lexington, KY. pp. 1-10.

Novotny, V. 1985. Discussion of "Probability Model of Stream Quality Due to Runoff", by D.M. Di Toro. Journal of Environmental Engineering, ASCE, 111(5):736-737.

Pisano, W.C. and C.S. Queiroz. 1977. Procedures for Estimating Dry Weather Pollutant Deposition in Sewerage Systems. EPA-600/2-77-120 (NTIS PB-270695), U.S. Environmental Protection Agency, Cincinnati, OH.

Roesner, L.A., H.M. Nichandros, R.P. Shubinski, A.D. Feldman, J.W. Abbott and A.O Friedland. 1974. A Model for Evaluating Runoff-Quality in Metropolitan Master Planning. ASCE Urban Water Resources Research Program, Technical Memo. No. 23 (NTIS PB-234312), ASCE, New York, NY.

Roesner, L.A. and S.A. Dendrou. 1985. Discussion of "Probability Model of Stream Quality Due to Runoff", by D.M. Di Toro. Journal of Environmental Engineering, ASCE, 111(5):738-740.

Roesner, L.A., J.A. Aldrich and R.E. Dickinson. 1988. Storm Water Management Model User's Manual Version 4: Addendum I, EXTRAN. EPA/600/3-88/001b (NTIS PB88-236658/AS). U.S. Environmental Protection Agency, Athens, GA.

Sartor, J.D. and G.B. Boyd. 1982. Water Pollution Aspects of Street Surface Contaminants. EPA-R2-72-081 (NTIS PB-214408), U.S. Environmental Protection Agency, Washington, DC.

Small, M.J. and D.M. Di Toro. 1979. Stormwater Treatment Systems. Journal of the Environmental Engineering Division. Proc. ASCE, 105(EE3):557-569.

Tasker, G.D. and N.E. Driver. 1988. Nationwide Regression Models for Predicting Urban Runoff Water Quality at Unmonitored Sites. Water Resources Bulletin, 24(5):1091-1101.

Terstriep, M.L. and J.B. Stall. 1974. The Illinois Urban Drainage Area Simulator, ILLUDAS. Bulletin 58, Illinois State Water Survey, Urbana, IL.

U.S. Forest Service. 1980. An Approach to Water Resources Evaluation of Non-Point Silvicultural Sources (A Procedural Handbook). EPA-600/8-80-012. U.S. Environmental Protection Agency, Athens, GA. 861 pp.

Water Pollution Control Federation. 1989. Manual of Practice on CSO Pollution Abatement, No. FD-17, WPCF, Alexandria, VA.

Williams, J.R. 1975. Sediment-Yield Prediction with Universal Soil Loss Equation Using Runoff Energy Factor. In: Present and Prospective Technology for Predicting Sediment Yields and Sources. U.S. Dept. of Agricultures. ARS-S-40.

Williams, J.R. and H.D. Berndt. 1977. Sediment Yield Prediction Based on Watershed Hydrology. Transactions of the ASAE. 20(6):1100-1104.

Williams, J.R. and A.D. Nicks. 1982. "CREAMS Hydrology Model – Option 1", Applied Modeling in Catchment Hydrology, Proceedings of the International Symposium on Rainfall-Runoff Modeling. Water Resources Publications, Littleton, CO. Singh, V.P. (Ed.), pp 69-86.

Williams, J.R., C.A. Jones and P.T. Dyke. 1984. A Modeling Approach to Determining the Relationship Between Erosion and Soil Productivity. Trans. ASAE 27(1):129-144.

Williams, J.R., A.D. Nicks and J.G. Arnold. 1985. Simulator for Water Resources in Rural Basins. ASCE J. Hydraulic Engineering. 111(6):970-986.

Wischmeier, W.H. and D.D. Smith. 1978. Predicting Rainfall–Erosion Losses: A Guide to Conservation Planning. Agricultural Handbook No. 537. U.S. Dept. of Agriculture, Agricultural Research Service.

Woodward-Clyde Consultants. 1989. Synoptic Analysis of Selected Rainfall Gages Throughout the United States. Report to EPA, Woodward-Clyde Consultants, Oakland, CA.

Young, R.A., C.A. Onstad, D.D. Bosch and W.P. Anderson. 1986. Agricultural Nonpoint Source Pollution Model: A Watershed Analysis Tool. Agriculture Research Service, U.S. Department of Agriculture, Morris, MN.

Zison, S.W., K. Haven and W.B. Mills. 1977. Water Quality Assessment: A Screening Methodology for Nondesignated 208 Areas. EPA-600/6-77-023. U.S. Environmental Protection Agency, Athens, GA.

Zison, S.W. 1980. Sediment-Pollutant Relationships in Runoff from Selected Agricultural, Suburban and Urban Watersheds. EPA-600/3-80-022, U.S. Environmental Protection Agency, Athens, GA.

CHAPTER 12

Eco-hydrological consequences of environmental degradation: Hydrology, ecology and environmental impacts

Paul H. Templet and Jens Sorensen

Abstract. Hydrologic and ecological systems are continually interacting and mutually shaping the landscape. Actions of man that change one or the other may have unanticipated effects, or impacts, which can be spread beyond the site by the hydrologic cycle. This chapter will present a way to recognize impacts to hydrology or anticipate changes from hydrologic changes. One useful technique is to link landforms and hydrologic processes through *Comparative Hydrology*, which defines eight major landforms and links them to a number of hydrologic processes. Changes to either hydrologic processes or landforms can result in impacts, and tables are given which allow some prediction. A process for developing an environmental assessment is presented along with techniques used by planners for impact assessment statements. The chapter provides a guide to those interested in impact assessment and refers the reader to additional sources for more detailed analyses.

12.1 Introduction

Professionals in the environmental management and hydrology fields are increasingly realizing the major role that the water cycle plays in modifying the landscape and spreading initial localized impacts over much larger areas over time, i.e., water has the capacity to distribute impacts. The converse is also true: modifying the landscape has impacts on the hydrologic cycle. Landforms and hydrology are a mutually evolving and reinforcing system and may exist in near steady-state in which change is slow or, less frequently, in states of rapid change. The continuous and interdependent nature of the hydrologic cycle insures that changes in one part

V. P. Singh (ed.), Environmental Hydrology, 427–461.
© 1995 *Kluwer Academic Publishers.*

of the system will elicit changes in a range of hydrological processes. In times of rapid change, usually brought on by man, the impacts are much more evident.

Many of the consequences of environmental change on hydrology involve changes in water distribution over space, time, quantity and quality. Most are well documented but others are subtle and only occur over long periods of time or far downstream. As with many environmental amenities, changes in hydrology that will occur in the future or in other places as a result of current activities are often given little consideration. Yet these predictable impacts may return to haunt us. Consider, for example, residential building in marginal flood plains, which leads to increased flooding, calls for public funding of flood control projects, followed by more building in flood plains (Dunne and Leopold, 1978) and so on as the cycle continues, at increasing public, private and environmental cost. Massive flood control projects on the Lower Mississippi River in Louisiana, while controlling flooding and enhancing navigation, have led to the U.S.' largest wetland loss rate (Templet et al., 1988; Day et al., 1989) and more public cost in terms of lost fisheries and recreation and new mitigation projects.

The undesired, and usually unintended, side effects can be partially treated by the relatively new disciplines of impact assessment and modeling that attempt to identify, predict, quantify and mitigate hydrological impacts and are presented elsewhere in this text. Considerable work has been done in hydrologic impact assessment and connecting impacts with particular project actions. Hydrology and environmental journals are filled with case studies, models, and theories that delve into great detail. There are also good efforts aimed at aggregating information (Leopold et al., 1971) using, for example, large matrices. One noteworthy attempt was the Environmental Evaluation System by Battelle Laboratories (Whitman et al., 1971), specifically for water resource projects, which used scalars to convert environmental measurements into common units. For a review of impact evaluation techniques see Westman (1985) or Spaling and Smit (1993) for Cumulative Environmental Change. What many methods lack, however, is a guide as to which impacts should be expected.

In this chapter we will attempt to identify and classify the consequences to hydrology from changes in the environment. We will take a holistic, or systems, view and try to establish a framework by which hydrologic changes can be anticipated from an action or set of actions which alter the environment The intent is to establish a workable process by which impacts to hydrology can be identified that can then be further explored in other works. We will then present a more formal impact assessment process for use in project planning. The reader is referred to the many excellent detailed texts replete in every university library for the in-depth analyses of a particular hydrological impact or issue. The goal of planning, as used here, is to maintain ecosystem health, which is in itself a subject for debate (Costanza et al., 1992). A working definition of ecosystem health for our purposes means that unwanted change is minimized or avoided.

A note of caution: this analysis is, of necessity, macro scale and generalizes considerably. It is presented as a means of familiarizing the reader with those aspects of the hydrologic environment that planners and scientists consider important in impact analysis. There is no substitute for site specific analysis by a trained hydrologist for more detailed and accurate results.

12.2 Environmental Change

The agent of change of interest is man as he manipulates and modifies the environment for greater socio-economic development. In the broadest sense resource or socio-economic development refers to any activity designed to improve one's condition. Water is often a constraint on development, both from the view of not enough or too much. It can also be an opportunity for development if planned properly with an understanding of the hydrology of the water resource. From an ecological or environmental perspective few, if any, of man's actions are beneficial to natural systems even though the change is helpful to man. Ecosystems have evolved over eons to optimum configurations, and it is unlikely that we can improve on their processes, efficiency or structure. What we can do is to appropriate more of the ecosystem flows of energy and materials, its structure or its very space for our own use. These modifications in the system for our benefit generally result in negative impacts, and they are the subject of this chapter. The role of planning, and we use the term very broadly, is to minimize the negative and maximize the positive changes caused by an action. There are other agents of change, e.g., fire, floods, storms, and earthquakes, but we do not normally initiate these natural changes, even though we try to minimize their impacts, and they will not be considered further.

12.2.1 A Partial List of the Agents of Change

Agriculture	Irrigation
Forest Management	Urban Area Expansion
Water Supply	Wastewater Disposal
Removal of Rainwater	Flood Control
Landscape Manipulation	Urban land Development
	Residential, Industrial
	Open Space, Commercial

There is another aspect of change to consider. The environment provides "free" services to development in the form of clean water and air, timber, fish, waste treatment, and others. If the environment is degraded by the development, its ability to provide those services may be impaired and the development itself may suffer. For example, unplanned tourist development has often led to polluted water, over crowded beaches, air pollution, reef exploitation and a reduction in fisheries. But these were the very amenities which the tourists were seeking in the first place. Over time the quality of the development will decline, and with it the economic benefits. Planning and management can help prevent the decline.

12.3 Comparative Hydrology and Landforms

In weaving paths through large amounts of information one is faced with the task of developing groupings that summarize, while not trivializing, and that can lead the reader to find plausible answers. There may be a number of ways to accomplish

this but one approach in considering hydrologic impacts is to group landforms according to similar hydrologic processes and responses. *Comparative Hydrology* (1989), a UNESCO publication, presents such a grouping and a framework for understanding the interactions between water and ecosystems in a number of different environments. In their words:

"... Comparative hydrology involves the study of the kind and extent of hydrological processes which act in one part of the earth, and differentiate it from another."

Landforms are classified by catchment response, i.e., flat or sloping, then by temperate or warm and further by dry or humid. Island, delta, and snow and ice hydrology are also discussed but are not carried over to this work because of their uniqueness. The classification into temperate/warm is based on annual potential evaporation (E_0) with warm being greater than 1000 mm and temperate lying between 500–1000 mm. The dry/humid classification is based on the ratio of annual precipitation to potential evaporation (P/E_0); various categories of arid lands lie below $P/E_0 = 0.5$ while the humid category is above 0.5. In their analysis, the interactions of climate, hydrology and vegetation form the basis for distinguishing various regions. But these factors are also important to the ecology of a region and thus this approach establishes linkages between hydrology and ecology. The classification leads to eight landforms: humid and dry, temperate and humid, and dry warm for both sloping and flat lands when deltas, cold lands and islands are omitted. Thus the number of categories are manageable and there are maps differentiating landforms in *Comparative Hydrology*.

The authors then analyze the hydrologic processes within each landform type and group them in two categories depending on whether the direction of water movement is mainly vertical or mainly horizontal. Vertical processes are more important in dry environments and flat land, while horizontal processes are more important in humid climates and sloping lands. The important processes are shown in Table 12.1 grouped by vertical or horizontal.

The addition of Transport and Reaction Processes to the usual hydrological processes is due to the desire to understand hydrology in relation to ecosystems, thus water's role as solvent and a means of transport of sediment, nutrients and pollutants becomes important. The combination of both type processes will constitute "hydrological processes" hereafter.

A cursory review of the list in Table 12.1 illustrates the importance of plants in hydrologic processes, which is in keeping with the emphasis that ecologists place on habitat. The most severe impacts of an action are often due to changes in land use that involves loss of vegetation and changes in the hydrology of an area that mutually reinforce each other and make the impacts proportionally worse. For example, deforestation of hillsides leads to increased runoff (volume and rate) and sedimentation that makes flooding predictably worse as channels fill with sediments and increased runoff leads to ever higher flood stages. Either sedimentation or increased runoff alone is a significant impact; together they can be disastrous. The 100-year flood plain is rising in many areas due to these combined effects with potential widespread flooding. The hydrologic cycle is the mechanism that ties the events together and spreads the impact far beyond the denuded hillside.

TABLE 12.1
Hydrologic, Transport and Reaction Processes

Vertical Processes	Horizontal Processes
Hydrology	
Precipitation	Surface Runoff Generation
Interception by Vegetation	Surface Flow Routing
Infiltration	Groundwater Recharge and Discharge
Soil Water Redistribution	
Evaporation and Transpiration	
Transport and Reaction Processes	
Precipitation	Water in Depressions
Interception by Vegetation	Overland Flow
Water in Plants	Stream Flow
Water in the Root Zone	Groundwater
Water below the Root Zone	

Note: Processes within snow and ice have been omitted from this table.

A partial summary matrix of some important process information by landform type has been extracted from the Comparative Hydrology analysis and is presented in Table 12.2. This should be used as a guide to begin the work of identifying the hydrologic processes of most interest to a particular landform type. A review of pertinent texts, e.g., Dunne and Leopold (1978), will fill out the process information and help to identify additional processes or impacts of importance.

12.4 Weighting Impacts

There is another way to determine which processes might be affected by the proposed action. Using the vertical (v) or horizontal (h) nature of hydrologic processes and the importance of either v or h to the eight landform types from *Comparative Hydrology*, we can develop a relative ranking (weighting) of processes per landform type. A simple ranking system consists of assigning a 1 for each like component (v or h) by hydrological process (Table 12.1) and landform (Table 12.2) and summing the 1's. For example, humid sloping land (whether temperate or warm) has 3 h's for horizontal processes for a score of 3. For vertical processes humid sloping land has 2 h's and 1 v for a score of 2. Thus horizontal processes are more important on sloping landforms, while vertical processes are more important on flat lands. Disruptions of horizontal processes are expected to cause greater impacts on sloping lands, and disruption to vertical processes would result in larger impacts on flat lands.

