Rainfall-Runoff Modelling

Rainfall-Runoff Modelling

The Primer

SECOND EDITION

Keith Beven

Lancaster University, UK

WILEY-BLACKWELL

A John Wiley & Sons, Ltd., Publication

This edition first published 2012 © 2012 by John Wiley & Sons, Ltd

Wiley-Blackwell is an imprint of John Wiley & Sons, formed by the merger of Wiley's global Scientific, Technical and Medical business with Blackwell Publishing.

Registered office: John Wiley & Sons, Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, UK

Editorial offices: 9600 Garsington Road, Oxford, OX4 2DQ, UK
The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, UK
111 River Street, Hoboken, NJ 07030-5774, USA

For details of our global editorial offices, for customer services and for information about how to apply for permission to reuse the copyright material in this book please see our website at www.wiley.com/wiley-blackwell.

Library of Congress Cataloging-in-Publication Data

Beven, K. J.
Rainfall-runoff modelling : the primer / Keith Beven. – 2nd ed.
p. cm.
Includes bibliographical references and index.
ISBN 978-0-470-71459-1 (cloth)
1. Runoff–Mathematical models. 2. Rain and rainfall–Mathematical models.
I. Title.
GB980.B48 2011
551.48′8–dc23

2011028243

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

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic books.

Set in 10/12pt Times by Thomson Digital, Noida, India

First Impression 2012

For the next generation of hydrological modellers

Contents

Preface to the Second Edition

Models are undeniably beautiful, and a man may justly be proud to be seen in their company. But they may have their hidden vices. The question is, after all, not only whether they are good to look at, but whether we can live happily with them.

A. Kaplan, 1964

One is left with the view that the state of water resources modelling is like an economy subject to inflation – that there are too many models chasing (as yet) too few applications; that there are too many modellers chasing too few ideas; and that the response is to print ever-increasing quantities of paper, thereby devaluing the currency, vast amounts of which must be tendered by water resource modellers in exchange for their continued employment.

Robin Clarke, 1974

It is already (somewhat surprisingly) 10 years since the first edition of this book appeared. It is (even more surprisingly) 40 years since I started my own research career on rainfall–runoff modelling. That is 10 years of increasing computer power and software development in all sorts of domains, some of which has been applied to the problem of rainfall–runoff modelling, and 40 years since I started to try to understand some of the difficulties of representing hydrological processes and identifying rainfall–runoff model parameters. This new edition reflects some of the developments in rainfall–runoff modelling since the first edition, but also the fact that many of the problems of rainfall–runoff modelling have not really changed in that time. I have also had to accept the fact that it is now absolutely impossible for one person to follow all the literature relevant to rainfall–runoff modelling. To those model developers who will be disappointed that their model does not get enough space in this edition, or even more disappointed that it does not appear at all, I can only offer my apologies. This is necessarily a personal perspective on the subject matter and, given the time constraints of producing this edition, I may well have missed some important papers (or even, given this aging brain, overlooked some that I found interesting at the time!).

It has been a source of some satisfaction that many people have told me that the first edition of this book has been very useful to them in either teaching or starting to learn rainfall–runoff modelling (even Anna, who by a strange quirk of fate did, in the end, actually have to make use of it in her MSc course), but it is always a bit daunting to go back to something that was written a decade ago to see just how much has survived the test of time and how much has been superseded by the wealth of research that has been funded and published since, even if this has continued to involve the printing of ever-increasing quantities of paper (over 30 years after Robin Clarke's remarks above). It has actually been a very interesting decade for research in rainfall–runoff modelling that has seen the Prediction in Ungauged Basins (PUB) initiative of the International Association of Hydrological Scientists (IAHS), the

implementation of the Representative Elementary Watershed (REW) concepts, the improvement of land surface parameterisations as boundary conditions for atmospheric circulation models, the much more widespread use of distributed conceptual models encouraged by the availability of freeware software such as SWAT, developments in data assimilation for forecasting, the greater understanding of problems of uncertainty in model calibration and validation, and other advances. I have also taken the opportunity to add some material that received less attention in the first edition, particularly where there has been some interesting work done in the last decade. There are new chapters on regionalisation methods, on modelling residence times of water in catchments, and on the next generation of hydrological models.

Going back to the original final paragraph of the 1st edition, I suggested that:

> The future in rainfall–runoff modelling is therefore one of uncertainty: but this then implies a further question as to how best to constrain that uncertainty. The obvious answer is by conditioning on data, making special measurements where time, money and the importance of a particular application allow. It is entirely appropriate that this introduction to available rainfall–runoff modelling techniques should end with this focus on the value of field data.

This has not changed in the last 10 years. The development, testing and application of rainfall–runoff models is still strongly constrained by the availability of data for model inputs, boundary conditions and parameter values. There are still important issues of how to estimate the resulting uncertainties in model predictions. There are still important issues of scale and commensurability between observed and model variables. There have certainly been important and interesting advances in rainfall–runoff modelling techniques, but we are still very dependent on the quantity and quality of available data. Uncertainty estimation is now used much more widely than a decade ago, but it should not be the end point of an analysis. Instead, it should always leave the question: what knowledge or data are required to constrain the uncertainty further?

In fact, one of the reasons why there has been little in really fundamental advances over the last decade is that hydrology remains constrained by the measurement techniques available to it. This may seem surprising in the era of remote sensing and pervasive wireless networking. However, it has generally proven difficult to derive useful hydrological information from this wealth of data that has become (or is becoming) available. Certainly none of the developments in field measurements have yet really changed the ways in which rainfall–runoff modelling is actually done. At the end of this edition, I will again look forward to when and how this might be the case.

I do believe that the nature of hydrological modelling is going to change in the near future. In part, this is the result of increased availability of computer power (I do not look back to the days when my PhD model was physically two boxes of punched cards with any nostalgia . . . programming is so much easier now, although using cards meant that we were very much more careful about checking programs before submitting them and old cards were really good for making to-do lists!). In part, it will be the result of the need to cover a range of scales and coupled processes to satisfy the needs of integrated catchment management. In part, it will be the result of increased involvement of local stakeholders in the formulation and evaluation of models used in decision making. In part, it will be the desire to try to constrain the uncertainty in local predictions to satisfy local stakeholders. The result will be the implementation of "models of everywhere" as a learning and hypothesis testing process. I very much hope that this will give some real impetus to improving hydrological science and practice akin to a revolution in the ways that we do things. Perhaps in another decade, we will start to see the benefits of this revolution.

It has been good to work with a special group of doctoral students, post-docs, colleagues and collaborators in the last 10 years in trying to further the development of rainfall–runoff modelling and uncertainty estimation methods. I would particularly like to mention Peter Young, Andy Binley, Kathy Bashford, Paul Bates, Sarka Blazkova, Rich Brazier, Wouter Buytaert, Flavie Cernesson, Hyung Tae Choi, Jess Davies, Jan Feyen, Luc Feyen, Jim Freer, Francesc Gallart, Ion Iorgulescu, Christophe Joerin, John Juston, Rob

Lamb, Dave Leedal, Liu Yangli, Hilary McMillan, Steve Mitchell, Mo Xingguo, Charles Obled, Trevor Page, Florian Pappenberger, Renata Romanowicz, Jan Seibert, Daniel Sempere, Paul Smith, Jonathan Tawn, Jutta Thielen, Raul Vazquez, Ida Westerberg, Philip Younger and Massimiliano Zappa. Many others have made comments on the first edition or have contributed to valuable discussions and debates that have helped me think about the nature of the modelling process, including Kevin Bishop, John Ewen, Peter Germann, Sven Halldin, Jim Hall, Hoshin Gupta, Dmiti Kavetski, Jim Kirchner, Mike Kirkby, Keith Loague, Jeff McDonnell, Alberto Montanari, Enda O'Connell, Geoff Pegram, Laurent Pfister, Andrea Rinaldo, Allan Rodhe, Jonty Rougier, Murugesu Sivapalan, Bertina Schaefli, Stan Schymanski, Lenny Smith, Ezio Todini, Thorsten Wagener, and Erwin Zehe. We have not always agreed about an appropriate strategy but long may the (sometimes vigorous) debates continue. There is still much more to be done, especially to help guide the next generation of hydrologists in the right direction . . . !

Keith Beven
Outhgill, Lancaster, Fribourg and Uppsala, 2010–11

About the Author

The author programmed his first rainfall–runoff model in 1970, trying to predict the runoff generation processes on Exmoor during the Lynmouth flood. Since then, he has been involved in many of the major rainfall–runoff modelling innovations including TOPMODEL, the Système Hydrologique Européen (SHE) model, the Institute of Hydrology Distributed Model (IHDM) and data-based mechanistic (DBM) modelling methodology. He has published over 350 papers and a number of other books. He was awarded the IAHS/WMO/UNESCO International Hydrology Prize in 2009; the EGU Dalton Medal in 2004; and the AGU Horton Award in 1991. He has worked at Lancaster University since 1985 and currently has visiting positions at Uppsala University and the London School of Economics.

List of Figures

1

Down to Basics: Runoff Processes and the Modelling Process

As scientists we are intrigued by the possibility of assembling our knowledge into a neat package to show that we do, after all, understand our science and its complex interrelated phenomena.

W. M. Kohler, 1969

Remember that the computer is a tool for simulation, and what is simulated becomes the reality of the user. In a society like ours – the post-modern society – there are no 'great stories' to justify a specific perception of reality like there were in the 19th century. We should rather see the situation thus: communication is based on a number of language games which are played according to specific sets of rules. Each group of society can 'play a game', and thus it is the efficiency of each game that justifies it. The computer medium should be seen as a technical device that allows its owner to play particularly efficient games. . . . A good program is one that creates the reality intended by the sender in the most efficient way.

P. B. Andersen and L. Mathiessen, 1987

1.1 Why Model?

As noted in the preface, there are many reasons why we need to model the rainfall–runoff processes of hydrology. The main reason is a result of the limitations of hydrological measurement techniques. We are not able to measure everything we would like to know about hydrological systems. We have, in fact, only a limited range of measurement techniques and a limited range of measurements in space and time. We therefore need a means of extrapolating from those available measurements in both space and time, particularly to ungauged catchments (where measurements are not available) and into the future (where measurements are not possible) to assess the likely impact of future hydrological change. Models of

Rainfall–Runoff Modelling: The Primer, Second Edition. Keith Beven.

different types provide a means of quantitative extrapolation or prediction that will hopefully be helpful in decision making.

There is much rainfall–runoff modelling that is carried out purely for research purposes as a means of formalising knowledge about hydrological systems. The demonstration of such understanding is an important way of developing an area of science. We generally learn most when a model or theory is shown to be in conflict with reliable data so that some modification of the understanding on which the model is based must be sought. However, the ultimate aim of prediction using models must be to improve decision making about a hydrological problem, whether that be in water resources planning, flood protection, mitigation of contamination, licensing of abstractions, or other areas. With increasing demands on water resources throughout the world, improved decision making within a context of fluctuating weather patterns from year to year requires improved models. That is what this book is about.

Rainfall–runoff modelling can be carried out within a purely analytical framework based on observations of the inputs to and outputs from a catchment area. The catchment is treated as a *black box*, without any reference to the internal processes that control the transformation of rainfall to runoff. Some models developed in this way are described in Chapter 4, where it is shown that it may also be possible to make some physical interpretation of the resulting models based on an understanding of the nature of catchment response. This understanding should be the starting point for any rainfall–runoff modelling study.

There are, of course, many hydrological texts that describe hydrological processes with varying degrees of mathematical analysis and numbers of equations. The more mathematical descriptions do not always point out the important simplifications that are being made in their analyses, but present the equations as if they apply everywhere. However, it is only necessary to sprinkle a coloured dye onto the soil surface and then dig to see where the dye has stained the soil (Figure 1.1) to realise the limitations of

Figure 1.1 *Staining by dye after infiltration at the soil profile scale in a forested catchment in the Chilean Andes (from Blume et al., 2009).*

hydrological theory (see also Flury *et al.*, 1994; Zehe and Flühler, 2001; Weiler and Naef, 2003; Kim *et al.*, 2006; Blume *et al.*, 2009). Whenever detailed studies of flow pathways are carried out in the field we find great complexity. We can perceive that complexity quite easily, but producing a mathematical description suitable for quantitative prediction is much more difficult and will always involve important simplication and approximation. This initial chapter will, therefore, be concerned with a perceptual model of catchment response as the first stage of the modelling process. This complexity is one reason why there is no commonly agreed modelling strategy for the rainfall–runoff process but a variety of options and approaches that will be discussed in the chapters that follow.

1.2 How to Use This Book

It should be made clear right at the beginning that this book is not only about the theory that underlies the different types of rainfall–runoff model that are now available to the user. You will find, for example, that relatively few equations are used in the main text of the book. Where it has been necessary to show some theoretical development, this is generally presented in boxes at the ends of the chapters that can be skipped at a first reading. The theory can also be followed up in the many (but necessarily selected) references quoted, if necessary.

This is much more a book about the concepts that underlie different modelling approaches and the critical analysis of the software packages that are now widely available for hydrological prediction. The presentation of models as software is becoming increasingly sophisticated with links to geographical information systems and the display of impressive three-dimensional graphical outputs. It is easy to be seduced by these displays into thinking that the output of the model is a good simulation of the real catchment response, especially if little data are available to check on the predictions. However, even the most sophisticated models currently available are not necessarily good simulations and evaluation of the model predictions will be necessary. It is hoped that the reader will learn from this book the concepts and techniques necessary to evaluate the assumptions that underlie the different modelling approaches and packages available and the issues of implementing a model for a particular application.

One of the aims of this book is to train the reader to evaluate models, not only in terms of how well the model can reproduce any data that are available for testing, but also by critically assessing the assumptions made. Thus, wherever possible, models are presented with a list of the assumptions made. The reader is encouraged to make a similar list when encountering a model for the first time. At the end of each chapter, a review of the major points arising from that chapter has been provided. It is generally a good strategy to read the summary before reading the bulk of the chapter. Some sources of rainfall–runoff modelling software are listed in Appendix A. A glossary of terms used in hydrological modelling is provided in Appendix B. These terms are highlighted when they first appear in the text.

This edition has been extended, relative to the first edition, particularly in respect of the chapters labelled “Beyond the Primer”. New material on the next generation of hydrological models, modelling ungauged catchments, and modelling sources and residence time distributions of water have been added. In addition, there has been a lot of research in the last decade on the use of distributed models and the treatment of uncertainty in hydrological predictions so that Chapters 5, 6 and 7 have also been substantially revised.

1.3 The Modelling Process

Most books on modelling start with the choice of model to be used for a particular application. Here, we start at an earlier stage in the modelling process with the *perceptual model* of the rainfall–runoff processes in a catchment (see Figure 1.2). The perceptual model is the summary of our perceptions of

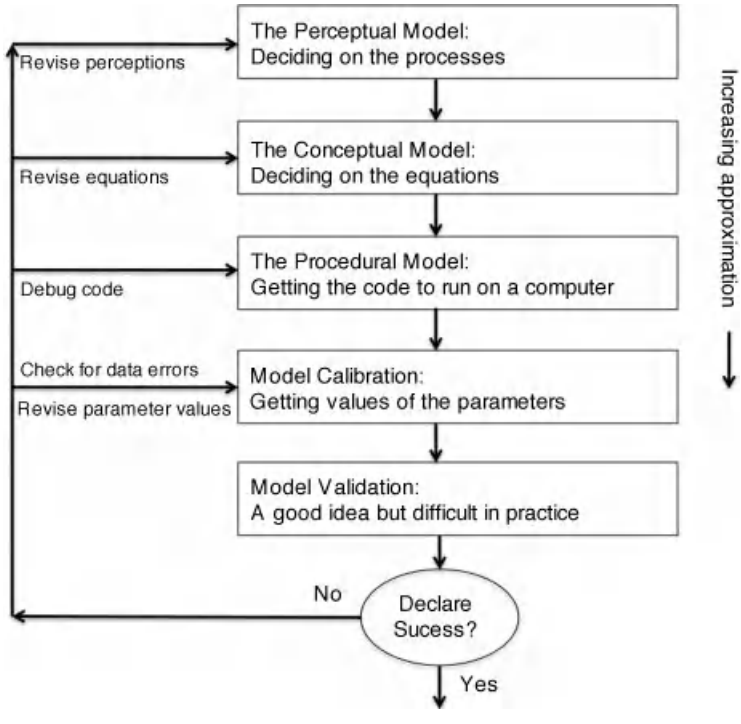


Figure 1.2 *A schematic outline of the steps in the modelling process.*

how the catchment responds to rainfall under different conditions or, rather, *your* perceptions of that response. A perceptual model is necessarily personal. It will depend on the training that a hydrologist has had, the books and articles he or she has read, the data sets that he or she has analysed and, particularly, the field sites and environments of which he or she has had experience. Thus, it is to be expected that one hydrologist's perceptual model will differ from that of another (for a typical personal example, see Section 1.4).

An appreciation of the perceptual model for a particular catchment is important. It must be remembered that all the mathematical descriptions used for making predictions will inevitably be simplifications of the perceptual model, in some cases gross simplifications, but perhaps still sufficient to provide adequate predictions. The perceptual model is not constrained by mathematical theory. It exists primarily in the head of each hydrologist and need not even be written down. We can *perceive* complexities of the flow processes in a purely qualititative way (see, for example, the experiments of Flury *et al.* (1994) and Figure 1.1) that may be very difficult indeed to describe in the language of mathematics. However a mathematical description is, traditionally, the first stage in the formulation of a model that will make **quantitative** predictions.

This mathematical description, we call here the *conceptual model* of the process or processes being considered. At this point, the hypotheses and assumptions being made to simplify the description of the processes need to be made explicit. For example, many models have been based on a description of flow through the soil using Darcy's law, which states that flow is proportional to a gradient of hydraulic potential (see Box 5.1). Measurements show that gradients of hydraulic potential in structured soils can vary significantly over small distances so that if Darcy's law is applied at the scale of a soil profile or

greater, it is implicitly assumed that some average gradient can be used to characterise the flow and that the effects of *preferential flow* through macropores in the soil (one explanation of the observations of Figure 1.1) can be neglected. It is worth noting that, in many articles and model user manuals, while the equations on which the model is based may be given, the underlying simplifying assumptions may not actually be stated explicitly. Usually, however, it is not difficult to list the assumptions once we know something of the background to the equations. This should be the starting point for the evaluation of a particular model relative to the perceptual model in mind. Making a list of all the assumptions of a model is a useful practice that we follow here in the presentation of different modelling approaches.

The conceptual model may be more or less complex, ranging from the use of simple mass balance equations for components representing storage in the catchment to coupled nonlinear partial differential equations. Some equations may be easily translated directly into programming code for use on a digital computer. However, if the equations cannot be solved analytically given some *boundary conditions* for the real system (which is usually the case for the partial differential equations found in hydrological models) then an additional stage of approximation is necessary using the techniques of numerical analysis to define a *procedural model* in the form of code that will run on the computer. An example is the replacement of the differentials of the original equations by finite difference or finite volume equivalents. Great care has to be taken at this point: the transformation from the equations of the conceptual model to the code of the procedural model has the potential to add significant error relative to the true solution of the original equations. This is a particular issue for the solution of nonlinear continuum differential equations but has been the subject of recent discussion with respect to more conceptual catchment models (Clark and Kavetski, 2010). Because such models are often highly *nonlinear*, assessing the error due only to the implementation of a numerical solution for the conceptual model may be difficult for all the conditions in which the model may be used. It might, however, have an important effect on the behaviour of a model in the calibration process (e.g. Kavetksi and Clark, 2010).

With the procedural model, we have code that will run on the computer. Before we can apply the code to make quantitative predictions for a particular catchment, however, it is generally necessary to go through a stage of parameter *calibration*. All the models used in hydrology have equations that involve a variety of different input and state variables. There are inputs that define the geometry of the catchment that are normally considered constant during the duration of a particular simulation. There are variables that define the time-variable boundary conditions during a simulation, such as the rainfall and other meteorological variables at a given time step. There are the state variables, such as soil water storage or water table depth, that change during a simulation as a result of the model calculations. There are the initial values of the state variables that define the state of the catchment at the start of a simulation. Finally, there are the model *parameters* that define the characteristics of the catchment area or flow domain.

The model parameters may include characteristics such as the porosity and hydraulic conductivity of different soil horizons in a spatially distributed model, or the mean residence time in the saturated zone for a model that uses state variables at the catchment scale. They are usually considered constant during the period of a simulation (although for some parameters, such the capacity of the interception storage of a developing vegetation canopy, there may be a strong time dependence that may be important for some applications). In all cases, even if they are considered as constant in time, it is not easy to specify the values of the parameters for a particular catchment *a priori*. Indeed, the most commonly used method of parameter calibration is a technique of adjusting the values of the parameters to achieve the best match between the model predictions and any observations of the actual catchment response that may be available (see Section 1.8 and Chapter 7).

Once the model parameter values have been specified, a simulation may be made and quantitative predictions about the response obtained. The next stage is then the *validation* or *evaluation* of those predictions. This evaluation may also be carried out within a quantitative framework, calculating one or more indices of the performance of the model relative to the observations available (if any) about the runoff response. The problem at this point is not usually that it is difficult to find an acceptable model,

particularly if it has been possible to calibrate the model parameters by a comparison with observed discharges; most model structures have a sufficient number of parameters that can be varied to allow reasonable fits to the data. The problem is more often that there are many different combinations of model structure and sets of parameter values that will give reasonable fits to the discharge data. Thus, in terms of discharge prediction alone, it may be difficult to differentiate between different feasible models and therefore to validate any individual model. This will be addressed in more detail in Chapter 7 in the context of assessing uncertainty in model predictions and testing models as hypotheses about how a catchment responds to rainfall.

On the other hand, the discharge predictions, together with any predictions of the internal responses of the catchment, may also be evaluated relative to the original perceptual model of the catchment of interest. Here, it is usually much more difficult to find a model that is totally acceptable. The differences may lead to a revision of the parameter values used; to a reassessment of the conceptual model; or even, in some cases, to a revision of the perceptual model of the catchment as understanding is gained from the attempt to model the hydrological processes.

The remainder of this chapter will be concerned with the different stages in the modelling processes. An example of a perceptual model of catchment responses to rainfall is outlined in Section 1.4; the additional information that might be gained from considering geochemical information in Section 1.5; the functional requirements of runoff production and runoff routing in Section 1.6; the definition of a conceptual model in Section 1.7; and model calibration and validation issues in Section 1.8.

1.4 Perceptual Models of Catchment Hydrology

There are many outlines of the processes of catchment response available in the literature. Most general hydrological texts deal, in greater or lesser detail, with the processes of catchment response. The volumes edited by Kirkby (1978), Anderson and Burt (1990), and Beven (2006d) are of particular interest in that the different chapters reflect the experience of a number of different hydrologists. Hydrological systems are sufficiently complex that each hydrologist will have his or her own impression or perceptual model of what is most important in the rainfall–runoff process so that different hydrologists might not necessarily agree about what are the most important processes or the best way of describing them. There are sure to be general themes in common, as reflected in hydrological texts, but our understanding of hydrological responses is still evolving and the details will depend on experience, in particular, on the type of hydrological environments that a hydrologist has experienced. Different processes may be dominant in different environments and in catchments with different characteristics of topography, soil, vegetation and bedrock. Part 10 in Volume 3 of the *Encyclopaedia of Hydrological Sciences* (Anderson, 2005) also gives a review of different types of runoff processes with contributions from different hydrologists; a review of runoff processes in semi-arid areas, for example, is provided by Beven (2002c) and in tropical areas by Bonell (2004).

One of the problems involved in having a complete understanding of hydrological systems is that most of the water flows take place underground in the soil or bedrock. Our ability to measure and assess subsurface flow processes is generally very limited. Most of the measurement techniques available reflect conditions only in the immediate area of the measurement probe. When the characteristics of the flow domain vary rapidly in space (and sometimes in time), the small-scale nature of such measurements can give only a very partial picture of the nature of the flow. Thus, there is much that remains unknown about the nature of subsurface flow processes and is, indeed, unknowable given the limitations of current measurement techniques. It is necessary to make inferences about the processes from the available measurements. Such inferences add information to the perceptual model of hydrological response, but they are inferences.

One way of gaining further understanding is to examine a part of the system in much greater detail. Many studies have been made of the flow processes on particular hillslopes or plots, or columns of

undisturbed soil brought back to the laboratory. It has generally been found in such studies that more detailed investigation reveals greater complexity and variability in the flow pathways. The same has generally been true of adding different types of information, such as the use of artificial or environmental tracers. Figure 1.1 is a good example of this (see also Section 1.5). Such complexity can be made part of the perceptual model. As noted above, it is not necessary that the perceptual model be anything more than a set of qualitative impressions, but complexity inevitably creates difficulty in the choice of assumptions in moving from the perceptual model to a set of equations defining a conceptual model. Choices must be made at this point to simplify the description and, as we will see, such choices have not always had a good foundation in hydrological reality.

Consider, briefly, one hydrologist's perceptual model. It is based on an outline set out in Beven (1991a), with some revision based on additional experience since then. In recession periods between storms, storage in the soil and rock gradually declines (Figure 1.3a). If there is a water table, the level and gradient will gradually fall. Storage will often be higher and water tables closer to the surface in the valley bottom *riparian areas*, partly because of downslope flow, particularly where there is convergence of flow in hillslope hollows. Storage in riparian areas may be maintained by return flows from deeper layers (e.g. Huff *et al.*, 1982; Genereux *et al.*, 1993), but also because soils tend to be deeper in valley bottoms (e.g. Dietrich *et al.*, 1995; Piñol *et al.*, 1997). Loss of water by evapotranspiration will have a greater or lesser effect on the profile of storage depending on season, climate and vegetation type and rooting depth. Many plants, however, may extract water from considerable depth with roots penetrating up to tens of meters into the soil and bedrock fractures and root channels also acting as pathways for infiltrating water (for example the Jarrah trees of Western Australia). Plants that are *phreatophytes* (such as the Cottonwoods of the western United States) will extract water directly from beneath the water table. These evapotranspiration and drainage processes will be important in controlling the *antecedent conditions* prior to a storm event.

The antecedent conditions, as well as the volume and intensity of rainfall (or snowmelt), will be important in governing the processes by which a catchment responds to rainfall and the proportion of the input volume that appears in the stream as part of the hydrograph (Figure 1.3b). Unless the stream is ephemeral, there will always be a response from precipitation directly onto the channel and immediate riparian area. This area, although a relatively small area of the catchment (perhaps 1–3%), may be an important contributor to the hydrograph in catchments and storms with low runoff coefficients. Even in *ephemeral streams*, surface flow will often start first in the stream channels. The extent of the channel network will generally expand into headwater areas as a storm progresses and will be greater during wet seasons than dry (e.g. Hewlett, 1974).

Rainfalls and snowmelt inputs are not spatially uniform, but can show rapid changes in intensity and volume over relatively short distances, particularly in convective events (e.g. Newson, 1980; Smith *et al.*, 1996; Goodrich *et al.*, 1997). The variability at ground level, after the pattern of intensities has been affected by the vegetation canopy, may be even greater. Some of the rainfall will fall directly to the ground as direct *throughfall*. Some of the rainfall will be intercepted and evaporated from the canopy back to the atmosphere. Some evaporation of intercepted water may occur even during events, especially from rough canopies under windy conditions, when the air is not saturated with vapour. Differences of up to 30% between incident rainfall and throughfall have been measured in a Mediterranean catchment (Llorens *et al.*, 1997). The remaining rainfall will drip from the vegetation of the canopy as throughfall or run down the branches, trunks and stems as *stemflow*. The latter process may be important since, for some canopies, 10% or more of the incident rainfall may reach the ground as stemflow resulting in local concentrations of water at much higher intensities than the incident rainfall. Some plants, such as maize, have a structure designed to channel water to their roots in this way.

Snowmelt rates will vary with elevation and aspect in that they affect the air temperature and radiation inputs to the snowpack. The water equivalent of the snowpack can vary dramatically in space, due particularly to the effects of wind drifting during snow events and after the snowpack has formed, as

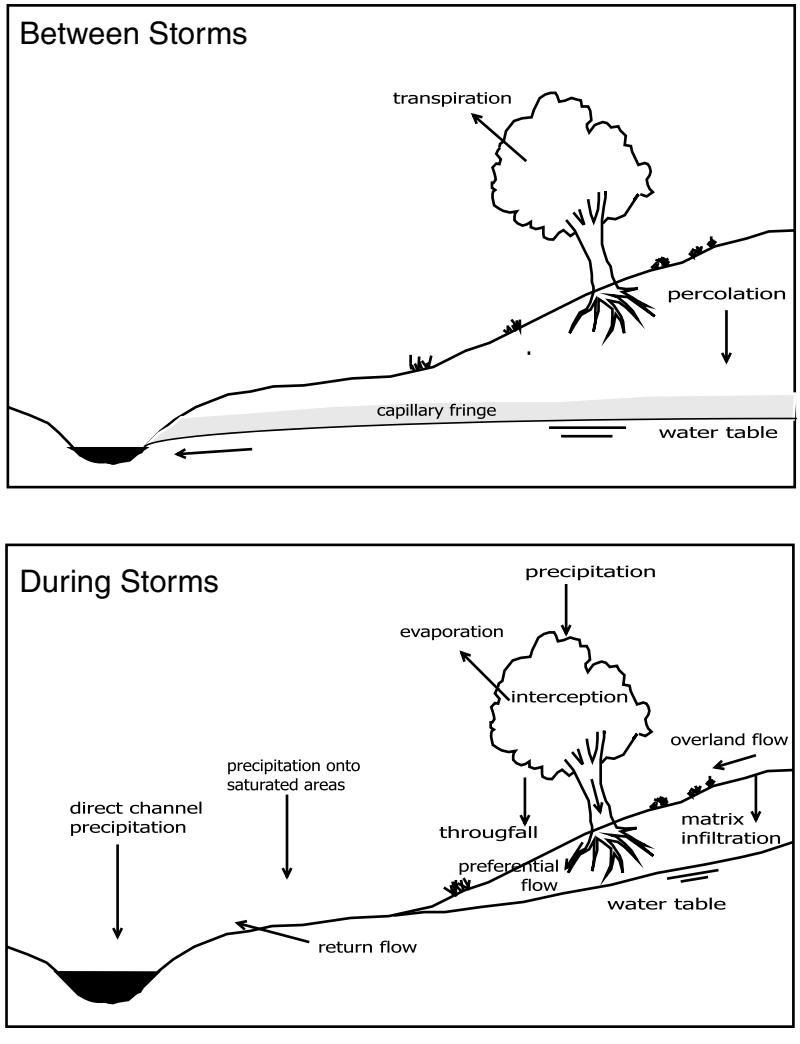


Figure 1.3 *The processes involved in one perceptual model of hillslope hydrology (after Beven, 1991a).*

affected by the topography and vegetation cover. Much deeper snow will often be found in the lee of ridges, a feature that has been well documented in the Reynolds Creek catchment in Idaho and elsewhere (see Bathurst and Cooley, 1996 and Section 5.3). There can also be feedback effects, in that deeper snow cover might lead to more water being available to the vegetation leading to greater growth and, in the case of trees, to greater trapping of snow drifting in the wind.

Once the rain or snowmelt water has reached the ground, it will start to infiltrate the soil surface, except on impermeable areas of bare rock, areas of completely frozen soil, or some man-made surfaces where *surface runoff* will start almost immediately. The rate and amount of infiltration will be limited by the rainfall intensity and the *infiltration capacity* of the soil. Where the input rate exceeds the infiltration capacity of the soil, *infiltration excess overland flow* will be generated. Soils tend to be locally

heterogeneous in their characteristics, so that infiltration capacities and rates of overland flow generation might vary greatly from place to place (Loague and Kyriakidis, 1997). In many places, particularly on vegetated surfaces, rainfalls only very rarely exceed the infiltration capacity of the soil unless the soil becomes completely saturated. Elsewhere, where infiltration capacities are exceeded, this will start in areas where soil permeabilities are lowest and, since infiltration capacities tend to decrease with inceased wetting, will gradually expand to areas with higher permeability. Bare soil areas will be particularly vulnerable to such *infiltration excess* runoff generation since the energy of the raindrops can rearrange the soil particles at the surface and form a surface crust, effectively sealing the larger pores (see Römkens *et al.*, 1990; Smith *et al.*, 1999). A vegetation or litter layer, on the other hand, will protect the surface and also create root channels that may act as pathways for infiltrating water. Bare surfaces of dispersive soil materials are particularly prone to crusting and such crusts, once formed, will persist between storms unless broken up by vegetation growth, freeze–thaw action, soil faunal activity, cultivation or erosion. Studies of crusted soils have shown that, in some cases, infiltration rates after ponding might increase over time more than would be expected as a result of the depth of ponding alone (see, e.g., Fox *et al.*, 1998). This was thought to be due to the breakdown or erosion of the crust.