TABLE 12.2 Landform and process matrix

Hydrological Processes	Approx. Global Location	Most Important Direction Of Movement	Precipitation	Interception by vegetation	Infiltration	Soil Water Redistribution	Evaporation & Transpiration
Direction of Process			(vertical)	(vertical)	(vertical)	(vertical)	(vertical)
Comparative Landform							
Sloping Land (well defined catchment)		Horizontal					
Humid Temperate	40-60 degs. lat.	Horizontal	Generally well supplied with water				
Dry Temperate	30-50 degs. lat., steppe, savanna & prairie	Vertical					
Humid Warm	40 degs. S to 40 N.; Tropical	Horizontal	2 Subregions based on continuous or seasonal rainfall	High, tree crowns and litter important	High Rates due to forest cover and litter	Good as long as litter and cover are undisturbed	High & stable, forest important for returning water to atmosphere
Dry Warm	40 degs. S to 40 N.; arid & semi arid	Vertical	Slight, seasonality increases with increasing latitude	Slight	High when water available in overland flow and along river beds	Slight	High, most available water is evaporated or transpired
Flatlands (Poorly defined Catchment)		Vertical					
Humid Temperate	Belt centered on 50 degs.	Horizontal			80% in sandy soil, 67% in loamy sand, 47% in silty clays		
Dry Temperate	40-50 degs. N., Eurasia	Vertical	Low to Moderate		Shallow Water Table	Vertical water exchange within soil important, capillary rise to plants or surface	Evapotranspiration near or exceeds precipitation
Humid Warm	40 degs. S to 40 N.; Tropical, Savannahs	Horizontal	Seasonal, Length of wet/dry seasons determines vegetation	Important as soil erodes easily	Seasonal	Dependent on wet/dry seasons, perched water table due to impermeable layer	Seasonal, Exceeds Precipitation in dry season, not in wet season, vegetation important regulator
Dry Warm	40 degs. S to 40 N.; arid & semi and	Vertical	Slight, seasonality increases with increasing latitude		Rates generally high, depending on local soil permeability	Most held at root level from which it is transpired	High, potential evapotranspiration exceeds rainfall

TABLE 12.2 Landform and Process matrix

Surface Runoff Generation (Horizontal)	Surface Flow Routing (Horizontal)	Groundwater Recharge and Discharge (Horizontal)	Water in Plants (Vertical)	Water in the Plant Root Zone (vertical)	Soil Water Below Plant Root Zone (vertical)	Water in Depressions (Horizontal)	Overland Flow-runoff (Horizontal)	Stream Flow (Horizontal)	Groundwater (Horizontal)
Variable source, increases with area		Recharge from Soil Water		Dry in Summer, Wet in Winter	Dry in Summer, Wet in Winter				
Decreases with Area		Recharge from Streams					Large Seasonal Variations	Large Seasonal Variations	
High, increases with Area		Recharge from Soil Water		Good	Good	Headwater Swamps stabilize regimes, delay surface runoff	Generally low to moderate if forest cover undisturbed, response rapid	Complex, large variation where rainfall is seasonal (monsoons), flooding	Accumulations are continuous, large fraction of rainfall infiltrates saturated zone
Wide variation but decreases with area	Overland to steambeds	Recharge primarily from Streams	Low to moderate plant cover	Root zone of plants deeper and more important		Low	Intermittent rapid flows most frequent depending on rainfall & slope, occasional floods	Often intermittent, high velocity, some flooding during high intensity rainfall	Rapid recharge and discharge, infiltration along river beds
Increases with Area, 30-60% of precip.		Recharge from Soil Water							
Decreases with Area	Sheet flow or shallow channels to streams, unstable and dependent on local gradient	Recharge from Streams, important as storage	Important as storage	Water Table near plant roots	Important as storage	Important as storage, overflow drains to streams	Sheet flow or small shifting channels from depressions to streambeds	Channels shifting and shallow, floods and droughts last longer	Generally near the surface
Increases with Area	Generally well defined streams, sheet flow	Recharge from Soil Water		Perched water table near plants	Perched water table near plants	Swamps, lakes and shallow depressions common	Generally well defined streams, sheet flow	Slow sluggish streams	
Decreases with Area	Poorly defined	Recharge from Streams	Low in Arid areas	Water Table far below surface	Very little	In impermeable areas rainfall stored in small depressions which may overflow to larger ones but evaporation very high except in subtropical areas	Small ephemeral channels or sheet flow, flows entering from uplands degenerate and losses from channels significant	Poorly organized channel network with large flood plains and low banks	Groundwater an important component of water resources but recharge low except when originating in rivers in the humid zone

TABLE 12.3 Process and Landform weighting factors

Process	Sloping Land (h) Humid Temperate (h)	Sloping Land (h) Dry Temperate (v)	Sloping Land (h) Humid Warm (h)	Sloping Land (h) Dry Warm (v)	Flat Land (v) Humid Temperate (h)	Flat Land (v) Dry Temperate (v)	Flat Land (v) Humid Warm (h)	Flat Land (v) Dry Warm (v)		
Precipitation (2v)	2	3	2	3	3	4	3	4		
Interception (2v)	2	3	2	3	3	4	3	4		
Infiltration (v)	2	2	2	2	2	3	2	3		
Soil Water Redistribution(v)	2	2	2	2	2	3	2	3		
Evaporation & Transpiration (v)	2	2	2	2	2	3	2	3		
Water in Plants (v)	2	2	2	2	2	3	2	3		
Water in the Root Zone (v)	2	2	2	2	2	3	2	3		
Water below the Root Zone (v)	2	2	2	2	2	3	2	3		
Surface Runoff Generation (h)	3	2	3	2	2	2	2	2		
Surface Flow Routing (h)	3	2	3	2	2	2	2	2		
Groundwater Re-&DisCharge (h)	3	2	3	2	2	2	2	2		
Water in Depressions (h)	3	2	3	2	2	2	2	2		
Overland Flow (h)	3	2	3	2	2	2	2	2		
Stream Flow (h)	3	2	3	2	2	2	2	2		
Groundwater (h)	3	2	3	2	2	2	2	2		
On Humid Sloping lands the Horizontal processes are more important, whereas on dry flat lands vertical processes predominate, in the other two cases, dry sloping and humid flat lands, there is some ambiguity.										

TABLE 12.4 Partial matrix of hydrologic process impact by various agents of change

Hydrologic	Precipitation	Interception	Infiltration	Soil Water	Evaporation &	Processes Within	Surface Runoff	Surface Flow	Groundwater Recharge	Water in
Processes		by vegetation		Redistribution	Transpiration	Snowpacks	Generation	Routing	and Discharge	Plants
Flow Direction	(vertical)	(vertical)	(vertical)	(vertical)	(vertical)	(vertical)	(Horizontal)	(Horizontal)	(Horizontal)	
Agents of Change										
Land Use Changes		Reduces	Modifies infiltration		Modifies evapotranspiration		Major increase in runoff & flooding			
Forest Clearance			Decreases with loss of forest		Decreases with loss of forest		Major increase in runoff		Recharge decreases with loss of forest cover	
Agriculture	Local droughts a risk	Reduced	Humid warm soils slake under impact of raindrops, develop crust which reduces infiltration	Changed due to surface soil changes, tilling or over grazing can create impervious horizons which decrease storage and productivity, increased salinization	Transpiration affected, surface evaporation increased		Major increase in runoff volume	Diversions for Ag. use a major impact, can lead to desertification in dry areas		
Urbanization	Acid precipitation	Reduced	Reduced due to man made impermeable structures		Transpiration reduced, evaporation may be increased		Major increase in runoff & flooding	Changed or held static by man made drainage	Reduced due to more impermeable structures such as roofs, roads, parking areas, etc.	
Industrialization	Acid precipitation	Reduced	Reduced due to man made impermeable structures		Transpiration reduced, evaporation may be increased		Major increase in runoff & flooding			
Channel Modification								Major changes		
Effluent Returns										
Extraction Activities								Diversions can cause major impacts		
Storage, Dams, impoundments, etc					Increased losses due to evaporation from larger surface area			Changed from Riverine to Lacustrine, thermal and chemical stratification occurs, fish type and distribution changed		

TABLE 12.4 Partial matrix of Hydrologic Process Impact by Various Agents of Change

Water in the	Soil Water Below	Water in	Overland Flow	Stream Flow	Groundwater	Water Quality	Erosion & Sediment	Management Implications
Plant Root Zone	Plant Root Zone	Depressions						
(vertical)	(vertical)		(Horizontal)	(Horizontal)	(Horizontal)	Both	Horizontal	
		Eutrophication increased	Major increases	Major increases		Sediment increases 5-10X if cover lost	Erosion & sediment low where cover is intact	Scale is important, large projects are the most disruptive, smaller but more numerous ones should have less impact, promote self sufficiency and conservation
		Eutrophication increased	Major increases	Major increases	Infiltration to groundwater will decrease substantially	Decreased quality, e.g. sediment increases 5-10X, Nitrate increases	Increases erosion & sedimentation, 1000 fold increases in stream sediment load not unusual	Do not remove forest cover or litter over large areas
Can be affected by changed soil horizons	Can be affected by changed soil horizons, humid warm soils unstable	Usually eliminated which leads to reduction in infiltration, an increase in runoff and increased severity of floods and droughts	Major increases	A major user of water for irrigation, increases in irrigation are a major impact, desertification and soil salinity increases a problem in dry areas	Overuse of groundwater for irrigation can cause depression of aquifers	Nutrients, sediment & BOD increases,	Increases erosion & sedimentation, soil loss can be very high in intensively cropped areas & overgrazed areas, desertification	Simulate local vegetative cover conditions, low till, integrate tree, livestock and row crops, reduce chemical use, reduce water losses through conservation, use carrying capacity to determine scale, reduce phosphorous loadings
Reduced	Reduced	Eutrophication increased	Major increases, roads, canals and railway embankments can affect flow adversely	Major increases	Overpumping of groundwater leads to subsidence and aquifer depression	Sediment (5-10X), organics, nutrients, salts, metals, sewage & BOD increase	Increases erosion	Conserve supply, treat sewage or use land disposal, flood protection, maintain vegetative cover, micro hydrologics can be useful in local areas, macro for large
		Eutrophication increased	Major increases, roads, canals and railway embankments can affect flow adversely	Major increases	Overpumping of groundwater leads to subsidence and aquifer depression	All Pollution Loading Increases		Recycle water, reduce pollution loads, implement pollution prevention practices, reduce phosphorous loadings
			Major changes	Major changes				
						Water Quality decreases		
				Mining can be a major user, overuse can cause shortages		Sediment, salinity, acid & metals increase	Sediment in rivers near mining operations generally very high	Recycle water, reduce pollution loads, implement pollution prevention practices, reduce phosphorous loadings
				Riverine ecological conditions changed to lacustrine		Decreased, chemical and biological balance changed	Siltation behind dam	Population resettlement difficult, minimize hydrologic disruption

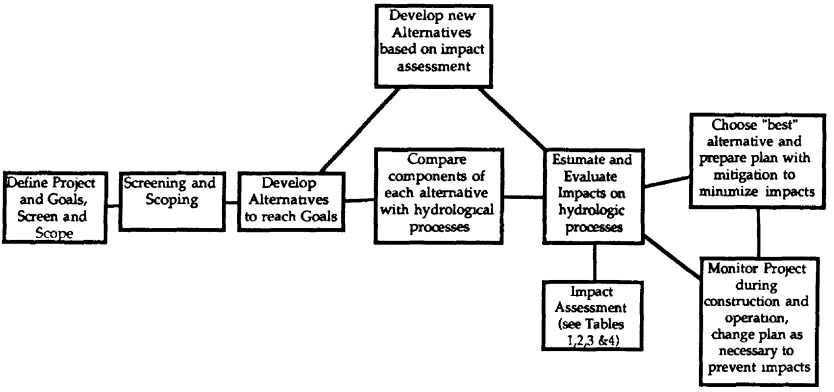


Fig. 12.1. Flow chart of the hydrologic planning process.

Precipitation and Interception processes are each assigned 2 v's since they appear twice in the list of Table 12.1. The results of the ranking process are shown in Table 12.3. The principle conclusion is that on humid sloping lands, both temperate and warm, horizontal processes are more important whereas on dry flat lands vertical processes dominate. In the other two cases, dry sloping and humid flat lands, the importance of the direction of the hydrologic process is more ambiguous. Specialized site-specific studies will have to determine the most important processes. The relative rankings in Table 12.3 can be used, with care, as importance factors in matrices that seek to aggregate impact information.

12.5 Overview of The Planning Process

The hydrologic planning process is covered thoroughly by Helweg (1985) and Westman (1985) and deals with general planning principles. Dunne and Leopold (1978) provide fairly detailed hydrologic information for the planning process. This section presents a simplified planning process and refers the reader to one of the references above for more details on either planning or hydrology. Figure 12.1 is a flow chart of the hydrologic planning process with boxes representing the steps in the process. It incorporates the critical components from a number of approaches and is consistent with the Environmental Impact Statement (EIS) procedure under the National Environmental Policy Act (NEPA). The basic approach is to define goals, e.g., reduce flooding or increase water supply, then develop alternatives to reach the goal(s), evaluate the alternatives for impacts and create new alternatives based on the impact analysis, choose a course of action and create a plan or project. The project is monitored during construction and operation and modified, if necessary, to further mitigate impacts. The reader should be aware that the planning process discussed here primarily concerns environmental factors and impacts and that economic, social and engineering feasibility concerns must be satisfied in other parts of the evaluation process.

It is important to note the difference between goals and projects. Goals generally allow more flexibility in choosing alternatives than do projects; for example,

if reduced flooding is the goal, then flood plain zoning or flood proofing may be the desired solution rather than building levees or dams. If the goal is defined as a project, i.e., building a dam, levee or other structure, then the non structural alternatives, such as local floodplain ordinances will be overlooked and opportunities missed. As Helweg (1985) points out "... Water resource planners tend to have engineering backgrounds, generally in civil engineering. They bring to planning a propensity to meet needs by building a structure of some kind. The technological solution is often emphasized at the expense of societal goals." What is required is a goal statement that is broad enough to allow many different types of alternatives to be formulated by not overly constricting the choices early in the planning process. The alternatives with poor characteristics will not survive the environmental, economic and engineering feasibility process that all projects endure in modern society. It is not appropriate to eliminate them before they have had a chance to be adequately tested.

The alternatives formulation step is, in many ways, the brainstorming or creative part of the project. There is little harm in allowing many alternatives, some far fetched, to creep into the process. All must pass fairly rigorous evaluation tests before making it to the final plan. The visionary idea from an engineering or environmental perspective may turn out to have economic advantages that are only evident in the final stages of evaluation. Once an alternative has been defined, its component parts must be examined for their impacts on hydrologic processes (Table 12.4). For example, if the goal is to increase agricultural production and one alternative is to increase irrigation then one component is finding a source of water. The source can be groundwater or surface and the impacts to each, if any, must be studied. It may turn out, and should, that the impacts can be reduced if conservation of water through drip irrigation, or other technique, is applied. This finding activates the first feedback loop in Figure 12.1, which provides a new alternative, i.e., conservation in irrigation. If the source of water for irrigation is surface water and the region is humid sloping land (warm or temperate) then an important hydrologic process, surface flow, may be affected (Tables 12.2 and 12.3). In a dry flat area soil salinization may become a problem because potential evapotranspiration exceeds precipitation. In this way a matrix of project components listed on one axis and the hydrologic processes shown in Table 12.2 on the other axis can be developed for each alternative.

Numbers in the boxes of the matrix can represent the magnitude of the impact and the importance factor (weighting). Table 12.4 is a matrix of potential impacts on hydrologic processes by various agents of change and can be used as a first approximation of potential impacts due to a project action. The importance number, i.e., the weighting of the impact, can be taken from Table 12.3 for particular landforms if desired or it can be determined in a similar manner as the magnitude. Both can also be determined by expert opinion, preferably with a panel of experts in a Delphi process.

The Delphi process uses procedures in which, for example, experts rate impacts individually, e.g., on a 1-10 scale, then the panel is presented with the range and mean of the cumulative ranking and asked for a new individual ranking. Convergence by the panel usually occurs quickly and the new mean is taken as the magnitude of the impact. Once the boxes in the matrix are filled and a matrix

has been done for each alternative, the numbers can be aggregrated to arrive at a grand index. This is done by multiplying the numbers within each box and then summing across and down each matrix. The lowest number should represent the alternative with the least impacts. There are advantages and disadvantages to summing in this way and they are covered in Westman (1985). Principal among these is the neglect of the equity issue, i.e., who gets the benefits and who gets the costs, second order effects, double counting, cumulative effects, and trying to convey too much information in a singe number, the grand index. There are also advantages and disadvantages to leaving the matrices to be read individually by the decision maker and somehow arriving at a decision. The principal advantage is that important information is not lost in the Grand Index; the principal disadvantage is that many decision makers do not want to deal with the large amount of information presented in the full matrices. Once a desired alternative has been selected, it is integrated in a plan with mitigation and monitoring components.

12.5.1 Water Conservation

Water scarcity is a growing problem in many countries, especially those in dry climates with rapidly growing populations (Postel, 1993). In many countries adequate available surface water supplies per capita are declining as pollution increases and groundwater is withdrawn faster than it can be replenished. The WorldWatch Institute (Postel, 1993) has developed quantitative measures of "water-stressed" (1,000–2,000 m^3/person yr) and "water-scarce" (< 1000 m^3/person yr). In those countries that are water scarce, a lack of water is considered to be a severe constraint on food production, economic development, and protection of natural systems. Twenty six countries, with a combined population of 230 million people, fell into the water scarce category in 1993 and more are being added every year. By the end of the nineties it is expected that over 300 million people will live in water scarce countries.

Environmental planning should incorporate water conservation features into all projects and plans. Water will assume more importance in the future as it becomes less available and conservation has other benefits that make the effort worthwhile. For example using water conserving toilets and shower heads in developed countries means that less water needs to be pumped and treated by central sewage plants, thus reducing costs. It is also easier to treat concentrated wastes than dilute wastes. In developing countries conserving water will reduce the burdens on public services that tend to be strained by high rates of growth and urbanization. Water conservation will also make more water available for higher priority items, like food production and manufacturing, than would ordinarily be available and perhaps avoid unnecessary competition for scarce water.

12.6 Detailed Planning and Impact Assessment Process

With this overview we will now discuss the parts of the process presented in Figure 12.1 in more detail.

12.6.1 Description of the Proposed Action

This step starts the whole process of impact assessment. The basic objective of this step should be that the description will provide adequate information to conduct the remainder of the process.

The proponent or initiator of the action must do the initial description, and in most cases, the proponent or initiator is also responsible for the complete and final description. If the impact assessment is to be prepared by an institution other than the unit that has described the action, it is obvious that description must provide the information needed to conduct an adequate assessment. The unit of government that conducts the impact assessment and is responsible for issuing the assessment report is usually termed the lead agency.

Similarly, and no less important, the description must provide reviewers unfamiliar with the proposed (or completed) action sufficient information to conduct an adequate independent review. This means that the action description has to provide enough information for reviewers to conduct their own impact assessment – if deemed necessary – that is independent from the impact assessment done by the lead agency.

Understandably, the proponent or initiator of an action usually wants to avoid including information in the action description that may adversely effect the desired outcome. Therefore, guidelines are necessary to specify what must be included in the description of an action. Guidelines prepared by governments with the longest history of impact assessment include the following specifications:

a. *Purpose.* What are the goals and objectives the action is designed to achieve? What socio-economic benefits will the action achieve?

b. *Integration with Higher or Co-lateral Actions.* What was the policy making cycle that lead to the action being proposed? How does the action integrate and support the higher or Co-lateral levels of plans and programs? Have impact assessments been conducted at the higher or co-lateral levels of plans or programs?

c. *Physical Description.* Describe the activities that will occur. The activities associated with the entire time span of the action should be included. For example, many projects have four phases: research and exploration, construction, operation, termination and restoration. Are there any activities that may be hazardous to health or safety? Indication of the magnitudes of the project should be given, such as area extent, number of personnel involved, equipment, and material requirements.

d. *Staging and Future Plans.* Are there future additions to the action that can be anticipated? Projects are commonly increased in size and or density several years after they receive initial government approval.

e. *Environmental and Socio-economic Setting.* Describe the area's environmental and socio-economic conditions. Maps should be used that indicate the appropriate conditions. Environmental conditions normally include geology,

soils, hydrology, flora, fauna, and hazards. Socio-economic conditions usually include present land use, employment, regional income, and demographic characteristics. Unusual or important elements of the existing situation should be pointed out. For example, the existence of rare species, landmarks, and unique community social characteristics should be identified.

A second major problem in action description is that the guidelines are often very general, such as the ones listed above. They cover all types of projects, all types of environments and all types socio-economic settings. If, however, the action description is being done by the same institution conducting the entire impact assessment process, the generality of guidelines may not be a problem. When the succeeding steps of the impact assessment process are completed, the significant impacts will be determined. Once the significant impacts have been determined, then it is possible to write the action description that contains the information necessary to conduct an impact assessment or impact review. In other words, the process is iterative.

12.6.2 SCREENING

The second step of the impact assessment (IA) process, screening, is to determine whether an IA is needed. Ideally this step will quickly and easily identify those actions that should have their impacts assessed and, equally important, identify those projects that do not warrant the IA process. In practice the screening step sorts actions into three categories:

a. those clearly requiring IA;

b. those clearly not requiring IA; and,

c. those for which the need of IA is unclear and therefore further analysis is required.

Six analytical methods have been developed to assist in making screening decisions. They are positive and negative lists; thresholds; checklists, matrices, and networks; overlay mapping and geographic information systems ; sensitive areas; and, compliance with an adopted plan. These six methods are described below in more detail.

12.6.2.1 Positive and Negative Lists

Positive and negative lists appear to be the simplest screening method. Positive lists specify types of development actions that would require an impact assessment. Negative lists indicate the categories of development actions that normally would be exempt from the IA process. The term "categorical exemption" is frequently applied to actions that comprise the negative list. Table 12.5 is an example of a positive list based on project types.

Some research is required to prepare lists, but they offer an easy-to-use system that is readily understood by all concerned. However, gaining acceptance by all institutions and individuals on all actions on both positive and negative lists will likely be a difficult and time consuming endeavor.

TABLE 12.5
List of those projects or activities in Thailand that require the preparation of an EIS

Type of Project/Activities	Sizes
Dam or reservoir	Storage volume greater than 100,000,000 cubic meters or storage surface area greater than 15 sq km
Irrigation	Irrigated area greater than 12,800 Ha
Commercial Airport	All sizes
Hotel or resort facilities in environmentally sensitive areas (adjacent to rivers and the coast)	Greater than 80 rooms
Mass transit system and expressway	All sizes
Mining	All sizes
Industrial estate	All sizes
Commercial port and harbor	With the capacity for vessels of greater than 500 ton gross
Thermal power plant	Capacity greater than 10 MW
Petrochemical Industry	Greater than 100 tons/day of raw materials required in production
Oil refinery	All sizes
Natural gas separation or processing	All sizes
Chlor-alkaline industry requiring NaCl as raw material for production of Na_2CO_3, NaOH, HCl, Cl_2, NaOCl and bleaching powder	Production capacity of each or combined product greater than 100 tons/day
Iron and/or steel industry	Requiring iron ore and/or scrap iron as raw materials for production greater than 100 tons/day or using furnaces with combined capacity greater than 5 tons/batch
Cement Industry	All sizes
Smelting industry other than iron and steel	Production capacity greater than 50 tons/day
Pulp industry	Production capacity greater than 50 tons/day

(Source: Sudara in Clark et al., 1914)

Positive lists may be compiled by reviewing existing developments and identifying those types of actions that routinely produce adverse impacts. Similarly, actions that seldom generate adverse impacts can be added to the negative list. In the case of those actions in which it is difficult to determine the appropriate list, other screening methods can be applied. For example, the most common problem with lists is that the same type of action can have considerable variation in size, design, and longevity. Also, an action that ordinarily produces no significant impacts may be proposed to locate in a particularly sensitive environment, and therefore should be subject to the IA process. To accommodate these two kinds of problems, the threshold method is frequently applied.

12.6.2.2 Thresholds

This method establishes thresholds for key features of the action, the environmental setting or the social setting. If it is anticipated that a threshold will be exceeded, then an IA is required. The positive and negative lists can be simply converted to thresholds by adding size dimensions, such as hotels with more than 25 rooms require an IA or all developments with a value in excess of $ 250,000 require an IA. Infrastructure requirements have also been used as a threshold.

A second means of setting thresholds is proximity to environmental and social conditions, such as all substantial development within 50 meters of the mean tide or in coastal wetlands require an IA.

Reliance on one threshold will often produce incorrect decisions. It is therefore normal practice to link a number of thresholds together, particularly action and environmental thresholds. For example, a hotel with more than 25 units within 500 meters of the mean high tide should require an IA. However, care will be needed to prevent a series of thresholds from becoming cumbersome and time consuming.

A common problem encountered when using thresholds is the development actions that are designed to just slip under the maximum levels. For example, if 25 rooms is the threshold level for an IA on a hotel, a common response by hotel developers is to divide a large project into segments of 24 units or less and submit these segments at different times: one 24 unit proposal each year over a four year period. This ploy is usually called segmentation. The requirement in the action description step to specify staging and future plans is designed to prevent this strategy of slipping under the threshold.

12.6.2.3 Checklists, Matrices and Networks

These three techniques provide a comprehensive framework that can be quickly used to systematically review an action's description to identify potential adverse impacts. Table 12.6 is a checklist based on types of environments and potential impacts. Table 12.7 provides an example of a matrix used for screening. Usually only one of the three techniques is used to screen actions since all three produce somewhat the same result. Each have their relative advantages. Normally checklists, matrices, or networks are not related to specific types of environments, and consequently it is not possible to determine if the adverse impact will be significant or not.

TABLE 12.6
Potential Impacts of Particular Concern According to Coastal Habitats

Coastal Habitat	Impacts of Particular Concern to the Sri Lanka Coast Conservation Department
1. Coral reefs	Physical damage to coral reefs and collection of reef organisms beyond sustainable limits. Increases in fresh- water runoff and sediments. Introduction of waterborne pollutants
2. Estuaries/lagoons	Encroachment. Changes in sedimentation patterns. Changes to the salinity regime. Introduction of water- borne pollutants. Destruction of submerged and fringing vegetation. Inlet modifications. Loss of fishery habitat
3. Mangroves	Changes in freshwater runoff, salinity regime and tidal flow patterns. Excessive siltation. Introduction of pollutants. Conversion of mangrove habitat and over- harvesting of resources
4. Seagrass beds	Physical alternations. Excessive sedimentaion of siltation. Introduction of excessive nutrients or pesticides
5. (tidal flats) Salt marshes	Degradation of birds habitat or seed fish collection sites. Obstruction of storm water runoff
6. Barrier beaches, sand dunes and spits	Sand mining, Erosion, Dune Migration

Source: Sri Lanka National Coastal Management Plan, 1989

12.6.2.4 Overlay Mapping and Geographic Information Systems

Overlay mapping has the capability of relating the action to its environmental and social setting. In its simplest form, overlay mapping is a series of single factor maps, such as geology, hazards, soils, vegetation, fauna, cultural sites, and landmarks. Overlays of these maps on the site of the proposed action may identify which of the factors that have been mapped will be affected. The maps and the information they portray should also provide some dimension on the significance of the impact.

During the last ten years, with the advent of computers, geographic information systems have become a standard practice in the field of environmental planning. All geographically referenced information can be digitized and entered into computer storage. Instead of laboriously overlaying maps with different characteristics, a geographic information system (GIS) can perform this function. In a GIS all the available types of information pertinent to environmental planning are referenced to their geographic location. This permits answering questions about potential impacts. For example, a GIS could quickly identify if a proposed development site and surrounding area contain habitats of rare or endangered species or include flood prone lands, archaeological sites or erosion prone soils. GIS' have demonstrated to be particularly useful in the identification of cumulative or synergistic impacts. In response to a proposal to clear forest for agriculture, a GIS could determine the cumulative impacts if all lands in a particular watershed that have similar

TABLE 12.7
Screening Questions

	Question	Rating	Score
1.	What is the approximate cost of the construction project?	High	10
		Medium	5
		Low	0
2.	How large is the area affected by the construction project?	High	10
		Medium	5
		Low	0
3.	Will there be a large, industrial type of project under construction?	Yes	10
		No	
4.	Will there be a large, water-related construction activity?	Yes	10
		No	0
5.	Will there be a significant waste discharge (in terms of quantity and quality) to natural waters?	Yes	10
		No	0
6.	Will there be a significant disposal of solid waste on land as a result of construction and operation of the project?	Yes	10
		No	0
7.	Will there be significant emissions (quantity and qualtiy) to the air as a result of construction and operation of the project?	Yes	10
		No	0
8.	How large is the affected population?	High	10
		Medium	5
		Low	0
9.	Will the project affect any unique resources (geological, historical/ archeological, cultural, ecological)?	Yes	10
		No	0
10.	Will the construction be on floodplains?	Yes	10
		No	0
11.	Will the construction and operation be incompatible with adjoining land use in terms of aesthetics/noise/odor/general acceptance?	Yes	10
		No	0
12.	Can the existing community infrastructure handle the new demands placed upon it during construction and operation of the project	No	10
		Yes	0

characteristics – soil type, slope, forest cover and in private ownership – were also allowed to be cleared for agricultural development.

12.6.2.5 Sensitive Areas

The environmental consequences of an action are a function of both the action and the setting, therefore, the sensitivity of the environmental and socio-economic conditions that characterize the setting provide another means to determine whether an impact assessment is required. A common practice in environmental planning is to map areas with environmental or socio-economic conditions that are particularly sensitive to adverse impacts. "Critical areas" or "sensitivity analysis" are other names commonly applied to such a program. The sensitive areas approach has been frequently used to define locations where cumulative impact is an issue. Watershed management and hazard prone areas are the two most common applications of sensitive area programs. Sensitive areas could also be considered as an overlay mapping technique.

12.6.2.6 Plan Compliance

If a comprehensive environmental management plan has been adopted for the area in which the action proposes to locate, the IA may be quickly reduced to simply determining if the action and its impacts will be in compliance with all aspects of the plan. If the action is in compliance with all aspects of the plan, no further impact assessment should be necessary since all the relevant environmental and socio-economic impact analyses were, presumably, done when the plan was prepared.

12.6.2.7 Summary

The screening method or methods should:

a. identify, at the lowest staff costs, the greatest number of actions having significant impacts;

b. minimize the number of actions likely to be incorrectly identified as having significant impacts;

c. minimize the number of actions exempted from the IA process that when completed generate significant unanticipated impacts; and,

d. identify other actions within the same system that may generate significant adverse impacts – particularly cumulative impacts.

12.6.3 SCOPING

The scoping step establishes the terms of reference for the impact assessment. The boundary of the scoping step is often difficult to distinguish from both screening and from the two steps that follow it. Many of the same methods used to screen actions can also be used to scope actions. The scoping step should:

a. identify the issues that should be addressed by the IA;

b. determine what appear to be significant impacts and what appear to be insignificant impacts;

c. provide some dimension on the relative importance of the potential significant impacts so that the depths of analyses can be apportioned;

d. identify appropriate measures to mitigate the potential significant impacts; and,

e. facilitate an efficient process that saves time and staff resources.

The scoping step is designed to overcome one of the inherent pitfalls in impact assessment – the initial tendency to identify all potential impacts without regard to what is important to decision making. It appears that impact assessment goes through an evolutionary phase when the effort is devoted to compiling massive compendia of scientific and technical data, much of which is of marginal relevance to the most significant issues. Scoping is important in preventing the costly collection of information that is not important to the ultimate decisions that have to be made on the proposed action.