In cold environments, the vegetation may also be important in controlling the degree to which a soil becomes frozen before and during the build up of a snowpack by controlling both the local energy balance of the soil surface and the drifting of snow cover. This may have important consequences for the generation of runoff during snowmelt, even though it may, in some cases, take place months later (Stadler *et al.*, 1996). It is worth noting that a frozen topsoil is not necessarily impermeable. There will usually be some reduction in potential infiltration rates due to freezing, but seasonal freeze–thaw processes can also lead to the break-up of surface crusts so as to increase infiltration capacities (Schumm, 1956). The effects of freezing will depend on the moisture content of the soil and the length of the cold period. Even where widespread freezing takes place, infiltration capacities may be highly variable.

It has long been speculated that during widespread surface ponding there could be a significant effect on infiltration rates of air entrapment and pressure build up within the soil. This has been shown to be the case in the laboratory (Wang *et al.*, 1998) and, in a smaller number of studies, in the field (Dixon and Linden, 1972). It has also been suggested that air pressure effects can cause a response in local water tables (e.g. Linden and Dixon, 1973) and that the lifting force due to the escape of air at the surface might be a cause of initiation of motion of surface soil particles. The containment of air will be increased by the presence of a surface crust of fine material but significant air pressure effects would appear to require ponding over extensive areas of a relatively smooth surface. In the field, surface irregularities (such as vegetation mounds) and the presence of macropores might be expected to reduce the build up of entrapped air by allowing local pathways for the escape of air to the surface.

In the absence of a surface crust, the underlying soil structure, and particularly the macroporosity of the soil, will be an important control on infiltration rates. Since discharge of a laminer flow in a cylindrical channel varies with the fourth power of the radius, larger pores and cracks may be important in controlling infiltration rates (Beven and Germann, 1981). However, soil cracks and some other *macropores*, such as earthworm channels and ant burrows, may only extend to limited depths so that their effect on infiltration may be limited by storage capacity and infiltration into the surrounding matrix as well as potential maximum flow rates. An outlier in the data on flow rates in worm holes of Ehlers (1975), for example, was caused by a worm still occupying its hole! The effects of such macropores on hillslope response might still, however, be significant, even in wet humid temperate environments (Marshall *et al.*, 2009). Some root channels, earthworm and ant burrows can also extend to depths of meters below the surface. The Jarrah trees of Western Australia are again a particularly remarkable example.

Overland flow may also occur as a *saturation excess* mechanism. Areas of saturated soil tend to occur first where the antecedent *soil moisture deficit* is smallest. This will be in valley bottom areas, particularly headwater hollows where there is convergence of flow and a gradual decline in slope towards the stream. Saturation may also occur on areas of thin soils, where storage capacity is limited, or in low permeability

and low slope areas, which will tend to stay wet during recession periods. The area of saturated soil will tend to expand with increased wetting during a storm and reduce again after rainfall stops at a rate controlled by the supply of water from upslope. This is the *dynamic contributing area* concept. Any surface runoff on such a saturated area may not all be due to rainfall but may also be due to a return flow of subsurface water. In this way, surface runoff may be maintained during the period after rainfall has stopped. When overland flow is generated, by whatever mechanism, some surface *depression storage* may need to be satisfied before there is a consistent downslope flow. Even then, surface flow will tend to follow discrete pathways and rills rather than occurring as a sheet flow over the whole surface.

A similar concept may be invoked in areas where responses are controlled by subsurface flows. When saturation starts to build up at the base of the soil over a relatively impermeable bedrock, it will start to flow downslope. The connectivity of saturation in the subsurface will, however, initially be important. It may be necessary to satisfy some initial bedrock depression storage before there is a consistent flow downslope. The dominant flow pathways may be localised, at least initially, related to variations in the form of the bedrock surface (McDonnell *et al.*, 1996). Some catchments, with high infiltration capacities and reasonably deep soils, may have responses dominated by *subsurface stormflow*. It is worth remembering that a 1 m depth of soil, with an average porosity of 0.4 has a storage capacity of 400 mm of water. Thus, if the infiltration capacity of the soil is not exceeded, a large 100 mm rainstorm could, in principle, be totally absorbed by that 1 m soil layer (ignoring the effects of any downslope flows), even if the antecedent storage deficit is only a quarter of the porosity. It has further been suggested that a certain degree of antecedent wetness is required before some threshold of connectivity is reached and significant downslope flow is achieved (the "fill and spill" hypothesis, see Tromp van Meerveld and McDonnell, 2006). Some soils are susceptible to piping for both natural and anthropogenic (field drainage) reasons. In the right conditions, such pipes can provide rapid conduits for downslope flows (see Jones, 2010; Chappell, 2010).

It is a common (and very convenient) assumption that the bedrock underlying small upland catchments is impermeable. This is not always the case, even in rocks that have little or no primary permeability in the bulk matrix. The presence of secondary permeability in the form of joints and fractures can provide important flow pathways and storage that may be effective in maintaining stream baseflows over longer periods of time. It is very difficult to learn much about the nature of such pathways; any characteristics are often inferred from the nature of recession curves and the geochemistry of baseflows since the bedrock can provide a different geochemical environment and long residence times can allow weathering reactions to provide higher concentrations of some chemical consitutuents (see, for example, the study of Neal *et al.*, 1997 in the Plynlimon research catchments in Wales).

There is an interesting possibility that connected fracture systems that are full of water could act as pipe systems, transmitting the effects of recharge very rapidly. Remember that if water is added to one end of a pipe full of water, there will be an almost instantaneous displacement of water out of the other end, whatever the length of the pipe and even if the velocities of flow in the pipe are relatively slow. The reason is that the transmission of the pressure effect of adding the water is very much faster than the actual flow velocity of the water. Such displacement effects are an explanation of rapid subsurface responses to storm rainfalls (see Section 1.5).

The perceptual model briefly outlined above represents a wide spectrum of possible hydrological responses that may occur in different environments or even in different parts of the same catchment at different times. Traditionally, it has been usual to differentiate between different conceptualisations of catchment response based on the dominance of one set of processes over another, for example, the *Hortonian model* in which runoff is generated by an infiltration excess mechanism all over the hillslopes (Figure 1.4a). This model is named after Robert E. Horton (1875–1945), the famous American hydrologist (he may be the only modern hydrologist to have a waterfall named after him) who worked as both hydrological scientist and consultant. I am not sure that he would have totally approved of such widespread use of the infiltration excess concept. Although he frequently used the infiltration excess concept as a

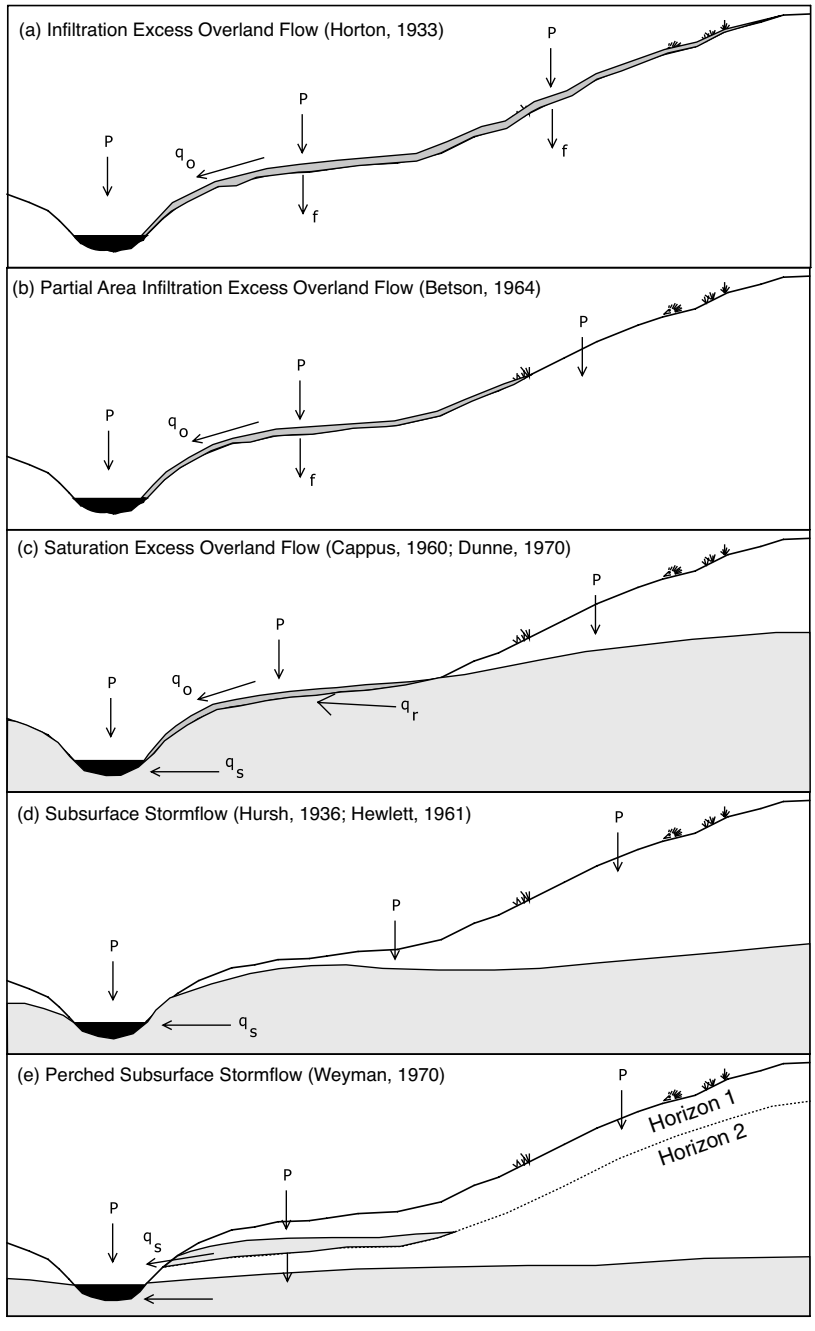


Figure 1.4 *A classification of process mechanisms in the response of hillslopes to rainfalls: (a) infiltration excess overland flow (Horton, 1933); (b) partial area infiltration excess overland flow (Betson, 1964); (c) saturation excess overland flow (Cappus, 1960; Dunne and Black, 1970); (d) subsurface stormflow (Hursh; Hewlett); (e) perched saturation and throughflow (Weyman, 1970).*

way of calculating the volume of runoff production from a particular rainfall (Horton, 1933), he also had a hydrological laboratory in his wooded back garden in Voorheesville, New York State (Horton, 1936) where he would surely not have observed infiltration excess overland flow very often. Beven (2004a), using data collected by Horton on infiltration capacities and local rainfall statistics, has suggested that overland flow might have occurred on the LaGrange Brook catchment only once in every 2–5 years. There is also some experimental evidence that suggests that some of the overland flow collected on runoff plots at the site might have been generated by a return flow mechanism (Beven, 2004a). Horton was an excellent scientist who published papers on a wide variety of hydrological and meteorological phenomena (see Box 1.1). His perceptual model of infiltration was different from the idea that infiltration is controlled by the hydraulic gradients within the soil profile. He thought that conditions at the surface were much more important and recognised seasonal effects due to cultivation, the redistribution of particles by rainsplash blocking larger pores, and the irregularity of the surface in allowing air to escape. His perceptual model surely involved a much wider range of processes than the model that now bears his name (see, for example, the work of Horton from 1936 and his process descriptions from 1942; see also the summary of some of his archived papers by Beven, 2004b, 2004c).

In the same period as Horton, however, Charles R. Hursh was working in the Coweeta watersheds in Georgia in the United States. These Southern Appalachian catchments are forested with soils that are deeply weathered and have generally high infiltration capacities. Surface runoff is restricted mainly to the channels, so here the storm runoff production must be controlled by subsurface responses (Figure 1.4d). Hursh published a number of articles dealing with subsurface responses to rainfall (see, for example, Hursh and Brater, 1941). A later director of the Coweeta laboratory, John Hewlett, was also influential in getting the importance of subsurface stormflow more widely recognised in the 1960s (Hewlett and Hibbert, 1967; Hewlett, 1974).

Independently in the 1960s, studies within the Tennessee Valley Authority (which at that time served as one of the major hydrological agencies in the eastern United States) were revealing that it was very difficult to predict runoff production in many catchments under the assumption that infiltration excess surface runoff was produced everywhere on the hillslopes. The information on infiltration capacities of the soils and rainfall rates could not support such a model. Betson (1964) suggested that it would be more usual that only part of a catchment would produce runoff in any particular storm and that, since infiltration capacities tend to decrease with increasing soil moisture and the downslope flow of water on hillslopes tends to result in wetter soils at the base of hillslopes, the area of surface runoff would tend to start close to the channel and expand upslope. This *partial area model* (Figure 1.4b) allowed for a generalisation of the Horton conceptualisation. It is now realised that the variation in overland flow velocities and the heterogeneities of soil characteristics and infiltration rates are important in controlling partial area responses. If runoff generated on one part of a slope flows onto an area of higher infiltration capacity further downslope it will infiltrate (the "run-on" process). If the high intensity rainfall producing the overland flow is of short duration, then it is also possible that the water will infiltrate before it reaches the nearest rill or stream channel. Bergkamp (1998), for example, estimated that for some plot scale experiments with artificial rainfalls at an intensity of 70 mm/h, the average travel distance for overland flow was of the order of 1 m!

Another form of partial area response was revealed by studies in a different environment by Dunne and Black (1970) working in Vermont. They observed surface runoff production, but on soils with high surface infiltration capacities. The surface runoff resulted from a saturation excess mechanism (Figure 1.4c), a type of response that had been previously suggested by Cappus (1960), working in France (and published in French).

These four major conceptualisations are all subsets of the more general perceptual model outlined previously. We now know that infiltration excess, saturation excess or purely subsurface responses might all occur in the same catchment at different times or different places due to different antecedent conditions, soil characteristics or rainfall intensities. In addition, an infiltration excess mechanism might take place

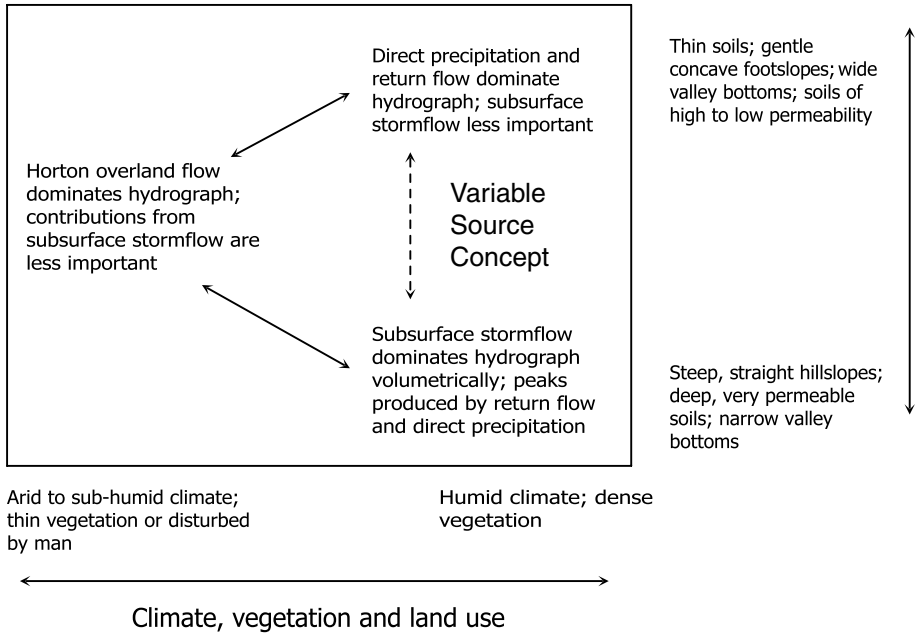


Figure 1.5 *Dominant processes of hillslope response to rainfall (after Dunne, 1978, with kind permission of Wiley-Blackwell).*

within the soil at a permeability break, perhaps associated with a horizon boundary. This might lead to the generation of a perched water table and even to saturation at the surface of a soil that may be unsaturated at depth (see Figure 1.4e; Weyman, 1970). Attempts have been made to suggest which mechanisms might be dominant in different environments (see Figure 1.5) but there may still be much to learn from direct observations of runoff processes in a catchment of interest. A selection of classic papers on runoff generation processes has been published by Beven (2006d), while Smakhtin (2002) has reviewed early process studies in Russia and Bonell (2004) gives a summary of work in tropical catchments. An interesting contribution is made by McGlynn *et al.* (2002) who discuss the way in which research has changed the perceptual model of runoff processes in the small Maimai catchment in New Zealand.

1.5 Flow Processes and Geochemical Characteristics

One of the most influential factors in revising hydrological thinking in the last 30 years has been the use of geochemical characteristics to provide additional information on flow processes. Some characteristics, in particular the use of artificial tracers, can provide direct information on flow velocities; others, such as the various environmental tracers, require a greater degree of inference. Even the results of artificial tracers may be difficult to interpret since most tracers are not ideal for following water movement over the wide range of time scales involved and it is difficult to apply artificial tracers at the catchment scale. Thus, any experiment will tend to sample only some of the possible flow pathways.

The environmental isotopes of oxygen and hydrogen are often used in catchment scale studies (see the review by Sklash, 1990). They have the advantage that they are part of the water molecule and will therefore follow the flow pathways of water in the catchment directly. There remain some difficulties of interpretation of the results due to spatial and temporal variations in the concentrations of the isotopes

in the rainfall inputs, the effects of vegetation on the input concentrations, and the spatial variability of concentrations of water stored in different soil horizons and parts of the catchment. However, at least in ideal conditions when there is a strong difference between the concentrations observed in rainfalls and the concentrations of water stored in the catchment before an event, the measured concentrations can be used in a simple two-component mixing model to differentiate between the contribution to the hydrograph for an event of the rainfall and the contribution of the water stored in the catchment prior to the event.

Some of the first hydrograph separations of this type were published by Crouzet *et al.* (1970) based on using tritium derived from atomic testing as a tracer (see also Stewart *et al.*, 2010) and revealed that the contribution of stored water (often called the "pre-event" or "old" water component) was surprisingly high (Figure 1.6). This result has been confirmed by many other studies for a wide range of different catchments

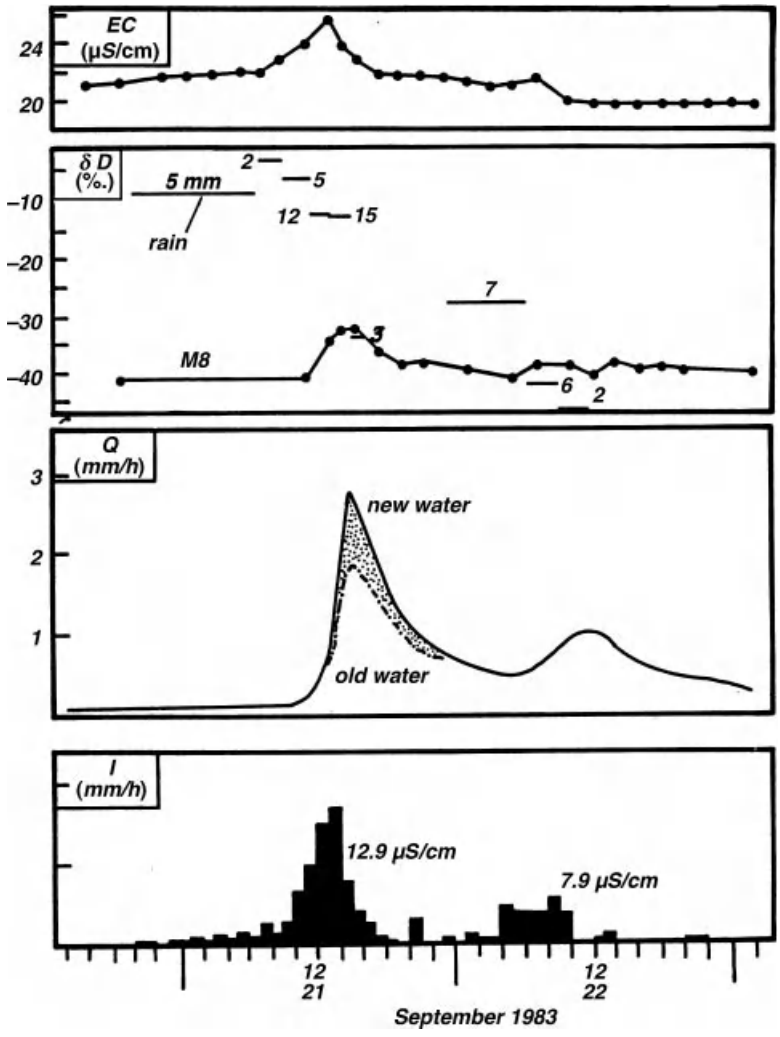


Figure 1.6 *Hydrograph separation based on the concentration of environmental isotopes (after Sklash, 1990, with kind permission of Wiley-Blackwell).*

and tracers, although the number of reports is dominated by results from humid temperate catchments and small to moderate rainfall events (Sklash, 1990). The study of Sklash and Farvolden (1979) is particularly interesting in showing that samples from surface runoff at a point were sometimes dominated by "old" water and sometimes by "event" water. The technique can be extended, using other environmental tracers, to three-component mixing to differentiate the rainfall contribution from "soil water" and "deep groundwater" components, where these components can be differentiated geochemically (see Bazemore *et al.*, 1994). Again, a major component of pre-event water is often found even, in some cases, for very rapidly responding processes such as pipe flows in wet soils (Sklash *et al.*, 1996).

The pre-event water is displaced from storage by the effects of the incoming rainfall. This must therefore necessarily involve subsurface flow processes. The fact that the rising limb of the hydrograph is often dominated by the pre-event water component reveals that this displacement can take place rapidly, despite the fact that subsurface flow velocities are generally assumed to be much slower than surface flow velocities. This perception is, in fact, one of the reasons for the continuing use of the Hortonian conceptualisation of runoff production, even now. If subsurface velocities are so slow, how can subsurface flow and pre-event waters make a major contribution to the hydrograph (Kirchner, 2003)?

The answer lies, at least in part, in the physics of the flow processes; in particular, in the saturated zone. It can be shown that there is a difference between the flow velocity of water and the velocity with which a disturbance to the saturated zone is propagated as a pressure wave, which is called the *wave speed* or *celerity*. The type of disturbance of interest here is the addition of recharge due to rainfall during an event. The theory suggests that an infinitely small disturbance at the water table will be propagated infinitely quickly. Larger disturbances will have a much smaller wave velocity, the magnitude of which is a function of the inverse of the *effective storage capacity* in the soil (the difference between the current soil moisture in the soil immediately above the water table and saturation). In a wet soil or close to a water table, the effective storage capacity may be very small so that the wave velocity may be very much faster than the actual flow velocity of the water (see Section 5.5.3). The increase in discharge to the stream during an event will then depend more on the response of the hydraulic potentials in the system, which will be controlled by the local wave velocities, than the actual flow velocities of the water. Thus if discharge starts to increase before the recharging water has had time to flow towards the channel, it will be water stored in the profile close to the stream that flows into the channel first. This water will be predominantly pre-event water, displaced by the effects of the rainfall. There may also be local exchanges between event water and pre-event waters that cause displacements into local surface runoff with higher velocities (Iorgulescu *et al.*, 2007).

Similar effects may take place in unsaturated soil, but here the picture or perception is made more complicated by the relative mobility of water stored in different parts of the pore space and by the effects of preferential flows within the structural voids of the soil. The important message to take from this section is that in many catchments, particularly in humid environments, an important part of the hydrograph may be made up of "old" water and may not be rainfall flowing directly to the stream. Certainly, it should not be assumed that fast runoff is always the result of overland flow or surface runoff on the hillslopes of a catchment. A more extensive discussion of the identification of runoff sources and modelling of residence times in catchments is to be found in Chapter 11.

1.6 Runoff Generation and Runoff Routing

The evidence discussed in the previous two sections has been primarily concerned with the processes of runoff generation, both surface and subsurface. Runoff generation controls how much water gets into the stream and flows towards the catchment outlet within the time frame of the storm and the period immediately following. There is also, however, a further component to consider, which is the routing of the runoff from the source areas to the outlet. The boundary between runoff generation and runoff routing

is not a very precise one. It would be more precise if it was possible to measure or predict the timing of inflows into the stream network itself accurately. Then the routing would only have to worry about the flow processes within the stream which can be reasonably well predicted on the basis of hydraulic principles (although in arid areas it may also be necessary to take account of the infiltration of some of the water into the stream bed). Unfortunately, it is generally not possible to predict the volume and timing of the inflows precisely, so the routing problem becomes one of the velocities of surface and subsurface flows on the hillslope as well as in the stream channel. It may be very difficult then to separate out the effects of the different possible flow pathways that different waters take on the timing of the hydrograph at the stream outlet.

However, every hydrological model requires two essential components: one to determine how much of a rainfall becomes part of the storm hydrograph (the *runoff generation* component), the other to take account of the distribution of that runoff in time to form the shape of the storm hydrograph (the *runoff routing* component). These two components may appear in many different guises and degrees of complexity in different models but they are always there in any rainfall–runoff model, together with the difficulty of clearly separating one component from the other.

In general, it is accepted that the runoff generation problem is the more difficult. Practical experience suggests that the complexities and nonlinearities of modelling the flow generation processes are much greater than for the routing processes and that relatively simple models for the routing may suffice (see discussion in Section 2.2).

1.7 The Problem of Choosing a Conceptual Model

The majority of hydrologists will be model users rather than model developers. Having said that, there has been no shortage of hydrologists, particularly those undertaking research for a doctorate, who have set themselves the task of developing a model. This is understandable; even now, the obvious approximation inherent in today's models suggests that it should be possible to do better! However, given the range of models consequently available in the literature or, increasingly, as modelling software packages, the problem of model choice is not so different for the model user as for a researcher wanting to develop a new and improved version. The question is how to decide what is satisfactory and what are the limitations of the models available. We will take a preliminary look at this question in this section and return to it in Chapter 12.

Let us first outline the "generic" choices in terms of a basic classification of model types. There are many different ways of classifying hydrological models (see, for example, Clarke, 1973; O'Connell, 1991; Wheater *et al.*, 1993; Singh, 1995). We concentrate on a very basic classification here. The first choice is whether to use a *lumped* or *distributed* modelling approach. Lumped models treat the catchment as a single unit, with state variables that represent averages over the catchment area, such as average storage in the saturated zone. Distributed models make predictions that are distributed in space, with state variables that represent local averages of storage, flow depths or hydraulic potential, by discretising the catchment into a large number of elements or grid squares and solving the equations for the state variables associated with every element grid square. Parameter values must also be specified for every element in a distributed model. There is a general correspondence between lumped models and the "explicit soil moisture accounting" (ESMA) models of O'Connell (1991) (see Section 2.4), and between distributed models and "physically-based" or process-based models. Even this correspondence is not exact, however, since some distributed models use ESMA components to represent different subcatchments or parts of the landscape as *hydrological response units* (see Section 6.2). Distributed models currently available must use average variables and parameters at grid or element scales greater than the scale of variation of the processes and are consequently, in a sense, lumped conceptual models at the element scale (Beven, 1989a, 2006b). There is also a range of models that

do not make calculations for every point in the catchment but for a distribution function of characteristics. TOPMODEL, discussed in Chapter 6, is a model of this type, but has the feature that the predictions can be mapped back into space for comparison with any observations of the hydrological response of the catchment. It could therefore be called, perhaps, a semi-distributed model. Chapter 9 discusses the next generation of hydrological models, which will, at least in part, start to resolve some of these issues.

A second consideration is whether to use a *deterministic* or a *stochastic* model. Deterministic models permit only one outcome from a simulation with one set of inputs and parameter values. Stochastic models allow for some randomness or uncertainty in the possible outcomes due to uncertainty in input variables, boundary conditions or model parameters. The vast majority of models used in rainfall–runoff modelling are used in a deterministic way, although again the distinction is not clear cut since there are examples of models which add a stochastic error model to the deterministic predictions of the hydrological model and there are models that use a probability distribution function of state variables but make predictions in a deterministic way. A working rule is that if the model output variables are associated with some variance or other measure of predictive dispersion the model can be considered stochastic; if the output values are single valued at any time step, the model can be considered deterministic, regardless of the nature of the underlying calculations.

There is one other modelling strategy, based on *fuzzy logic* methods, that looks highly promising for the future. The number of fuzzy models is currently few (see, for example, Bárdossy *et al.*, 1995; Hundecha *et al.* 2001; Özelkan and Duckstein, 2001) but the range of application would appear to be large. In particular fuzzy models would appear to offer the potential for a more direct translation from the complexity of the perceptual model into a procedural model. Applications to date, however, have often used an intermediate conceptual model to formulate the fuzzy rules and have defuzzified the results so as to run as essentially deterministic solutions.

So, these are the broad generic classes of rainfall–runoff model. Within each class there is a range of possible model structures. How then to go about choosing a particular model structure for a particular application? The following suggested procedure is based, in essence, on considerations of the **function** of possible modelling structures:

1. Prepare a list of the models under consideration. This list may have two parts: those models that are readily available and those that might be considered for a project if the investment of time (and money!) appears to be worthwhile.
2. Prepare a list of the variables predicted by each model. Decide whether the model under consideration will produce the outputs needed to meet the aims of a particular project. If you are interested in the rise in the water table in valley bottoms due to deforestation, for example, a model predicting the lumped response of the catchment may not fulfill the needs of the project. If, however, you are only interested in predicting the discharge response of a catchment for real-time flood forecasting, then it may not be necessary to choose a distributed modelling strategy.
3. Prepare a list of the assumptions made by the model (see the guides in the chapters that follow). Are the assumptions likely to be limiting in terms of what you know about the response of the catchment you are interested in (your perceptual model)? Unfortunately the answer is likely to be yes for all models, so this assessment will generally be a relative one, or at best a screen to reject those models that are obviously based on incorrect representations of the catchment processes (i.e. any reasonable hydrologist should not try to use a model based on Hortonian overland flow to simulate the Coweeta catchments mentioned in Section 1.4).
4. Make a list of the inputs required by the model, for specification of the flow domain, the boundary and initial conditions and the parameter values. Decide whether all the information required can be provided within the time and cost constraints of the project.

5. Determine whether you have any models left on your list. If not, review the previous steps, relaxing the criteria used. If predictions are really required for an application, one model at least will need to be retained at this stage!!

1.8 Model Calibration and Validation Issues

Once one or more models have been chosen for consideration in a project, it is necessary to address the problem of parameter calibration. It is unfortunate that it is not, in general, possible to estimate the parameters of models by either measurement or prior estimation. Studies that have attempted to do so have generally found that, even using an intensive series of measurements of parameter values, the results have not been entirely satisfactory (Beven *et al.*, 1984; Refsgaard and Knudsen, 1996; Loague and Kyriakidis, 1997). Prior estimation of feasible ranges of parameters also often results in ranges of predictions that are wide and may still not encompass the measured responses all of the time (Parkin *et al.*, 1996; Bathurst *et al.*, 2004).

There are two major reasons for these difficulties in calibration. The first is that the scale of the measurement techniques available is generally much less than the scale at which parameter values are required. For example, there may be a hydraulic conductivity parameter in a particular model structure. Techniques for measuring hydraulic conductivities of the soil generally integrate over areas of less than $1\,m^2$. Even the most finely distributed models however, require values that effectively represent the response of an element with an area of $100\,m^2$ or, in many cases, much larger. For saturated flow, there have been some theoretical developments that suggest how such effective values might change with scale, given some underlying knowledge of the fine scale structure of the conductivity values. In general, however, obtaining the information required to use such a theory at the hillslope or catchment scale would be very time consuming and expensive and would result in a large number of holes in the hillslope! Thus it may be necessary to accept that the small scale values that it is possible to measure and the effective values required at the model element scale are different quantities (a technical word is that they are *incommensurate*) – even though the hydrologist has given them the same name. The effective parameter values for a particular model structure still need to be calibrated in some way. It is also often the case that the time and space scales of model-predicted variables may be different from the scale at which variables of the same name can be measured (for example, soil water content). In this case, the variables used in calibration may also be incommensurate.

Most calibration studies in the past have involved some form of *optimisation* of the parameter values by comparing the results of repeated simulations with whatever observations of the catchment response are available. The parameter values are adjusted between runs of the model, either manually by the modeller or by some computerised optimisation algorithm until some "best fit" parameter set has been found. There have been many studies of optimisation algorithms and measures of goodness of fit or *objective functions* in hydrological modelling (see Chapter 7). The essence of the problem is to find the peak in the *response surface* in the *parameter space* defined by one or more objective functions. An example of such a response surface is shown in Figure 1.7. The two basal axes are two different parameter values, varied between specified maximum and minimum values. the vertical axis is the value of an objective function, based on the sum of squared differences between observed and predicted discharges, that has the value 1 for a perfect fit. It is easy to see why optimisation algorithms are sometimes called "hill climbing" algorithms in this example, since the highest point on the surface will represent the optimum values of the two parameters. Such a response surface is easy to visualise in two-parameter space. It is much more difficult to visualise the response surface in an N-dimensional parameter hyperspace. Such surfaces can often be very complex and much of the research on optimisation algorithms has been concerned with finding algorithms that are robust with respect to the complexity of the surface in an N-dimensional space and find the *global optimum* set of parameter values. The complexity of the surface might also depend on

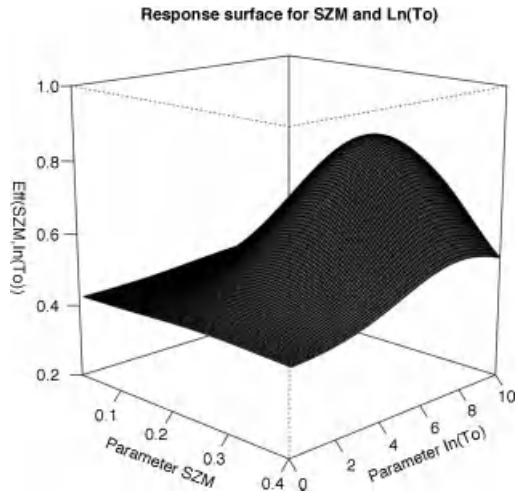


Figure 1.7 *Response surface for two TOPMODEL parameters (see Chapter 6) in an application to modelling the stream discharge of the small Slapton Wood catchment in Devon, UK; the objective function is the Nash–Sutcliffe efficiency that has a value of 1 for a perfect fit of the observed discharges.*

the nature of the model equations, especially if there are thresholds involved, and the correct numerical integration of the equations in time (Kavetski and Clark, 2010).

However, for most hydrological modelling problems, the optimisation problem is ill-posed in that if the optimisation is based on the comparison of observed and simulated discharges alone, there may not be enough information in the data to support the robust optimisation of the parameter values. Experience suggests that even a simple model with only four or five parameter values to be estimated may require at least 15 to 20 hydrographs for a reasonably robust calibration and, if there is strong seasonal variability in the storm responses, a longer period still (see, for example, Kirkby, 1975; Gupta and Sorooshian, 1985; Hornberger *et al.*, 1985; Yapo *et al.*, 1996). For more complex parameter sets, much more data and different types of data may be required for a robust optimisation unless many of the parameters are fixed beforehand.

These are not the only problems with finding an optimum parameter set. Optimisation generally assumes that the observations with which the simulations are compared are error free and that the model is a true representation of that data. We know, however, at least for hydrological models, that both the model structure and the observations are not error free. Thus, the optimum parameter set found for a particular model structure may be sensitive to small changes in the observations, to the period of observations considered in the calibration, to the way in which the model predictions are evaluated, and possibly to changes in the model structure (such as a change in the element discretisation for a distributed model).

There may also be an issue about whether all the observed data available for use in a model calibration exercise are useful. Where, for example, there is an inconsistency between measured rainfalls and measured discharges, or incommensurability between observed and predicted soil moisture variables, then not all the observed data may be informative (Beven and Westerberg, 2011). The issue of identifying *disinformation* in calibration data does not, however, appear to have received much attention. In some cases, inconsistencies might be obvious, such as hydrographs that are recorded when no rainfall has been measured in any of the rain gauges in a catchment or, more generally, when runoff coefficients appear to be more than 100%. However, heavy rainfalls observed in one or more raingauges

with no hydrograph response might be more ambiguous: it could be hydrologically consistent even if difficult to model and thus be a result of model structure error rather than disinformative data. Even widely used data sets might show inconsistencies and introduce disinformation into the inference process. An example is provided by the Leaf River catchment data that has been used in many model calibration studies. Beven (2009b) suggested that there might be a period of inconsistent data in the validation period used in Vrugt *et al.* (2009). We return to the issue of disinformation in calibration in Chapter 7.

There are a number of important implications that follow from these considerations:

- The parameter values determined by calibration are effectively valid only inside the model structure used in the calibration. It may not be appropriate to use those values in different models (even though the parameters may have the same names) or in different catchments.
- Care should be taken not to include inconsistent or disinformative data in the calibration process as this will lead to biased estimates of parameter values. This is one example of a more general problem of errors and uncertainties in the modelling process that result from a lack of knowledge rather than statistically random errors.
- The concept of an optimum parameter set may be ill-founded in hydrological modelling. While one optimum parameter set can often be found there will usually be many other parameter sets that are very nearly as good, perhaps from very different parts of the parameter space. It is highly likely that, given a number of parameter sets that give reasonable fits to the data, the ranking of those sets in terms of the objective function will be different for different periods of calibration data. Thus to decide that one set of parameter values is the optimum is then a somewhat arbitrary choice. Some examples of such behaviour will be seen in the "dotty plots" of Chapter 7 where the possibility of rejecting the concept of an optimum parameter set in favour of a methodology based on the *equifinality* of different model structures and parameter sets will be considered.
- If the concept of an optimum parameter set must be superceded by the idea that many possible parameter sets (and perhaps models) may provide acceptable simulations of the response of a particular catchment, then it follows that *validation* of those models may be equally difficult. In fact, rejection of some of the acceptable models given additional data may be a much more practical methodology than suggesting that models might be validated.

The idea of equifinality is an important one in what follows, particularly from Chapter 7 onwards. It suggests that, given the limitations of both our model structures and the observed data, there may be many representations of a catchment that may be equally valid in terms of their ability to produce acceptable simulations of the available data. In essence then, different model structures and parameter sets used within a model structure are competing to be considered acceptable as simulators. Models can be treated and tested as hypotheses about how the catchment system is functioning in this sense (see Beven, 2010a). Some may be rejected in the evaluation of different model structures suggested in Section 1.7; even if only one model is retained, the evaluation of the performance of different parameter sets against the observed data will usually result in many parameter sets that produce acceptable simulations.

The results with different parameter sets will not, of course, be identical either in simulation or in the predictions required by the modelling project. An optimum parameter set will give only a single prediction. Multiple acceptable parameter sets will give a range of predictions. This may actually be an advantage since it allows the possibility of assessing the uncertainty in predictions, as conditioned on the calibration data, and then using that uncertainty as part of the decision making process arising from a modelling project. A methodology for doing this is outlined in Chapter 7.

The rest of this book builds upon this general outline of the modelling process by considering specific examples of conceptual model and their application within the context of the types of evaluation procedure,

for both model structures and parameter sets as hypotheses, outlined above. We take as our starting point that, *a priori*, all the available modelling strategies and all feasible parameter sets within those modelling strategies are potential models of a catchment for a particular project. The aims of that project, the budget available for the project, and the data available for calibrating the different models will all limit that potential range of simulators. The important point is that choices between models and between parameter sets must be made in a logical and scientifically defensible way allowing for data uncertainties and limitations. It is suggested, however, that at the end of this process, there will not be a single model of the catchment but a number of acceptable models (even if only different parameter sets are used within one chosen model structure) to provide predictions.

There are clearly implications for other studies that depend on models of rainfall–runoff processes. Predictions of catchment hydrogeochemistry, sediment production and transport, the dispersion of contaminants, hydroecology, and, in general, integrated catchment decision support systems depend crucially on good predictions of water flow processes. To keep my task manageable, I have not chosen to address the vast literature that deals with these additional topics but each additional component that is added to a modelling system will add additional choices in terms of the conceptual representation of the processes and the values of the parameters required (but see the discussion of residence time distributions in Chapter 11). In that all these components depend on the prediction of water flows, they are subject to the types of uncertainty in predictive capability that have been outlined in this chapter and are discussed in more detail later. This is not only a research issue. In the UK, uncertainties in model predictions have already played a major role in decisions made at public inquiries into proposed developments. The aim of this volume is to help provide a proper basis for rainfall–runoff modelling across this range of predictive contexts.

1.9 Key Points from Chapter 1

- There are important stages of approximation in the modelling process in moving from the *perceptual model* of the response of a particular catchment, to the choice of a *conceptual model* to represent that catchment and the resulting *procedural model* that will run on a computer and provide quantitative predictions.
- One particular perceptual model has been outlined as a basis for comparing the descriptions of different models that are given in the following chapters.
- Studies of the geochemical characteristics of runoff, and particularly the use of environmental tracers, have resulted in an increased appreciation of the importance of subsurface storm runoff in many catchments.
- A basic classification of modelling strategies has been outlined, differentiating between *lumped* and *distributed* models and *deterministic* and *stochastic* models.
- Some preliminary guidelines for the choice of a conceptual model for a particular project have been outlined. This problem will be reconsidered in Chapter 10.
- The problem of the calibration of parameter values has been outlined. The idea of an optimum parameter set has been found to be generally ill-founded in hydrological modelling and can be rejected in favour of the concept of the *equifinality* of different models and parameter sets.
- It is expected that, at the end of the model evaluation process, there will not be a single model of the catchment but a number of acceptable models (even if only using different parameter sets within one chosen model structure) to provide predictions.
- Prediction of other processes, such as hydrogeochemistry, erosion and sediment transport, and ecology, that are driven by water flows will introduce additional choices about conceptual model structures and parameter values and will be subject to the uncertainty arising in the rainfall–runoff predictions.

Box 1.1 The Legacy of Robert Elmer Horton (1875–1945)

Robert Elmer Horton was one of the major figures of 20th century hydrology, the foremost scientific hydrologist of his time in the USA and perhaps the only hydrologist to have a waterfall named after him, near to his long-time home in Voorheesville, New York State. A more detailed account of his life and publications has been compiled by Hall (1987). When Horton died in 1945, his scientific papers were bequeathed to the American Geophysical Union and are now housed in the National Archives in College Park, Maryland. There are 95 boxes of them, representing the broad range of his interests, the experimental data collected by him and his assistants (Richard van Vliet, Howard Cook, Harry Leach and James Erwin) and his work as a "hydraulic consulting engineer".

As far as I know, only a small proportion of the boxes of papers have been examined since they were archived by Walter Langbein more than 50 years ago, but they contain a range of fascinating material concerned with infiltration processes, surface runoff, the importance of subsurface flows and hydrological predictions (Beven, 2004a, 2004b, 2004c). They include records of data collected in the Horton Hydrological Laboratory (in his garden) and nearby LeGrange Catchment which was monitored using a thin plate weir. Horton emerges as an impressively careful experimenter, demanding of both himself and his assistants (and very demanding of others in the Geological Survey, Soil Conservation Service and National Weather Service with whom he had to deal), though not without humour.

In many hydrological textbooks, Horton is primarily remembered for being the originator of the concept of infiltration excess overland flow. Indeed this is very often now called "Hortonian overland flow". His final monumental paper, presenting a theory for the development of hillslopes and drainage basins by surface erosion, is based on this concept (Horton, 1945). In addition, his infiltration equation (see Box 5.2) has been widely used to predict effective rainfall and surface runoff. His papers reveal, however, that his conception of hydological processes in catchments was much more complex than this. He recognised the importance of partial surface runoff areas, of rapid subsurface responses, and of macropores, air pressure and surface effects in infiltration (Beven, 2004b, 2004c). In addition, his experimental data on rainfall rates and infiltration rates suggest that the occurrence of widespread surface runoff might have been rather rare on the LeGrange catchment (Beven, 2004a).

He made his living, however, as a hydrological consultant. Part of his work involved analysing rainfall–runoff relationships and predicting the response of catchments, effectively using rainfall–runoff models. The infiltration equation provided Horton with the means to make predictions of runoff volumes (at least if the changes of infiltration capacities due to soil, seasonal and land treatment effects could be estimated). Elsewhere, he treats the problem of allowing for depression storage and the hydraulics of overland flow and channel flow, including transitions from laminar to turbulent flow, in routing that runoff to the basin outlet (Horton, 1935). He realized, however, that in applying this theory he would need to deal with the fact that there might be different infiltration capacities in different parts of a basin and that this would not be a problem in prediction (he proposed a distributed approach based on the division of a basin into "meshes" of different shapes and characteristics; Horton, 1938). It would, however, be a problem in the analysis of rainfall–runoff data in a complex basin. He realised that it would then only be possible to derive some average infiltration capacity over a whole catchment based on the observed rainfalls and discharges.

Horton does not, however, seem to have had too many doubts about applying average infiltration capacities derived under one set of rainfall conditions to the prediction of runoff under other conditions. He did observe that there were seasonal changes in infiltration capacities derived in this way (Figure B1.1.1), and that minimum values might be more robustly estimated than maximum values (which might still underestimate the potential maximum values achievable). He suggests, however, that minimum values might be useful in the estimation of the "maximum flood intensity" to be expected on a given area (Horton, 1937, p. 385).

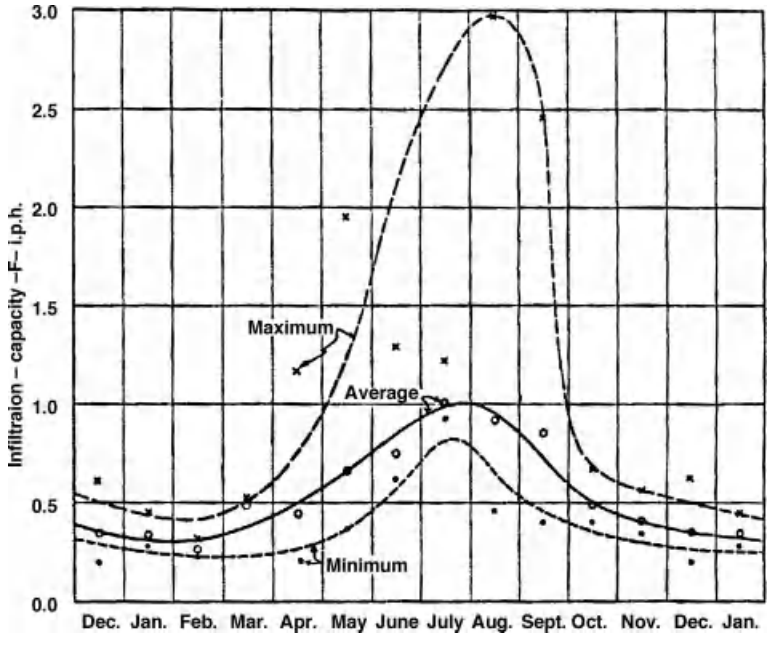


Figure B1.1.1 *Seasonal changes in catchment average infiltration capacities derived from analysis of rainfall and runoff data for individual events (after Horton, 1940; see also Beven, 2004b, with kind permission of the American Geophysical Union).*

This issue of the difference between simulation and analysis (model calibration) continues to tax rainfall–runoff modelling and hydrological science today. The complexity of models that can be calibrated robustly is restricted by the limitations of the data that are generally available for any application. We are therefore in a similar situation to Horton: we would like to take full account of the heterogeneity and changing characteristics of the processes that control runoff in our rainfall–runoff models, but for practical applications it is always necessary to compromise by making some simplifying assumptions.

2

Evolution of Rainfall–Runoff Models: Survival of the Fittest?

Everything of importance has been thought of before by someone who did not invent it.

Alfred North Whitehead, 1920

2.1 The Starting Point: The Rational Method

It is worth remembering that rainfall–runoff modelling has a long history and that the first hydrologists attempting to predict the flows that could be expected from a rainfall event were also thoughtful people who had insight into hydrological processes, even if their methods were limited by the data and computational techniques available to them. We can go back nearly 150 years to the first widely used rainfall–runoff model, that of the Irish engineer Thomas James Mulvaney (1822–1892), published in 1851 (and reproduced in Loague, 2010). The model was a single simple equation but, even so, manages to illustrate most of the problems that have made life difficult for hydrological modellers ever since. The equation was as follows:

$$Q_p = CA\overline{R} \tag{2.1}$$

The Mulvaney equation does not attempt to predict the whole hydrograph but only the hydrograph peak Q_p. This is often all an engineering hydrologist might need to design a bridge or culvert capable of carrying the estimated peak discharge. The input variables are the catchment area, A, a maximum catchment average rainfall intensity, $\overline{R}$, and an empirical coefficient or parameter, C. Thus, this model reflects the way in which discharges are expected to increase with area and rainfall intensity in a rational way. It has become known as the *Rational Method*. In fact, variations on Equation (2.1) have been published by a variety of authors based on different empirical data sets (see Dooge (1957) for a summary) and are still in use today (try searching on "peak discharge rational method").

The scaling parameter C reflects the fact that not all the rainfall becomes discharge, but here the method is not quite so rational since it makes no attempt to separate the different effects of runoff generation and

Rainfall–Runoff Modelling: The Primer, Second Edition. Keith Beven.

runoff routing that will control the relationship between the volume of rainfall falling on the catchment in a storm, effectively $A\overline{R}$, and the discharge at the hydrograph peak. In addition, the coefficient C is required to take account of the nonlinear relationship between antecedent conditions and the profile of storm rainfall and the resulting runoff generation. Thus C is not a constant parameter, but varies from storm to storm on the same catchment and from catchment to catchment for similar storms. The easiest way to get a value for C is to back-calculate it from observations of rainfall and peak discharge (the very simplest form of model calibration). Predicting the correct value for a different set of conditions, perhaps more extreme than those that have occurred before, or for a catchment that has no observations is a much more difficult task.

Similar difficulties persist to the present day, even in the most sophisticated computer models. It is still difficult to take proper account of the nonlinearities of the runoff production process, particularly in situations where data are very limited. It is still easiest to obtain effective parameter values by back-calculation or calibration where observations are available; it remains much more difficult to predict the effective values for a more extreme storm or ungauged catchment. There are still problems of separating out the effects of runoff generation and routing in model parameterisations (and in fact this should be expected because of the real physical interactions in the catchment).

However, it is not impossible to make predictions, even with such simple models. Even in the pre-computer era, the Rational Method evolved into the Graphical Estimation technique (see the work of Linsley, Kohler and Paulhus (1949) or Chow (1964) for full details). This was an attempt to summarise a wide range of analyses carried out for catchments in the USA into a set of graphs or *nomograms* that could be used to predict peak discharges under different rainfall and antecedent conditions (Figure 2.1). This approach has been used as a design tool for many years and has been put into mathematical form by, for example, Plate *et al.* (1988).

2.2 Practical Prediction: Runoff Coefficients and Time Transformations

In Chapter 1 and Section 2.1, the problem of separating the effects of runoff generation and runoff routing have been raised. This differentiation of two sets of processes was the essence of the first attempts to model hydrographs, starting back in the 1890s. It must be remembered that all the calculations at that time had to be done by hand, without benefit even of electronic calculators. At this time, the word "computer" meant a human being who did calculations. The calculating aids available were limited to log tables. Thus the calculations had to be simple.

In a paper recently re-discovered by Charles Obled, a French engineer, Édouard Imbeaux (1892), working on floods in the catchment of the Durance in south-east France, was perhaps the first to attempt to use a distributed hydrological model. His idea was to split the catchment up into zones on the basis of travel time to the catchment outlet. Zone 1 would be the area for which runoff could reach the outlet within one time step (e.g. one hour). Zone 2 would be the area with a travel time of two time steps, and so on (see Figure 2.2). Imbeaux argued that if the production of runoff could be calculated for each area then it was a relatively simple matter to route that runoff to the catchment outlet to obtain a prediction of the hydrograph. Snowmelt is also an important issue and Imbeaux came up with an early version of the degree-day method of predicting snowmelt, taking account of the effect of elevation on temperatures in zones at different distances from the outlet. Different antecedent conditions, different rainfall rates, and different snowmelt conditions would result in different amounts of runoff production and, after the routing, different hydrographs. The resulting time–area diagram represents the delays for runoff from each portion of the catchment. A similar concept was used in the USA by Ross (1921, also reproduced in Loague, 2010), Zoch (1934), Turner and Burdoin (1941), and Clark (1945), and in the UK by Richards

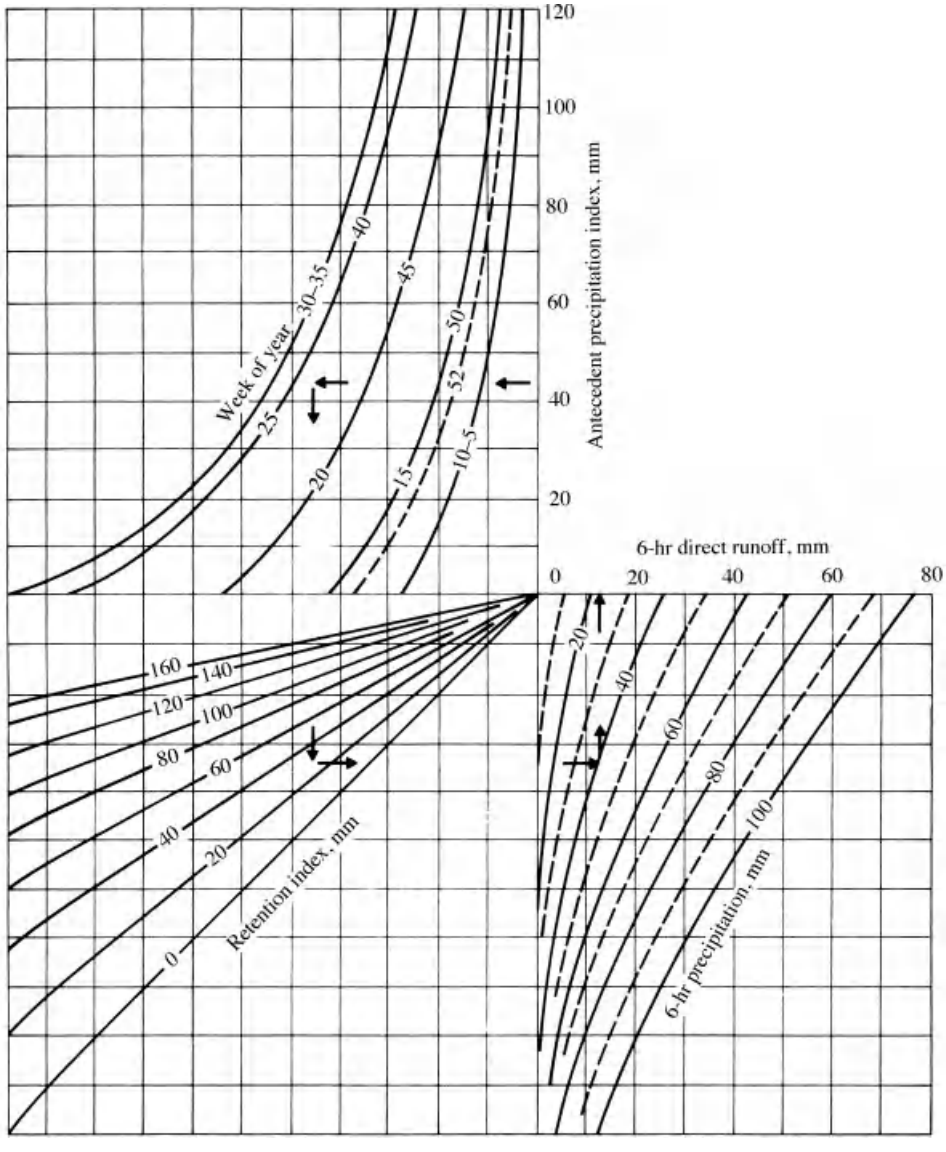


Figure 2.1 *Graphical technique for the estimation of incremental storm runoff given an index of antecedent precipitation, the week of the year, a soil water retention index and precipitation in the previous six hours; arrows represent the sequence of use of the graphs (after Linsley, Kohler and Paulhus, 1949).*

(1944) in one of the first books to be published on rainfall–runoff modelling and flood estimation. These ideas still underlie some of the distributed models being used today and Kull and Feldman (1998), for example, have demonstrated how the time-area method can be used with distributed rainfall inputs derived from the NEXRAD radar system.

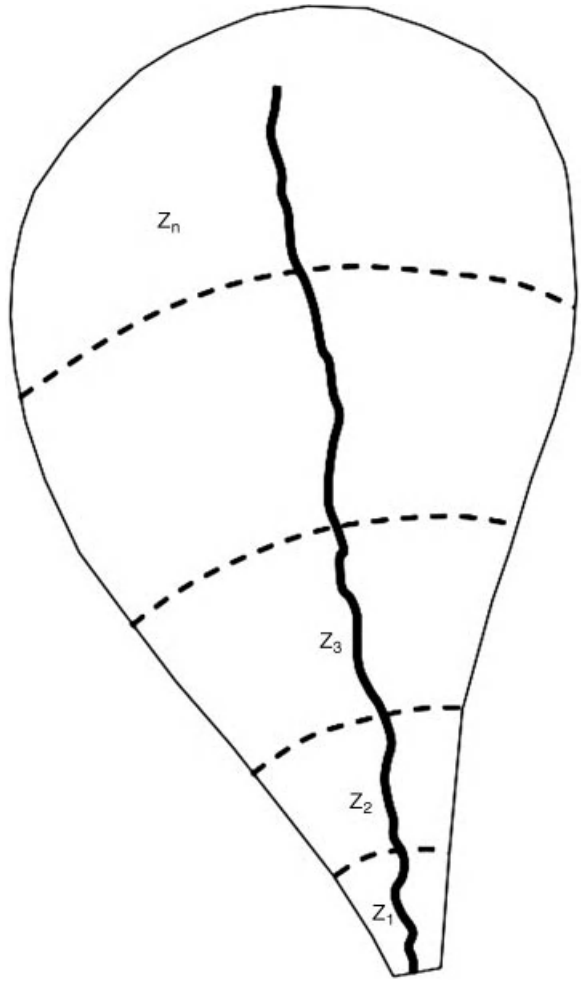


Figure 2.2 *Creating a time–area histogram by dividing a catchment into zones at different travel times from the outlet (from Imbeaux, 1892).*

Note that these early studies made an assumption of *linearity* in routing the runoff. Linearity means that the routing times for the different zones are always the same, regardless of the amount of runoff being routed: the routing process is then mathematically a linear operation (see Box 2.1). This was an approximation. It has been known for some centuries that flow velocities change in a nonlinear way with flow rate or flow depth. The assumption of linearity, however, makes the computations very much easier.

It also works surprisingly well, as will be shown in Chapter 4. Any inaccuracies due to the linearity assumption for routing the runoff are generally less than the inaccuracies associated with deciding how much of the rainfall to route, i.e. the problem of estimating the *effective rainfalls* or *runoff coefficient* for an event. Effective rainfall is that part of a rainfall event that is equal in volume to the runoff generated by the event. The runoff coefficient is that proportion of the total rainfall in an event that becomes runoff.

The way in which runoff generation is predicted is generally nonlinear, with a runoff coefficient that depends on both antecedent conditions and rainfall.

The major problem with the time-area concept was much more the difficulty of deciding which areas of the catchment would contribute to the different zones, since there was little information on velocities of flow for all the possible surface and subsurface flow pathways. This problem was avoided by Sherman (1932, reproduced in Loague, 2010), who used the idea that the various time delays for runoff produced on the catchment to reach the outlet could be represented as a time distribution without any direct link to the areas involved. Because the routing procedure was linear, this distribution could be normalised to represent the response to a unit of runoff production, or effective rainfall, generated over the catchment in one time step. He initially called this function "the unitgraph" and later the *unit hydrograph*; it has become one of the most commonly used hydrograph modelling techniques in hydrology, being simple to understand and easy to apply (especially with the benefit of modern computers). The unit hydrograph represents a discrete *transfer function* for effective rainfall to reach the basin outlet, lumped to the scale of the catchment.

The unit hydrograph remains a linear routing technique such that the principle of *superposition* can be applied. Thus, two units of effective rainfall in one time step will produce twice as much predicted runoff in the hydrograph at the catchment outlet as one unit, with the same time distribution (Box 2.1). The calculated outflows from effective rainfalls in successive time steps can be distributed in time by appropriately delayed unit hydrographs and added up to calculate the total hydrograph at the outlet. It is also generally assumed that the form of the unit hydrograph does not change over time.

There remains the more difficult problem of how to determine the amount of effective rainfall to route. This is definitely a nonlinear problem that involves a variety of hydrological processes and the heterogeneity of rainfall intensities, soil characteristics and antecedent conditions in the same way as the coefficient C of the rational formula of Section 2.1. Thinking about the problem of estimating effective rainfalls was the start of thinking about modelling the rainfall–runoff process on the basis of understanding of hydrological processes. It is not yet, however, a solved problem and there remain a number of competing models for estimating effective rainfalls based on different assumptions.

A major step in tackling this problem was made when, just a year after Sherman introduced his unitgraph, Robert Horton published his paper on the generation of runoff when the infiltration capacity of the soil is exceeded (Horton, 1933). Horton's work was based on experiment and he used an empirical function to describe the decrease in infiltration capacity over time that he found in his experiments (e.g. element A in Figure 2.3), even though simplified solutions of the Darcy flow equation for flow through

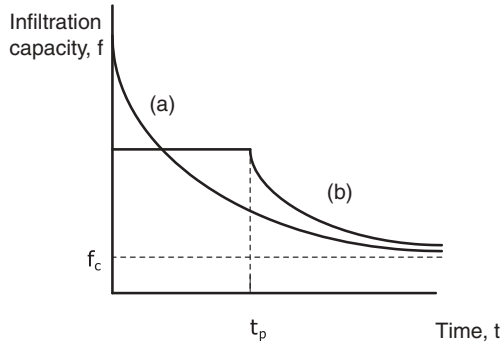


Figure 2.3 *Decline of infiltration capacity with time since start of rainfall: (a) rainfall intensity higher than initial infiltration capacity of the soil; (b) rainfall intensity lower than initial infiltration capacity of the soil so that infiltration rate is equal to the rainfall rate until time to ponding, t_p; f_c is final infiltration capacity of the soil.*

soils had been available at least since the paper by Green and Ampt (1911). Horton argued, in fact, that his experiments suggested that infiltration could not be profile controlled, as assumed in the theory underlying the Green–Ampt equation, but was controlled by surface effects such as the redistribution of fine particles (Beven, 2004b). Since then, many other infiltration equations have been proposed, mostly based on various simplifications of the nonlinear Darcy flow problem (see, for example, the review of Parlange and Haverkamp (1989) and Box 5.2).

All of these equations provide estimates of the local limiting infiltration capacity of the soil over time. When rainfall during a storm exceeds the infiltration capacity, then water will start to pond at the surface and, perhaps after the storage in local depressions has been filled, may start to run downslope as overland flow. Comparison of rainfall rates with infiltration capacities therefore provides a means of estimating the effective rainfall for a storm (e.g. element B in Figure 2.3) – if runoff is actually being generated by an infiltration excess mechanism. However, as we saw in Chapter 1, this is not always the case, and even when surface runoff does occur, infiltration rates may show a high degree of heterogeneity in space. There is little doubt that this type of approach to estimating effective rainfall has often been misapplied, and probably continues to be misapplied (or at least misinterpreted) 60 years after the original formulation of the concepts. Robert Horton understood some of these issues (see Beven, 2004a) but also understood the value of the infiltration excess approach as an engineering tool.