When mitigation measures are being suggested, it is important to focus the study on workable, acceptable solutions to the problems. It is easy for the study to waste time considering measures that are impracticable or totally unacceptable to the developer or the government.

The most important factor in the scoping step appears to be who participates. It is common for scoping to be done only by the lead agency. This is particularly true of projects that are not large-scale in size or impact. Impacts from almost any type or scale of action will be multidisciplinary in nature. If one individual is scoping potential impacts, the impacts identified will likely reflect the individual's professional education and experience. If the assessor is a biologist, the impacts identified will be mostly biological. Many significant impacts outside the assessor's professional knowledge will likely be missed.

A second option for broadening the scoping perspective is to create an ad-hoc team drawn from disciplines. This approach is generally used for only large-scale projects. Members of the team reflect the range of impacts of the proposed action. A variation of the second option is to draw members from other government agencies. This inter governmental variation also has the advantage of drawing in other agencies at the beginning of the impact assessment process. Early involvement of agencies with expertise on the potential impacts of an action is important because these same agencies are often required to participate in the review step of impact assessment, particularly if an environmental impact statement is prepared. Another common observation in the IA literature is that early review and critique is far more effective and efficient than at the conclusion of the assessment process. At the scoping stage the options should be broader and more flexible, therefore easier to incorporate cost-effective redesign or mitigation measures.

It has been repeatedly demonstrated that public involvement can produce acceptable terms of reference for the impact assessment and reduce the likelihood of controversy once the assessment has been prepared. However, public involvement as well as interagency involvement has also caused delays and escalated the costs

of assessment. It is not possible to specify the extent to which interests outside the lead agency should be involved in scoping because the answer depends on the type and scale of the proposed action, as well as type, range, intensity and social distribution of impacts. It is clear that effective scoping depends on the exchange of information and concerns among interested parties, including the public. This will be achieved if the organizational arrangement for scoping is designed to achieve the five objectives of this step.

A number of procedures have been designed to conduct the scoping step as well as the subsequent steps of considering alternatives and estimating impacts. Two procedures that are gaining popular acceptance are joint fact finding and adaptive environmental assessment. Because of the time and cost involved in setting up these procedures they have generally been applied to large scale projects.

12.6.4 Identification and Selection of Alternatives

The identification and consideration of alternatives is one of the key aspects of impact assessment. This step provides the means by which the action's assumptions, goals and needs can be examined. A range of alternatives provides the basis for a comparative assessment of the different means to achieve the stated objective of the action. In an assessment of alternatives, decision makers should be provided with the information on how each option compares in respect to the relative costs and benefits for each impact.

One common difference between large scale projects and small scale projects is that large scale projects usually do pose a full range of alternatives. If any alternative is proposed by a small scale project, it is usually the "no action" alternative.

The identification and selection of alternatives usually involves three questions. How should alternatives be identified? What is the reasonable range of alternatives to be considered? What level of examination should be applied to each alternative? Few methods exist to assist in the identification of alternatives. The methods that do exist are associated with types of social objectives, such as new transportation corridors or flood control. For example, alternative transportation routes are commonly identified by using a series of overlay maps that portray various environmental and socio-economic factors.

Alternatives can be organized into a number of categories. Governments have required consideration of the following categories:

a. Demand Alternatives. Tourism and residential development are common demand alternatives for shorelands on a high quality recreational beach.

b. Activity Alternatives. impoundments, river channelization, levees and flood plain zoning are all activity alternatives for minimizing the hazards of river floods.

c. Location Alternatives. On what other sites could the activity occur?

d. Process Alternatives. Many industrial or public service facilities have a number of alternative design configurations, each of which may have different input, output, and pollution dimensions.

e. Mitigation Measures. What alternatives exist to mitigate the adverse impacts identified in the scoping step (see Section 12.6.6)

Guidelines on the range of alternatives usually declare that all reasonable options should be considered, including the no action alternative. A reasonable alternative is one that is capable of accomplishing the objectives of the proposed action within a reasonable period of time, taking into account economic, social, environmental and technological factors. One of the inherent problems of IA is making the distinction between reasonable and unreasonable alternatives. The concern is not wasting time and money on the assessment of alternatives that are not economically or politically realistic.

The identification of the full range of reasonable alternatives often requires a multidisciplinary perspective. Many alternatives frequently are beyond the scope of the proponent of the action or the lead agency. For example, the proponent of a new power plant usually does not have the authority or the expertise to consider energy conservation or changing charges to power users as alternatives to meeting projected increases in energy demands. Because the identification of alternatives is best done as a multi-disciplinary exercise, inter-agency or public participation is often advocated. A number of manuals on impact assessment combine scoping and the identification of alternatives as one step because they can both be facilitated by interagency review and public participation.

Ensuring that the full range of reasonable alternatives is identified is important for at least two reasons. One reason is to find the option that ensures all possible joint gains have been secured. Impact assessment can be seen as a process for resolving conflicts among institutions or individuals with different interests. Identification of the full range of reasonable alternatives to resolve a conflict among competing interests means that no feasible options for maximizing benefits and minimizing costs have been missed.

The second reason why a full range of alternatives is important is to prevent costly delays. Failure to include a reasonable alternative often results in a challenge that the impact assessment is inadequate. Such a challenge can produce a significant delay, particularly if a new impact assessment must be conducted in order to expand the comparative analysis to include the new alternative.

The range of alternatives is larger at the higher levels of planning. By the time the project cycle or cone of decision gets to the project level, there is relatively little flexibility of alternative location, technology types or the design of the facility.

12.6.4.1 Estimation

Once the significant impacts have been determined by the scoping step, an estimation should be made of each impact. Making a quantitative or qualitative estimation of impact usually requires data and one or more predictive methods. Ideally the estimation would provide quantification of the following dimension:

a. *Probability of Occurrence.* To what extent is impact likely to occur? The probability can range from 1 – absolute certainty that the impact will occur – to an infinitely small percentage, such as the likelihood of coastal development being destroyed by a tsunami.

b. *Duration.* Will the impact occur just during a phase of the proposed action, such as soil erosion during the construction phase, or will the impact be permanent, such as the extinction of a species?

c. *Magnitude.* What will be the spatial dimensions of the impact? What will be the intensity of the impact, such as decibels of noise, parts per thousand turbidity, or the number of new employees?

d. *Social Distribution or Incidence.* What social groups and interests are positively or negatively affected by the impact?

12.6.4.2 Estimation Methods

In 1984, a survey of 140 impact studies drawn from a variety of different countries, but mainly North America and West Europe, identified 150 different estimation methods. In addition, 200 further estimation methods were identified by the investigation in a subsequent search of environmental evaluation literature. Evidently the number of estimation methods that might be used in IA is very large. However, critical and comparative reviews of estimation methods have been done for different types of impacts (e.g., impoundments, irrigation systems, roads), different types of environments (mangroves, coral reefs), and different types of impacts, such as air or water pollution. Estimation methods are usually classified according to type. The three most common types used in IA are statistical models, physical models and experimental methods.

Statistical models specify relationships between cause and effect in the form of one or more mathematical functions. They vary greatly in their complexity and information requirements. The simplest describe a fairly straightforward, direct relationship between input and output (e.g., emission and air quality, effluent and water quality) that is based upon the statistical analysis of the direct input-output relationship observed in similar environmental circumstances. More complex models also specify the environmental mechanisms through which inputs are processed and transformed into different impacts (e.g., a dynamic mathematical model of an estuary or floodplain).

Physical models are illustrative or working-scale models which have been built to replicate some component of the environment. For example, three dimensional models have been constructed of rivers, estuaries, and bays to predict their likely flooding dynamics or sediment impacts. Similarly, wind tunnels and wave chambers are used to help in predicting the impact of pollutant discharges on air and water quality.

Field and laboratory experimental methods are used to determine the number and type of environmental or socio-economic components that may be lost, disturbed or damaged by a project in circumstances where the magnitude of change in the environment cannot be reliably predicted. The data required to construct these measures may be derived by using either an existing inventory or the findings of a special survey.

Other estimation approaches exist that are less formalized than the three types described. Whether or not they are less useful can only be judged in the context in

which they are applied. One such approach is to identify similar (or analog) situations to that of the action being assessed. Through site visits and/or examination of the relevant literature, one deduces by analogy, what the action's impacts might be. Another common means of estimation is expert opinion, which is often used to predict both the likely scale and significance of an impact. This opinion may be based upon the use of any of the methods already described and, to this extent, is no different from those methods. However, expert opinion may also be in the form of a "judgment" in which the estimation is not expressed in quantitative terms nor is its method of derivation revealed. This is sufficiently different to be included as a different type of estimation method.

The surveys of IA practice that have been undertaken suggest that the estimation methods most frequently used are the less formal approaches, notably the use of expert opinion. This is particularly true in developing nations. Physical models, specially commissioned experiments, and mathematical models are not used frequently in impact assessment, particularly for small and medium scale projects. The apparent lack of sophistication in the use of estimation methods might be interpreted as evidence of serious deficiency in IA practice. However, there may be some justification for this state of affairs, if:

a. the requirements of much estimation work, particularly in the early stages of the IA process, can be satisfied using simpler methods;

b. the time and resource constraints imposed upon the operation of the IA process are too severe to enable the data collection and/or calibration, of the more complex methods to be completed.

The simplest methods are mainly used for such purposes as initial screening, comparisons of alternatives where the similarities and differences in impacts are very clear-cut, and preliminary estimation of impacts where limited project and base-line environmental data are available. By contrast, complex models are used in exceptional circumstances where there is concern by decision makers that particular activities may have major, irreversible environmental effects.

There are a number of general criteria to use when choosing between different estimation methods – appropriateness of the task to be undertaken, replication, consistency, and economy in the use of time and other resources. These may be elaborated into a fuller checklist for assessing the suitability of estimation methods.

Be clear at the outset about the nature of the information which is sought as output from the estimation method. What types of environmental impact are to be predicted. In what form? For whose use? On what assumptions? For what geographical area? For what duration and at what frequency? To what degree of accuracy?

Can the method be used to produce the information needed? If not, can it be modified to do this? Can the method be legitimately applied to the particular activity and environment under investigation and is it appropriate to the particular use to which it is to be put? Are any limitations of the method, and the assumptions implicit in its use, acceptable if used in this particular context?

Are the data inputs available which are needed to use the method? If not, can they be collected using the available resources of time, manpower, and equipment?

Are the resources available to apply the method? Include, for example, computing time, laboratory facilities, field study resources, skilled staff.

Are the outputs from the use of the method in a suitable form to use as inputs for estimation of higher-order impact (i.e., socio-economic values that may change)?

Does the method provide a sufficiently detailed and reliable estimation of the impact for the particular stage in the IA process in which it is being used?

Are the outputs from the use of the method likely to be credible and acceptable to decision-makers, as well as to scientists and other users? Can the outputs be presented in a form which is understandable and useful to the different users?

Failure to account for risk and uncertainty is one of the largest pitfalls of most estimation methods. Environmental and social systems are affected by random factors. Usually it is possible to estimate a range of probabilities for impacts because some key factors may be unknown. In the long run, the range of probabilities can be narrowed through scientific research or experimentation. Unfortunately impact assessors rarely have the luxury of gathering much original data.

The term risk has been used in many different ways. Confusion can arise when risk is defined as a compound measure of both probability and magnitude. We use risk to indicate an estimated probability. Uncertainty can never be eliminated because it is characteristic of dynamic systems and hydrology is very dynamic. Nevertheless, people commonly make decisions, often without explicitly analyzing their implications. Development decisions – particularly in developed nations – are commonly not completed until cautious scientists are able to reach unanimity. The best that impact assessment can do about uncertainty is to indicate the ranges of uncertainty about key causal relationships, estimate risks, and include a sensitivity analysis to explore the effects of the alternative assumptions of the estimation method.

12.6.4.3 Data

All estimation methods require data, and data requirements are particularly substantial to operate the complex models that are used to describe or predict functioning of hydrologic systems. Baseline models that describe both past and present functions of such systems are especially data-intensive. In order to accurately describe the long and short-term fluctuation, data must be periodically collected over the duration of the fluctuation period. For example, several years of rainfall and runoff data are needed to identify areas likely to be inundated during floods.

The task of obtaining data to assess impacts is a particularly difficult challenge in developing countries. The existing data base may be sparse, scattered, or of questionable validity. Often the data base does not include basic topographic maps or resource inventories such as soil surveys. Maps and aerial photographs may be restricted by military security. The cost of acquiring new data can be prohibitive. A related concern is the set of opportunity costs that may be associated with delaying the project in order to collect time series data. For example, it may not be either socially or economically possible to postpone the construction of an impoundment for several years in order to obtain reliable runoff data for forecasting floods.

It is also important, although understandably difficult, to avoid allowing available data banks to control the design and execution of the assessment. Data can be

TABLE 12.8
A Simplified Scoring-Weighting Comparison of Three Alternatives

Impact Category	Weight	Alternative A		Alternative B		Alternative C	
		Score	Score x Weight	Score	Score x Weight	Score	Score x Weight
Air Pollution	4	1	4	5	20	10	40
Noise	1	6	6	2	2	10	10
Landscape	5	4	20	4	20	7	35
Total			30		42		85

used to provide a list of environmental characteristics in the vicinity of a proposed development action. But merely a listing of existing flora, fauna, soils, geology, hydrology and land use has limited value to IA. Data should be used more profitably to define impact networks and to validate models of environmental or hydrologic systems. As an example of the later application, models of eutrophication require very specific hydrological, chemical and biological data; it is not easy to use observations collected for other purposes.

12.6.4.4 The Comparison of Alternatives

Once estimates have been made, practitioners are faced with information in all sorts of units of measurement referring to widely different systems. Some effects are easy to quantify, others are inevitably descriptive. The more elaborate the estimation step, the more difficult it is to synthesize the results.

There is an explicit trade-off between keeping estimated impacts in their original units of measure and presenting the information in a form that facilitates the selection among alternatives by decision-makers. A recurring tendency in IA is to aggregate all impacts into composite or unifying measurements such as a number or money. If every impact of each alternative can be added together to produce one number, the selection of alternatives simply becomes an exercise in picking the option with the best number.

In order to aggregate the scores into a single figure for each alternative, weighting must be placed on each of the impact types according to their relative degree of importance. Table 12.8 places the highest importance on landscape with a weighting of 5, then air pollution with 4 and noise with 1. It is obvious that weighting is largely an arbitrary exercise that depends on the subjective opinion of the individual or group that sets the weights.

Once the magnitudes of estimated effects are transformed into some unified measure, they tend to lose their practical meaning and the new numbers take on a significance of their own. Generally, reduction of all impacts to a single number usually degrades the information because it calls for subjective judgments about scaling and weighting the impact types. Furthermore, when impacts are transformed into weighted sums, it is tempting to perform statistical operations,

regardless of the levels of measurement used in scaling each estimation. These may be invalid and produce unsound and biased ranking of alternatives.

If impacts are aggregated and the implicit preferences and the judgment are to be valid, the scoring and weighting should represent the preferences of all affected by the action, not only the IA practitioners. Aggregation has been successfully used as a public participation technique. Individuals or groups with different interests can place their own scaling and weighting on the impacts to compare how their aggregation compares with others. It is desirable, but rarely seen, for judgment of impact significance to be expressed in terms of the different vantage points of those affected by the action.