The reasons for this are functional. The infiltration excess model of effective rainfall and the unit hydrograph together provide the necessary functional components of a hydrological model, i.e. a way of estimating how much of the rainfall becomes runoff and a means of distributing that effective rainfall through time to predict the shape of the hydrograph. It is not therefore necessary to apply this method under the assumption that it is actually surface runoff in excess of the infiltration capacity of the soil that is being routed by the unit hydrograph (as in Figure 2.4a). The simplest models of effective rainfall, assuming either that there is a constant "loss" rate (the Φ-index method) (Figure 2.4b) or that a constant proportion of the rainfall is effective rainfall (Figure 2.4c), are also still widely used but are less obviously surface runoff generation models; rather, they are very simple ways of deriving an approximate runoff coefficient. This type of estimation of effective rainfall serves the functional requirement of having a loss function that is nonlinear with respect to total rainfall if the parameter is allowed to vary with antecedent state of the catchment, regardless of whether the runoff generation process is actually due to an infiltration excess mechanism. Both methods involve only one parameter but they lead to different patterns of effective rainfalls in time for the same event. Other ways of calculating effective rainfalls, with similar functionality, are also commonly used.

Another interesting method for estimating effective rainfalls in the USDA Soil Conservation Service (SCS) curve number approach (McCuen, 1982). This is also often interpreted as an infiltration equation (e.g. the recent papers of Yu, 1998, and Mishra and Singh, 1999), but in fact has its origins in the analysis of runoff volumes from small catchments by Mockus (1949, reproduced in Loague, 2010) and which may not only have included infiltration excess overland flow as the runoff generating mechanism (see Box 6.3). The critical assumption of the SCS method is that the ratio of the actual runoff to the potential runoff (rainfall less some initial abstraction) is equal to the ratio of the actual retention to the potential retention. There is no physical justification for this assumption. Mockus himself suggested only that it produced rainfall–runoff curves of the type found on natural watersheds. It is therefore a purely empirical function for estimating a runoff coefficient and any process interpretation equating the retention to infiltration and the runoff to surface runoff has been made since the original work. This illustrates how deeply the Hortonian concepts of runoff generation have permeated the development of rainfall–runoff modelling. The SCS method remains widely used within a number of current distributed models and is covered in more detail in Chapter 6.

The calculation of effective rainfalls is a major problem in the use of the unit hydrograph technique, especially since it is inherently linked to decisions about hydrograph separation to determine the amount of storm runoff for an event. However, since use of the unit hydrograph is a linear operation, given

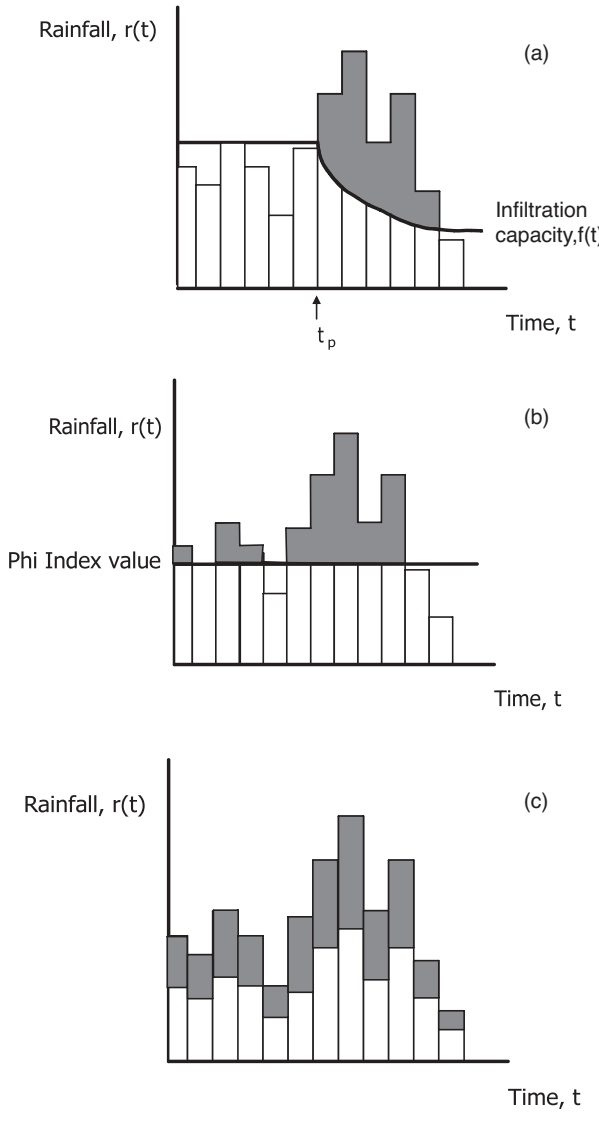


Figure 2.4 *Methods of calculating an effective rainfall (shaded area in each case): (a) when rainfall intensity is higher than the infiltration capacity of the soil, taking account of the time to ponding if necessary; (b) when rainfall intensity is higher than some constant "loss rate" (the ϕ index method); (c) when effective rainfall is a constant proportion of the rainfall intensity at each time step.*

a sequence of rainfalls and a separated storm hydrograph, once a unit hydrograph is available for a catchment, it can be used in an inverse way to estimate a pattern of effective rainfall. Indeed, by using an iterative process starting with some initial estimate of the form of the unit hydrograph, both effective rainfall sequences and the unit hydrograph can be calibrated without making any assumptions about the nature of the runoff generation processes (see Box 4.2). Unfortunately, this does not appear to make it

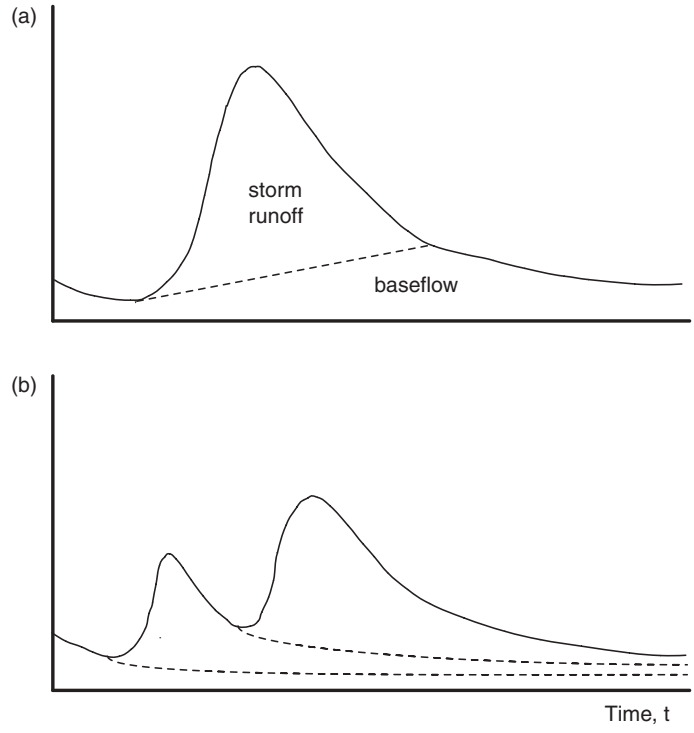


Figure 2.5 *Hydrograph separation into "storm runoff" and "baseflow" components: (a) straight line separation (after Hewlett, 1974); (b) separation by recession curve extension (after Reed et al., 1975).*

any easier to interpret the effective rainfalls derived in this way so that it is easier to predict the effective rainfalls for other storms.

There is one further problem in applying the unit hydrograph technique. In any storm hydrograph that occurs in a perennial stream, some of the stream discharge would have occurred even without the rainfall event that causes the hydrograph response. This is usually called the *baseflow component* of the stream and, if there has been a dry period since the previous storm event, is usually assumed to be derived from subsurface flow. Early on it was found that the effective rainfall is more linearly related to stream discharge if the total hydrograph is separated into a baseflow component and a *storm runoff* component (Figure 2.5a). Hydrograph separation then became an important component of the application of the unit hydrograph model: the problem is that there are no satisfactory techniques for doing hydrograph separation. Indeed, some very strange methods of hydrograph separation are reported in the literature (see review in Beven, 1991). About the only physically justifiable technique for hydrograph separation is to try to estimate the flow that might have occurred if the storm had not happened. However, such a procedure tends to lead to storm runoff hydrographs with very long tails and can get quite complicated in the case of several storms in quick succession (Figure 2.5b) so is not commonly used (but see Reed *et al.*, 1975). In fact, the best method of dealing with hydrograph separation is to avoid it all together, as discussed in Section 2.3.

In many texts, the storm runoff component is called the "surface runoff" component. This is correct in that the total stream hydrograph is measured as surface runoff in the channel, but the name also subtly

perpetuates the generally incorrect conceptual link with the idea that runoff is generated by an infiltration excess mechanism. This nomenclature should be discouraged as should the idea that the effective rainfall is the same water that forms the discharge hydrograph. Tracer information generally suggests that it is not (as discussed in Section 1.5 and Chapter 11).

Despite all these limitations, the unit hydrograph model works for predicting discharges. As noted above, it has the basic functional components needed. There are modern variants of the technique that work extremely well both in short-term flood forecasting and in prediction over longer periods (see Chapter 4). There are also now variants linked into Geographical Information Systems (GIS) that have returned to concepts similar to the original Ross time-area diagram formulation. The approach could be considered as the "model to beat" for discharge prediction and we will return to it a number of times.

2.3 Variations on the Unit Hydrograph

As experience in the application of the unit hydrograph approach was gained, a number of difficulties with the approach came to be appreciated, both in calibration to a particular catchment and in prediction. Two major problems arise in calibration. One has already been mentioned, that of hydrograph separation. However, once a technique has been chosen, the volume of storm runoff arising from the calibration storms can be calculated. Because of the linearity assumption, comparison of this runoff volume with the storm rainfall volume means that the runoff coefficient for each calibration storm can be calculated exactly. Thus, in calibration at least, the choice of a method for separating the effective rainfall from the total rainfall is not such a critical issue, because it can be ensured that the volumes match. This is one reason why the very simple Φ index method continues to be used today.

Once effective rainfall and storm runoff time series are available, a second problem in calibration arises from numerical difficulties in calculating the unit hydrograph. If the unit hydrograph is treated as a histogram (Figure 2.6a), then each ordinate of the histogram is an unknown to be determined, effectively a parameter of the unit hydrograph. However, the histogram ordinates are strongly correlated, especially on the recession limb; with the errors inherent in the time series of effective rainfalls and storm runoff, this makes for a mathematically ill-posed problem. Attempts at a direct solution tend to lead to oscillations, sometimes extreme oscillations, in the unit hydrograph ordinates that are physically unacceptable as a representation of the catchment routing.

A number of ways of avoiding such oscillations have been tried, including imposing various constraints on the shape of the hydrograph (e.g. Natale and Todini, 1977), and superimposing data from many storms and determining an average unit hydrograph by a least squares procedure (e.g. O'Donnell, 1966). This latter approach is still used in the DPFT-ERUHDIT model of Duband *et al.* (1993).

Another approach is to reduce the number of parameters to be determined. This can be achieved by specifying a particular mathematical form for the unit hydrograph. The simplest possible shape, with only two parameters, is the triangle (if the base time and time to peak are specified, then the mass balance constraint of having a hydrograph equal to the unit volume means that the peak height of the triangle can also be calculated). The triangle was chosen as a simple model in procedures for predicting the response of ungauged catchments in the UK *Flood Studies Report* (NERC, 1975). It has been retained in the revised procedures in the UK *Flood Estimation Handbook* (IH, 1999), see also Shaw *et al.* (2010).

This is not the only two parameter model that can be used, however. One of most well-known and widely used models is the so-called Nash cascade, which can be visualised as a sequence of N *linear stores* in series, each with a mean residence time of K time units (Nash, 1959). The resulting mathematical form for the unit hydrograph $h(t)$ is equivalent to the Gamma distribution:

$$h(t) = \left(\frac{t}{K}\right)^{N-1} \frac{\exp(-t/K)}{K\Gamma(N)} \tag{2.2}$$

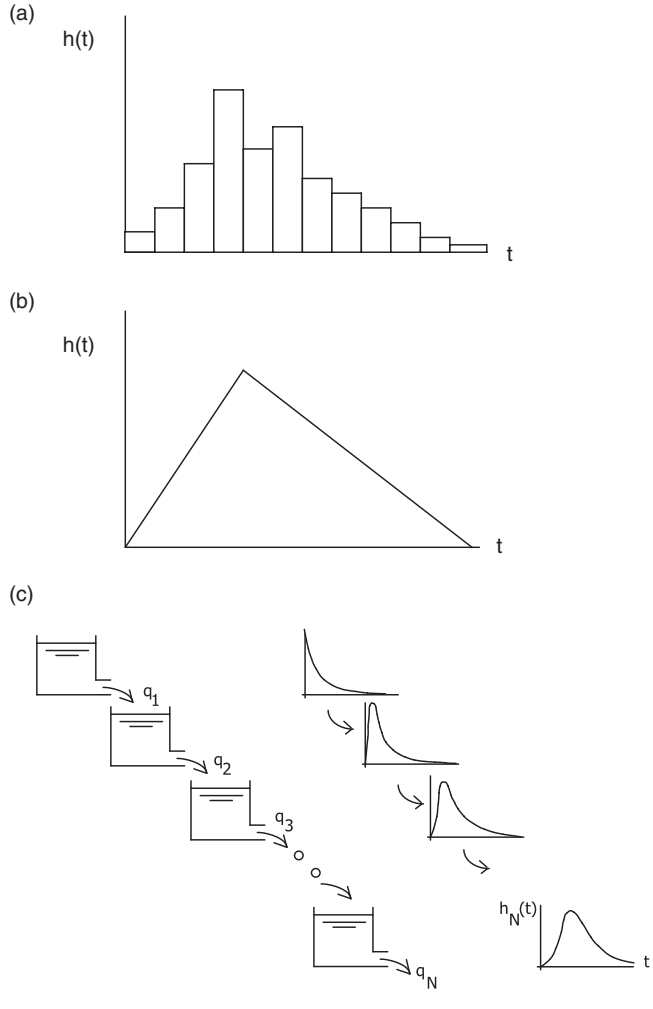


Figure 2.6 *The unit hydrograph as (a) a histogram; (b) a triangle; (c) a Nash cascade of N linear stores in series.*

where $\Gamma(N)$ is the Gamma function ($\Gamma(N) = (N-1)!$ for integer values of N). For different values of N and K, the Gamma distribution has quite a flexible range of forms (Figure 2.6c). Mathematically, N does not have to be an integer number of stores but can also take on fractional values to give a wider range of shapes in fitting the observed data. Dooge (1959, reproduced in Loague, 2010) provided a summary of a number of other simple linear models that could be used, including those with time delays (see also Dooge and O'Kane, 2003).

The advantage of these functions is that, in general, much more stable estimates of the parameters are obtained, while retaining a flexibility in shape for representing a variety of different catchment hydrographs. Attempts have been made to relate the resulting parameters to different variables representing catchment characteristics, but it must be remembered that the parameter values will be dependent on the

procedures used in deriving the parameters, and particularly on the hydrograph and rainfall separation techniques used.

In the last few years, however, there have been some successful attempts to avoid the problems inherent in these separation techniques and derive a model that relates the total rainfall to the total discharge, not only for single storms but also for continuous simulation. These models stem from developments in general linear systems analysis pioneered by Box and Jenkins (1970). The general linear model that allows explicit hydrograph separation to be avoided is discussed in detail in Chapter 4. A physical interpretation of the range of models is as one or more linear storage elements arranged in series or in parallel (as, for example, the series of equal stores in the Nash cascade above). Given time series of inputs and outputs that are related in a reasonably linear way there are now robust algorithms available for the estimation of the parameters.

The critical phrase here is "related in a reasonably linear way", since, as we have already noted, total rainfall is not related linearly to total discharge. In the past, there have been a number of attempts to use nonlinear transfer functions based on Volterra series (Amorocho and Branstetter, 1971; Diskin and Boneh, 1973) but still requiring a prior estimation of the effective rainfalls. More recent approaches have attempted to relate total rainfalls directly to total discharge. Such models clearly need to have some form of nonlinear transformation of the rainfall inputs but it has proven possible to retain the linearity assumption in the routing component while still maintaining an adequate representation of the long time delays associated with the baseflow component. The result is often a parallel model structure, with part of the rainfall being routed through a store with a short mean residence time to model the storm response and another part through a store with a long mean residence time to model the baseflow. Examples are the IHACRES model of Jakeman *et al.* (1990) and the bilinear power model of Young and Beven (1994), which are very similar in their modelling of the routing component but differ in their approach to modelling the catchment nonlinearity. A more detailed examination of this type of general *transfer function* model is given in Section 4.3.

These approaches are based on letting an analysis of the data suggest the form of the model. Another line of recent development of unit hydrograph theory has been the attempt to relate the unit hydrograph more directly to the physical structure of the catchment and, in particular, to the channel network of the catchment with the aim of developing models that will provide accurate simulations of ungauged catchments. Two lines of approach may be distinguished (see Section 4.5.2), one which is based on analysis of the actual structure of the network (using the *network width function*) and one which uses more generalised geomorphological parameters to represent the network (the *Geomorphological Unit Hydrograph (GUH)* approach). Both are concerned primarily with the routing problem, not with the estimation of the effective rainfall.

The availability of modern GIS databases has also allowed a return to the original Imbeaux/Ross concept of the time-area diagram representation of the unit hydrograph. Overlays of different spatial databases of soil, vegetation and topography data within a GIS results in a classification of parcels of the landscape with different functional responses. Amerman (1965) called these parcels "unit source areas" but they are now more commonly known as *hydrological response units* (HRUs, e.g. Figure 2.7) or "hydrotopes". The topography of the catchment can also be used to define flow directions and distances to the outlet for each hydrological response unit which can provide the basis for a routing algorithm (either linear or nonlinear). A representation of the response for each HRU allows a calculation of the effective rainfall to be routed to the outlet to form the predicted hydrograph.

This type of distributed model is not often presented within the context of unit hydrograph models but the similarities with the time-area diagram concept are clear, particularly if a linear algorithm is used to route runoff generated on each HRU to the outlet. The technology used has changed dramatically, of course, with the availability of GIS databases and modern computer graphics for pre- and post-simulation data processing. Figure 2.7 is clearly much more impressive than Figure 2.2 but, underneath, the approaches are conceptually very similar. It also has to be remembered that the definition of "response

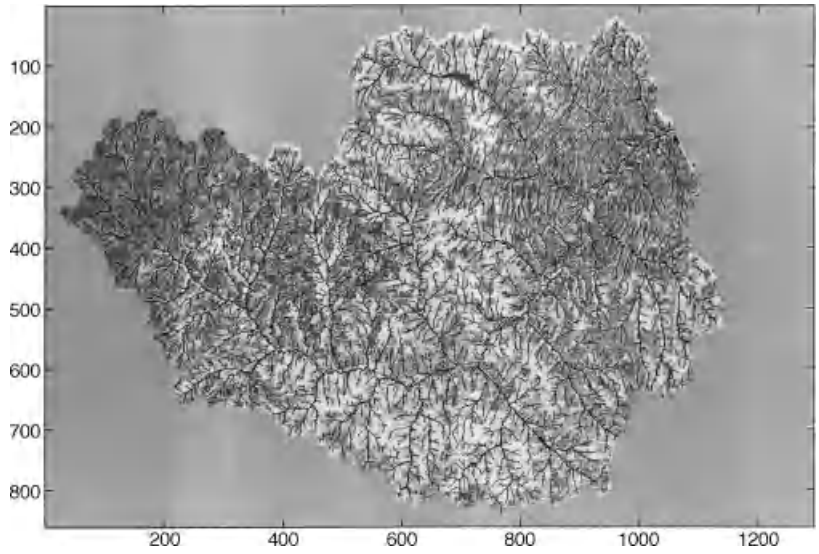


Figure 2.7 *A map of hydrological response units in the Little Washita catchment, Oklahoma, USA, formed by overlaying maps of soils and vegetation classifications within a raster geographical information system with pixels of 30 m.*

units" by GIS overlays does not solve the problem of determining how much of the rainfall becomes runoff, since the classification of the landscape by its soil and vegetation characteristics does not give the hydrological parameters needed to describe the processes operating at the response unit scale directly. Many such models use simple conceptual components to describe each HRU, similar to the type of explicit soil moisture accounting models that have been used very widely at the catchment scale since the very earliest days of rainfall–runoff modelling on digital computers. Such an approach can also be used to predict spatial variations in both evapotranspiration and snowmelt (e.g. Gurtz *et al.*, 1999).

2.4 Early Digital Computer Models: The Stanford Watershed Model and Its Descendants

The computational constraints on rainfall–runoff modelling persisted until the 1960s when digital computers first started to become more widely available. Even so, those computers that were available were expensive, very slow by today's standards, and had very limited memory available. Even the biggest and most expensive were much less powerful than a simple portable PC of today. The types of program that could be run were limited in size and complexity. During this period there was, however, a very rapid expansion in the number of hydrological models available. For the most part they were of similar form: a collection of storage elements representing the different processes thought to be important in controlling the catchment response with mathematical functions to describe the fluxes between the thresholds. One of the first and most successful of these models was the Stanford Watershed Model developed by Norman Crawford and Ray Linsley at Stanford University (a version of the model from 1960 is reproduced by Loague, 2010). The Stanford Watershed Model later evolved into the Hydrocomp Simulation Program (HSP) and was widely used in hydrological consulting. The model survives, with the addition of water

quality components, in the form of the US EPA's Hydrological Simulation Program – Fortran (HSPF, Donigian *et al.*, 1995). Models of this type, called explicit soil moisture accounting (*ESMA) models* by O'Connell (1991), varied in the number of storage elements used, the functions controlling the exchanges, and consequently in the number and type of parameters required. The Stanford Watershed Model had up to 35 parameters, although it was suggested that many of these could be fixed on the basis of the physical characteristics of the catchment and only a much smaller number needed to be calibrated.

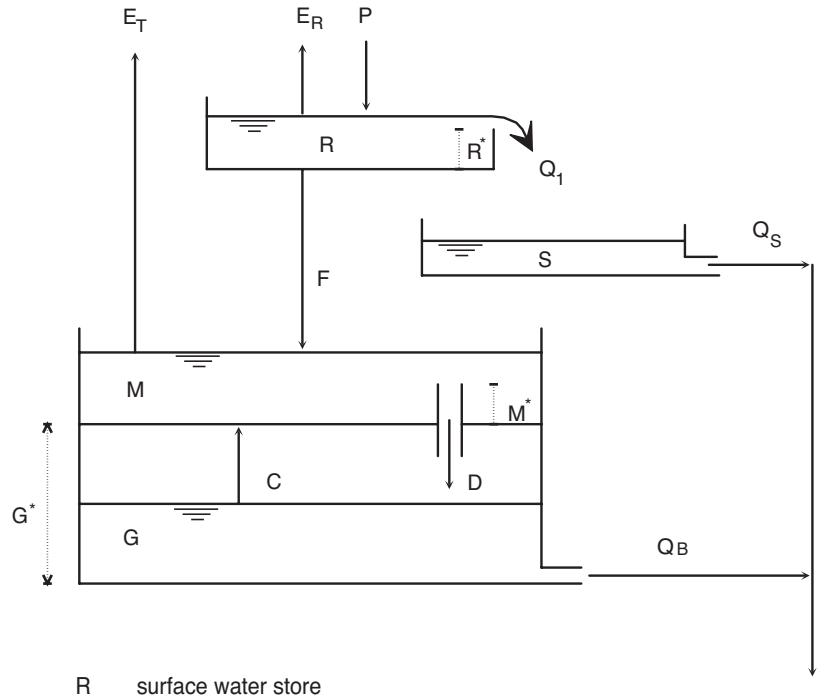
In the early years of digital computing, there was a tendency for every hydrologist with access to a computer (there were no personal computers then) to build his own variant of this type of model. It was not, after all, a difficult programming exercise. It was one of the ways that I learned how to program computers as an undergraduate by writing a model to try to "hindcast" runoff generation on Exmoor during the Lynmouth flood event. This model was written in the Algol programming language, in 1971, stored on punched cards and run on the Bristol University Elliot 503 computer with 16 kB of memory, with all output to paper from a lineprinter. This is an indication of how rapidly the resources available to the modeller have changed in the last four decades.

Most of these models had sufficient number of parameters and flexibility to be able to produce a reasonable fit to the rainfall–runoff data after some calibration. Indeed it was all too easy to add more and more components (and more associated parameters) for different processes. The potential for a confusing plethora of models was soon recognised, and Dawdy and O'Donnell (1965) tried to define a relatively simple "generic" model structure, with just a few parameters (Figure 2.8). This did not, however, stop a continued expansion in the number of models of this type published in the hydrological literature (see, for example, the review by Fleming, 1975). A number of examples that are still in current use include the HSPF, SSARR and Sacramento models from the USA, the HBV model from Sweden, the Tank model from Japan, the UBC model from Canada, and the RORB and AWBM models from Australia (Singh, 1995; Singh and Frevert, 2005; Boughton, 2011).

A comparison of different models of this type reveals the subjectivity involved in defining a particular model structure, albeit that there is often similarity in some components. An example in current use, variously known as the Xinanjiang model, Arno model or Variable Infiltration Capacity (VIC) model, is described in Box 2.2. This model is of interest in that, although it can be classed as an ESMA-type model, the surface runoff generation component can also be interpreted in terms of a distribution function of catchment characteristics (see Section 2.6). It has also been implemented as a macroscale hydrological model or *land surface parameterisation* in some global climate models.

Functionally, there are also similarities between the modern IHACRES package, mentioned in Section 2.3, and ESMA-type models, since in both cases the runoff generation and runoff routing components are based on storage elements. The difference lies in the modern approach of trying to find the simplest model structure supported by the data (see the discussions by Jakeman and Hornberger, 1993, and Young, 2000, 2003) and in not necessarily fixing the model structure beforehand. Instead, an analysis of the data should be allowed to suggest what the appropriate structure should be, as in the data-based mechanistic approach of Young and Beven (1994) (see Chapter 4).

Providing some data are available to calibrate parameter values, the results from even simple ESMA models can be quite acceptable, both in modelling discharges (Figure 2.9) and in soil moisture deficit modelling (Figure 2.10). The performance demonstrated in Figure 2.10 is particularly impressive if it is remembered that 1976 was one of the driest summers on record in the UK. A number of comparisons of ESMA models have been published, although the limitations both of model structures and of input and output data means that it has not generally been possible to conclude that one model consistently performs better than another after the model parameters have been calibrated (see, for example, the studies of Franchini and Pacciani, 1991; Chiew *et al.*, 1993; and Editjatno *et al.*, 1999). More recently, software has been developed that allows different model components to be put together to allow appropriate model structures to be found for particular applications. This software includes the USGS Precipitation-Runoff Modelling System (PMRS) of Leavesley *et al.* (2002), the Imperial College Rainfall–Runoff Modelling



R surface water store
R* surface water storage threshold
S surface water routing store (with parameter K_S)
F infiltration to soil moisture store (with parameters fo, fc, k)
M soil moisture store
M* soil moisture storage threshold
D recharge to groundwater store
C maximum rate of capillary rise
G groundwater store (with parameter Kg)
G* groundwater storage threshold

Figure 2.8 *Schematic diagram of the Dawdy and O'Donnell (1995) conceptual or explicit soil moisture accounting (ESMA) rainfall–runoff model.*

Toolbox of Wagener *et al.* (2004), the FUSE system of Clark *et al.* (2008) and the FLEX system of Fenicia *et al.* (2008a, 2008b).

ESMA models are also being used to predict the impact of climate change in different countries (see also Chapter 8). In this context, their use is more problematic because the accuracy of the predictions rely heavily on the availablity of data for calibration. Thus, if the parameter values of a model are successfully calibrated under current conditions, and then the model is used with a different range of climatic inputs, representing perhaps one possible scenario for conditions some time in the 21st century, there is no guarantee that the current accuracy will be maintained, especially if the changed conditions are more extreme. No data are available, however, for calibration under the changed conditions. Thus, the impact predictions should be expected to be more uncertain than current-day simulations and any impact predictions should be associated with an estimation of uncertainty. This is not yet commonly done.

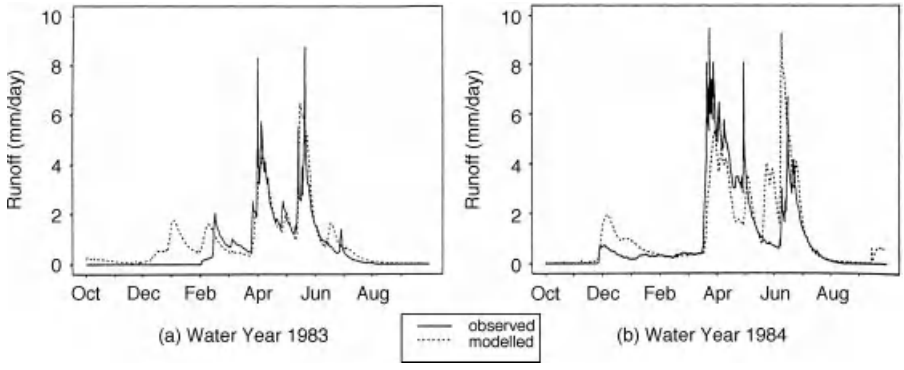


Figure 2.9 *Observed and predicted discharges for the Kings Creek, Kansas (11.7 km^2) using the VIC-2L model (see Box 2.2); note the difficulty of simulating the wetting up period after the dry summer (after Liang et al. 1996, with kind permission of Elsevier).*

ESMA-type models are also still used for representing GIS-derived hydrological response units. Models of this type include the group response units WATFLOOD model (Cranmer *et al.*, 2001); the PMRS (Leavesley and Stannard, 1995); the USDA-ARS SWAT model of Arnold *et al.* (1998); the Arc Hydro modelling system of Maidment (2002) and the recent model of Vinogradov *et al.* (2011). We run into a few model classification problems here: such models aim to represent hydrological processes in a distributed manner, but using functional components in the style of ESMA models at the scale of the GIS-derived HRU, rather than attempting full process descriptions. It is a modelling technology that has been driven

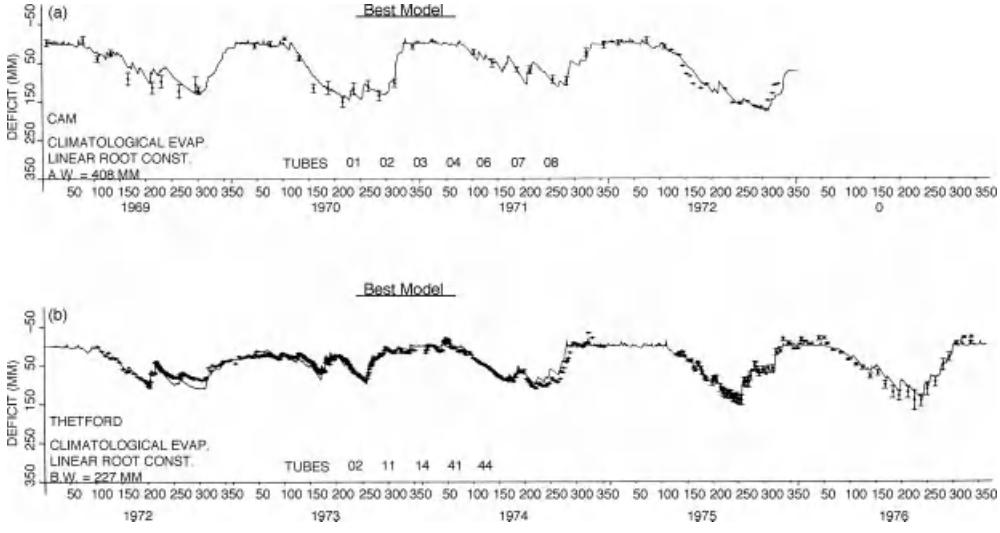


Figure 2.10 *Results from the prediction of soil moisture deficit by Calder et al. (1983) for sites in the UK: (a) the River Cam and (b) Thetford Forest; observed soil moisture deficits are obtained by integrating over profiles of soil moisture measured by neutron probe; input potential evapotranspiration was a simple daily climatological mean time series (with kind permission of Elsevier).*

by the availability of GIS and remote-sensing data rather than any real advances in the understanding of how to represent hydrological processes. They can perhaps be considered most appropriately with the class of semi-distributed models, and we return to them in Chapter 6.

2.5 Distributed Process Description Based Models

It will be realised from the discussion in Section 2.4 that there is considerable subjectivity associated with the definition of an ESMA-type model, even if the hydrologist is attempting, as far as possible, to reflect her or his perceptual model of how catchments work. Hence the wide range of ESMA models available. Another early (in the digital computer era) response to this subjectivity was the attempt to produce models based directly on equations describing all the surface and subsurface flow processes in the catchment. A blueprint for such a model was described in the article by Freeze and Harlan, 1969, reproduced by Loague, 2010).