There are a number of methods more sophisticated than the simple aggregation method just described. Most claim to take the subjective judgment out of assessment by developing numerical equivalents for the magnitude and significance impact. Others try to establish systematic and replicable methods for incorporating the judgment of experts or interested parties. Much has been written about these methods and their relative attractions or faults. In fact, both the simple and more complex aggregation methods are used rarely and most are impractical for the requirements of project planning and solving environmental and socio-economic problems. When choosing between alternative sites, an aggregation method can be used to structure choice, but rarely is it prudent to convert multi-dimensional impacts into a composite measure.

Matrices without aggregated sums can be useful as ways of displaying information about comparative effects. Indeed this is the most commonly used method for assembling and relatively comparing impacts for judgment of significance. The impacts are not aggregated but just given excellent, good, fair, poor and very poor ratings. In choosing between alternatives or in identifying problems that need attention in project design, it is often few key factors that make the difference about which decision-makers need the full information provided by estimation and evaluation.

12.6.4.5 Comparison of Impacts to Adopted Plans or Standards

A relatively easy means of evaluation is to compare the estimated impacts with adopted plans or standards. Plan compliance was previously described as a means of screening development actions. The use of adopted plans, however, to evaluate impacts is limited by two factors: the specificity of the plan in respect to the estimated impacts, and the extent to which there is strong support for the plan. Many environmental plans are based on a set of vague policies. For example, discharges from land uses adjacent to Turtle Bay should not significantly degrade mangrove systems of Turtle Bay. The policy needs to include measurable standards of degradation, such as "effluent discharges into Turtle Bay shall not be in excess of x ppt of phosphorous, y ppt nitrogen, and z biochemical oxygen demand." Specific impacts cannot be evaluated against vague policies.

Public support is the other problem commonly associated with using plans to evaluate impacts. Many plans are just documents and maps that were created by professional staff without broad-based participation from the interest groups that will be affected by implementation of the plan. The interest groups that support

the proposed development action often find little or no opposition from supporters of the plan. The development proposal is the real world of tangible socioeconomic benefits and the plans are often in the dream world of imagined benefits. In a conflict between a plan and the impacts of a development action, the impacts often prevail and the plan is revised.

12.6.4.6 Incidence of Impact

The evaluation step should also determine who gains and who loses as a result of an action. To put this another way, every physical effect has socio-economic consequences that in themselves are impacts to be concerned with, or are usually distributed unevenly among groups in the population and thus affect equity. Whether or not one defines IA as incorporating the assessment of socio-economic and cultural effects is hardly important in the end because it is the consequences for human development, income generation, and long-term productivity that should be the factors considered by decision makers. The key point here is whether the expected socio-economic and distributive consequences are part of the IA process. If they are not, are they taken into consideration in the ultimate decisions about the proposed action?

Conventional project appraisal is focused on the expected benefits of a project and sometimes takes into account the distribution of the benefits, applying some preferred social welfare function that may weigh benefits to poorer groups higher. IA often starts from the assumption that in trying to achieve socio-economic benefits, there may be some undesirable distributive effects unless a more comprehensive, systematic appraisal is followed. There is also a presumption that in most cases the population groups in developing nations most likely to suffer the adverse costs do not control the decision-making process. Usually they have relatively little influence on the decision making process.

12.6.5 PRESENTATION

The objective of most impact assessments is to ensure that environmental and socio-economic problems are foreseen and addressed by decision makers. To achieve this, decision-makers must fully understand the IA's conclusions. Lack of sufficient attention to the presentation and communication of findings has been one of the reasons for the limited usefulness of impact assessments and impact statements. Most decision-makers are unlikely to use information, no matter how important it is, unless it is presented in formats that are immediately meaningful. Some ways to improve the presentation are:

a. Briefly present "hard" facts and estimations about impacts. Comment on the reliability of this information, and summarize the consequences of each of the proposed options.

b. Write in the terminology and vocabulary that is used by decision makers and the community affected by the action.

c. Present the essential findings in a concise document, supported by separate background materials if necessary.

d. Make the document easy to use, providing graphics when possible such as diagrams, tables, photographs, and maps.

Whatever the process of analysis, assessment of significance, and economic evaluation, decision-makers need to know what the serious problems are, how they can be solved, and how much it will cost. In addition, it should be clear how this information was obtained, what the basis for judgment was, and whose vantage point it represents. Where significant impacts are highlighted, it is preferable to express them in their original units of measurement and to suggest practical measures for mitigation and management.

If the action is complex and has a broad array of impacts, the presentation should contain a summary, as well as a complete record of the analysis. Pictures, graphs and maps should be used to simplify and summarize important information. For example, information on the impacts of particular development opinions can be displayed on a matrix that summarizes relevant information. More detailed descriptions of likely effects are appended to the summary. Care should be taken not to degrade the information in summarizing it, to indicate levels of certainty, and the differential effects on groups of the population. The crucial factor is to provide the decision-maker with the information needed to balance the choice involved in selecting an alternative or approving a proposal.

12.6.6 Mitigation of Adverse Impacts

Mitigation has a variety of meanings in the IA literature. It can mean avoiding the impact altogether by not undertaking a certain development action. The term is used more commonly to refer to either eliminating adverse impacts or reducing them to acceptable levels by taking one or more of the following steps;

a. limiting the degree of magnitude of the development action and its implementation, such as reducing the size of area that will be converted from mangroves to aquaculture ponds;

b. changing the design of the development action (for instance having a road elevated on pilings over a wetland instead of a solid fill road base);

c. changing the location of the action (this could mean either another location on the site or different site);

d. compensation for the impact by repairing, rehabilitating, or restoring the environmental or cultural system (for example, restoring degraded wetlands adjacent to wetlands that will be filled by a development action, such as a tourist resort);

e. compensation provided by replacing or providing substitute resources or environments (e.g., fisherman are provided with affordable dock space in port development to compensate them for the loss of beach landing area taken over by a resort development);

f. preservation and maintenance operations during the life of the action, such as setting aside habitats that will continually protect rare and endangered species);

g. restoring the environment to its former condition after the development action is completed (the reconstruction of dunes and the planting of vegetation after beach mining activities are terminated); and,

h. relocation or recording cultural and environmental features before the site is developed (for example relocating historic buildings or recording cultural practices on film and tape).

12.6.7 Review and Comment

One common distinction between large- and small-scale projects is the requirement for external review of the impact assessment. Almost all legislation or executive orders require a draft statement be circulated to affected units of government for review and comment. Many EIS procedures in developed countries also require the draft statement be broadly circulated to members of the public whose interests may be affected by the proposed action. By contrast, impact assessments of actions that do not require or merit an EIS usually are not circulated outside the lead agency for review and comment and particularly not to the public. Many or most projects could benefit from external review. In that respect the review steps are very similar to the scoping step since impacts are multi-disciplinary in nature, and the review of an impact assessment should be multidisciplinary. If the lead agency has the multi-disciplinary expertise, then external review might produce little additional information. Also the review process takes time (usually at least thirty days) and usually imposes additional costs by delaying initiation of an action.

The lack of early liaison between an agency that is considering a development action and institutions or individuals that have jurisdiction or interest relating to potential impacts is likely to lead to one or more of the following situations:

a. the range of alternative actions will be unduly restricted;

b. key environmental data and information that may already be available in one agency will be overlooked and/or duplicated by another agency;

c. the assessment of cumulative impacts will be incomplete due to the failure to learn of projects and programs that are under consideration by other agencies;

d. assessments completed by other agencies for similar projects or for projects in the same or similar local areas will be overlooked; and,

e. the potential for cooperative actions among different agencies with respect to collecting base-line data, monitoring impacts, and designing and implementing ameliorative measures will be minimized.

Many analysts of impact assessment believe that public review of assessment efforts should begin in the earliest phase of action planning and continue throughout the development phase. These analysts argue that there are two rationales for

public participation: efficiency and legitimacy. The efficiency rationale suggests that participation will result in better designed government services and a greater ability to gain the support of diverse groups in fulfilling policies after their adoption. The legitimacy rationale sees participation as a way to increase people's confidence in officials and make government more stable and accountable. Government administrators usually accept efficiency more readily than legitimacy as a rationale.

On the other hand, lack of public participation often has the following effects:

a. increasing alienation among the public at large due to the public's perception of its powerlessness to contribute meaningfully in key phases of environmental decision-making by governmental agencies; and,

b. increasing alienation among the public as the perception grows that some interest groups and individuals have more influence in decision-making than other groups and individuals.

12.6.8 THE DECISION

The impact assessment process is designed to improve decision making. One distinction that can be made among governments that practice IA is the extent to which decision making is a "professionalised" process or an interactive process. A "professionalized" decision making process involves only mid- to high-level bureaucrats with technical expertise (technocrats) to analyze the problems and make decisions. The participation of top level bureaucrats and elected officials is limited to getting occasional feedback from the decisions made by the technocrats. Non-governmental groups that have a vested interest in a decision are excluded from the process.

The interactive decision making process has all three groups in communication with each other. The elected officials and top level bureaucrats make the decisions after consultation with the technocrats and non-government interests. Problems are continuously monitored by all three groups.

One of the major objectives of initiating an impact assessment process has been to open up the decision making process – particularly to those affected by the proposed action. The interactive decision making process is the one usually mandated by laws that establish an EIA requirement.

Another distinction among institutions that practice IA is the extent to which the final analysis and recommendations are the decision document. Many governments do not use the impact assessment as the decision document, but as one input to be considered along with other inputs such as benefit-cost analysis, economic feasibility studies, and market analysis.

How the IA is used in the decision depends on at least six factors:

a. How well the information is communicated, and whether the decision maker understands the IA conclusions and recommendations;

b. Are there significant adverse impacts, and if so, can these impacts be adequately mitigated – at least to the satisfaction of those whose interests will be adversely affected by the impact? How politically influential are those whose interests will be adversely affected?

c. Are the IA conclusions supported by those who will benefit from the impacts of the action? How politically influential are those who will benefit from the impacts of the action?

d. Is there sufficient information to reach conclusions about the probabilities of impact, the degree of uncertainty, and the amount of risk, or is there insufficient information to reach conclusions about critical potential impacts?

e. Is there a consensus among those with expertise on each of the significant impacts that the assessment was based on good data and used the appropriate estimation methods?

f. Are the IA conclusions supported by other analysis such as benefit-cost analysis and economic feasibility studies?

To the extent that all of the above factors are positive, the action will probably be approved. Conversely, to the extent that all the factors are negative, the action will probably be denied or stringent mitigation measures will be imposed. Of course, decision makers have a number of options in addition to simple denial or approval. These include:

deny the action with conditions for reconsideration;
approve the action with mitigation measures;
defer decision on the action until further information is obtained;
defer decision until a better IA is conducted; or
defer decision until a plan or programmatic IA is prepared.

For most proposed development projects it is not a question of whether the adverse impacts will warrant denial. Rather, the two important questions are: whether mitigation of adverse impacts will be economically feasible and whether indirect (or secondary) impacts can be adequately controlled. On the latter point, in developing nations it is common for IA not to address the indirect or secondary consequences of a development action, such as the changes in land use and land tenure that will result from the construction of a new highway or new power plant.

Existing conflicts will be mitigated if it can be shown that the mitigation measures will generate more socio-economic benefits than costs. Potential conflicts will be mitigated if the proposed project can support the additional costs. Ideally, however, the rate and magnitude of adverse impacts will continue to decrease over time as government agencies become more familiar with systems planning, develop their data bases, and build economies that enable consideration of longer time horizons in policy making, as well as the consideration of non utilitarian benefits.

12.6.9 Auditing or Monitoring Impacts

Auditing in the context of IA usually refers to monitoring the actual impacts of the action undertaken. The term, monitoring, has also been used to describe this step. Moreover, monitoring is sometimes used as a term to mean a critical assessment of the entire IA program

Government units involved in the IA process should establish an auditing step both to ensure that their decisions are carried out according to the conditions of the decision and to determine the actual impacts of the action as implemented. In fact auditing can serve six purposes:

a. to ascertain that mitigation measures being implemented as agreed;

b. to warn agencies of unanticipated adverse impacts or sudden changes in impact trends;

c. to provide immediate warning whenever a preselected impact indicator approaches a preselected critical level (this will also be the means to ensure that the legal standard for effluents are not exceeded);

d. to provide information that could be used by agencies to control the timing, location, and level of impacts of a project. Control measures involve preliminary planning as well as the possible implementation of regulation and enforcement measures. If an intergovernmental monitoring system is used, it would facilitate appropriate response measures;

e. to provide information that could be used for evaluating the effectiveness of implemented mitigation measures; and,

f. to provide information that could be used to verify estimated impacts and thus validate impact prediction techniques. Based on these findings the techniques, e.g., statistical models, could be modified or adjusted as appropriate.

Despite the six purposes that can be served by auditing impacts, most IA programs do not include this step as a routine procedure. Usually, the IA administrative agency can not afford to bear the costs and provide the expertise needed to conduct audits.

12.7 Conclusion

The hydrologic cycle has a great effect on landscapes and ecosystems over human time scales. It sculpts the land and helps determine the type of ecosystem that will exist on that land. For these reasons it is important that environmental planners and hydrologists recognize the interactions between hydrology and the ecosystem and do what they can to maintain and enhance both systems. This chapter provides concepts from *Comparative Hydrology* that seek to link particular landforms with the most relevant hydrologic processes, thus providing a focus for hydrologic planners. It also provides impact information for hydrologic processes operating within particular landforms and suggests management implications.

A planning process is presented in overview format followed by a detailed impact assessment procedure. The process will allow a planner to draft an environmental impact statement that should be adequate for most planning endeavors. Very large projects or plans will require more information than this chapter can provide and the reader is referred to other sources.

References

Comparative Hydrology. 1989. Falkenmark, M. and Chapman, T. (eds.). UNESCO, Paris.

Costanza, Robert 1992. Toward an Operational Definition of Ecosystem Health, in Ecosystem Health, New Goals for Environmental Management, ed. by Costanza, R., Norton, B.G. and Haskell, B.D., Island Press, Wash. D.C.

Day, J. W., Jr. and Templet, P. H. 1989. 'Consequences of Sea Level Rise: Implications from the Mississippi Delta,' Coastal Management, Vol 17, pp 241-257, 1989.

Dunne, T. and Leopold L. B. 1978, Water in Environmental Planning, W.H. Freeman and Co. San Francisco., p.18.

Helweg, Otto J. 1985. Water Resources Planning and Management. John Wiley and Sons, New York, NY.

Leopold, L.B., Clarke, F.E., Hanshaw, B.B. and Balsley, J.R. 1971. A Procedure for Evaluating Environmental Impact. Geological Survey Circular 645. U.S. Dept. Interior, Wash. D.C.

Postel, S. 1993. 'Water Scarcity Spreading', in Vital Signs 1993, ed. by Lester R. Brown, WorldWatch Institute; W.W. Norton and Co. , NY,NY, pp106,7.

Spaling H. and Smit B. 1993. 'Cumulative Environmental Change: Conceptual Frameworks, Evaluation Approaches, and Institutional Perspectives,' Environmental Management, Vol. 17, #5, Sept./Oct. 1993, pp. 587-600.

Templet, P.H. and Meyer-Arendt, K. 1988. 'Louisiana Wetland Loss: A Regional Water Management Approach to the Problem.' Environmental Management, Vol 12, #2, pp 181-192, 1988.

Whitman, .L., Dee, N., McGinnis, J.T., Fahringer, D.C. and Baker, J.K. 1971. Design of an Environmental Evaluation System. Rep. PB-201 743. Battelle Columbus, Columbus, Ohio.

Westman, W.E. 1985. Ecology, Impact Assessment, and Environmental Planning, A Wiley-Interscience publication, John Wiley and Sons, New York, NY.

Glossary of terms

Glossary of terms for Chapter 2

Accuracy: the difference between the approximate solution obtained using a numerical model and the exact solution of the governing equations, divided by the exact solution.