In their seminal paper, Freeze and Harlan wrote down the equations for different surface and subsurface flow processes and showed how they could be linked by means of common boundary conditions into a single modelling framework. Their analysis is still the basis of the most advanced distributed rainfall–runoff modelling systems today. The equations are all nonlinear partial differential equations, that is to say differential equations that involve more than one space or time dimension (see Box 2.3). For the type of flow domains and boundary conditions of interest in rainfall–runoff modelling, such equations can normally only be solved by approximate numerical methods which replace the differential terms in the equations by an approximate discretisation in time and space (see Chapter 5). However, the descriptive equations that they used for each process required, in all cases, certain simplifying assumptions. Thus, for subsurface flow, it is assumed that both saturated and unsaturated flows can be described by Darcy's law (that flow velocity is proportional to a hydraulic conductivity and a gradient of total potential, see Box 5.1), while for surface flows it was assumed that the flow could be treated as a one-dimensional, cross-section averaged flow either downslope over the surface or along a reach of the channel network in a catchment (leading to the St Venant equations, see Box 5.6).

Freeze and Harlan (1969) discuss the input requirements and boundary conditions for these equations in some detail. Meteorological data are required to define rainfall inputs and evapotranspiration losses. Their **model definition input** includes necessary assumptions on the extent of the flow domain, assumptions about prescribed potential or prescribed flow boundaries (especially zero flow at impermeable or divide boundaries) and the way in which the flow domain is divided to create space and time increments for solving the process equations. **Flow parameter inputs** are then required for each element of the solution grid, allowing consideration of heterogeneity of catchment characteristics to be taken into account.

This type of distributed model allows the prediction of local hydrological responses for points within the catchment. The first applications of this type of model were made for hypothetical catchments and hillslopes by Freeze (1972). The calculations required the largest computers available at the time (Al Freeze was then working at the IBM Thomas J. Watson Research Centre at Yorktown Heights), and even then only a limited flow domain and coarse mesh of points could be solved. The first application to a field site of this type of model was published by Stephenson and Freeze (1974), who attempted to model a single hillslope at the Reynolds Creek catchment in Idaho (see Section 5.3). The results were not particularly successful but, as they pointed out, it was a complex hillslope underlain by fractured basalts with complex flow pathways, They had limited knowledge of the inputs and initial conditions for the simulation and computing constraints limited the number of simulations they could actually try. Arising from these difficulties, they were also the amongst the first to discuss the difficulties of validating hydrological models.

Distributed models of this type have the possibility of defining parameter values for every element in the solution mesh. There may, even with continuing limitations imposed by computational constraints, be

many thousands of such elements. In addition, the process equations require many different parameters to be specified for each element. With so many parameter values, parameter calibration by comparison with the observed responses in a catchment becomes very difficult. Which parameter values should be changed to try to improve the simulation? The choice is not always obvious because of the interactions between the effects of different parameters that follow from the physical basis of the model.

In principle, of course, parameter adjustment of this type should not be necessary. If the process equations are valid, it should follow that the parameters should be strongly related to the physical characteristics of the surface, soil and rock. Techniques are also available for measuring such parameters, although, as noted in Chapter 1, there are problems of scale in such measurements. Most measurement techniques can only be used to derive values at scales much smaller than the element mesh used in the approximate solution. The model requires effective values at the scale of the elements. If the soil was homogeneous, this would not matter too much, but unfortunately soils and surface vegetation tend to be very heterogeneous at the measurement scale such that establishing a link between measurements and element values is difficult, even theoretically. Indeed, for the case of coupled surface and subsurface flow processes, it has been suggested that the concept of effective values of element scale parameters may not be valid (Binley *et al.*, 1989). This remains an important topic in distributed modelling that requires further research.

Despite these difficulties, there has been a strong surge in the use of distributed modelling over the last two decades. This has partly been because increases in computer power, programming tools and digital databases have made the development and use of such models so much easier and partly because there is a natural tendency for a model development team to try to build in as much understanding from their perceptual model of the important processes as possible. Thus, there is an obvious attraction of distributed process modelling. There are also very good scientific reasons underlying the effort. One is the need for distributed predictions of flow pathways as a basis for other types of modelling, such as the transport of sediments or contaminants. It may not be possible to make such predictions without a distributed model of some sort.

Another reason for the surge is the use of models for impact assessment. Changes in land use, such as deforestation or urbanisation, often affect only part of a catchment area. With a distributed model, it is possible to examine the effects of such piecemeal changes in their correct spatial context. There is also an argument that, because of the physical basis of the model, we might be able to make a better assessment of the effects of changing characteristics of a catchment because it will be easier to adjust parameter values that have physical meaning. The difficulties of specifying effective parameter values at the element scale, however, rather undermines this argument.

Examples of distributed process-based models include the Système Hydrologique Européen (SHE) model, originally a joint project between the Institute of Hydrology in the UK, the Danish Hydraulics Institute and SOGREAH in France, but now being developed separately (see Abbott *et al.*, 1986a; Bathurst *et al.*, 1995; Refsgaard and Storm, 1995). Both the Danish (MIKE SHE) and UK (SHETRAN) versions of SHE have been extended with full three-dimensional subsurface components (but also with simplified subsurface storage elements), and sediment and water quality components. Refsgaard *et al.* (2010) provide a summary of 30 years of experience of using the MIKE SHE model. The UK Institute of Hydrology also developed the Institute of Hydrology Distributed Model (IHDM, Calver and Wood, 1995); in Australia, there are the THALES model (Grayson *et al.*, 1995) and the CSIRO TOPOG-dynamic model (Vertessy *et al.*, 1993); and there are a number of models developed in the USA including the Integrated Hydrologic Model (InHM) of VanderKwaak and Loague (2001) and the Gridded Surface/Subsurface Hydrologic Analysis model (GSSHA) of Downer *et al.* (2005). GSSHA was an extension of an earlier infiltration and surface runoff model, CASC2D (Downer *et al.*, 2002). These models differ primarily in the way they discretise a catchment and solve the process equations (sometimes with simplifications), but all are essentially based on the original Freeze and Harlan blueprint from 1969 as a description of the flow processes (see Chapter 5 for more details).

It has to be said that neither the process descriptions nor the questions demanding further research have changed very much in the 40 years since the Freeze and Harlan blueprint was published. Certainly the tremendous advances in computer power, numerical methods for solving partial differential equations and advances in programming techniques have allowed more robust solutions with finer spatial and temporal resolutions to be implemented in applications to larger catchments, but computational constraints on the full implementation of three-dimensional solutions still remain.

It must be remembered that the blueprint remains a major simplification of the perceptual model discussed in Chapter 1 (Beven, 1989, 2001, 2002a). Some processes, such as preferential flow in soil macropores, are omitted altogether, for good reason: there is no adequate descriptive equation to integrate the effects of macropore flow at the element scale (although see Bronstert and Plate, 1997; Faeh *et al.*, 1997; and Klaus and Zehe, 2010, for attempts to do so). Similarly, the idea of describing surface runoff as a sheet flow of uniform depth and velocity across the slope, as used in most process models, is clearly a gross simplification of reality. Few attempts have been made to take account of variations in depth on surface flow velocities and infiltration rates (Dunne *et al.*, 1991; Tayfur and Kavvas, 1998). These models are based on flow physics in name, but it is a very approximate physics and is likely to remain so until measurement and visualisation techniques for studying flow processes in the field improve to allow better descriptions, particularly of subsurface flow. The issues involved in this type of distributed modelling remain contentious (Beven, 2010).

2.6 Simplified Distributed Models Based on Distribution Functions

The fully distributed models discussed above are complex and both computationally and parametrically demanding. The mathematical descriptions of processes on which they are based are, however, still simplifications of reality. A separate strand in rainfall–runoff modelling can be distinguished that attempts to maintain a distributed description of catchment responses but which does so in a much simpler way, without the detailed process representations of SHE and the other fully distributed models. This type of model generally uses a form of distribution function to represent the spatial variability of runoff generation. The distribution may be based on a purely statistical description, as in the Probability Distributed Moisture (PDM) model of Moore and Clarke (1981); on a simple functional form, as in the Xinanjiang/ARNO/VIC model (see Box 2.2); on GIS-derived hydrological response units (described in Section 2.4); or on some simplified physical reasoning leading to a distribution of an index of hydrological similarity, as in TOPMODEL (the topography-based model of Beven and Kirkby (1979), discussed in Section 6.3). It is worth noting that a uniform distribution function for infiltration capacity was included as one of the components of the original Stanford Watershed Model and that many of these models have ESMA-type components.

In all these models, the distribution function component is an attempt to make allowance for the fact that not all of the catchment can be expected to respond in an exactly similar way. Volumes of runoff generation, for example, should be expected to vary with position in the catchment. This is in keeping with a generally held perceptual model of hydrological responses. In representing these different functional responses as a distribution function, it is recognised that this type of variability is important but that it might be very difficult to specify exactly where in the catchment runoff generation is happening in any particular event. The aim is to get the bulk responses at the catchment scale modelled correctly. TOPMODEL does allow for the distribution function calculations to be mapped back into the catchment, but it is not expected that the predictions will be any more than approximately correct in space.

One of the advantages of this distribution function approach is that some important nonlinearities of the runoff generation process can be reflected in the distribution function but without introducing the large number of parameter values needed for fully distributed models. This will generally make model calibration much easier where some observed data are available for comparison with model predictions.

Such models are also now being used as the basis for the type of distributed hydrological response unit models noted earlier, such as the Grid to Grid (G2G) model (Moore *et al.*, 2006; Bell *et al.*, 2009).

2.7 Recent Developments: What is the Current State of the Art?

Computers continue to get more powerful. As a result there is absolutely no doubt that distributed hydrological models will become more detailed, more complex and more closely coupled to geographical information systems for the input of data and display of results (see, for example, Maidment, 2002; Vieux, 2004; Refsgaard *et al.*, 2010). This, in one sense, is the current start of the art. There is a question, however, as to whether this type of development will lead to better hydrological predictions. The answer to this question is not at all clear. More complexity means more parameters; more parameters mean more calibration problems; more calibration problems often mean more uncertainty in the predictions, particularly outside the range of the calibration data.

This still leaves open the possibility that other, parametrically simpler, models may have much to offer. If the interest is in discharge prediction only, then it would appear that, where calibration data are available, simple lumped parameter models, such as IHACRES, can provide just as good simulations as complex physically based models. For distributed predictions, no study has yet demonstrated that a fully distributed model can do better in predicting the distributed responses in a catchment than a much simpler distribution function model, such as TOPMODEL, where the assumptions of the simpler model are reasonably valid (see Franchini and Pacciani (1991) and discussion of Chapter 7).

It is probable that these various categories of rainfall–runoff model will start to be subsumed into a single framework in the next generation of software products. Two important concepts that have been introduced in the decade since the first edition of this book will be important in shaping the next generation of models. These are the Representative Elementary Watershed (REW) concept first introduced by Reggiani *et al.* (1999) and the Models of Everywhere concept introduced by Beven (2006c). These ideas will change the way that modellers approach the modelling problem and so have been given a discussion of their own in Chapter 9.

Something that has not changed is that the application of all types of model is limited by the available data on how hydrological systems work. Models are data constrained because of the strong limitations of current measurement techniques. However, even if improved measurement techniques lead to a better understanding of complex flow processes, it appears that it will be necessary for the foreseeable future to distinguish between models developed for understanding (which describe those processes in detail at small scales) and models developed for prediction at catchment scales. The former will certainly depend on detailed (even if statistical) descriptions of the geometry of the flow domain. The latter will not be able to demand such inputs because they will not be measurable practically or economically at the larger scales at which predictions are required. Models for prediction will necessarily reflect the types of data that are readily available.

In Chapter 3, we look at the data available for modelling in more detail. Chapter 4 then takes a look at modern lumped catchment scale models, Chapter 5 at fully distributed physically based models, Chapter 6 at distribution function and semi-distributed models, and Chapter 7 at model calibration and uncertainty in the practical application of models.

2.8 Where to Find More on the History and Variety of Rainfall–Runoff Models

There are now many more sources of information about hydrological modelling than a decade ago, in the form of written texts, encyclopaedias, e-books, and Internet resources including sources of free and

commercial software. Where these are relevant to the current text they are referenced appropriately, including (as far as is possible) up-to-date web sites. A summary of web resources is also given in Appendix A.

A useful background volume is the IAHS Benchmark volume of classic papers in rainfall–runoff modelling prepared by Keith Loague (2010). The volumes edited by Vijay P Singh (1995) and Singh and Frevert (2002a, 2002b, 2005), which collected chapters written by developers of models for large and small catchments, are also valuable sources of information. Distributed models of various types are discussed in the texts by Mike Abbott and Jens Christian Refsgaard (1996), David Maidment (2002) and Baxter Vieux (2004); lumped conceptual ESMA-type models by Thorsten Wagener *et al.* (2004); and linear systems models by Jim Dooge and Philip O'Kane (2003). There are also many relevant articles in the *Encyclopaedia of Hydrological Sciences* edited by Malcolm Anderson (2005), including a section devoted to rainfall–runoff modelling in Part 11 of Volume 3.

2.9 Key Points from Chapter 2

- Any hydrological model must include functional components that account for the relationship between total rainfall and runoff generation in an event, and the routing of the generated runoff to the catchment outlet.
- The volume of rainfall equivalent to the generated runoff is called the effective rainfall. It has a strongly nonlinear dependence on the antecedent state of the catchment. Once the effective rainfall has been calculated, linear routing methods, such as the unit hydrograph, often work quite well.
- Following the work of Robert Horton, early applications of the unit hydrograph technique assumed that all storm runoff was generated by an infiltration excess mechanism. This is not generally true but the methods continue to be applied successfully, despite difficulties of hydrograph and rainfall separation, because they have the functionality needed for discharge prediction at the catchment scale.
- With calibration of parameter values, even simple storage element (ESMA) models can produce good predictions of streamflow hydrographs and soil moisture deficits.
- Modern transfer function models aim to overcome some of the problems of defining the storage elements of an ESMA model correctly by letting the data available determine an appropriate structure and level of complexity while avoiding the problem of hydrograph separation (see Chapter 4).
- The earliest distributed models were based on the time-area concept. Recent work based on the definition of distributed hydrological response units by overlays of different data types in a GIS system is based on essentially similar concepts.
- Fully process-based distributed models allow the prediction of local hydrological responses within a catchment but have many parameter values that must be specified for every grid element. This makes parameter calibration difficult but direct measurement or estimation of effective parameter values at the grid scale is also difficult due to heterogeneity of catchment characteristics and the limitations of the available measurement techniques (see Chapter 5).
- Simpler models, based on distributions of responses within a catchment may still have much to offer for prediction at the catchment scale and some, such as TOPMODEL, have the potential to map those responses back into the catchment to allow additional evaluation of the simulations (see Chapter 6).
- The accuracy of the predictions of rainfall–runoff models in application to specific catchments remains limited by the availability of data as well as process representations. The resulting uncertainties should, where possible, be estimated (see Chapter 7).
- Recent concepts of the Representative Elementary Watershed (REW) and Models of Everywhere are beginning to blur the boundaries between the different types of model. The next generation of models will change the way in which the modelling problem is tackled (see Chapter 9).

Box 2.1 Linearity, Nonlinearity and Nonstationarity

Unit hydrograph and linear transfer function models in general are based on an assumption of linearity and stationarity in time. In this context this may be simply understood in terms of the relationship between inputs and outputs. A response that is stationary in time means that a unit of input will always produce the same output response (remember that for the unit hydrograph, the inputs are *effective rainfalls* not total rainfalls). A response that is linear means that if two units of input fall in the same time step, we would expect twice the output response. If two units of input fall in successive time steps, the two associated output responses, suitably delayed, can be simply added together to produce a total output. More complex input sequences can be dealt with as the simple addition of unit responses. This is called the principle of *superposition*. It follows directly from the assumption of a linear model.

In a *nonlinear* model, the principle of superposition breaks down since it cannot be assumed that a unit of input will always produce the same output. In rainfall–runoff modelling, nonlinear responses are primarily due to two causes. The most important is the effect of antecedent conditions. Thus the relationship between total rainfall and runoff is generally considered to be nonlinear because the wetter the catchment prior to a unit input of rainfall, the greater the volume of runoff that will be generated.

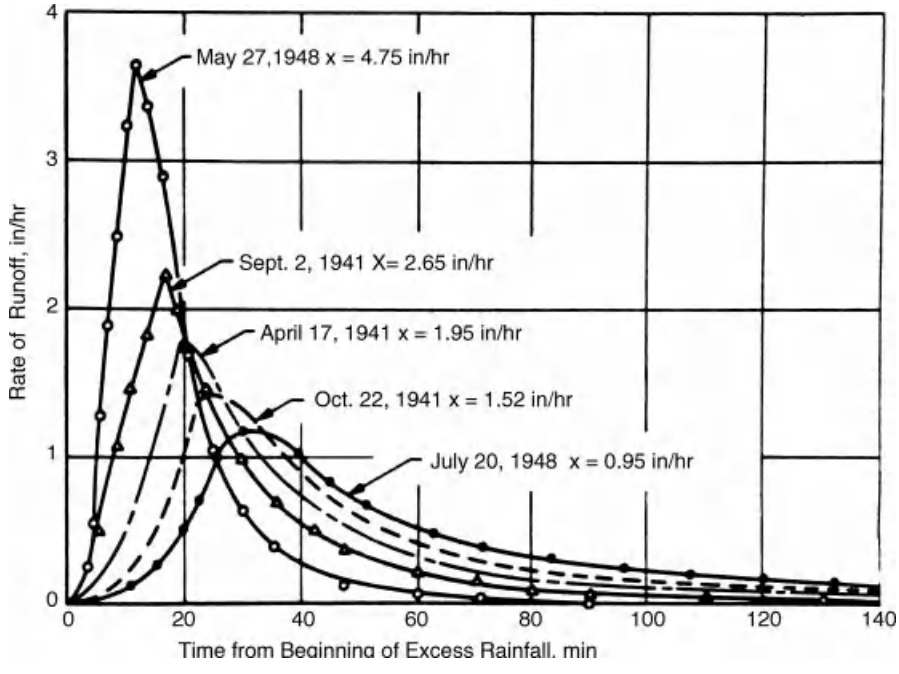


Figure B2.1.1 *Nonlinearity of catchment responses revealed as a changing unit hydrograph for storms with different volumes of rainfall inputs (after Minshall, 1960).*

A secondary cause of nonlinearity is due to the change of flow velocities and celerities with discharge. In Chapter 1, it was explained that inputs to the system will cause a response in the output faster than the pore water velocity (see also Section 5.5.3) due to the way in which pressure changes associated with disturbances to the system propagate with a wave celerity. In

general, for both surface and subsurface flow processes, average flow velocities and celerities will increase with flow in a nonlinear way. Faster flow velocities mean that the runoff will get to a measurement point more quickly; faster celerities mean that the time distribution of runoff (e.g. the shape of the unit hydrograph) will change as runoff increases. This was shown for a small catchment in the classic study of Minshall (1960). Minshall showed a dependence of the shape of the unit hydrograph derived for a small catchment on the volume of effective rainfall. Larger storm sizes resulted in a faster time to peak and higher peak discharge in the unit hydrograph (Figure B2.1.1). In larger catchments, the effects of the routing nonlinearity are not always easy to distinguish in data analysis. The nonlinearity is not always a greater than linear increase with increasing inputs. For example, when a river overtops its banks in a flood, the slow-moving water on the flood plain may lead to a decrease in the average velocity of the discharge and celerity of the flood wave. In semiarid areas, transmission losses due to infiltration into a dry channel bed may lead to responses that become more nonlinear with increasing catchment area (Goodrich *et al.*, 1997).

It is sometimes also difficult to distinguish between simple nonlinear and *nonstationary* responses, and indeed the difference may only be one of interpretation. A nonstationary response is one for which the relationship between inputs and outputs is changing over time. One obvious reason for this would be long-term changes in the characteristics of a catchment due to changes in land use, such as urbanisation or the installation of field drainage. However, the effects of antecedent conditions might also be considered as a nonstationary effect, especially if the nature of the processes involved in runoff generation is changing. Such a relationship is called "nonstationary" if it cannot be represented as a simple nonlinear function of the inputs or other available variables.

Box 2.2 The Xinanjiang, ARNO or VIC Model

A description of the class of models variously named the Xinanjiang (Zhao and Liu, 1995), ARNO (Todini, 1996) or Variable Infiltration Capacity (VIC) models (Wood *et al.*, 1992; Liang *et al.*, 1994; Lohmann *et al.*, 1998a) is included here as one example of an explicit soil moisture accounting (ESMA) or "conceptual" rainfall–runoff model.

ESMA models are typically constructed from connected storage elements, with parametric functions controlling the exchanges between elements, losses to evapotranspiration and discharges to the stream. In general, all the parameters are effective catchment scale parameters and are calibrated on the basis of a comparison of observed and predicted discharges, adjusting the values of the parameters until a best fit is obtained (see Chapter 7 for a discussion of this type of model calibration). This class of models has been chosen over all other ESMA models because of one interesting feature: in VIC-type models, there is a function that attempts to allow for the heterogeneity of fast runoff production in the catchment. Hence the name "variable infiltration capacity", although it should be noted that there is no necessary inference that the fast runoff is produced by an infiltration excess mechanism (but see Zhao and Liu, 1995, for a process interpretation in this way).

The original idea for this class of models originated in the 1970s in China, where it has been widely applied (see Zhao *et al.*, 1980; Zhao, 1992). The idea was later adapted for use in a flood forecasting system for the Arno river in Italy by Todini (1996). Its simplicity has also seen it used in large-scale hydrological modelling and as the land-surface component in atmospheric circulation and global climate models (GCMs) (see Dumenil and Todini, 1992; Liang *et al.*, 1994; Lohmann *et al.*, 1998a, 1998b; Wood *et al.*, 1992) and predictions of the impacts of future climate changes (Christensen *et al.*, 2004). The essentials of the VIC hydrological model have also recently been incorporated into a community land surface model (Wang *et al.*, 2008).

The intention of applying the model in this macroscale context was to try to improve the prediction of runoff production and routing in large-scale land surface parameterisations in GCMs. It had two advantages in this respect: firstly in at least attempting to take account of some of the heterogeneity in soil moisture storage and runoff generation at the large scale (which had been lacking in earlier land surface parameterisations); and secondly the more realistic routing of runoff enabled more rigorous comparisons with observed discharges from large basins. It is also important in this context that these are achieved in a computationally efficient way, since the complexity of individual components of GCMs is still constrained by the limitations of currently available supercomputers. Applications to the Mississippi basin (Liston *et al.*, 1994), Arkansas Red River basin (Abdulla *et al.*, 1996; Abdulla and Lettenmaier, 1997), Columbia River Basin (Nijssen *et al.*, 1997) and Weser River (Lohmann *et al.*, 1998b) have been reported. VIC was included in the Project for the Intercomparison of Landsurface Parameterisation Schemes (PILPS) (see, for example, the comparison of runoff predictions reported in Lohmann *et al.*, 1998c) and in the Model Parameter Estimation Experiment (MOPEX) (Andreassian *et al.*, 2006).

The version of the model described here is the VIC-2L (two-layer) structure of Liang *et al.* (1994, see also Lohmann *et al.*, 1998a). The addition of a thin upper soil layer produces the VIC-3L model (Liang and Xie, 2001). The form of the fast runoff production function is perhaps seen most clearly in Figure B2.2.1. The curved function represents the distribution of local total storage capacities in the basin. As rainfall is added more and more of the storage capacities are filled and, once filled, any excess rainfall on that part of the catchment is assumed to become fast runoff. Between rainstorms, it is assumed that all the storages gradually drain, thereby setting up the antecedent conditions prior to the next storm. It is interesting to note that one of the earliest ESMA models, the Stanford Watershed Model (Figure 2.7), had a similar storage capacity function for the prediction of fast runoff but assumed that the distribution was always uniform between some minimum and maximum capacities across the basin. The variable infiltration capacity form allows for a non-uniform distribution according to the power function:

$$i = i_m \left[1 - (1 - A_i)^{\frac{1}{b}}\right] \tag{B2.2.1}$$

where i is the infiltration storage capacity, i_m is a maximum infiltration storage capacity for the area, A_i is the fraction of the area with infiltration capacity less than i, and b is a shape parameter controlling the form of the distribution. For $b = 1$ the infiltration capacity is uniformly distributed, as assumed in the Stanford Watershed Model. This distribution function has two parameters, i_m and b.

This equation is applied to the upper soil layer. For any level of storage in the upper soil layer, an equivalent threshold for saturation, i_o, and the equivalent area of the catchment that will be saturated, A_s can be calculated (see Figure B2.2.1). The distribution of local storage deficits for the area that is not yet saturated can be then defined as:

$$d_i = i - i_o = i_m \left[1 - (1 - A_i)^{\frac{1}{b}}\right] - i_o; \quad i > i_o \tag{B2.2.2}$$

During a rainstorm, average catchment rainfall in excess of any canopy interception losses is added to the upper soil layer storage at each time step. The saturated area is calculated as that part of the catchment for which upper layer storage exceeds i for that time step. Any rainfall in excess of saturation is assumed to reach the stream as fast runoff. Evapotranspiration is assumed to take place at the potential rate for the area that is saturated and at a reduced rate depending on storage deficit for the remaining part of the area. At the end of the time step, the upper layer storage is depleted by drainage to a lower soil water storage, assuming that drainage is due to gravity alone and that the unsaturated hydraulic conductivity in the upper layer can be described as a function of storage in that layer by the Brooks–Corey relation (see Box 5.4). This requires three further parameters to be specified: a hydraulic conductivity, K_s, a residual moisture content, θ_r, and a pore size distribution index, B.

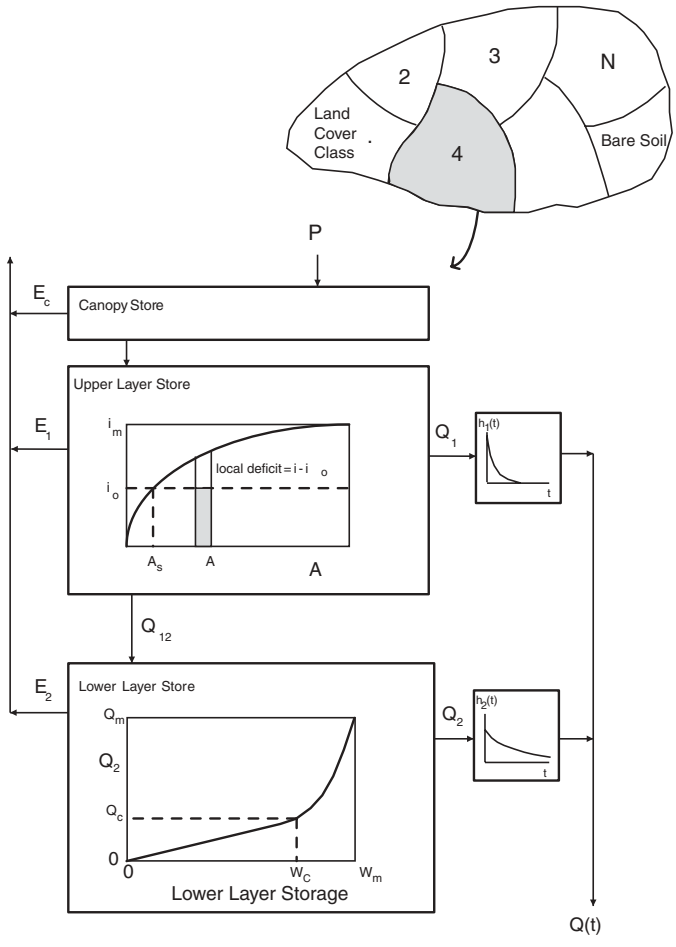


Figure B2.2.1 *Schematic diagram of the VIC-2L model (after Liang* et al.*, 1994, with kind permission of the American Geophysical Union).*

The lower storage is recharged by this drainage from the upper layer and loses water according to a baseflow function that is linear for low storage values but becomes nonlinear at higher storage values so as "to represent situations where substantial subsurface stormflow occurs". The function has four parameter values: the maximum water content in the lower store, W_m, the maximum flow rate from the lower store, Q_m, and the storage and flow rate at the upper limit of the linear part of the function, W_c and Q_c.

Three types of evapotranspiration are included in the model: evaporation from a wet vegetation canopy, transpiration from the vegetation and evaporation from bare soil. The interception capacity of the canopy is taken as a linear function of leaf area index (LAI) with a parameter K_L. The evapotranspiration component uses a Penman–Monteith formulation (see Box 3.2) which also requires specification of aerodynamic and architectural resistance parameters, r_a and r_o, and canopy resistance. Canopy resistance is allowed to vary as a simple function of soil moisture, with five parameters: the minimum value of canopy resistance when water is not limiting, r_{oC}, the moisture content at which transpiration starts to decline as a result of water

stress in each layer, W_1^c and W_2^c, and the wilting point at which transpiration has declined to zero for each layer, W_1^w and W_2^w. The proportion of roots active in transpiration can also be divided between the two layers, adding one extra parameter. Thus, this far the model has 17 parameters for a single land cover element. It may be argued that many of these parameters may be physically meaningful and can be estimated on the basis of soil and vegetation characteristics but one of the limitations of using this type of model at large scales is knowing how local variability should be reflected in effective values of the parameters (see Section 1.8).

An additional snow storage and melt component can be added to the model at the expense of additional parameters (Andreadis and Lettenmaier, 2006). Liang *et al.* (1994) also show how the model can be applied separately to each of a number of different land cover classes. Land cover classes vary in their leaf area index, aerodynamic resistance, architectural resistance, minimum canopy resistance, soil layer storage capacities and relative fraction of roots in the two soil water stores. Increasing the number of land cover classes therefore means that more parameter values must be specified.

The model is completed by routing of the fast runoff production and the baseflow to the catchment outlet. This was originally carried out using two linear storage elements with different mean residence times, that for the baseflow being longer than that for the fast runoff. These parameters are currently difficult to estimate from physical measurement or reasoning and are generally calibrated by comparison of observed and predicted discharges. An improved runoff algorithm for the VIC model has been implemented by Liang and Xie (2001).

As well as the link back in time to the Stanford Watershed model noted above, there are some interesting similarities between features of this model and others to be discussed in Chapters 4, 5 and 6.

This model contains the following assumptions:

A1 Infiltration storage capacity is distributed in space according to a power law distribution.
A2 Vertical recharge to the lower store is calculated as gravity drainage depending on storage in the upper layer store using a Brooks–Corey hydraulic conductivity function.
A3 Drainage from the lower layer store is calculated from a function that has both linear and nonlinear segments.
A4 Calculated actual evapotranspiration takes account of wet and dry canopies, storage in upper and lower soils, and the proportion of active roots in each store.
A5 Runoff generation from the upper store and baseflow drainage from the lower store are routed through parallel linear stores with different time constants.
A6 Different vegetation types may be modelled separately, regardless of their spatial distribution in the catchment (and at the expense of introducing more parameters).

Box 2.3 Control Volumes and Differential Equations

An easy way of understanding the development of the type of differential equations that are used in the physically based descriptions of surface and subsurface flow processes is through a control volume approach. Here, we start by considering the mass balance equation for a one-dimensional channel flow, with x representing distance in the downstream direction and t representing time. We define a control volume of length Δx and look at the changes that take place over a small time step Δt (see Figure B2.3.1). Over the time step, the change in storage in the control volume depends on the net balance of the input from upstream Q_i, the lateral recharge into the channel per unit length of channel q (which might be negative if there is infiltration into the channel bed), and the output discharge from the control volume Q_o.

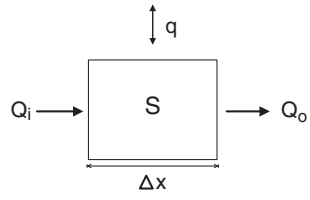


Figure B2.3.1 *A control volume with local storage S, inflows Q_i, local source or sink q, output Q_o and length scale Δx in the direction of flow.*

Thus the mass balance for the control volume may be written:

$$\begin{aligned}\Delta S \Delta x &= \Delta t Q_i - \Delta t Q_o + \Delta t \Delta x q \\ &= -\Delta t\,(Q_o - Q_i) + \Delta t \Delta x q \\ &= -\Delta t \Delta Q + \Delta t \Delta x q\end{aligned}$$

where ΔS is the incremental change of storage per unit length of channel over the time step Δt and ΔQ is the incremental change in discharge across the control volume in the downstream direction (i.e. if the inflow is greater than the outflow, ΔQ will be negative and $-\Delta Q$ will be positive). Dividing through by $\Delta t \Delta x$ gives

$$\frac{\Delta S}{\Delta t} = -\frac{\Delta Q}{\Delta x} + q$$

If we make the increments very small, this may be written in differential equation form as:

$$\frac{\partial S}{\partial t} = -\frac{\partial Q}{\partial x} + q$$

This type of equation is called a partial differential equation because it involves differentials with respect to more than one variable (here x and t).