Acid anion: negatively charged ion that does not react with hydrogen ion in the pH range of most natural waters.

Acid-base chemistry: the reaction of acids (proton donors) with bases (proton acceptors). This means the reactions of natural and anthropogenic acids and bases, the result of which is described in terms of pH and acid neutralizing capacity of the system.

Acid cation: hydrogen ion or metal ion that can hydrolyse water to produce hydrogen ions, e.g., ionic forms of aluminium, manganese, and iron.

Acid neutralizing capacity (ANC): the equivalent capacity of a solution to neutralize strong acids. The components of ANC include weak bases (carbonate species, dissociated organic acids, aluminohydroxides, borates, and silicates) and strong bases (primarily, OH-). ANC was measured by the Gran titration procedure.

Acidic deposition: transfer of acids and acidifying compounds from the atmosphere to terrestrial and aquatic environments via rain, snow, sleet, hail, cloud droplets, particles, and gas exchange.

Acidic episode: an episode in a water body in which acidification of surface water to an acid neutralizing capacity less than or equal to 0 occurs.

Acid lake or stream: a lake or stream in which the acid neutralizing capacity is less than or equal to 0.

Acidification: the decrease of acid neutralizing capacity in water or base saturation in soil caused by natural or anthropogenic processes.

Acidified: pertaining to a natural water that has experienced a decrease in acid neutralizing capacity or a soil that has experienced a reduction in base saturation.

Alkalinity: The equivalent sum of $HCO_3^- + CO_2^{-3} + OH^-$ minus H^+, i.e., buffering conferred by the bicarbonate system; the terms ANC and alkalinity are some times used interchangeably. ANC includes alkalinity plus additional buffering from dissociated organic acids and other compounds.

Base cation: an alkali or alkaline earth metal cation (Ca_2^+, Mg_2^+, K^+, Na^+, etc).

Base cation buffering: the capacity of a watershed soil or a sediment to supply base cations (Ca_2^+, Mg_2^+, K^+, Na^+) to receiving surface waters in exchange for acid cations (H^+, Al3+); may occur through cation exchange in soils or weathering of soil or bedrock minerals.

Base saturation: the proportion of total soil cation exchange capacity that is occupied by exchangeable base cations, i.e., by Ca_2^+, Mg_2^+, K^+, and Na^+.

Buffer intensity: the change in pH per unit addition of strong acids or strong bases.

Cation: a positively charged ion.

Cation exchange: the interchange between a cation in solution and another cation on the surface of any surface-active material such as clay or organic material.

Cation exchange capacity: the sum total of exchangeable cations that a soil can adsorb.

Cation leaching: movement of cations out of soil, in conjunction with mobile anions in soil solution.

Cation retention: the physical, biological, and geochemical processes by which cations in watersheds are held, retained, or prevented from reaching receiving surface waters.

Chelator: a class of compounds, organic and inorganic, that can bind metal ions and change their acid-base chemistry and biological availability.

Circumneutral: close to neutrality with respect to pH (neutral pH = 7); in natural waters, pH 6–8.

Conceptual model: simplified or symbolic representation of prototype or system behaviour and responses.

Decomposition: the microbially mediated reaction that converts solid of dissolved organic matter into its constituents.

Denitrification: biologically mediated conversion of nitrate to gaseous forms of nitrogen (N_2, NO, N_2O); denitrification occurs during decomposition of organic matter.

Deterministic model: mathematical model in which the behaviour of every variable is completely determined by the governing equations.

Detritus: dead and decaying organic matter originating from plants and animals.

Dimensional analysis: an analysis based on the requirement that all functional relationships (for the model and prototype) be dimensionally homogeneous.

Dissolved inorganic carbon: the sum of dissolved carbonic acid, bicarbonate, and carbonate in water sample.

Dissolved organic carbon: organic carbon that is dissolved or unfilterable in a water sample.

Dry deposition: transfer of substances from the atmosphere to terrestrial and aquatic environments via gravitational settling of large particles and turbulent transfer of trace gasses and small particles.

Dynamic model: a mathematical model in which time is included as an independent variable.

Empirical model: representation of a real system by a mathematical description based on experimental or observational data.

Episodes: a subset of hydrological phenomena known as events. Episodes, driven by rainfall or snowmelt, occur when acidification takes place during a hydrologic event. Changes in other chemical parameters, such as aluminium and calcium, are frequently associated with episodes.

Episodic acidification: the short term decrease of acid neutralizing capacity from a lake or stream. This process has a time scale of hours to weeks and is usually associated with hydrological events.

Equivalent: unit of ionic concentration, a mole of charge; the quantity of a substance that either gains or loses one mole of protons or electrons.

Evapotranspiration: the process by which water is returned to the air through direct evaporation or transpiration by vegetation.

Filtering: procedure of removing or smoothing undesired perturbations in a modelling context.

Forecast: to estimate the probability of some future event or condition as a result of rational study and analysis of available data.

Gran analysis: a mathematical procedure used to determine the equivalence points of a titration curve for acid neutralizing capacity.

Groundwater: water in a saturated zone within soil or rock.

Heuristic model: representation of a real system based on a reasoned, but unproven, argument.

Hindcast: to estimate the probability of some past event or condition as a result of rational study and analysis of available data.

Hybrid model: mathematical model with elements of more than one type of model.

Hydraulic residence time: a measure of the average amount of time water is retained in a lake basin. It can be defined on the basis of inflow/lake volume, represented as 'RT', or on the basis of outflow (outflow/lake

volume) and represented as Tw. The two definitions yield similar values for fast-flushing lakes, but diverge substantially for long-residence time seepage lakes.

Hydrological event: pertaining to increased water flow or discharge resulting from rainfall or snowmelt.

Hydrological flow paths: surface and subsurface routes by which water travels from where it is deposited by precipitation to where it drains from a watershed.

Initial conditions: given values of dependent variables or relationships between dependent and independent variables at the initiation of a computation.

Inorganic aluminum: the sum of free aluminum ions (Al3+) and dissolved aluminum bound to inorganic ligands, operationally defined by labile monomeric aluminum.

Intensity effect: a change in the concentration of solutes in soil solution or other natural water, or in the case of H^+, the solution pH.

Ionic strength: a measure of the interionic effect resulting primarily from electrical attractions and repulsions between various ions. I = 1/2 CiZ2i for all ions I (cations and anions), where Ci is the concentration of ion i and Zi is the valence.

Linear model: mathematical model based entirely on linear equations.

Littoral zone: the shallow, near shore region of a body of water, often defined as the band from the shoreline to the outer edge of the occurrence of rooted vegetation.

Long term acidification: the decrease of acid neutralizing capacity in a lake or stream on a seasonal or longer time frame.

Macropore flow: flow of water through large pores or voids in soil or rock (macropores), in response to the force of gravity.

Mathematical model: model using mathematical relationships to represent the prototype.

Matrix flow: flow of water through micropores or mesopores in soil or rock, in response to gravitational and capillary forces.

Mechanistic model: mathematical model in which actual ecosystem processes are explicitly included.

Mineral acids: inorganic acids, e.g., $H2SO_4$, HNO_3, HCL, H_2CO_3.

Mineral weathering: dissolution of rocks and minerals by chemical and physical processes.

Mobile ions: anions that flow in solutions through watershed soils, wetlands, streams, or lakes without being adsorbed or retained through physical, biological, or geochemical processes.

Model: an abstraction or representation of a prototype or system, generally on a smaller scale.

Natural acids: acids produced within terrestrial or aquatic systems through natural, biological, and geochemical processes; i.e. not a result of acidic deposition or deposition of acid precursors.

Nitrification: oxidation of ammonium to nitrate or nitrite by microorganisms; a by-product of this reaction is H^+.

Nitrogen fixation: biological conversions of elemental nitrogen (N_2) to organic N.

Nonlinear model: mathematical model using one or more nonlinear equations.

Numerical model: mathematical model in which the governing equations are not solved analytically. The governing equations are solved approximately, using discrete numerical values to represent the variables involved and using arithmetic operations.

Nutrient cycling: the movement or transfer of chemicals required for biological maintenance or growth among components of the ecosystem by physical, chemical, or biological processes.

Organic acids: acids possessing a carboxyl (–COOH) group or phenolic (C–OH) group; includes fulvic and humic acids.

Organic aluminium: aluminum bound to organic matter, operationally defined as that fraction of aluminum determined by colorimetry after sample is passed through a strong cation exchange column.

Parameter: (a) a characteristic factor that remains at a constant value during the analysis, or(b) a quantity that describes a statistical population attribute.

pH: the negative logarithm of the hydrogen ion activity. The pH scale runs from 1 (most acidic) to 14 (most alkaline); a difference of one pH unit indicates a tenfold change in hydrogen ion activity.

Physiography: the study of the genesis and evolution of land forms; a description of the elevation, slope, and aspect of a study area.

Precision: a measure of the capacity of a method to provide reproducible measurements of a particular analysis (often represented by variance).

Predict: to estimate some current or future condition on the basis of historical or current observation, experience, or scientific reason.

Probabilistic model: mathematical model in which the behaviour of one or more of the variables is either completely or partially described by equations of probability.

Reduction - oxidation (redox) reactions: reactions in which substances gain or lose electrons, i.e., in which substances are converted from an oxidized to a reduced oxidation state and vice versa.

Regolith: the layer of loose, weathered rock material that forms the surface of the land and rests on the bedrock.

Respiration: the biological oxidation of organic carbon with concomitant reduction of external oxidant and the production of energy. In anaerobic respiration, O_2 is reduced to CO_2. Anaerobic respiration processes utilize NO_3^- (denitrification), sulphate (sulphate reduction), or CO_2.

Retention time: the estimated mean time (usually expressed in years) that water resides in a lake prior to leaving the system. (See hydraulic residence time).

Roundoff error: cumulative error introduced by rounding results from individual arithmetic operations.

Salt effect: the process by which hydrogen ions are displaced from the soil exchange complex by base cations (from neutral salts). The result is a short-term increase in the acidity of associated water; also referred to as sea-salt effect.

Saturated flow: flow of water through the voids in rock or soil at a pressure greater than atmospheric, i.e., under a head of pressure.

Scenario: one possible deposition sequence following implementation of a control or mitigation strategy and the subsequent effects associated with this deposition sequence.

Sensitive: tending to experience capacity effects or intensity effects, as a result of changes in acidic deposition; or biological effects as a result of a change in acid-base chemistry.

Sensitivity: the tendency of a system to experience capacity effects or intensity effects as a result of changes in acidic deposition, or biological effects as a result of a change in acid-base chemistry; a measure of the degree to which a model's output responds to small changes in an input value.

Sensitivity analysis: test of a model in which the value of a single variable or parameter is changed, and the impact of this change on the dependent variable is observed.

Simulation: description of a prototype or system response to different conditions or inputs using a model rather than actually observing the response to the conditions or inputs.

Simulation model: mathematical model that is used with actual or synthetic input data, or both, to produce long-term time series or predictions.

Specific conductance: the conductivity between two plates with an area of 1cm2 across a distance of 1 cm at 250C.

Steady state: the condition that occurs when a property (e.g. mass, volume, concentration) of a system does not change with time. This condition requires that sources and sinks of the property are in balance (e.g., inputs equal outputs; production equals consumption)

Steady state model: a model in which the variables under investigation are assumed to reach equilibrium and are independent of time.

Stochastic: an occurrence that is probabilistic in nature.

Stochastic model: see probabilistic model.

Stream order: a method of categorizing streams based on their position in the drainage network. First-order streams are permanent streams with no permanent tributaries. Higher order streams are formed by the confluence of two or more streams of the next lower stream order.

Strong acid anion sum (SAA or CA): refers to the equivalent sum of SO_4^{2-}, NO_3^-, Cl^-, and F-. The term specifically excludes organic acid anions.

Strong acids: acids with a high tendency to donate protons or to completely dissociate in natural waters, e.g. NaOH.

Sulphate adsorption: the process by which sulphate is chemically exchanged (for OH) or adsorbed onto positively charged sites on the soil matrix; under some conditions this process is reversible, and the sulphate may be desorbed.

Sulphate reduction: (1) the conversion of sulphate to sulphide during the decomposition of organic matter under anaerobic conditions (dissimilatory sulphate reduction) and (2) the formation of organic compounds containing reduced sulphur compounds (assimilatory sulphate reduction).

Sulphate retention: the physical, biological, and geochemical processes by which sulphate in watersheds is held, retained, or prevented from reaching receiving surface waters.

Sum of base cations: (SBC or CB)- refers to equivalent sum of Ca_2^+, Mg_2^+, Na^+, and K^+. The term specifically excludes cations Aln+ and Mn2+.

Surficial geology: characteristics of the earth's surface, especially consisting of unconsolidated residual, colluvial, alluvial, or glacial deposits lying on the bedrock.

Synoptic: relating to or displaying conditions as they exist at a point in time over a broad area.

Systematic error: a consistent error introduced in the measuring process. Such error commonly results in biased estimations.

Theoretical model: representation of a real system by a mathematical description based on general physical laws rather than experimental data.

Total monomeric aluminum: the simple unpolymerized form of aluminum present in inorganic or organic complexes.

Turnover: the interval of time in which the density stratification of a lake is disrupted by seasonal temperature variation, resulting in entire water mass becoming mixed.

Uncertainity analysis: the quantification of error or uncertainity in results-analytical or simulated. This quantification might or might not partition total error into its constituents such as intrinsic natural variability, parameter error, input error, model error, or other error sources.

Unsaturated flow: flow of water through the voids in rock or soil at a pressure less than atmospheric. i.e. flow in response to gravitational and capillary forces through the unsaturated zone between the land surface and the water table.

Validation: comparison of model results with a set of prototype data not used for validation. Comparison includes the following: (1) using a data set very similar to the verification data to determine the validity of the model under conditions for which it was designed; (2) using a data set quite different from the verfication data to determine the validity of the model under conditions for which it was not designed but could possibly be used; and (3) using post-construction prototype data to determine the validity of the predictions based on model results.

Variable: a quantity that may assume any one of a set of values during analysis.

Verfication: check of the behavior of an adjusted model against a set of prototype conditions.

Watershed: the geographic area from which surface water drains into a particular lake or point along a stream.

Weak acids: acids with a low proton-donating tendency that tend to dissociate only partially in natural waters, e.g. H_2CO_3, H_4SiO_4, and most organic acids.

Weak bases: bases with a low proton-accepting tenedency that tend to dissociate only partially in natural waters, e.g. HCO_3^-, $Al(OH)_4$.

Wet deposition: transfer of substances from the atmosphere to terrestrial and aquatic environments via precipitation, e.g. rain, snow, sleet, hail, and cloud droplets. Droplet deposition is sometimes referred to as occult deposition.

Glossary of terms for Chapter 8

Advection: The process by which contaminants are transported by the bulk motion of flowing ground water.

Analytical model: A mathematical model whose solution can be obtained using common analytical methods. Usually simplifying assumptions are required.

Anisotropic: A term used to describe the property at a specific point in a geologic formation where hydraulic conductivity is dependent on the direction of measurement.

Aquiclude: A formation which, although porous and capable of absorbing water, does not transmit it at rates sufficient to furnish an appreciable supply for a well or spring.

Aquifer: A water bearing formation that provides a ground water reservoir; a formation, a group of formations, or part of a formation that contains sufficient saturated permeable material to yield significant quantities of water to wells and springs.

Aquitard: A term used to describe the less permeable beds in a stratigraphic sequence. These beds may be permeable enough to transmit water in quantities that are significant in the study of regional ground water flow, but their permeability is not sufficient to allow the completion of production wells within them.

Area of influence of a well: The area surrounding a well within which the piezometric surface has been lowered when pumping has produced the maximum steady rate of flow.

Artesian: Synonymous with confined. Artesian water and artesian water body are equivalent respectively to confined ground water and confined water body.

Attenuation: Attenuation is the retardation or fixation of contamination by or to the porous medium. It includes the processes of mechanical filtration, microbial decomposition, chemical precipitation and coprecipitation, sorption, and gaseous exchange.

Attenuation isotherm: An attenuation isotherm is a graphical representation of the relationship between attenuated and solution phases. It is dependent both on the nature of the porous medium and the solute.

Average linear velocity: The velocity in a specified direction whose magnitude is equal to (1) the volumetric flux of a specified cross sectional area divided by (2) the void area for that specified cross section. See seepage velocity.

Boundary conditions: Information required to obtain a unique solution of a model's governing equation. This information concerns the physical state of the process and includes the geometry of the boundary as well as the values of the dependent variable at the boundary of its derivative normal to the boundary.

Capillary fringe: The lower subdivision of the zone of aeration, immediately above the water table in which the interstices are filled with water under pressure less than that of the atmosphere, above it by surface tension. Its upper boundary with the intermediate belt is indistinct but is sometimes defined arbitrarily as the level at which 50 percent of the interstices are filled with water.

Cation exchange capacity: Cation exchange capacity of a soil is the amount of exchangeable cations in meg per gram of solid material measured at pH 7.0. Its value is dependent on temperature, the type and concentration of ions present, and the type of soil present. Clays and colloids have high cation exchange capacities.

Coefficient of hydrodynamic dispersion : The term used in contaminant transport modelling for simulating the dispersion process. Two components are included in this term: one concerns the characteristic properties

of the porous medium (dispersivity), the other deals with the process of molecular diffusion (coefficient of molecular diffusion).

Coefficient of molecular diffusion: A component of the dispersion coefficient which deals with the effects of the molecular diffusion process on contaminant transport.

Conceptual model: Formulation of the cause- effect relationships which govern how a system operates.

Cone of depression: The depression, in the shape of an inverted cone, of the piezometric (potentiometric) surface of a body of ground water which defines the area of influence of a well.

Convection: See advection.

Darcy's law: A derived formula for the flow of fluids through porous media based on the assumptions that the flow is laminar and that inertia can be neglected.

Deterministic models: Models which define cause and effect relationships based on an understanding of the physical systems.

Diffusion: Molecular diffusion is the spreading out of solutes due to a concentration gradient.

Discharge: Discharge is that water which is discharged from the dynamic ground water flow system by such means as baseflow, springs, seepage areas, and evapotranspiration.

Discharge area: A discharge area is an area where the direction of ground water flow is toward the watertable. The hydraulic head increases with depth.

Discharge velocity: The rate of discharge of water through a porous medium per unit of total area perpendicular to the direction of flow.

Dispersion, hydrodynamic: The organized spreading or distribution during advective transport of contaminants in ground water systems which results from the inherent heterogeneity of soil matrix geometry of all natural soil systems and the thermal-kinetic energy of the contaminant material. This term reflects the cumulative effects of microscopic dispersion and macroscopic dispersion and molecular diffusion.

Dispersion, macroscopic: Hydrodynamic dispersion resulting from the effect of small soil bodies or lenses present in larger soil bodies of significantly different hydraulic conductivity.

Dispersivity: A characteristic property of the porous medium which is used in the determination of the coefficient of hydrodynamic dispersion.

Distribution coefficient: A term which is a component of the retardation factor and is defined as the ratio of (1) the mass of solute on the solid phase per unit mass of solid phase to (2) the concentration of solute in solution.

Drawdown: The vertical distance to which the free water elevation is lowered or the reduction of the pressure head due to the removal of free water.

Effective porosity: Effective porosity refers to the amount of interconnected pore space available to fluid transmission..

Error, conceptual: Error which results from the application of a valid model to an inappropriate field situation.

Error, Data: Error which results from the inability to accurately or completely determine aquifer description data such as transmissivities, storage coefficients and distribution heads.

Error, Round-Off: Error which results from the finite accuracy of computer calculations.

Error, Truncation: Error which results from replacement of model differential equations by a set of algebraic equations. The exact solution of the algebraic equations differs somewhat from the solution of the original differential equations.

Finite difference method: A numerical method used for solution of model governing equations by a set of algebraic equations. The exact solution of the algebraic equations differs somewhat from the solution of the original differential equations.

Finite Element Method: A numerical method used for solution of model governing equations which is based on discretization of the continuum and results in an approximation of actual circumstances. Under certain conditions application of this method has advantages over the finite difference method.

Flow Line: The path that a particle of water follows in its course of seepage under laminar flow conditions.

Flow net: A graphical representation of flow lines and equipotential lines used in the study of seepage phenomena.

Flow, Steady: Steady flow occurs when at any point the magnitude and direction of the specific discharge are constant in time.

Flow, Uniform: A property is uniform if, at a given instant, it is the same at every point. Thus, uniform flow occurs if at every point the specific discharge has the same magnitude and direction.

Flow, Unsteady: Occurs when at any point the magnitude or direction of the specific discharge changes with time. The word transient is used in reference to the temporary features of unsteady flow.

Flow Velocity: The volume of water transferred per unit of time and per unit of area normal to the direction of the net flow.

Fluid Potential: The mechanical energy per unit mass of a fluid at any given point in space and time with regard to an arbitrary state and datum. Loss of fluid potential incurs as the fluid moves from a region of high potential to one of low potential.

Fluidity: Used in conjunction with the term hydraulic conductivity. An attribute of a fluid which incorporates the properties of density and viscosity.

Free water: Water that is free to move through a soil or rock mass under the influence of gravity.

Free Water Elevation: (water table, ground water surface, free water surface, ground water elevation) Elevations at which the pressure in the water is zero in relation to the atmospheric pressure.

Gaseous exchange: Certain volatile pollutants and decomposition products can escape to the atmosphere.

Governing Equations: The central mathematical equation of a ground water flow or solute transport model which includes all the aspects of the conceptual model.

Gravitational Water: Water which moves into, through, or out of the soil under the influence of gravity.

Ground water, Confined: Ground water under pressure significantly greater than atmospheric, its upper limit is the bottom of a bed of hydraulic conductivity distinctly lower than that of the material in which confined water occurs.

Ground water, Unconfined: Water in an aquifer that has a water table.

Ground water basin: A ground water basin is a three dimensional, closed system which contains the entire flow paths followed by all the water within the basin. The basin may contain one or more recharge discharge areas.

Head, Hydraulic: The elevation in regard to a specified reference level at which water stands in a piezometer connected to the point in question in the soil. Its definition can be extended to soil above the water table if the piezometer is replaced by a tensiometer. The hydraulic head is a potential expressed in terms of the height of a water column. More specifically, it can be identified with the sum of gravitational and capillary potentials and may be termed the hydraulic potential.

Head, Static: The height above a standard datum of the surface of a column of water that can be supported by the static pressure at a given point. The static head is the sum of the elevation head, h, and the pressure head, H_p, that is, $H = H_e + H_p$. Under conditions to which Darcy's law may be applied, the velocity of ground water is so small that the velocity head, $H_v = v^2/2g$, is negligible. Head, when used alone, is understood to mean static head. Static head is proportional to the fluid potential; therefore, the head is a measure of the potential.

Head, Total: The total head of a liquid at a given point is the sum of three components: (1), the elevation head, H_e, which is equal to the elevation of the point above a datum, (2) the pressure head, H_p, which is the height of a column of static water that can be supported by the static pressure at the point, and (3) the velocity head, h, which is the height to which the kinetic energy of the fluid is capable of lifting the fluid. Thus, the total head can be expressed as $H = H_e + H_p + H_v$.

Heterogeneous porous medium: A medium is heterogenous or non-homogeneous if its properties of isotropy or anisotropy vary from point to point in the medium.

Homogeneous porous medium: A medium is homogeneous if its properties of isotropy or anisotropy are constant over the medium.

Hydraulic conductivity: Hydraulic conductivity is a measure of the ability of a fluid to move through a porous medium. Hydraulic conductivity is the proportionality factor in Darcy's law as applied to the viscous flow of water in soil, that is, the flux of water per unit gradient of hydraulic potential.

Hydraulic gradient: The change in static head per unit of distance in a given direction. If not specified, the direction generally is understood to be that of the maximum rate of decrease in head.

Hydraulic Pressure: The pressure exerted by the water at any given point in a body of water at rest. The hydrostatic pressure of ground water is generally due to the weight of water at higher levels in the zone of saturation.

Infiltration: The downward entry of water into the soil.

Infiltration Rate: A soil characteristic determining or describing the maximum rate at which water can enter the soil under specified conditions, including the presence of an excess of water. It has the dimensions of velocity.

Infiltration velocity: The actual rate at which water is entering the soil at any given time. It may be less than the maximum because of a limited supply of water. It has the same units as the infiltration rate.

Initial Conditions: Information required to obtain a unique solution of a model's governing equation under unsteady-state conditions. This information is comprised of the values of dependent variable specified everywhere inside the boundary.

Interstice: An opening or space between one thing and another, as an opening in a rock or soil that is not occupied by solid matter. Synonym: void; pore.

Interestial Water: Subsurface water in an interstice. Synonym: pore water.

Isotropic porous medium: A medium is isotropic if its properties at any point are the same in all directions from that point.

Iterative Methods: Mathematical solution methods for handling systems containing large numbers of simultaneous equations. The procedure is based on making an initial determination of a set of solution values followed by successive improvement until the correct solution is obtained. There are a number of iterative methods available. Successive over relaxation (SOR) is generally the most efficient for ground water problems.

Mathematical Model: A simulation in mathematical terms, including appropriate assumptions, which represents a simplified version of reality. Mathematical models can be analytical or numerical.

Mechanical filtration: Mechanical filtration is the physical removal of suspended material from solution by constrictions within the porous medium.

Microbial decomposition: Microbial decomposition is the removal of organic constituents in solution of microorganisms.

Model calibration: The process where model input data is adjusted until computed results match the field values.

Model verification: A procedure following model calibration where the goal is to demonstrate that the model is capable of simulating some historical event for which field data is available. Generally, some additional refinement of parameters will be necessary during verification.

Molecular Diffusion: The process where ionic or molecular constituents move under the influence of their kinetic activity in the direction of their concentration gradient.

Numerical Model: A mathematical model whose solution requires that the model's partial differential equation (governing equation) be approximated using a numerical technique such as the finite element or finite difference method.

Percolation : The movement of gravitational water through soil. The downward movement of water through soil; especially, the downward flow of water in saturated or nearly saturated soil at hydraulic gradients of the order of 1.0 or less.

Permeability, Intrinsic, k: A measure of the relative ease with which a porous medium can transmit a liquid under a potential gradient. It is a property of the medium alone and is independent of the nature of the liquid and of the force field causing movement. It is a property of the medium that is dependent upon the shape and size of the pores.

Piezometer: Piezometer is an instrument set into the ground to measure hydraulic head.

Piezometric surface: The surface at which water will stand in a series of piezometers.

Pore: A small to minute opening or passageway in a rock or soil; an interstice.

Porosity: The ratio, usually expressed as a percentage, of (1) the volume of voids of a given soil or rock to (2) the total volume of the soil or rock mass. In regard to the movement of water, only the system of interconnected interstices is significant.

Porosity, Effective: Refers to the amount of interconnected pore space available for fluid transmission. It is expressed as a percentage of the total volume occupied by the interconnecting interstices; the ratio of (1) the volume of the voids of a soil or rock mass that can be drained by gravity to (2) the total volume of the mass.

Potential Drop: The difference in total head between two equipotential lines.

Potentiometric Surface: A surface which represents the static head. In relation to an aquifer, it is defined by the levels to which water will rise in tightly cased wells. The water table is a particular potentiometric surface.

Radius of influence of a well: The distance from the center of the well to the closest point at which the piezometric surface is not lowered when pumping has produced the maximum steady rate of flow.

Recharge: Recharge is that water which percolates down through the unsaturated zone to the water table and enters the dynamic ground water flow system.

Recharge area: A recharge area is an area where the direction of ground water flow is away from the water table. The hydraulic head decreases with depth.

Retardation Factor: This term is used to describe the retardation of a contaminant front relative to the bulk mass of water in cases where the partitioning of a contaminant can be adequately described by the distribution coefficient.

Reynolds Number: Reynolds number is a dimensionless number indicating the magnitude of the ratio of inertial to viscous forces.

Saturated Thickness: A term used in two-dimensional areal ground water modelling applications to define the thickness of saturated zone.

Seepage Velocity: The rate of discharge of seepage water through a porous medium per unit area of void space perpendicular to the direction of flow. See average linear velocity.

Sensitivity Analysis: A process for determining what parameters or data requirements model solutions are most sensitive to.

Sorption: Sorption is the removal of a portion of solute from the liquid phase due to electrical interactions. It includes the processes of adsorption, absorption, and ion exchange. Absorption occurs if ions in solution enter within the structural unit of a soil particle. Adsorption occurs if ions in solution are held onto the surface of a soil particle. Ion exchange occurs if ions in solution replaces ions on or within a soil particle.

Specific discharge: Specific discharge or specific flux is the rate of discharge of ground water per unit cross sectional area measured at right angles to the direction of flow. It is sometimes called the bulk velocity or the Darcian velocity.

Specific Storage: In problems of three dimensional transient flow in a compressible ground water body, it is necessary to consider the amount of water released from or taken into storage per unit volume of the porous medium. The specific storage, Ss , is the volume of water released from or taken into storage per unit volume of the porous medium per unit change in head.

Specific Yield: The specific yield of a rock or soil is the ratio of (1) the volume of water which the rock or soil, after being saturated, will yield by gravity to (2) the volume of the rock or soil. The definition implies that gravity drainage is complete.

Steady flow: Steady flow occurs when at any point the magnitude and direction of the specific discharge are constant in time.

Storage Coefficient: The storage coefficient is the volume of water an aquifer releases from or takes into storage per unit surface area of the aquifer per unit change in head. In a confined water body the water derived from

storage with a decline in head comes from expansion of the water and compression of the aquifer; similarly, water added to storage with a rise in head is accommodated partly by compression of the water and partly by expansion of the aquifer. In an unconfined water body, the amount of water derived from or added to the aquifer by these processes generally is negligible compared with that involved in gravity drainage or filling of pores. Hence, in an unconfined water body, the storage coefficient is virtually equal to the specific yield.

Successive Over Relaxation (SOR): An iterative method often used in solution of ground water models which generally requires significantly fewer iterations than other methods to reach convergence.

Tortuosity: A dimensionless geometric parameter of porous media which is defined as the average ratio of the actual 'roundabout' path to the apparent, or 'straight' flow path; i.e. it is the ratio of (1) the distance traversed by an average parcel of water flowing through the pores of a soil specimen to (2) the length of that specimen.

Transmissivity: The rate at which water of the prevailing kinematic viscosity is transmitted through a unit of the aquifer under a unit hydraulic gradient. Although spoken of as a property of the aquifer, it embodies also the saturated thickness of the aquifer and the properties of the contained liquid. It is equal to an integration of the hydraulic conductivities across the saturated part of the aquifer perpendicular to the flow paths.

Unsaturated Flow: The movement of water in a soil which is not filled to capacity with water.

Void: Space in a soil or rock mass not occupied by solid mineral matter. This space may be occupied by air, water, or other gaseous or liquid material.

Void ratio: The ratio of (1) the volume of void space to (2) the volume of solid particles in a given soil mass.

Water table: The water table is an imaginary surface beneath the surface of the ground at the upper bound of the saturated zone where the pressure is atmospheric.