The control volume approach may be used to derive all the differential equations used in process descriptions in hydrology (such as the channel flow momentum balance equation of Box 5.6). If we extend this mass balance to a control volume in three spatial dimensions, x, y, and z, using local velocity v rather than discharge as the solution variable, we would have

$$\frac{\partial S}{\partial t} = -\frac{\partial v}{\partial x} - \frac{\partial v}{\partial y} - \frac{\partial v}{\partial z} + s$$

where s is a local source or sink flux. This can also be written as

$$\frac{\partial S}{\partial t} = -\nabla v + s$$

where ∇ is called the differential operator.

All these forms of the mass balance equation involve more than one unknown and cannot be solved without more information. This is usually to assume that the local inflow or source/sink flux is known and that understanding of the flow process can be used to relate Q to S, either directly or through some intermediate variable. Many such process descriptions result in non-linear partial differential equations that are not easily solved analytically but must be solved by approximate numerical methods (see Box 5.3).

3
Data for Rainfall–Runoff Modelling

It may seem strange to end a review of modelling with an observation that future progress is very strongly linked to the acquisition of new data and to new experimental work but that, in our opinion, is the state of the science.

George Hornberger and Beth Boyer, 1995

Ultimately, the success of a hydrological model depends critically on the data available to set it up and drive it. In the first two chapters of this book there have been several references to the fact that hydrology is limited as a science by data availability and measurement techniques. In some areas of data collection relevant to rainfall–runoff modelling, techniques have improved in recent years. We now have a much better idea of spatial rainfall variations due to the development of rainfall radar; improvements in transducers and data loggers have led to more reliable and more continuous measurements of water levels, water tables and soil moisture; ultrasonic devices have led to faster (if not necessarily more accurate) measurements of flow velocities in streams and estimation of discharges; there are techniques available for the direct estimation of evapotranspiration rates; and remote-sensing techniques have led to a range of spatial datasets being available for use in modelling. This book is not about data collection for hydrological purposes and does not cover measurement techniques in detail but the rest of this chapter considers the issues associated with the main types of data that are available to the rainfall–runoff modelling process.

3.1 Rainfall Data

Rainfall–runoff modelling still depends heavily on the records from point raingauges, both recording raingauges giving estimates of rainfall intensities at time steps of one hour or better, and daily raingauges. In large catchments, models using a daily time step may be perfectly adequate for application purposes: the spatial variation in inputs is generally more important than the temporal variation. In small catchments, a daily time step may be longer than the storm response time of the catchment and finer time resolution may be required for adequate modelling of the dynamics of the response and the hydrograph peak. Recording

Rainfall–Runoff Modelling: The Primer, Second Edition. Keith Beven.

raingauges become more important but they are more expensive to operate and much fewer in number. Thus it may still be necessary to use daily gauges to get an estimate of the total volume of rainfall over a catchment, using the nearest recording raingauge to give an approximate idea of the distribution of rainfall in time, the *storm profile*.

Even on small catchments, daily rainfalls may be adequate for obtaining acceptable predictions of runoff volumes (rather than hydrograph peaks), especially when volumes over longer times, such as monthly steps, are required. This is implied by the successful simulation of *soil moisture deficits* shown in Figure 2.9, since discharge will complete the water balance for these sites. Hydrograph prediction, however, is particularly difficult when a storm spans two daily measurements, since the fixed daily measurement period (often 9 a.m. to 9 a.m.) is hydrologically arbitrary.

Raingauge-measured volumes may be subject to error. In particular, they depend on the design of the raingauge in relation to wind conditions at the site and rainfall intensities. The best design is thought to be a raingauge with the orifice set at ground level and surrounded by an anti-splash grid but this is not always practical, particularly in environments with frequent snow. A variety of designs of wind shield have been used in different countries to try to mitigate this wind effect. The wind effect can be large; estimates of reductions of up to 20% have been reported at windy sites for gauges only 30 cm above the ground compared to ground-level gauges (Rodda and Smith, 1986). High rainfall intensities can also cause problems for some types of recording raingauge, such as the tipping bucket; if the tips start to occur too rapidly the buckets start to bounce, so that high intensities may require a specific calibration.

Rainfall volumes and intensities can vary rapidly in space and time, particularly in convective rainfall events (see Figure 3.1). Thus, as well as an interpolation of rainfall volumes in time to produce the storm profile, it may also be necessary to interpolate in space since raingauge measurements represent only

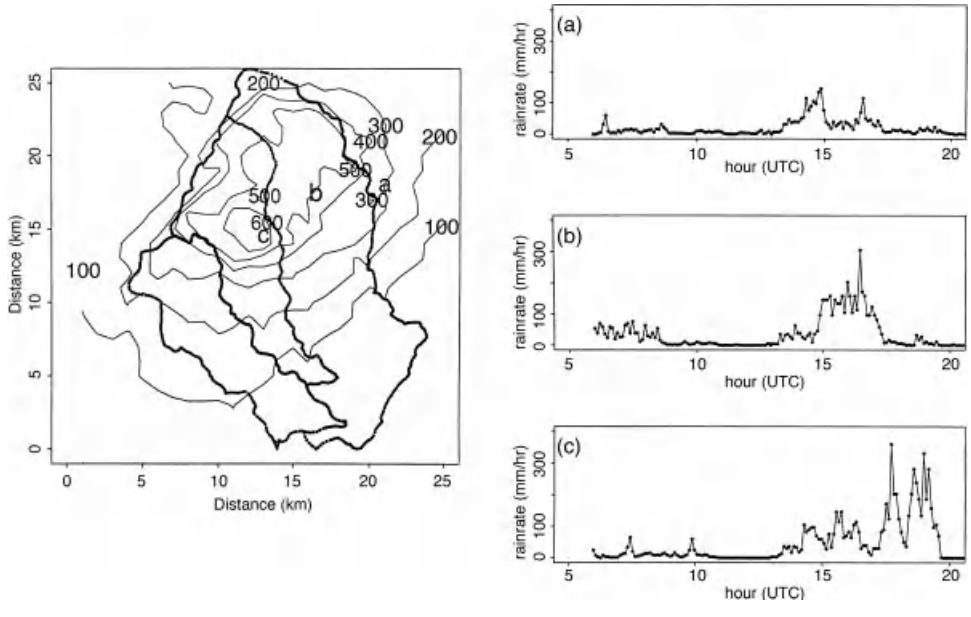


Figure 3.1 *Variations in rainfall in space and time for the storm of 27 June 1995 over the Rapidan catchment, Virginia (after Smith et al., 1996, with kind permission of the American Geophysical Union).*

point measurements. A number of techniques are available for such spatial integration: simple averaging, Thiessen polygons and inverse distance weighting and a variety of others (see, for example, Shaw *et al.*, 2010). None of these techniques can be more than an approximation to the actual volume of rainfall over the catchment and accuracy of a particular technique is likely to change from storm to storm.

The development of radar rainfall measurement has led to a much greater appreciation of the temporal and spatial variation of rainfall intensities than was previously available from raingauge measurements alone. Much of Europe and large areas of the USA are now routinely monitored by ground-based radar rainfalls. The radar has a revolving antenna that sends regular electromagnetic pulses at a low upward angle into the atmosphere. A detector measures the strength (and, in some cases, the frequency attenuation) of the return signal. The principle is that the return signal to the radar is strongly dependent on the intensity of the precipitation in the path of the radar beam at different distances from the measurement site. A calibration function then allows the intensity of rainfall at each distance to be estimated; the estimated intensities are then normally interpolated onto a square grid, commonly with a resolution of 2 or 4 km for operational radars.

This would appear to be a very important development in the data available for rainfall–runoff modelling, and indeed it is, but there are some important limitations that must be recognised. The first is that the radar does not measure the rainfall at ground level but at some distance above the ground (often hundreds of metres and increasing away from the radar station). There is thus potential for changes in the patterns of intensities at ground level, particularly where winds are strong and where there is a strong orographic effect. Secondly, the calibration of the radar depends not only on rainfall intensity but also on the type of precipitation, particularly the drop size distribution and whether the rain is all liquid water or a mixture of water and ice (which produce very different return signals). Thus it may be necessary to "correct" the basic calibration of the radar in different ways. This is most often done by continuously adjusting the estimates of intensity produced by the radar using online data from recording raingauges at the ground surface. In some sense then, the radar becomes an expensive (but very effective) spatial interpolation technique. It is effective because it can give an indication of cells of high rainfall intensity that might be completely missed by the network of ground-level gauges. This may be particularly important in high magnitude, localised rainfall events (such as that of Figure 3.1; from Smith *et al.* (1996) for an example of a high-magnitude event in Virginia).

Direct measurement of drop size distributions using disdrometers has also been used to try to improve calibration (but is subject to the limitation that the size distributions measured at ground level might not be the same as those at the height of the radar beam, even allowing for wind drift effects). It might also be possible to implement local radar systems with overlapping coverage of a catchment to resolve some of the problems with the existing network of weather radars making use of cheaper X-band radars (see, for example, Anagnostou *et al.* 2004; Matrosov *et al.* 2010; van de Beek *et al.*, 2010). X-band radar has a much more limited areal coverage but can give more local detail that might be particularly useful, for example, in forecasting runoff in urbanised catchments (Maki *et al.*, 2005).

An alternative method of obtaining spatial rainfall information is to make use of the attentuation of existing microwave signals, such as the extensive network of mobile phone transmitter–receiver links. This requires no additional equipment, but can only estimate rainfalls averaged over the length of a linear link. It also requires negotiations with mobile phone companies to obtain the data. This work is still in its early stages but might prove useful in future (see, for example, Leijnse *et al.*, 2007, 2008).

The estimation of rainfall is very important in rainfall–runoff modelling, since no model, however well-founded in physical theory or empirically justified by past performance, can produce accurate hydrograph predictions if the inputs to the model do not adequately characterise the rainfall inputs (the well-established GIGO principle of Garbage In Garbage Out applies). A good example is reported by Hornberger *et al.* (1985). In their study, a calibrated rainfall–runoff model for the 5 km^2 White Oak Run catchment in Virginia, which had previously performed well in reproducing observed discharges,

failed completely to predict a later "validation" storm. In fact, the volume of rainfall recorded by the raingauges was far less than the volume of discharge recorded in the stream (there was a line of raingauges at different elevations on one side of the catchment, but the intense storm was centred on the other side). It is difficult in such circumstances for any model to predict the response accurately. Inadequate estimation of the rainfall inputs to a catchment must therefore increase the uncertainty of runoff predictions.

A number of models reported in the literature have included a rainfall multiplier as a parameter to be calibrated as one way of trying to allow for the fact that the raingauge data available might not be a good characterisation of the rainfall inputs to a catchment area. It is not clear that this is generally a good strategy in that it may be only some events for which the rainfall inputs are not well estimated. For extreme events, as in the White Oak Run example above, it may be quite obvious that there is a problem. For more moderate events, it may be suspected that some events are not well estimated, but it will not be clear which events might be problematic. Thus, a constant rainfall multiplier would not be an appropriate way of adjusting the catchment inputs. It would, then, be better to implement some quality controls (see Section 3.2) and, if necessary, exclude some periods of data from the modelling exercise as unreliable or *disinformative* (see Beven and Westerberg, 2011).

However, there may be cases when an adjustment might be justified. It is a fairly common situation, for example in mountainous terrain, that one or more raingauges might be available in the valley bottoms, but none in the higher elevations where it is expected that precipitation inputs might be greater. The average catchment rainfall input might then be consistently higher than that recorded by the valley bottom gauges. Some adjustment would be necessary to achieve a reasonable water balance. Even in such a case, however, calibration of a rainfall multiplier might not be the best solution, since there would be a distinct possibility that the calibration process would result in interaction with other parameters being calibrated at the same time. This is definitely the case when event by event rainfall multipliers are used to try to correct for input errors (see Section 7.8). Thus, it might be better to make a prior adjustment on the basis of physical reasoning rather than allowing the multiplier to vary during calibration. This would also allow for the possibility of making different adjustments in different periods, although it will be rare that there would be sufficient information on which to base such a variable adjustment. However, calibration of any other parameter values would still be conditional on the adjusted inputs. The general case, as noted in Section 1.8, is that any calibrated parameter values must be conditional on the sequence of inputs used, even if no such adjustments are made.

The estimation of precipitation inputs in the form of snow raises a whole range of additional problems. The hydrologist is interested in the water equivalent of the snow, which depends on both its depth and profile of density, both of which change over time as the snowpack structure evolves and ripens. Snow water equivalent may be measured directly at a point or on transects known as *snow courses* by field measurements of snow depths and density profiles but this can be arduous and expensive to maintain at frequent intervals. The best continuous measurements method available is to measure the weight of snow above a point using a pressure measurement device, such as a snow pillow. An increase in the pressure indicates a new fall of snow; a decrease, loss by *sublimation* or, more importantly, melting. The continuous measurement of pressure can then give a good indication of the rates of melt that are required for hydrograph modelling.

Unfortunately, such installations are expensive and remain relatively rare. They also, like a raingauge, only give an indication of conditions at a single point and snow packs are renowned for their variability both in terms of water equivalent and rates of melt, particularly in mountainous terrain and where vegetation extends above the pack. The redistribution of snow by wind; the effects of topography and vegetation on snow collection, temperature and insolation conditions; freeze–thaw cycles; and changing pack albedos over time are all factors that affect this variability and make modelling snowmelt very difficult indeed (see, for example, the study by Bathurst and Cooley, 1996). This is one area of hydrology where remote sensing has proven especially useful (see Section 3.7).

3.2 Discharge Data

The availability of discharge data is important for the model calibration process. Discharge data are, however, generally available at only a small number of sites in any region. It is also an integrated measure in that the measured hydrograph will reflect all the complexity of flow processes occurring in the catchment. It is usually difficult to infer the nature of those processes directly from the measured hydrograph, with the exception of some general characteristics such as mean times of response in particular events. Rainfall–runoff modelling for sites where there are no discharge data is a very much more difficult problem. This ungauged catchment problem is one of the real challenges for hydrological modellers and has been the subject of the ten-year Prediction of Ungauged Basins (PUB) research initiative of the International Association of Scientific Hydrologists (see Chapter 10).

There are many different ways of measuring discharges (see, for example, Herschy, 1995). Except for very small flows, it is difficult to make a direct measurement. The level of water in a channel is, however, relatively easy to measure and most methods for estimating discharges require a conversion of a water level measurement to flow. If it is done as the water flows through a well-maintained weir or flume structure, this conversion can be accurate to better than 5%. If there is no such structure, or if a structure is overtopped in a high flow, then the accuracy may be very much worse than this, particularly where the cross-section and effective roughness at the gauging site might change over time as a result of either sediment transport or seasonal vegetation growth (e.g. Westerberg *et al.*, 2010a). In the worst case of extreme floods, the water-level measuring device may itself be washed away and then the only resort is to try to estimate the maximum flow using the *slope-area* method, in which the cross-sectional area of the flow and the slope of the water surface are estimated from the trash lines indicating the maximum extent of the flow and a uniform flow roughness equation is used to determine an average velocity. Since at the crest of a flood, the flow may be non-uniform, highly turbulent, and with a high sediment load in a dynamically changing cross-section, it may be difficult to estimate an effective roughness coefficient and cross-sectional area, and hence the average velocity and discharge. Errors in discharge estimates will then be much higher.

These potential errors tend to get forgotten when the discharge data are made available as a computer file for use in rainfall–runoff modelling. There is always a tendency for the modeller to take the values as perfect estimates of the discharge. To some extent this is justified: the data are the only indication of the true discharges and the best data available for calibrating the model parameters. However, if a model, any model, is calibrated using data that are in error, then the effective parameter values will be affected and the predictions for other periods, which depend on the calibrated parameter values, will be affected. This is an additional source of uncertainty in the modelling that we return to in Chapter 7.

For now, it is worth stressing that, prior to applying any model, the rainfall–runoff data should be checked for consistency. Some errors, of course, may not be obvious, but the following types of simple check can be made (see also Section 7.17).

- If possible, check the discharge rating curve for consistency, and for how far the "observed" discharges are the result of extrapolation of the rating curve far beyond the range of the directly measured discharge values. It is very difficult to obtain direct measurements of discharges during flood conditions (even at sites with a gantry or cableway extending across the flood plain) so high discharges determined by extrapolation might be very uncertain.
- Calculate the total volumes of rainfall and runoff for different periods in the record, choosing periods separated by similar low flows where possible so that the calculated volumes are not greatly affected by recession discharges. Is the runoff coefficient (the ratio of runoff to rainfall volume) consistent with expected seasonal changes? Lower values would be expected in the summer, higher values in the winter.

- Are the runoff coefficients consistent for increasing storm volumes (allowing for seasonal variations)? For example, are any runoff coefficients greater than 100%? This would indicate that one or other of the measurements is in error since mass balance makes it difficult for a catchment to produce more runoff outputs than rainfall inputs.
- If more than one discharge gauge or raingauge are available, check for consistency between the gauges (normalising for differences in area for discharges). Compare runoff coefficients or use *double mass curves* to check for changes in slope in the accumulated volumes at different gauges.
- Check for any obvious signs that infilling of missing data has taken place. A common example is where measured rainfall intensity is apparently constant for a period of 24 hours, suggesting that a volume from a daily raingauge has been used to fill in a period where the recording raingauge was not working. Hydrographs with long flat tops are also often a sign that there has been a problem with the measurements.

These types of simple check are easy to make and at least allow some periods of data with apparently unusual behaviour to be checked more carefully or eliminated from the analysis. There is a danger, of course, of rejecting periods of data on the basis that a chosen model cannot be made to give a good simulation of that period. Unless there is some other reason for rejection, this should not be considered good practice, since it is normally the case that the modeller learns more about the limitations of a model from situations where it cannot give good simulations than where it does. The reader should remain aware, however, that most of the hydrological modelling literature tends (quite naturally) to report the best simulations with any given model rather than the worst!!

3.3 Meteorological Data and the Estimation of Interception and Evapotranspiration

3.3.1 Estimating Potential Evapotranspiration

In many environments, evapotranspiration makes up a larger proportion of the catchment water balance than stream discharge. Thus, for longer periods of rainfall–runoff simulation, it will generally be necessary to estimate actual evapotranspiration from a catchment in order to have an adequate representation of the antecedent state of the catchment prior to each rainfall event.

We must distinguish here between estimates of *potential evapotranspiration* and *actual evapotranspiration*. Potential evapotranspiration is the loss expected over a surface with no limitation of water. It is a function of the *atmospheric demand*, that is the rate at which the resulting water vapour can be moved away from the surface. The atmospheric demand depends primarily on the energy available to convert liquid water to vapour from net radiation, the humidity gradient in the lower atmosphere, the wind speed, and the roughness of the surface. Rough surfaces, such as forests, will have higher potential evapotranspiration rates than smooth surfaces, such as a lake, given similar radiation, humidity and wind conditions. In general, actual evapotranspiration rates will be at the potential rate until the water supply from the soil becomes limiting.

Methods for estimating potential evapotranspiration range from using a simple annual sine curve to the more physics-based Penman–Monteith equation detailed in Box 3.1. A simple seasonal sine curve for daily potential evapotranspiration, regardless of the variations in the weather, would appear to be far too simplistic to be a useful model of potential evapotranspiration. The study of Calder *et al.* (1983), however, showed that such a curve could give results equally as good as more complex formulations requiring more data in modelling soil moisture deficits for several sites in the UK. They treated the mean daily potential evapotranspiration, the only parameter required by the model, as a parameter to be calibrated and found that similar values could be used for all their study sites. A seasonal sine curve can be defined

generally in the form:

$$E_p = \bar{E}_p \left(1 + sin \left\{ \frac{360i}{365} - 90 \right\} \right) \tag{3.1}$$

where $\bar{E}_p$ is the climatological mean daily potential evapotranspiration rate in mm/day and i is the day of the year. Thus, the only parameter needed to apply the sine curve method is the mean annual daily potential evapotranspiration rate. If the diurnal variation in evapotranspiration is also of interest, then the daily values can be redistributed in time using a separate sinusoidal variation over the potential sunshine hours each day.

This method has the advantage that it requires no meteorological variables to be available, but clearly can therefore take no account of the effects of changing temperatures, humidity and cloudiness from day to day and hour to hour on the potential evapotranspiration estimates. The effective mean daily evapotranspiration rate, $\bar{E}_p$, will then become a parameter to be estimated, although the Calder *et al.* (1983) study suggested that, at least in a humid temperate environment, this might be a relatively conservative quantity in space.

A number of empirical approaches for estimating potential evapotranspiration have been suggested based on different levels of data availability. For example, if only mean daily temperature data are available, the empirical equation of Hamon (1961) can be used to estimate daily potential evapotranspiration rates. There are many others based on the Dalton evapotranspiration law that require data on the humidity deficit, i.e. both wet and dry bulb temperatures (see Bras, 1990; Calder, 1990; Singh and Yu, 1997). All empirical methods are conditional on the range of conditions used in their calibration and care should be taken not to use them outside that range. Comparisons of a variety of methods for estimating potential evapotranspiration rates have been provided, for example, by Federer *et al.* (1996), Xu and Singh (2002) and Lu *et al.* (2005).

The best simple physics-based approach available is the Penman–Monteith equation (Monteith, 1965; see Box 3.1). This equation attempts, in a simple way, to take into account the energy balance of the surface and the way in which turbulence in the low atmosphere controls the movement of vapour away from the surface, but is more demanding in terms of both data requirements and parameter values. It requires meteorological data on net radiation, air (dry bulb) temperature, humidity (or wet bulb temperature) and wind speed. It also requires an estimate of two resistance coefficients: *aerodynamic resistance*, r_a, is an expression of the roughness of the canopy surface; *canopy or surface resistance*, r_c, is an effective parameter for the surface as a whole, expressing how easily water vapour moves from the stomata or leaves, or the pores of a bare soil surface, into the air. For a wet canopy or surface, the surface resistance is zero. For a dry canopy, but without a limitation on supply of water, a typical value of r_c might be 50 ms^{-1}. This implies that in fact there should be different potential rates of evapotranspiration depending on whether the canopy is wet or dry (see next section).

These are representative of the methods available for the calculation of potential evapotranspiration rates. However, as we discussed in Chapter 1, a limitation on the supply of water available for evapotranspiration during long dry periods will mean that the actual evapotranspiration rate may be much smaller than the potential rate. In the Penman–Monteith equation; this would be reflected in an increasing canopy resistance as the soil dries. Most rainfall–runoff models contain components, more or less sophisticated, that attempt to simulate this reduction in actual evapotranspiration as the soil dries and water becomes limiting. Models intended primarily for the prediction of discharges have tended to use relatively simple components based on relating soil moisture storage to a ratio of actual to potential evapotranspiration; but it has been suggested that a better approach is to predict actual evapotranspiration rates directly through the use of the effective canopy resistance (Wallace, 1995).

Recent developments in so-called SVAT (soil–vegetation–atmosphere–transfer) or LSP (land surface parameterisation) models, which aim to predict the fluxes of latent and sensible heat to the atmosphere

as a boundary condition for atmospheric circulation models, have resulted in highly complex model structures with multiple soil layers and multiple vegetation layers. In effect, these models aim to predict the way in which the Penman–Monteith canopy resistance changes with water availability and other factors such as solar radiation, leaf temperature, carbon dioxide concentration, vapour pressure deficit and position in the canopy. Some recent examples are the Community Land Model (CLM) (Bonan *et al.*, 2002; Oleson *et al.*, 2008); the UK Met Office MOSES model (Smith *et al.*, 2006; Rooney and Claxton, 2006); and ISBA-TOPMODEL used by MeteoFrance (Vincendon *et al.* 2010). Such models have a very large number of parameters for each of the soil and vegetation layers that may be very difficult to estimate *a priori* (and may, in fact, change over time). Interestingly, the runoff generation components of such models tend to be rather simple (see, for example, Lohmann *et al.*, 1998c). This is a good example that what is considered important in a model depends, initially at least, on the perceptions and background of the modeller.

It is worth noting that the availability of meteorological data may be a problem in applying some of the more demanding methods, including the Penman–Monteith equation. Net radiation, temperature, humidity and wind speed data necessitate either an automatic weather station to be installed within the catchment of interest or, at the very least, a high-quality meteorological station to be nearby. When this is not the case, some of the simpler methods, even the simple sine curve approach, may still have value.

At many principal meteorological stations, an evaporation pan may also be available, with measurements of the depth of water lost from a reservoir open to the atmosphere, usually on a daily basis. There are several different sizes of pans in use, even within the USA. Such measurements can give an index of the rate of *potential evapotranspiration* at a site but the measured rate does depend on the way in which the pan is exposed and the nature of the surroundings at the site. In general, such pans evaporate more water than would be lost from the surrounding surface, even for non-limiting water conditions. Thus, in general, pan evaporation estimates must be multiplied by an empirical pan coefficient to improve the estimate for potential evapotranspiration for a particular type of surface. For a US "Class A" pan, for example, the coefficient is of the order of 0.7. Pan coefficients are tabulated in many hydrological texts (e.g. Bras, 1990) and the UN FAO have proposed a widely used set of coefficients to adjust measured pan evaporation to water use by different crops (see the review by Pereira *et al.*, 1999).

3.3.2 Evaporation of Water Intercepted by the Vegetation Canopy

The very low canopy resistance for wet canopy conditions noted above is why, given a source of energy, the loss of intercepted water from a wet canopy tends to be at higher rates than transpiration from a dry canopy. The effect is particularly marked for rough forest canopies in windy conditions. This can be taken account of within the Penman–Monteith formulation of Box 3.1 by allowing high rates of evaporation at $r_c = 0$ from a conceptual interception storage until that storage is dry, after which transpiration rates are predicted using dry canopy r_c values. The evaporation of intercepted water from leaf surfaces in rough canopies can be very efficient and is a significant component of the total water balance in some environments (Calder, 1990). It has also been suggested that there could even be significant losses during some storm events where the rain falls through unsaturated air such that there is still a humidity deficit above the canopy. A number of models of the *interception* process have been proposed, of varying degrees of complexity (Rutter *et al.*, 1975; Gash, 1979; Calder, 1986). The most widely used is probably the Rutter model which is described in detail in Box 3.2 together with the Calder stochastic model. In general, without specific measurements of throughfall and stemflow below a vegetation canopy, it is not possible to identify the parameters of an interception model independently, so that estimating the parameters of the model depends on finding a study of a similar vegetation type reported in the literature, although extrapolation from one site to another should be done with care (see Chapter 10).

3.3.3 Direct Estimation of Actual Evapotranspiration

There are now methods available for the direct measurement of actual evapotranspiration rates over a surface using the *eddy correlation* technique (see, for example, Shuttleworth *et al.*, 1988) but, although a global network of eddy correlation measurement stations is expanding, the main use of such instruments has been during short field campaigns studying land–atmosphere interactions (Figure 3.2). More generally, indirect methods of estimating evapotranspiration are used.

Another promising technique is the use of the laser scintillometer. This uses measurements of the disturbance of a laser beam by rapid variations in the density of the lower atmosphere to estimate the transfers of sensible heat away from the surface. Coupled with a closure of the energy budget for the surface, this can also give an estimate of the latent heat fluxes and evapotranspiration. One nice feature of this technique is that it can integrate the fluxes over the length of the laser beam and therefore provides a larger scale measurement of evapotranspiration than eddy correlation measurements at a site

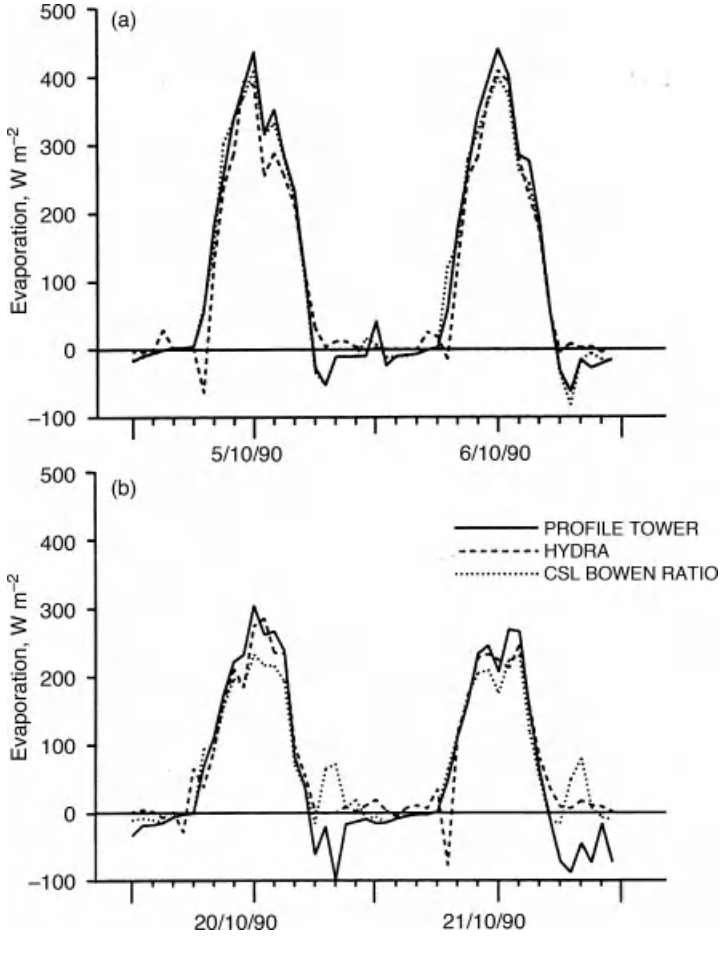


Figure 3.2 *Measurements of actual evapotranspiration by profile tower, eddy correlation and Bowen ratio techniques for a ranchland site in Central Amazonia (after Wright et al., 1992, with kind permission of John Wiley and Sons).*

(although the latter will also be affected by the nature of the evapotranspiration over some variable upwind fetch length). The method has been shown to provide good estimates of sensible heat fluxes over both homogeneous (De Bruin *et al.*, 1995; McAneny *et al.*, 1995) and heterogeneous (Chehbouni *et al.*, 1999) surfaces and has been used to ground truth other methods of obtaining large scale estimates of actual evapotranspiration (e.g. Schüttemeyer *et al.*, 2007).

The evapotranspiration at a point is necessarily affected by the nature of the surrounding surface. Following original work by Bouchet (1963), Morton (1978) has suggested that pan measurements can be used to derive estimates of the actual evapotranspiration of the surrounding area since, under given energy inputs, the lower the actual evapotranspiration of the surrounding area, the drier the air and the greater the resulting pan measurement will be. This *complementarity approach* has been used in a number of catchment models and extensively tested by Morton (1983a, 1983b). Venturini *et al.* (2008) have applied the complementarity approach at large scales for different land uses using satellite data. The complementarity approach has been criticised as inaccurate (for example by LeDrew, 1979, and Lhomme and Giulioni, 2006) but has a certain intrinsic appeal and may yet require re-evaluation as the difficulties of the more physically based but parameter-rich approaches become more widely appreciated. Recent modifications of the method have been suggested by Crago and Crowley (2005) and Szilagyi and Jozsa (2008).

3.4 Meteorological Data and The Estimation of Snowmelt

In many environments, snowmelt may be the source of the annual maximum discharge in most years and may be a major cause of flooding. Meteorological data will also be required in the modelling of snow accumulation and melt. Again, different types of snow models demand different types of data. The very simplest snow model is the temperature index or *degree-day method*. This model, in its simplest form, is based on the hypothesis that snowmelt is proportional to the difference between air temperature and a threshold melt temperature (Box 3.3). Thus data on air temperature are required as an input; the threshold temperature is effectively a parameter.

A typical modern variant for snowmelt–runoff modelling, the Swiss SNOW1-ETH4 model, has been proposed by Hottelet *et al.* (1993, see also Ambroise *et al.*, 1996). This still only requires temperature as an input but tries to take account of whether precipitation is falling as rain or snow and the heat deficit of the pack that must be satisfied before significant melt will occur. This variant increases the number of parameters that must be specified. The degree-day method is applied within the Snowmelt Runoff Model (SRM) of Rango and Martinec (1995), which has been widely applied in both the USA and Europe (e.g. Mitchell and DeWalle, 1998). The SRM makes calculations for different elevation bands within the catchment and takes account of the depletion of the snow-covered area as the melt season proceeds (see Box 3.3).

The degree-day method is obviously a very simple approach, but has the advantage of demanding only temperatures as an input. The method is most accurate when melt is dominated by heat input due to radiation and the pack is ripe at 0°C and ready to melt. The method is least accurate when melt is dominated by heat advected by an air mass (see, for example, Braun and Lang, 1986). A combination energy budget approach to modelling the snowpack and snow melt, similar to that used for evapotranspiration in the Penman–Monteith equation, was originally proposed by Anderson (1968). This is much more demanding in both meteorological data (net radiation, temperatures, humidity and wind speed) and parameter values. Models of the changing structure of the snowpack as a result of freezing and thawing conditions have also been proposed (e.g. Morris, 1991), requiring even more parameters. A number of distributed energy budget models for predicting snowmelt have been developed (e.g. Blöschl *et al.*, 1991; Marks and Dozier, 1992). However, an intercomparison of snowmelt models carried out by the World Meteorological Organisation

(WMO, 1986) showed, at that time, no clear advantage of the more complex models when comparisons were made for a variety of catchments and over a number of years of data.

3.5 Distributing Meteorological Data within a Catchment

One of the problems in applying all these evapotranspiration and snowmelt models at the catchment scale is taking account of the variability of meteorological conditions within the catchment. Insolation depends on the angle and aspect of different slopes; wind speeds depend on wind direction and pressure gradients in relation to the form of the topography; temperatures depend on elevation; humidities depend on evapotranspiration upwind. Distributed rainfall–runoff models have the potential to take such variations into account, but this requires a further model to distribute the meteorological data measured at one point or, at best, a small number of points, to other points in the catchment using digital elevation and other distributed data. This problem is greatest for hilly and mountainous catchments with a wide elevation range. It is a particular problem for the estimation of snowmelt early in the season, since snowmelt generally starts at lower elevations on slopes with a southerly aspect and may be significantly delayed at higher elevations. Models for predicting distributed meteorological data within a catchment have been suggested, for example, in the SAFRAN-CROCUS snowmelt model of Brun *et al.* (1993), in RHESSys (Band *et al.*, 1991; Hartman *et al.*, 1999), and by Blöschl *et al.* (1991).

Remember that there is also the problem in snowmelt modelling of knowing how much snow is there to melt in the first place, since it is very difficult to obtain information on spatial patterns of snow depth and density to get estimates of snow water equivalent. It is possible to use remote sensing to estimate changing patterns of snow covered area which can be used as a constraint on snow melt models (e.g. Blöschl *et al.*, 1991; Rango, 1995).

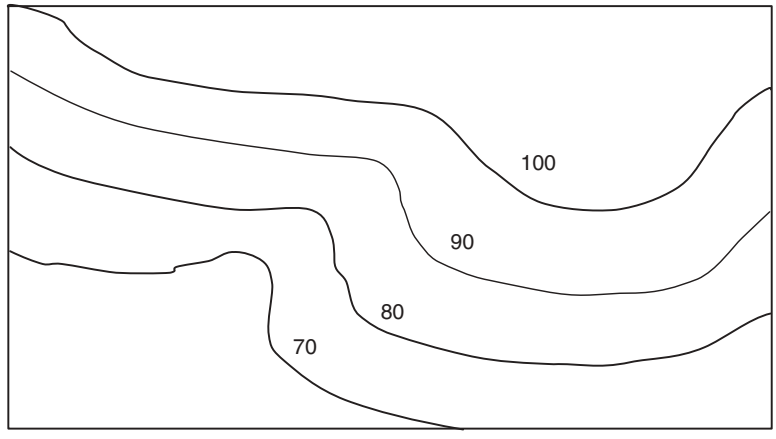
3.6 Other Hydrological Variables

Rainfalls and discharge are the measured hydrological variables most often available to the modeller and certainly are the most useful to the rainfall–runoff modeller. However, in some catchments, other types of hydrological measurement may be available, such as measurements of standing water levels in wells, profiles of soil moisture and spatial patterns of near surface soil moisture. Such data clearly gives more information on the hydrological behaviour of a catchment but the amount of information may be limited since, with the exception of a few research catchments, the number of measurement sites is likely to be small. The scale of the measurements is also important in this context. Such internal measurements tend to be small scale or "point" measurements, reflecting the hydrological conditions only in the immediate vicinity and, to some extent, up-gradient. Thus, it may be difficult to compare such measurements with the predictions of even the most distributed rainfall–runoff models available. Geophysical methods such as ground-penetrating radar, electrical resistance tomography and cross-borehole tomography can give a better indication of moisture patterns in space, but are still only relatively local in scale. The use of such internal measurements in model calibration and evaluation are considered again in Sections 5.4 and 6.4.

3.7 Digital Elevation Data

In many developed countries of the world, digital elevation maps (DEM) or terrain maps (DTM) are becoming available at a resolution fine enough to broadly represent the form of hillslopes (50 m in the UK and France; 30 m in the USA; 25 m in Switzerland). DEMs with a fixed grid size are known as *raster* data. Digitised contour maps (*vector* DEMs) may also be available (Figure 3.3a). In fact, to date, most

(a)



(b)

95	103	103	104	104	105	105	105	104	103	103
86	91	95	95	97	100	102	103	103	103	100
78	80	82	84	84	90	95	100	103	100	95
71	72	72	70	78	86	86	95	96	95	86
69	69	67	68	75	80	82	84	84	82	80
69	66	68	69	70	72	75	76	78	77	76

(c)

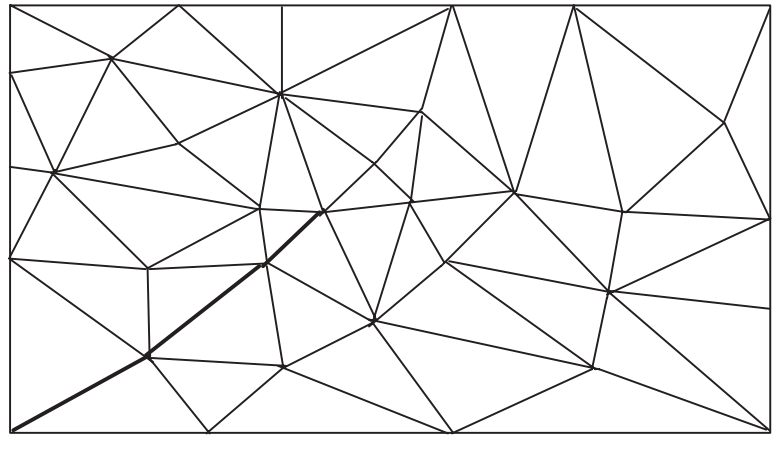


Figure 3.3 *Digital representations of topography: (a) vector representation of contour lines; (b) raster grid of point elevations; (c) triangular irregular network representation.*

raster DEMs have been built by interpolating from digitised contours (Figure 3.3b) and, as a result, may, in places, be subject to significant error, particular where in flat topography where there are few contours (particularly where, in some cases, databases store only integer elevation values) or where there are short steep slopes. Topography can also be efficiently represented as a *triangular irregular network* (TIN, Figure 3.3c). There is potential for developing good resolution topographic data from photogrammetric analysis of aircraft- or satellite-derived stereo images or directly from aircraft-borne laser altimetry (e.g. Weltz *et al.*, 1994) or synthetic aperture radar (SAR). Aircraft platform techniques can give elevations on a grid of down to better than 2 m with elevations of an accuracy of 0.1 m, although it must be pointed out that these will be elevations of the surface seen by the sensor, which is not necessarily the same as the ground surface where there is significant tall vegetation cover or buildings.

NASA's Shuttle Radar Topography Mission (SRTM) provided a digital elevation model derived from SAR technology at a global resolution of 90 m and a refined resolution of 30 m in the USA. Since 2009, this has been superceded by the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) mission global DEM which provides a wider coverage at a resolution of 30 m between latitudes of 83 degrees north and south. This was calculated from 1.3 million near-infrared images using stereoscopic methods. It has been suggested that there are some artifacts in the ASTER DEM and that it is inferior to the SRTM data at the same resolution. It is worth remembering that, as hydrologists, we are generally interested in relative elevations and gradients to derive flow pathways but this is a difference operation that is sensitive to uncertainties in the data (e.g. Endreny and Wood, 2001). Thus, it will be worth investigating the data retrieved from these types of database for unusual features that might be artifacts of the way they have been produced. Topographic data are also useful, however, in modelling patterns of evapotranspiration and snowmelt (e.g. Kafle and Yamaguchi, 2009)

Water does have a tendency to flow downhill, at least for shallow hydrological systems, so that knowing something about the form of the topography should have some utility in hydrological modelling. Distributed models can clearly use this type of data directly and there are also models, such as TOPMODEL, that are based on a prior analysis of the catchment topography (see Section 6.3 and Box 6.1). Resolution is clearly an issue here. Coarse resolution DEMs will not be able to provide an adequate description of hillslope flow pathways; distributed models may not be able to use all the information in a fine scale DEM because of computational constraints. Variables derived from topographic data, calibrated parameter values, and model predictions within distributed models based on DEMs are known to be sensitive to grid resolution (e.g. Zhang and Montgomery, 1994; Bruneau *et al.*, 1995; Quinn *et al.*, 1995; Saulnier *et al.*, 1997a).

The analysis of a DEM to derive apparent flow pathways has been an interesting topic of research in itself. The methods available depend on whether a raster or vector DEM is available. For raster DEMs, a comparison of methods was published by Tarboton (1997). For each grid cell, there are eight possible flow directions. There may be several surrounding grid elements with elevations lower than the cell being considered. The problem is how to distribute the potential flow to these different possible pathways.

Some inaccuracy is clearly inevitable, but the methods that give the best results, at least visually, appear to be the multiple flow direction algorithm of Quinn *et al.* (1995) (Figure 3.4a) and the resultant vector method of Tarboton (1997) (Figure 3.4b). An analysis program based on the multiple direction algorithm is available as freeware (see Appendix A).

With vector data, the problem is how to derive the lines of greatest slope or *streamlines* for flow on the hillslopes. The idea is that water will follow the same direction of flow as a ball running down the same (smoothed!!) surface topography. Since, under this assumption, water should not cross a streamline, it may then be possible to represent the flow between two streamlines, in a *stream tube*, as a one-dimensional flow of varying width in the downslope direction (two dimensions if the vertical is taken into account, as in Figure 3.5). This is the basis for distributed models such as TOPOG (Vertessy *et al.*, 1993) and the Institute of Hydrology Distributed Model (Calver and Wood, 1995). The streamlines should always be at right angles (orthogonal) to the contours. If the contours are available in digital form, calculating the

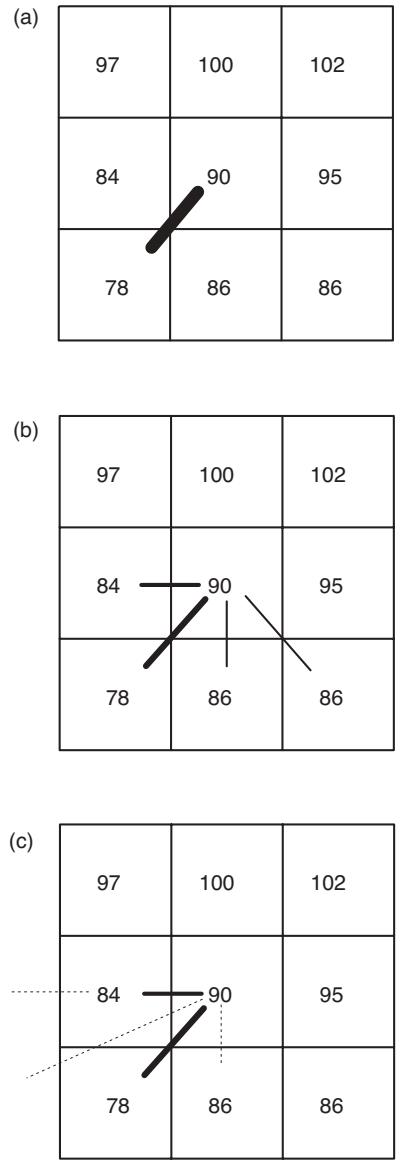


Figure 3.4 *Analysis of flow lines from raster digital elevation data: (a) single steepest descent flow direction; (b) multiple direction algorithm of Quinn et al. (1995); (c) resultant vector method of Tarboton (1997).*

streamlines automatically is a complex problem, but at least one package, TAPES-C, associated with the Australian THALES and TOPOG models, is available (e.g. O'Loughlin, 1986; Grayson *et al.*, 1995).

TIN DEMs are widely used in GIS systems, visualising a three-dimensional topography on screen, and have been used as the basis for a number of distributed hydrological models, since within each slope facet represented within the TIN, the aspect, slope angle and downslope flow direction are easily

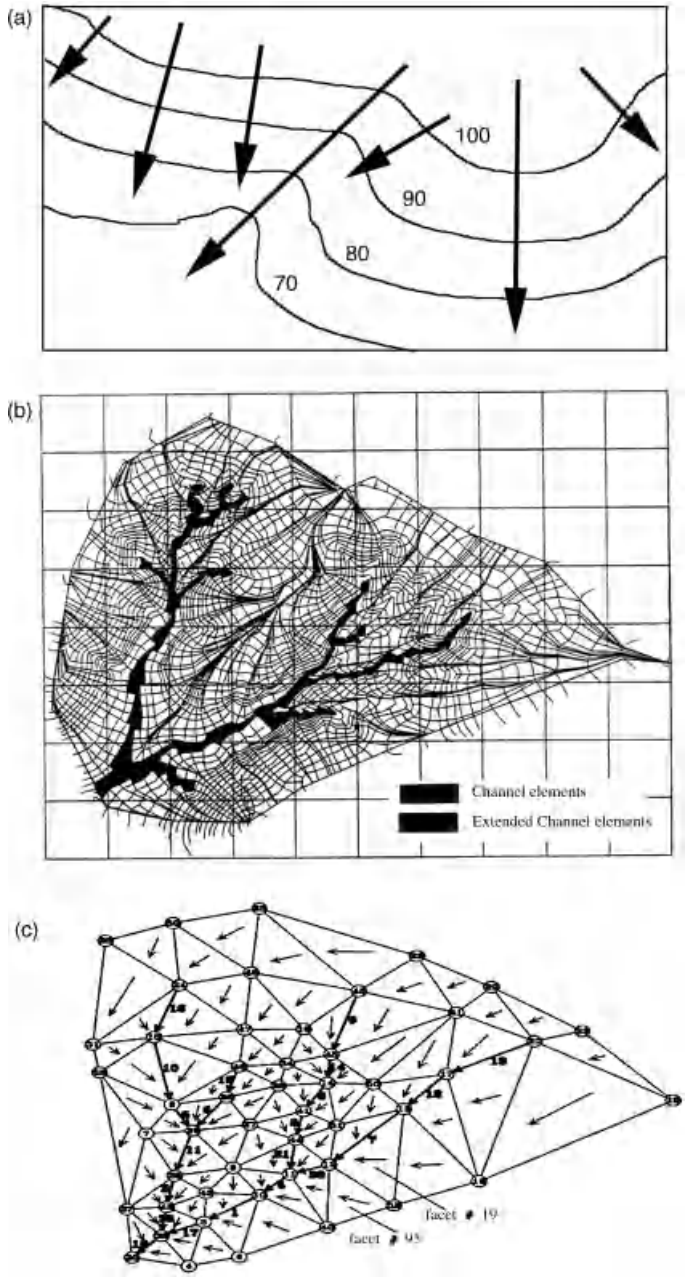


Figure 3.5 *Analysis of flow streamlines from vector digital elevation data: (a) local analysis orthogonal to vector contour lines; (b) TAPES-C subdivision of streamlines in the Lucky Hills LH-104 catchment, Walnut Gulch, Arizona (after Grayson et al. (1992a), with kind permission of the American Geophysical Union); (c) TIN definition of flow lines in the Lucky Hills LH-106 catchment (after Palacios-Velez et al., 1998, with kind permission of Elsevier).*

calculated. The main issue in the construction of a TIN is the best discretisation or *tessellation* of the space to represent the topographic form of a catchment most efficiently. The study by Nelson *et al.* (1999) provides a technique for automatically generating a TIN representation of the topography from elevation points or a vector DEM. Once a TIN is defined, algorithms are also available for automatically delineating the river network and catchment area for any point on the network (Palacios-Velez and Cuevas-Renaud, 1986; Jones *et al.*, 1990). A number of distributed rainfall–runoff models have been based on a TIN representation of topography (such as the tRIBS model of Ivanov *et al.*, 2004, 2008, and the Pennsylvania Integrated Hydrological Model of Qu and Duffy, 2007).

The idea of analysing the topography of the catchment to give an indication of flow pathways is clearly an attractive one and can result in some attractive computer graphics when model predictions are superimposed back onto a three-dimensional picture of the topography. However, there are some limitations to such analyses that the user must be aware of. Regardless of the algorithms or type of DEM used, all DEM analyses depend crucially on the assumption that the flow pathways will be controlled predominantly by the topography of the catchment. This will only be a good assumption for catchments with relatively shallow soils underlain by impermeable or near impermeable bedrock. If there are deeper flow pathways they may deviate significantly from those suggested by an analysis of the surface topography. Recent work has also shown that even in shallow systems, the bedrock topography may have a greater control over downslope saturated flow than the surface topography, at least in some catchments (McDonnell *et al.*, 1996). Finally, it is worth repeating that, even under ideal hydrological circumstances, an adequate representation of the flow pathways will require a DEM of fine enough resolution to define the shapes of the hillslopes. Analyses suggest that, for raster data, resolutions coarser than 100 m will not suffice.

3.8 Geographical Information and Data Management Systems

Topographic data are just one type of distributed data that are becoming more readily available to the hydrological modeller in the form of digital geographical information systems (GIS). These are software packages that allow different types of spatial data to be overlain and manipulated. Most GIS do not easily handle data variables that change over time but some, such as the Institute of Hydrology Water Information System (WIS), have been specifically designed for time variable data. Several have facilities for the analysis of flow directions, although they are generally limited to single flow direction algorithms, such as in the ARC-VIEW or GRASS packages. There are now many hydrological modelling packages with links to GIS packages such as ArcHydro (Maidment, 2002), Green Kenue from Environment Canada and the open source RHydro.

Other variables that can be easily stored and manipulated within a GIS are maps of vegetation type, soil type and geology. The data can then be used in different ways. The characteristics of each element of a raster grid of arbitrary size could be derived, for example. Or, by overlaying different layers of information, irregular *hydrological response units* (HRUs) of different characteristics could be identified (see Section 6.6). A map could be displayed of all grid elements or HRUs having similar characteristics and so on (see, for example, Figure 2.7).

The problem for the hydrological modeller is that the types of information that are generally available in GIS form are only indirectly relevant to the rainfall–runoff processes. Knowing the soil and vegetation type classifications of an HRU is certainly informative, but what parameter values should be used for each classification? In principle, such parameter values could be stored directly within a GIS but only if the values are known. Soil type, as mapped by the soil scientist, for example, may not be the same as that required by the hydrologist. In the UK, a hydrological classification of soil types is now available in GIS form (the HOST classification, Boorman *et al.*, 1995) but this is based on expectations of soil hydrological behaviour and does not directly give parameter values required in a hydrological model. In

the USA, the USDA STATSGO database defines soil characteristics for the whole of the country, but at a scale of 1 km^2 (USDA SCS, 1992). In Europe, the European Soils Database (ESDB) is also available in vector form or as a raster 1 km^2 grid.

This is also true of the other GIS data types. In general, another model is required to interpret the GIS data in a form that can be used in a hydrological model. An example is the type of *pedotransfer function* models suggested by Rawls and Brakensiek (1989). Regression analysis is used to provide relationships between soil texture and soil hydraulic parameters. Soil texture is a common characteristic reported in soil surveys and a soil classification can normally be associated with a texture. A pedotransfer function can then be used to derive values for parameters such as porosity and hydraulic conductivity (see Box 5.5).

However, values derived in this way should be interpreted with care. We have already noted that measured soil hydraulic parameters may be highly variable in space, even within a single soil unit, and that the effective parameter values required for different models may be model structure and scale dependent (Section 1.8). Thus, GIS-derived values of parameters may be associated with a considerable degree of uncertainty which is often ignored. Having said that, a number of studies have reported success in rainfall–runoff modelling based on GIS data (see Section 6.6).

Within this GIS framework, rainfall–runoff modelling may be just one component of a larger catchment management or decision support system (DSS). Examples are the WATERSHEDSS package developed by the USDA (Chaubey *et al.*, 1999) and the UK NERC-ESRC Land-use Programme (NELUP) DSS (Dunn *et al.*, 1996). Few GIS packages, however, allow flexible modelling structures to be built directly onto the spatial framework of the GIS. One exception is PC-RASTER which has a built-in programming language that allows model structures, together with input and output data, to be readily incorporated into the GIS. Other general programming languages, such as MATLAB and PV-WAVE, can also be used with large raster spatial databases for modelling (see, for example, Clapp *et al.*, 1992; Romanowicz, 1997).

3.9 Remote-sensing Data

Another source of distributed data for hydrological modelling comes from remote sensing. A full review of the types of data that can be provided by remote sensing is beyond the scope of this book but excellent coverage is given in Section 5 of the *Encyclopaedia of Hydrological Sciences* (Anderson, 2005). Remote sensing may be used in the estimation of input data (including topography, rainfall and evapotranspiration rates), state variables (including soil moisture, snow cover, snow water equivalent and areas of flood inundation) and model parameter values (mostly derived through the classification of soil and vegetation types from remote sensing). In fact, topography or land cover distribution derived from remote sensing may contribute to the database stored in a GIS. Many of the same problems apply as for GIS data: remote sensing does not generally give information that is directly hydrologically relevant; a model is required to interpret the remote-sensing signal into a form that is hydrologically useful (as discussed earlier, for example, for the case of radar-derived estimates of rainfall). It is not often recognised that the interpretation model can be a significant source of uncertainty in the resulting images supplied to the user. Corrections for atmospheric effects for satellite sensors, for example, involve empirical coefficients or parameters that are not known precisely and that may be time variable. These uncertainties are not often quantified and the user has little option but to assume that they are small. However, remote sensing will increasingly be an important source of spatial information and future hydrological models will make increasing use of different types of imaging in model calibration, evaluation of predictions and data assimilation to improve short-term forecasting.

The main uses of remote sensing in hydrological modelling to date have been in the estimation of precipitation (primarily using ground-based radar), land cover types and vegetation parameters, soil moisture and snow cover. Snow cover mapping from satellite images is now used operationally in the

USA, in conjunction with hydrological modelling and ground surveys, for water resources assessment, flood forecasting and dam regulation. The utility of the technique remains subject to the limitations of the spatial and temporal resolution of the satellites used, the requirement for generally cloud-free conditions in the image, and the fact that only cover can be easily assessed, rather than snow water equivalent, although attempts have been made to use passive microwave remote sensing to estimate the spatial patterns of water equivalent of snow packs. It is, however, the only method for getting the widespread large-scale coverage that is required for major river basins. Ground-based photographic measurements have also been used in snowmelt modelling (e.g. Blöschl *et al.*, 1991).

There are some other remote-sensing techniques and interpretation models that may become more useful in the future. Work in this area has become increasingly important with the development of work in "macroscale" hydrology, which has been encouraged by the needs of global atmospheric circulation modelling for hydrological predictions at large scale. Active and passive microwave techniques for the measurement of soil moisture have been studied for some time. Active sensors transmit a signal to the ground and measure the return signal; passive systems measure only the natural microwave transmission from the surface. The first satellites with active microwave sensors that can be used for soil moisture estimation (ERS1, ERS2, JERS1 and Radarsat) have been orbiting since the early 1990s. They provide good spatial resolution of 30–100 m but with repeat coverage periods of the order of 30 days. Active microwave systems have also been used from ground- and aircraft-based platforms. The wavelengths normally used penetrate, at most, only the first few centimetres of the soil surface. The return signal then depends on the dialectric constant of the surface soil layer. The dialectric constant varies with moisture content, most strongly for soils that are neither too wet or too dry. Sensitivity of the measurement is much reduced at both high and low moisture contents and it is difficult to differentiate, for example, between a wet and saturated soil surface, particularly when the vegetation cover may also be wet. Passive microwave satellite sensors (SSM/I, AMSR and SMOS) provide much higher frequency measurements (1–3) days but much lower spatial resolution (5–50 km) than would be considered useful in rainfall–runoff modelling.

The problem is that, for both active and passive microwave sensors, the signal also depends on the water content of the vegetation, the roughness of the surface and the state of the atmosphere. In most images from active radar systems, for example, the most obvious features are those associated with the topography and roughness of the surface, such as different vegetation covers. Extracting the soil moisture content signal means extracting a second-order effect and often relies on the availability of ground measurements that can be used to calibrate the interpretation of the radar image. The approach will work best, therefore, where there is a uniform surface, particularly if there is little vegetation cover. There have, however, been some very interesting results presented from airborne and satellite microwave sensors, using both active and passive microwave systems. The techniques may well be improved in the future. The wavelength of the signal remains a limitation, in that only surface soil moisture can currently be detected in this way. There is, then, a problem in relating that surface soil moisture content to the moisture profile and flow processes or to the results of a hydrological model. One of the few studies that has attempted such a comparison at a large scale is that of Wood *et al.* (1993).

There is a remote-sensing technique for estimating the changes of water storage in the landscape over time. Changes in storage will change the gravity anomaly and data from the GRACE satellite, launched in 2002, are processed to produce monthly estimates of patterns of water storage. The spatial scale of the estimates is, however, 60 km so that the information will be most useful in constraining macroscale models of large catchments or global hydrology (e.g. Güntner, 2008) or changes in large scale groundwater bodies (e.g. Rodell *et al.*, 2007).

Satellite information might also be useful in the estimation of river stage and discharges. At present, the various satellite altimetry methods do not have a sufficiently high spatial or time resolution to be useful except on the very largest rivers (Birkett *et al.* 2002; Trigg *et al.*, 2009). This might change in the future, however, when the planned SWOT satellite is launched some time in the next decade. This will

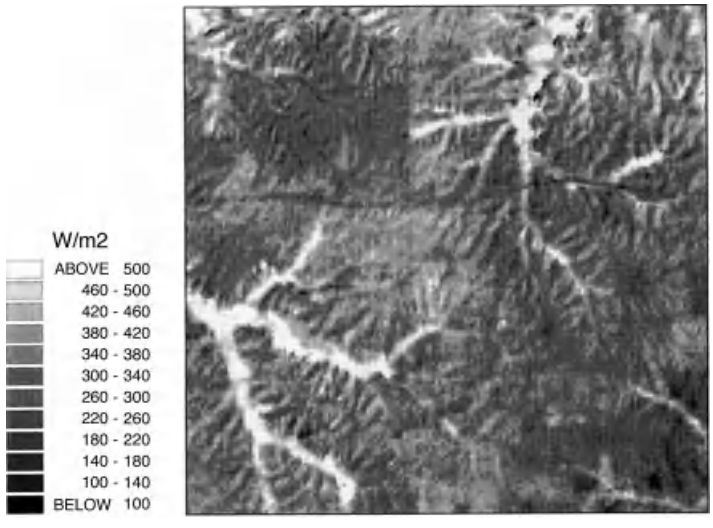


Figure 3.6 *Predicted spatial pattern of actual evapotranspiration based on remote sensing of surface temperatures; note that these are best estimates of the evapotranspiration rate at the time of the image; the estimates are associated with significant uncertainty (after Franks and Beven, 1997b, with kind permission of the American Geophysical Union).*

have the capability of measuring water levels with a resolution of 100 m to an accuracy of the order of 0.1 m and a slope accuracy of 1 cm/km over 90% of the globe and a maximum repeat time of 22 days. This will still be most useful in larger rivers with seasonal changes in discharges but has the potential to improve modelling of flood waves in those rivers.

There have also been some advances in the use of remote sensing for the estimation of spatial patterns of evapotranspiration. Studies such as those of Holwill and Stewart (1992), Bastiaanssen *et al.* (1994, 1998), Xinmei *et al.* (1995), Franks and Beven (1997b), Su (2002), Calcagno *et al.* (2007), Schüttemeyer *et al.* (2007) and Kalma *et al.* (2008) have used single and multiple images of remotely sensed surface temperatures with simple energy balance modelling to separate sensible and latent heat fluxes to calculate patterns of actual evapotranspiration flux (Figure 3.6). The patterns have, in every case, revealed significant and interesting heterogeneity in fluxes, but error calculations suggest that the absolute values of the calculations may be subject to large uncertainty (Franks and Beven, 1997b; see also, Kalma *et al.*, 2008).

3.10 Tracer Data for Understanding Catchment Responses

The use of tracer information to improve understanding of catchment responses goes back to the late 1960s and 1970s when the pioneering studies of Dincer *et al.* (1970) on snowmelt and Pinder and Jones (1969), Crouzet *et al.* (1970) and Sklash and Farvolden (1979) on rainfall–runoff processes first started to reveal the extent to which the hydrograph is made up of pre-event or "old" water stored in a catchment prior to a hydrograph-producing event (see Section 1.4). Since then, a variety of environmental and artificial tracer techniques have been used in such studies. Artificial tracers are generally limited to profile and plot scale experiments (for example, the multiple tracer experiments of Abdulkabir *et al.*, 1996) since it

is difficult to manipulate the inputs for a whole catchment (but see the Gårdsjön catchment experiment Rodhe *et al.*, 1996). Thus catchment scale analyses are generally limited to environmental tracers and particularly the isotopes of oxygen and hydrogen/deuterium/tritium ($^{18}O/^{16}O$, $^{2}H/^{1}H$ and ^{3}H) that form part of the water molecule and should consequently be good tracers of water flow pathways. Tritium is radioactive and mostly derives from atmospheric nuclear bomb tests in the 1950s and 1960s. Such tests were banned in 1963 (although France continued testing until 1974 and China until 1980), so that the levels of tritium in the atmosphere have since declined, accelerated by radioactive decay of the remaining tritium (which has a half life of 12.32 years). Until relatively recently, isotope analyses were expensive and time-consuming so that many other chemical characteristics have been used to infer water sources and pathways from the early study of Pinder and Jones (1969) onwards. A new generation of laser spectrometers is now bringing down the cost of isotope analyes (e.g. Berman *et al.*, 2009) and we can expect to see many more intensive studies of isotope behaviours in hydrological systems in future. It has been argued in a number of papers that tracer information, where it is available, can provide useful additional constraints for use in model calibration to test whether rainfall–runoff models are getting the right results for the right reasons (e.g. Seibert and McDonnell, 2002; Kirchner, 2006; McGuire *et al.*, 2007; Vaché and McDonnell, 2006), although the need to reproduce tracer concentrations as well as discharge hydrographs generally means that more model parameters need to be calibrated. This is discussed further in Chapter 11.

3.11 Linking Model Components and Data Series

There are increasing demands to link rainfall–runoff models to other model components that might be concerned with sediment transport, water quality, hydroecology, urban drainage control systems, etc. In the past, this has tended to be done within specific modelling packages, though there were some early attempts at a more general solution, such as the ANNIE database system used by the USGS (Leavesley and Stannard (1995) include a brief description in their work) to link hydrological models to spatial and temporal data files, and the UK Institute of Hydrology Water Information System (WIS). In the last decade, there have been significant developments in working towards standards for such exchanges (and not only in hydrology – this is a problem common to many environmental modelling domains).