Water-Holding Capacity: The smallest value to which the water table in which all voids, large and small, are filled with water under pressure greater than atmospheric.

Zone, Saturated: That part of the earth's crust beneath the water table in which all voids, large and small, are filled with water under pressure greater than atmospheric.

Zone, Unsaturated: Replaces the terms zone of aeration and vadose zone, and is the zone between the land surface and the water table. It includes the capillary fringe. Generally, water in this zone is under less than atmospheric pressure, and some of the voids may contain air or other gases at atmospheric pressure.

Contributors

Dr. R.B. Ambrose, Jr.
Environmental Research Laboratory
U.S. Environmental Protection Agency
Athens, GA 3605, USA

Dr. D. A. Barry
Department of Environmental Engineering
University of Western Australia
Perth, WA 6009, Australia

Dr. Ken Bencala
Water Resources Division
U.S. Geological Survey
Mail Stop 421
345 Middlefield Rd.
Menlo Park, CA 94025, USA

Mr. A. G. Bobba
Environment Canada
National Water Research Institute
River Research Branch
Burlington, Ontario
Canada L7R 4A6

Dr. J. Boll
Dept. of Agricultural and Biological Engineering
Cornell University
Ithaca, NY 14853, USA

Dr. W. J. Bond
CISRO Division of Soils
GPO Box 639
Canberra, ACT 2601, Australia

Dr. W. G. Booty
Environment Canada
National Water Research Institute
River Research Branch
Burlington, Ontario
Canada L7R 4A6

Dr. M. L. Brusseau
Department of Soil and Water Science
University of Arizona
Tucson, AZ 85721, USA

Dr. D. M. Dolan
International Joint Commission
Great Lakes Regional Office
100 Ouellette Ave., 8th Floor
Windsor, Ontario N9A 6T3, Canada

Mr. A. S. Donigian, Jr.
Aqua-Terra, Inc.
2672 Bayshore Pkwy., Suite 1001
Mountain View, CA 94043-1011, USA

Dr. B. L. Finlayson
Centre for Environmental Applied Hydrology
University of Melbourne
Parkville, Victoria 3052, Australia

Mr. J. C. Imhoff
Aqua-Terra, Inc.
2672 Bayshore Pkwy., Suite 1001
Mountain View, CA 94043-1011, USA

Dr. D. S. Jeffries
Environment Canada
National Water Research Institute
River Research Branch
Burlington, Ontario
Canada L7R 4A6

Professor T. A. McMahon
Centre for Environmental Applied Hydrology
University of Melbourne
Parkville, Victoria 3052
Australia

Professor M. Mimikou
Department of Civil Engineering
Technical University of Athens
Athens, Greece

Professor J. Y. Parlange
Dept. of Agricultural and Biological Engineering
Cornell University
Ithaca, NY 14853, USA

Dr. Bruce Pivetz
Dept. of Agricultural and Biological Engineering
Cornell University
Ithaca, NY 14853, USA

Dr. R. L. Runkel
Water Resources Division
U.S. Geological Survey
Box 25046, Mail Stop 415
Denver Federal Center
Denver, CO 80225-0046, USA

Professor V. P. Singh
Department of Civil and Environmental Engineering
Louisiana State University
Baton Rouge, LA 70803-6405, USA

Dr. J. Sorenson
Harbors and Coastal Center
University of Massachusetts at Boston
Morrissey Boulevard
Boston, MA 02125-3393, USA

Dr. F. Stagnitti
School of Computing and Mathematics
Deakin University
P.O. Box 423
Warrnambool 3280
Australia

Dr. T. S. Steenhuis
Dept. of Agricultural and Biological Engineering
Cornell University
Ithaca, NY 14853, USA

Professor P. H. Templet
Institute of Environmental Studies
Louisiana State University
Baton Rouge, LA 70803, USA

Dr. Y. Tan
Centre for Environmental Mechanics
CSIRO
GPO Box 821
Canberra, ACT 2601
Australia

Professor P. J. Wierenga
Department of Soil and Water Science
University of Arizona
Tucson, AZ 85721, USA

Professor T.C. Young
Department of Civil and Environmental Engineering
Clarkson University
Potsdam, NY 13699, USA

Subject Index

activity coefficients 155–156
adsorption, chemical 331–332
adsorption, hydrophobicity 331–332
adsorption, isotherms 334–335
adsorption, microorganisms 331–332, 334–336
adsorption, physical 331–332
adsorption, physico–chemical 331–332
adsorption, polymer bridging 331–332
advection 139–144, 147, 151
advection-dispersion equation 144-145, 147, 150, 162, 322, 339
Andrew's curves 127
annual exceedance series 116–117
annual maximum flood series 115, 117, 122
aquatic ecology 107
attachment, adsorption 334–335
attachment coefficient 335–336
attachment equilibrium 334–335
attachment, interception 334
attachment, kinetic adsorption 335–336
attachment, microorganisms 323, 331–337
attachment, non–equilibrium 335–336
attachment rate 334–336
attachment, sedimentation 333
attachment, straining 332–333
average recurrence interval 116–118, 122, 124

bacteria 321–348
bacteria growth 327–330
bacteria decay 330–331
bacteria, transport or movement 321–348
bacteria, attachment and detachment 331–337
bankfull discharge 117, 123, 126
baseflow 110, 112–113, 120–121, 129, 133
batch reactor 158
best management practices (bmps) 380, 384, 388, 401
biodegradation 213
biofilm 323, 329–330, 336–337, 340–342, 345–346, 348
biomass, biofilm 323, 329–330, 336–337, 340–342, 345–346, 348
biomass distribution 323, 348
biomass, macroscopic 323, 327, 348
biomass, microcolony 323
Brownian diffusion 326–327
buildup/washoff functions 388, 390–394, 399, 406

CE–QUAL 409–415
change 72
change, climatic 69
channel form 107–108, 124, 131
channel network 107
chemical equilibrium 154
chemical reactions, heterogeneous 152
chemical reactions, homogeneous 151
chemotaxis 325, 338
chemotaxis flux 325, 338
chemotaxis velocity 325, 338
Colwell's indices 120
comparative hydrology 427
complimentary error function 150
conservation of mass 141–142
contaminant loads 359
continuous release 151–152
continuous simulation 389, 394, 402, 405
continuous sources 149
continuum, environmental 1
continuum, water 2
control volume 141–143
conventional pollutants 378, 396, 399, 408, 409, 413–415
CREAMS 399–406

data errors 109
databases 384, 415–417, 420, 421
decay, bacteria 330–331
decay, microorganisms 330–331, 337–338
decay rate 331, 337–338
decay, viruses 330–331, 340
decision support systems 415, 419, 421
desorption 153, 157, 160, 162
detachment 334–340
detachment coefficient 336
detachment rate 336
detachment, shear loss 336–337
discharge 108, 110, 112–113, 115–118, 122–126, 130, 131, 135
dispersion 139–147, 151
dissolution 153, 155, 160
dissolved oxygen 137, 152–153, 160
distribution coefficient 158

ecological 427
environmental 427
environmental flow 108, 124, 130–131
environmental hydrology 107, 124, 131, 133
equilibrium approach 153–154
equilibrium constant 155-156
eutrophication 357

evapotranspiration 82
event mean concentrations (emcs) 388–391
EXAMS 409–415

Fick's law of Diffusion 143
filtration of microorganisms 335
first-order analysis (variance propagation) 360
flood frequency analysis 117
floods 115–118, 122–123, 127
flow duration curves 113, 115, 121
fluvial geometry 123, 131
flux 143
fluxes 324–327, 337–338,
fluxes, advective 324–335, 338
fluxes, brownian diffusion 326–327
fluxes, chemotactic 325, 338
fluxes, diffusion, dispersion and motility 326–327, 337–338
fluxes, microorganisms 324–327, 337–338
fluxes, water 324–325
forcing function 358
Freundlich adsorption isotherm 334–336
Freundlich isotherms 158, 162

GIS 403, 418
global 75
graphics 395, 410, 416, 417, 421
greenhouse 69
groundwater 4, 9, 85
growth models, bacteria 327–330
growth models, general approach 328
growth models, monod 327–330
growth models, moser 328
growth models, multiple monod 328
growth models, structured 327
growth models, unstructured 327
growth, bacteria 327–330, 337–338, 341–342
growth rate, specific 327–328
growth saturation constant 327–328

habitat 108, 128, 130–131, 136
heterogeneous reactions 152, 160
high flow 115
homogeneous reactions 151, 160
HSPF 387, 389–394, 399–401, 406, 409–415
hydrograph 107, 110–115, 117–118, 120–121, 134
hydrograph separation 112
hydrological 77
hydrology 427
hydrous ferric oxides 157

impacts 427
in-place pollutants 358
instantaneous release 147, 150–152
instantaneous sources 148–149
insteam flow incremental methodology 130
integral water management 5, 7–8
interception of microorganisms 334
interflow 3
ion activity product (IAP) 156
irrigation 124, 127–128, 132–133

kinetic approach 153–154, 162

Langmuir isotherm 158
lateral inflow 144, 147, 149
loading functions 395–399, 406
loading 358ff
log-pearson type iii distribution 116
longitudinal dispersion coefficient 143
low flow 108, 115, 118–119, 125, 128, 131, 134–135

management 88
mass action equation 155
mass balance equation 141–142
mathematical model 203
mean annual flow 109, 114, 119, 126, 128
microcolony 323
model 358
model development 377–379, 387, 406, 416–422
model development support 378, 394, 397, 406, 416
model user support 382, 394, 397, 405, 411, 413, 415–416, 419–421
moisture 84
molecular diffusion 139–140
monte carlo simulation 360
motility 326–327
motility coefficient 326–327

non-urban runoff 381, 387, 395, 399
nonpoint source 148–149
nonpoint source pollution 377–381, 384–385, 387, 394, 395–399, 402, 405–409, 413, 416, 419
nutrients 378, 380, 384–386, 397, 399–403, 409, 410, 414–416, 418

overland flow 3

parameterization 360
partial duration series 116–117, 124
pesticide transport 210
pesticides 378, 380, 381, 383, 385–387, 397, 398, 401–405, 407, 415, 418
pH 159ff
phosphorus 365
plankton 385–386, 401, 409–410, 412, 414–415
planning 88
point source 147–148
precipitation 82, 153, 155, 160
preferential transport 195
primary models 383–384
process 427

quickflow 110, 112–113, 121

rainfall-runoffmodel 110, 129
randomization 359
receiving water 377, 378, 380, 381, 384–387, 389, 391, 398, 407, 408, 415, 419
recession 3
redundancy 367
regional 75
regression analysis 382, 388–390, 392–394, 398, 406–407
regulated rivers 118, 124, 127, 130–131
river regulation 127, 134
runoff 84

sample collection 367
sampling designs 367–369
saturated areas 3
saturated zones 3
saturation index (SI) 156–157
scenarios 78
screening models 382–384, 388, 396–398, 406
sea-level 86
seasonal regime 120–121, 127–129
secondary models 383, 384
sediment 107, 115, 124–125, 132, 377–378, 380, 383, 385–386, 390–391, 397–401, 402–406, 409, 410, 414, 415
sedimentation of microorganisms 333
shear stress 139, 141
simulation models 380, 382, 389, 392, 393, 395–396, 398–399, 401, 407, 411–412, 416
soil structure 200
solubility product 156
sorption 153, 157–158, 160, 162
sorption isotherms 158
specific growth rate of bacteria 327–328
spell analysis 118–119
stage 106–107, 117
steady flow 141
storm runoff 4
storm hydrograph 110, 112
straining of microorganisms 332–333
streamflow hydrograph 3
substrate 327–328, 340–342
substrate consumption 341
substrate transport 340–341
subsurface flow 3
surface complexation models 158
surface flow 110, 112–113
surface runoff 3
survival of microorganisms 330–331
SWMM 389–395, 406, 416
synthetic data 109

total maximum daily loads (TMDLS) 378, 380, 387, 407
toxic chemical 357ff
toxic pollutants 378, 385–386, 396, 399, 405, 408, 410, 413–415
tracer experiment 145
transient storage 144–147
transport of microorganisms 321–348

uncertainty 358ff, 409, 420–421
ungauged streams 124
urban runoff 380, 387, 389–390, 392
user interaction 378, 417, 419, 421
USLE 396, 398–399, 403

variability 72, 107, 114, 119–123
viruses 330–336, 340, 346
viruses attachment and detachment 331–336
viruses decay or inactivation 330–331
viruses transport 340, 346

warming 69
WASP 409–415
water flow velocity 325
water quality management 358
watershed modeling 377–378, 416–418
watershed response 9
wetlands 127–128

Water Science and Technology Library

1. A.S. Eikum and R.W. Seabloom (eds.): *Alternative Wastewater Treatment.* Low-Cost Small Systems, Research and Development. Proceedings of the Conference held in Oslo, Norway (7–10 September 1981). 1982
   ISBN 90-277-1430-4
2. W. Brutsaert and G.H. Jirka (eds.): *Gas Transfer at Water Surfaces.* 1984
   ISBN 90-277-1697-8
3. D.A. Kraijenhoff and J.R. Moll (eds.): *River Flow Modelling and Forecasting.* 1986 ISBN 90-277-2082-7
4. World Meteorological Organization (ed.): *Microprocessors in Operational Hydrology.* Proceedings of a Conference held in Geneva (4–5 September 1984). 1986 ISBN 90-277-2156-4
5. J. Němec: *Hydrological Forecasting.* Design and Operation of Hydrological Forecasting Systems. 1986 ISBN 90-277-2259-5
6. V.K. Gupta, I. Rodríguez-Iturbe and E.F. Wood (eds.): *Scale Problems in Hydrology.* Runoff Generation and Basin Response. 1986
   ISBN 90-277-2258-7
7. D.C. Major and H.E. Schwarz: *Large-Scale Regional Water Resources Planning.* The North Atlantic Regional Study. 1990 ISBN 0-7923-0711-9
8. W.H. Hager: *Energy Dissipators and Hydraulic Jump.* 1992
   ISBN 0-7923-1508-1
9. V.P. Singh and M. Fiorentino (eds.): *Entropy and Energy Dissipation in Water Resources.* 1992 ISBN 0-7923-1696-7
10. K.W. Hipel (ed.): *Stochastic and Statistical Methods in Hydrology and Environmental Engineering.* A Four Volume Work Resulting from the International Conference in Honour of Professor T. E. Unny (21–23 June 1993). 1994
    10/1: Extreme values: floods and droughts ISBN 0-7923-2756-X
    10/2: Stochastic and statistical modelling with groundwater and surface water applications ISBN 0-7923-2757-8
    10/3: Time series analysis in hydrology and environmental engineering
    ISBN 0-7923-2758-6
    10/4: Effective environmental management for sustainable development
    ISBN 0-7923-2759-4
    Set 10/1–10/4: ISBN 0-7923-2760-8
11. S.N. Rodionov: *Global and Regional Climate Interaction: The Caspian Sea Experience.* 1994 ISBN 0-7923-2784-5
12. A. Peters, G. Wittum, B. Herrling, U. Meissner, C.A. Brebbia, W.G. Gray and G.F. Pinder (eds.): *Computational Methods in Water Resources X.* 1994
    Set 12/1–12/2: ISBN 0-7923-2937-6
13. C.B. Vreugdenhil: *Numerical Methods for Shallow-Water Flow.* 1994
    ISBN 0-7923-3164-8
14. E. Cabrera and A.F. Vela (eds.): *Improving Efficiency and Reliability in Water Distribution Systems.* 1995 ISBN 0-7923-3536-8

Water Science and Technology Library

15. V.P. Singh (ed.): *Environmental Hydrology.* 1995 ISBN 0-7923-3549-X

Kluwer Academic Publishers – Dordrecht / Boston / London