One such major initiative is the Open Geospatial Consortium (OGC) which has worked since 1994 to define standards for geospatial data and services. The US-based Consortium of Universities for the Advancement of Hydrologic Science Inc. (CUAHSI, www.cuahsi.org) has been developing open source standards for exchange of data in its WaterML and CUAHSI Hydrologic Information System (CUAHSI-HIS) that are intended to be consistent with the OGC standards. The European Community has also funded a number of projects of this type. The OpenMI Association (see www.openmi.org/reloaded) is developing open source software designed to provide common interfaces between models and data on a time step by time step basis during simulations (see Gregerson *et al.*, 2007). If model components are programmed to be compatible with the interface definitions then different components should be able to automatically exchange information, regardless of the spatial and temporal time steps. The standards support distributed model components with different grid sizes. The Interoperability and Mapping project (INTAMAP, www.intamap.org) is developing a web-based framework for the real-time mapping of critical environmental variables using geostatistical methods. It can be used to exchange data, carry out statistical analysis and visualise results. OGC-compatible standards for communicating uncertainty in variables and model results are also being developed using UncertML (www.uncertml.org). This is a rapidly developing area that is also taking advantage of developments in *cloud computing*, such as GoogleMaps and the Google Earth Engine, that allow the presentation of model results in a highly visual form (Figure 3.7).

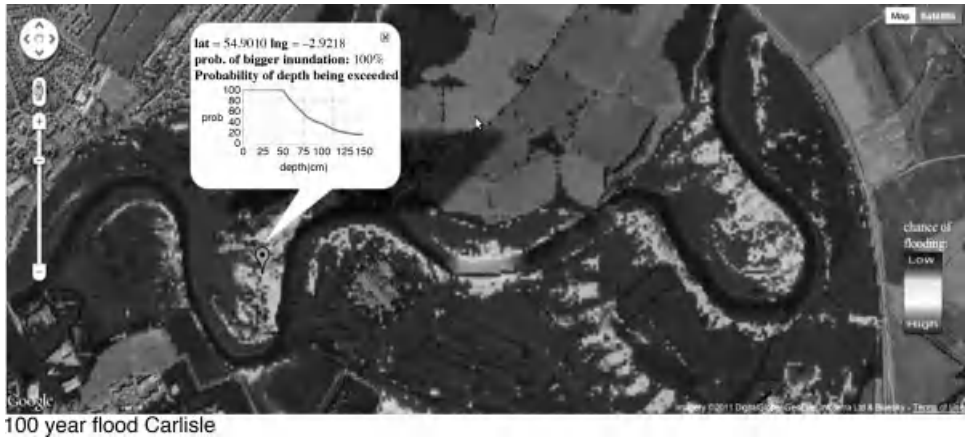


Figure 3.7 *Estimates of uncertainty in the extent of inundation of the 100-year return period flood for the town of Carlisle, Cumbria, UK, superimposed on a satellite image of the area using GoogleMaps facilities; the inset shows the exceedance probabilities for depth of inundation at the marked point (after Leedal et al., 2010).*

3.12 Key Points from Chapter 3

- The data available for rainfall–runoff modelling are generally point data and may not be error free, even though they must often be treated as error free in applications.
- Data should be checked for consistency before being used in a rainfall–runoff modelling study. Some simple checks can be used to identify periods of unusual behaviour that can be checked more carefully or eliminated from the analysis.
- Methods for directly measuring evapotranspiration rates are not yet used routinely. A number of different methods for estimating potential and actual evapotranspiration demand different levels of data availability.
- Spatial data, such as radar rainfalls and satellite images at different wavelengths, are increasingly becoming available through remote sensing, including active and passive microwave sensors used in the estimation of surface soil moistures. In general, such data require some interpretative model to provide hydrologically useful information. This interpretative model may be a source of error in this information.
- Geographical information systems are increasingly being used to store catchment data and interact with distributed hydrological models in setting up model runs and displaying the results. The information stored in a GIS (e.g. soil type and vegetation type) may also require an interpretative model before being useful in hydrological modelling.
- Digital elevation data, in either raster or vector form, can be the basis of distributed modelling of both model inputs and rainfall–runoff processes. The latter may require the derivation of hillslope and channel flow pathways from the digital elevation data. Different methods of analysis and different resolutions of data will give different apparent flow pathways.
- Environmental and artificial tracer data might provide valuable information to help in evaluating whether rainfall–runoff models are getting the right results for the right reasons (see also Chapter 11).
- There are developing software standards for interfacing different model components and data sets, such as those provided by OpenMI, INTAMAP and the CUAHSI Hydrologic Information System. In the future, there will be increasing use of cloud computing facilities in accessing data, running models and visualising results.

Box 3.1 The Penman–Monteith Combination Equation for Estimating Evapotranspiration Rates

The Penman–Monteith equation is based on a combination of a simplified energy balance equation for the surface and equations for the transport of sensible heat and latent heat away from the surface. It is what has been called a *big leaf* model, in that there is an assumption that a complex vegetation canopy can be represented as if it were acting as a single transpiring surface at some effective height above the ground. The energy balance equation, illustrated in Figure B3.1.1, can be written as

$$H = R_n - A - G - S \tag{B3.1.1}$$

where H is the total energy available for evapotranspiration, R_n is net radiation (ranging from $-50\,\mathrm{Wm^{-2}}$ on a clear night to more than $500\,\mathrm{Wm^{-2}}$ at midday in summer), A is heat loss due to *advection* ($\approx 1\,\mathrm{Wm^{-2}}$ for a downwind temperature gradient of $1^{\circ}\mathrm{C\ km^{-1}}$), G is heat loss into the ground (usually positive during the day and negative at night) and S is the energy flux into physical and biochemical storage in the vegetation (up to $15\,\mathrm{Wm^{-2}}$ during the day, and perhaps $3\,\mathrm{Wm^{-2}}$ at night).

It is assumed that total available energy can be partitioned into two components: the transport of sensible heat to or from the surface (that is energy directly involved in heat or cooling of the air above the surface by conduction and convection) and the transport of latent heat (that is energy used in vaporising water lost from the surface by evaporation or transpiration). Thus:

$$H = C + \lambda E \tag{B3.1.2}$$

where C is the sensible heat flux and λE is the latent heat flux as a product of the latent heat of vaporisation, $\lambda (= 2.4710^6, \mathrm{Jkg^{-1}})$ and the evapotranspiration rate ($\mathrm{kgm^{-2}s^{-1}} \approx \mathrm{mms^{-1}}$)

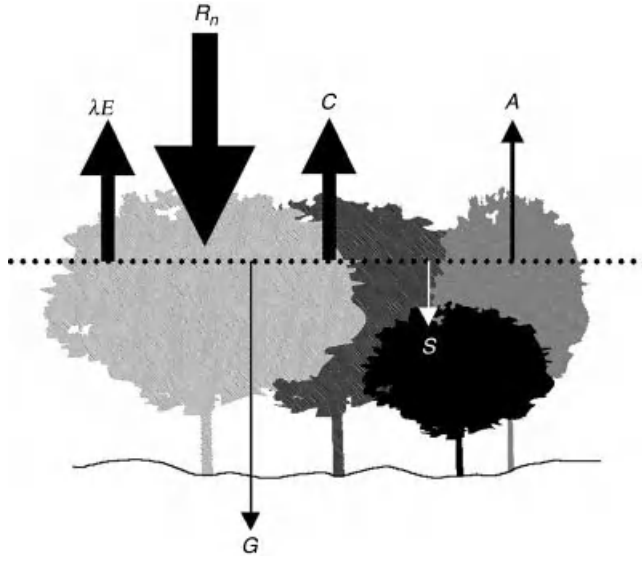


Figure B3.1.1 *Schematic diagram of the components of the surface energy balance. R_n is net radiation, λE is latent heat flux, C is sensible heat flux, A is heat flux due to advection, G is heat flux to ground storage, S is heat flux to storage in the vegetation canopy. The dotted line indicates the effective height of a "big leaf" representation of the surface.*

Then

$$\lambda E = H - C = \frac{H}{1+\beta} \tag{B3.1.3}$$

where $\beta = C/\lambda E$ is known as the Bowen ratio. Experimental evidence suggests that the Bowen ratio will often have a fairly constant value for a surface, at least for clear sky conditions without soil water limitations on evapotranspiration (Brutsaert and Sugito, 1992; Nichols and Cuenca, 1993; Crago and Brutsaert, 1996).

The sensible heat flux is a function of the temperature gradient in the air above the vegetation canopy, whereas the latent heat flux is a function of the humidity or vapour pressure gradient above the canopy. Both are also dependent on factors such as the roughness of the canopy and wind speed (expressed as an aerodynamic resistance to transport). Rough canopies and higher wind speeds (low values of aerodynamic resistance) result in much more efficient mixing of the air and faster rates of transport. The transport equations are generally assumed to be of the form:

$$C = \frac{1}{r_{a,H}} \rho_a c_p \left(T_o - T_z\right) \tag{B3.1.4}$$

where $r_{a,H}$ is the aerodynamic resistance to transport of heat, ρ_a is the density of the air, c_p is the specific heat capacity of the air, T_o is the temperature of the surface and T_z is the temperature of the air at some reference height z. The *big leaf* assumption here becomes apparent in the use of the surface temperature, T_o, which must represent some effective value for all the different surfaces of the canopy as a whole.

For the latent heat flux, the equivalent transport equation is

$$\lambda E = \frac{1}{r_{a,V}} \frac{\rho_a c_p}{\gamma} \left(e_o - e_z\right) \tag{B3.1.5}$$

where $r_{a,V}$ is the aerodynamic resistance to the transport of vapour, e_o is the vapour pressure at the effective canopy surface, e_z is the vapour pressure at the reference height z and γ is called the psychrometric constant ($= 66$ PaK^{-1}). The problem with these equations so far is that the temperature and vapour pressure at the surface are not easily measured. To make the system of equations solvable, John Monteith came up with the idea of using an additional conceptual expression for the transport of vapour (Monteith, 1965) from the interior of the stomata of the leaf surfaces to the free air, as

$$\lambda E = \frac{1}{r_c} \frac{\rho_a c_p}{\gamma} \left(e_s\left(T_o\right) - e_o\right) \tag{B3.1.6}$$

where r_c is an effective stomatal resistance for the canopy as a whole, generally known as the canopy resistance, and $e_s\left(T_o\right)$ is the saturated vapour pressure at the surface temperature T_o. Combining these expressions allows the unknown vapour pressure at the conceptual *big leaf* surface to be eliminated such that

$$\lambda E = \frac{1}{r_{a,V} + r_c} \frac{\rho_a c_p}{\gamma} \left(e_s\left(T_o\right) - e_z\right) \tag{B3.1.7}$$

There is still the problem of estimating $e_s\left(T_o\right)$. This is done by assuming that $e_s\left(T_o\right)$ can be approximated by the expression $e_s\left(T_z\right) + \Delta_e \left\{T_o - T_z\right\}$ where Δ_e represents the slope of the saturation vapour pressure versus temperature curve. The original form of the Penman–Monteith equation uses this linear interpolation of the saturation vapour pressure curve. Milly (1991) has suggested that a higher order approximation will produce more accurate predictions.

In most applications; a further approximation is made: that the aerodynamic resistances $r_{a,V}$ and $r_{a,H}$ can both be assumed equal to the equivalent resistance for momentum transport in a well mixed neutral boundary layer r_a for which a value can be derived from assumptions about the wind speed profile. In particular, for a logarithmic wind speed profile, turbulence theory suggests that

$$r_a = \frac{\ln\left\{\left(z-d\right)/z_o\right\}^2}{\kappa^2 u_z} \tag{B3.1.8}$$

where d is called the zero plane displacement, z_o is the roughness height, u_z is the wind speed at the measurement height, z, and κ is the von Karman constant. This expression assumes a well mixed boundary layer above the vegetation canopy. Corrections may be required for stable atmospheric conditions.

After these approximations

$$\lambda E = \frac{1}{r_a + r_c}\frac{\rho_a c_p}{\gamma}\left(e_s\left(T_z\right) - e_z + \Delta_e\left\{T_o - T_z\right\}\right) \tag{B3.1.9}$$

But from the expression for sensible heat flux

$$\left\{T_o - T_z\right\} = Cr_a/\rho_a c_p = r_a\left[H - \lambda E\right]/\rho_a c_p \tag{B3.1.10}$$

so that

$$\lambda E = \frac{1}{r_a + r_c}\left\{\frac{\rho_a c_p}{\gamma}\left(e_s\left(T_z\right) - e_z\right) + \frac{\Delta_e r_a}{\gamma}\left[H - \lambda E\right]\right\} \tag{B3.1.11}$$

Rearranging this equation gives

$$\lambda E\left(1 + \frac{\Delta_e r_a}{\gamma\left(r_a + r_c\right)}\right) = \frac{1}{r_a + r_c}\left\{\frac{\Delta_e r_a}{\gamma}R_n + \frac{\rho_a c_p}{\gamma}\left(e_s\left(T_z\right) - e_z\right)\right\} \tag{B3.1.12}$$

or

$$\lambda E = \frac{\Delta_e H + \rho_a c_p\left(e_s\left(T_z\right) - e_z\right)/r_a}{\Delta_e + \gamma\left(1 + r_c/r_a\right)} \tag{B3.1.13}$$

This is the Penman–Monteith equation. Use of the equation requires measurements of temperature, humidity and wind speed at the reference height z, available energy H, and estimates of the two resistance coefficients r_a and r_c. The approximation $H \approx R_n$ is often made. The equation can be applied with hourly data to provide estimates of the diurnal pattern of evapotranspiration rates.

The resistance coefficients have an important control on predicted evapotranspiration rates, particularly when the resistances are low (Beven, 1979a). The variation of predicted evapotranspiration with r_a and r_c for a particular set of meteorological conditions is shown in Figure B3.1.2. Typical values for a dry grass canopy would be $r_a = 50sm^{-1}$ and $r_c = 50sm^{-1}$, while for a dry tree canopy $r_a = 10sm^{-1}$ and $r_c = 50sm^{-1}$. The highest actual evapotranspiration rates will be predicted for a rough canopy (low r_a) with intercepted water on the leaf surfaces ($r_c = 0$).

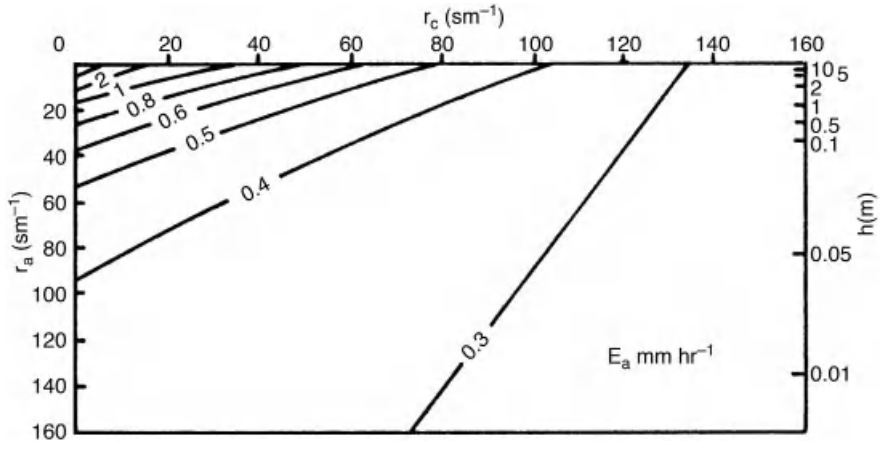


Figure B3.1.2 *Sensitivity of actual evapotranspiration rates estimated using the Penman–Monteith equation for different values of aerodynamic and canopy resistance coefficients (after Beven, 1979a, with kind permission of Elsevier).*

The effects of drying of the soil on evapotranspiration rates can be reflected in an increase in r_c with decreasing soil moisture, although it is known that other factors, such as leaf temperature, carbon dioxide concentration, insolation and even chemical signalling in the plant can play a role in determining the effective canopy resistance. Calder (1977) suggested an empirical relationship for the change in canopy resistance for transpiration that was a product of a seasonal sinusoid and a function of vapour pressure deficit. Other, more complex, relationships have been proposed by Jarvis (1976), Sellers (1985) and Tardieu and Davies (1993). These types of relationship, and their links to carbon dioxide exchanges, are included in many SVAT or LSP models linked to atmospheric circulation models. Such models, however, require many more parameter values and there has also been a move, led by John Monteith, to investigate the possiblity of using simple models with fewer parameters (Monteith, 1995a, 1995b).

More recently, it has been suggested that it might be possible to introduce additional constraints on rates of evapotranspiration by coupling carbon to the water budget of the plant canopy. Evapotranspiration losses from plants are, after all, really only a biproduct of the need for plants to exchange carbon dioxide with the atmosphere through their stomata for the purposes of photosynthesis. Thus, if the net productivity of carbon and its use could be predicted, then this would provide an important constraint on actual evapotranspiration. There are a number of vegetation growth models available in the literature, with many parameters to be specified. These do not always produce good predictions of productivity (e.g. Mitchell *et al.*, 2009). It has been suggested, however, that it might be possible to simplify this approach by invoking an optimality principle whereby a plant optimises its carbon production given the prevailing boundary conditions of energy and water availability (Schymanski *et al.*, 2007). Schymanski *et al.* (2008) have implemented a model that optimises the carbon cost of maintaining its root system in the face of changes in the profile soil water content and have shown that the resulting time evolution of the rooting profile provides a better prediction of evapotranspiration than a model using a fixed rooting profile. They have also shown that invoking the optimality principle in constraining actual evapotranspiration can help guide the calibration of catchment scale models (Schymanski *et al.*, 2009; see also Section 9.7).

The assumptions of the Penman–Monteith model for evapotranspiration can be summarised as follows:

A1 The energy available for evapotranspiration, H, can be estimated from knowledge of net radiation, ground heat flux, advective heat flux and storage of heat in the vegetation canopy. In applications, this is often reduced to an estimate based on the net radiation alone, with other terms neglected as small.
A2 The latent heat and sensible heat fluxes can be estimated from first-order gradient flux relationships with the gradients of vapour pressure and temperature, respectively, being determined between some measurement height and some conceptual height within the canopy (the "big leaf" assumption).
A3 The aerodynamic resistance for the movement of vapour and heat away from the canopy are similar and can be estimated from the aerodynamic resistance for the downwards flux of momentum as determined under the assumptions of a logarithmic velocity profile.
A4 The controls of stomata on evapotranspiration rates can be represented by some effective canopy resistance for the movement of vapour from stomata to the air within the canopy.
A5 The saturated vapour pressure of the air within the stomata is assumed to be consistent with the effective canopy temperature and may be predicted by extrapolating from the measurement height using the slope of the known theoretical saturated vapour pressure–temperature relationship.

Box 3.2 Estimating Interception Losses

In Section 3.3, we have made the point that, in many environments, evapotranspiration is a greater proportion of the catchment water balance than stream discharge. In turn, an important component of the total actual evapotranspiration from a catchment can be due to evaporation from water intercepted on the vegetation canopy, particularly from rough canopies that are frequently wetted. In forest canopies subject to frequent wetting in the windy environment of upland UK, *interception* may amount to more than 20% of the rainfall inputs (Calder, 1990). In such circumstances, the vegetation canopy may have an important effect on the amount and pattern of intensity of rainfall reaching the ground.

B3.2.1 Regression Models of Throughfall and Stemflow

There have been many experimental studies of interception, many of which have reported results in the form of regression equations for storm throughfall and stemflow as

$$V_{TF} = B_{TF}P_{storm} - C_{TF}$$
$$V_{SF} = B_{SF}P_{storm} - C_{SF}$$

where P_{storm} is the storm rainfall total, the subscripts TF and SF refer to throughfall and stemflow respectively, V_{TF} and V_{SF} are the volumes of throughfall and stemflow in the storm, C_{TF} and C_{SF} are minimum storage capacities for throughfall and stemflow, and B_{TF} and B_{SF} are coefficients. This type of regression relationship can obscure considerable scatter in the measurements for individual storms so that, in making predictions, the uncertainty in estimated throughfall and stemflow predictions should be assessed. The storage capacities and coefficients vary with vegetation type and seasonal vegetation growth. Not all such studies have differentiated between throughfall and stemflow.

Most of these experiments were based on volumetric collection of throughfall and stemflow with storm by storm measurements. Increasingly, however, hydrologists have made measurements on a more continuous basis and this has allowed the development of more dynamic interception models.

B3.2.2 The Rutter Model

Perhaps the most widely used model of interception is that proposed by Jack Rutter (Rutter *et al.*, 1971, 1975; Calder, 1977; Gash and Morton, 1978). A schematic diagram of the model is shown in Figure B3.2.1.

The model is based on two storage components, one for canopy interception and one for stemflow. Incoming precipitation is partitioned into direct throughfall that reaches the ground without interacting with the canopy, input to the canopy interception and stemflow stores. The partition coefficients, p and p_t, shown in Figure B3.2.1 will be vegetation type and seasonally dependent, but are often assumed constant in applications. Drainage from the stemflow store begins when the minimum storage capacity is exceeded. For the canopy store, drainage takes place at the rate:

$$D_t = D_s \exp(b\{S_{TF} - C_{TF}\}) \tag{B3.2.1}$$

where D_s is the drainage rate when the storage depth S_{TF} just equals the capacity C_{TF} and b is a coefficient. Note that, if applied in this form, the results are dependent on the time steps used and it is better to integrate the equation over the required time step to calculate a storage at the end of that time step, from which the integrated drainage may be obtained from the change in storage. For the stemflow store, all storage in excess of the capacity C_{SF} is assumed to drain rapidly to the ground. The values of the storage minimum capacities C_{TF} and C_{SF} can be determined by the type of regression analysis on storm volumes described above.

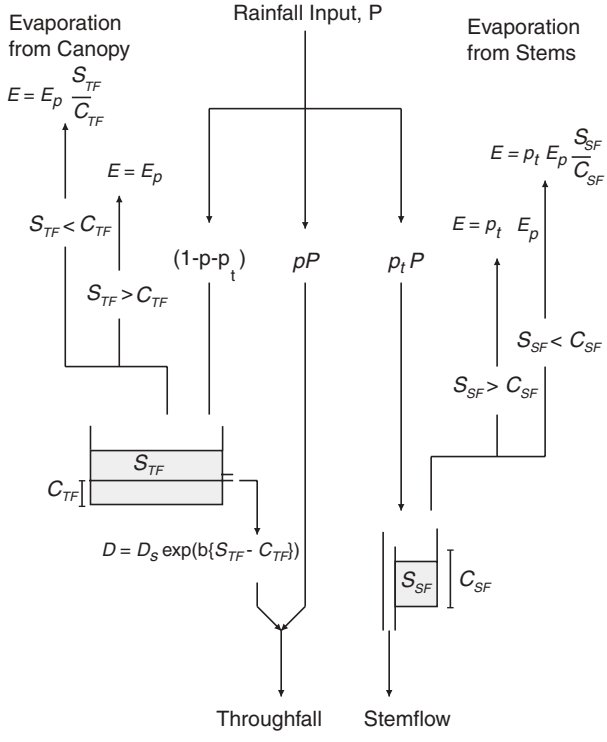


Figure B3.2.1 *Schematic diagram of the Rutter interception model (after Rutter et al., 1971, with kind permission of Elsevier).*

The Rutter model also takes account of evaporation from the two stores based on a potential evapotranspiration rate for a wet canopy calculated using the Penman–Monteith equation (see Box 3.1) with a canopy resistance of zero. If either store is greater than its storage capacity, then evaporation takes place at the potential rate. If either storage is below its minimum capacity then it is assumed that part of the canopy is dry and evaporation is reduced proportionally as

$$E_a = E_p \frac{S_{TF}}{C_{TF}} \tag{B3.2.2}$$

where E_a is the actual evaporation rate and E_p is the estimated potential rate from the interception store. A similar form is used for the stemflow store, allowing that this may account for a proportion of the total potential evaporation (usually taken as equal to p_t) until that store is dry.

The Rutter model therefore requires the specification of six parameters, p, p_t, D_s, b, C_{TF} and C_{SF} together with estimates of rainfalls and potential evaporation as inputs. In many vegetation types, notably crops and deciduous trees, values of the parameters are not constant throughout the year but change with the pattern of leaf development. Care should also be taken with adopting values found in the literature, since the values quoted may depend on the particular form of the model used. Values for different types of tree canopy are given in Calder (1977); Gash and Morton (1978); Dolman (1987); Gash *et al.* (1980); and Lloyd *et al.* (1988). The most important parameter is generally the total storage capacity, $C_{TF} + C_{SF}$, as this has dominant control of the amount of evaporation that takes place at the higher wet canopy rates.

The assumptions of the Rutter model may be summarised as follows:

A1 Rainfall inputs can be apportioned between direct throughfall, canopy interception and stemflow storage.
A2 Drainage from the canopy interception store is an exponential function of storage in excess of a minimum storage capacity, once that capacity is exceeded.
A3 Drainage from the stemflow store takes place as soon as the storage capacity is filled.
A4 Evaporation from the canopy and stemflow stores is at the wet canopy potential rate if storage is greater than the minimum capacity, but is linearly reduced below the potential rate if the canopy or stems are partially wet. Note that, as the canopy dries, if the evaporation from the canopy is not sufficient to satisfy the potential rate, then it is possible that there will be some additional transport of water to the atmosphere due to transpiration or soil evaporation).

A simplified analytical variant of the Rutter model was proposed by John Gash (1979) and has met with reasonable success in a number of different environments (Gash *et al.*, 1980; Lloyd *et al.*, 1988; Navar and Bryan, 1994).

B3.2.3 The Calder Stochastic Model of Interception

The way in which net throughfall is related to canopy storage in the original Rutter model (which allowed drainage even for a canopy storage less than the capacity value) means that the canopy reaches maximum storage only after much more rainfall has fallen than the capacity volume, while the exponential drainage function predicts a small amount of throughfall even when the storage is zero. To avoid this problem, Ian Calder (1986) proposed an alternative stochastic model of interception based on the probabilities of raindrops striking elemental areas making up canopy surfaces. In the original formulation, a Poisson distribution was assumed to relate the mean number of drops retained to the mean number of drops striking the element. The result was a model in which interception was dependent on both storm volume and the mean drop size of the rainfall with the effective canopy storage also being a function of drop size.

In a later development, Calder (1996) extended the stochastic model to account for secondary drops, that is drops falling from upper parts of the canopy that strike lower elemental areas, and for a kinetic energy dependence effect that is a function of drop size and rainfall intensity. This version of the model has been tested by Calder *et al.* (1996) and Hall *et al.* (1996). This model requires seven parameters to be specified. This is too many to be determined from measurements of throughfall rates alone and calibration requires some detailed measurements during wetting experiments, including drop sizes (Calder *et al.*, 1996).

Box 3.3 Estimating Snowmelt by the Degree-Day Method

As noted in the main text, there are many problems and complexities in modelling the accumulation and melting of snowpacks on a complex topography. While process-based energy budget and pack evolution models are available (e.g. Anderson, 1968; Blöschl *et al.*, 1991; Morris, 1991; Marks and Dozier, 1992), it has proven difficult to demonstrate that they can produce generally more accurate operational predictions than simpler empirical models, at least without assuming that parameter values, such as albedo of the pack, are time variable (e.g. Braun and Lang, 1986). Of the empirical models, the most widely used is the degree-day method. An early version of the degree-day method was presented by Imbeaux (1892), including the use of different air temperatures in different elevation bands in predicting snowmelt floods in the Durance catchment in France. There are now many modern variants of the approach (e.g. Bergstrom, 1975; Martinec and Rango, 1981; Hottelet *et al.*, 1993; Moore *et al.*, 1999; Hock, 2003; Zhang *et al.*, 2006). An example of the accuracy of snowmelt discharge predictions in a large catchment for two years using a degree-day snowmelt model with a daily time step is shown in Figure B3.3.1.

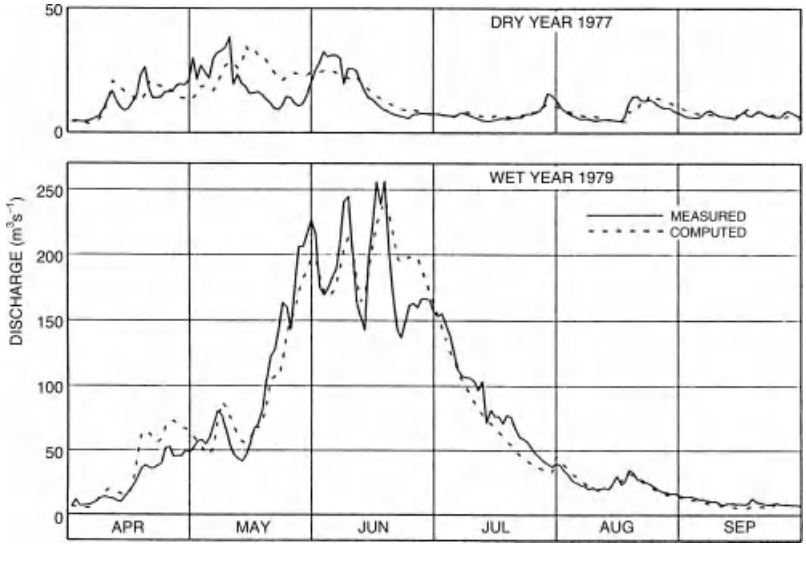


Figure B3.3.1 *Discharge predictions for the Rio Grande basin at Del Norte, Colorado (3419 km^2) using the Snowmelt Runoff model (SRM) based on the degree-day method (after Rango, 1995, with kind permission of Water Resource Publications).*

In its simplest form, the degree-day method predicts a daily melt rate from

$$M = F \max(0, \bar{T} - T_F) \tag{B3.3.1}$$

where M is melt rate as a water equivalent per unit area [LT^{-1}], F is the degree-day factor [$LT^{-1}K^{-1}$], $\overline{T}$ is mean daily air temperature [K], and T_F is a threshold temperature [K] close to the freezing point of water. Thus, in its simplest form, the degree-day method takes no account of the temperature of the snowpack; variations in local radiation balance due to slope, aspect, cloudiness, changing albedo, partial snow cover and vegetation projecting above the snow; inputs of heat associated with rainfall, changing snow pack area, etc., except in so far as these can be accounted for by adjusting either F or the temperature values. The degree-day method works best when the snow pack has ripened to the melting temperature, but even then the diurnal changes in air temperature and day-to-day weather conditions can lead to variations in melt rates in both time and space. It is not expected, however, that the degree-day method will be accurate at all points in space, only that it will produce reasonable estimates of melt and the lifetime of a pack over an area, given some knowledge of the snow water equivalent at the end of the accumulation period.

Various modifications to the degree-day method have been made to try to extend the basic concept of a temperature-dependent melt. In the ETH-4 version of the method (Hottelet *et al.*, 1993), a continuous balance of water equivalent and snowpack temperature is maintained throughout the winter period. This version of the degree-day method was used by Ambroise *et al.* (1996b) in modelling the small Ringelbach catchment in the Vosges, France.

In the ETH-4 model, when precipitation occurs, it is added to the pack water equivalent as liquid water, snow or a mixture of the two depending on the air temperature at the time relative to two threshold temperatures. Above T_{rain}, all precipitation is assumed to be in liquid form; below T_{snow}, all precipitation is assumed to occur as snow. Between the two, the proportion of snowfall is given as the ratio $(T_j^a - T_{snow})/(T_{rain} - T_{snow})$ where T_j^a is the air temperature at time step j. Any precipitation in the form of snow is added directly to the pack.

At each time step, the temperature of the pack, T_j^s, is updated according to the equation

$$T_j^s = C_T^1 T_j^a + (1 - C_T^1) T_{j-1}^s \tag{B3.3.2}$$

Rainfall is added to the pack at a rate depending on the pack surface temperature. If T_j^s is less than a threshold temperature T_c (which need not be the same as T_F), all rainfall is assumed to be frozen into the pack. Otherwise, if $T_j^s > T_c$ then the rainfall is added to the pack as liquid water content. Part of this liquid water content may freeze at a rate proportional to the temperature difference between the pack temperature and the threshold temperature. If there is some liquid water content, the increase in snowpack water equivalent is calculated as

$$S_j = S_{j-1} + C_T^2 (T_j^s - T_c) \tag{B3.3.3}$$

The model is completed by conditions to ensure that there is not more melt or freezing than available water equivalent.

This simple snowmelt model already contains a number of parameters that must be calibrated. These are the coefficients F, C_T^1, and C_T^2 and the threshold temperatures T_F and T_c. These might vary with location; it has long been recognised that the melt coefficient F is not generally constant but increases during the melt season (Figure B3.3.2). The World Meteorological Organisation (WMO, 1964) recommends values of the degree-day coefficient that increase through the melt season. Hottelet *et al.* (1993) represent this change as a gradually increasing sine curve defined by maximum and minimum values of F, resulting in one additional parameter to be estimated.

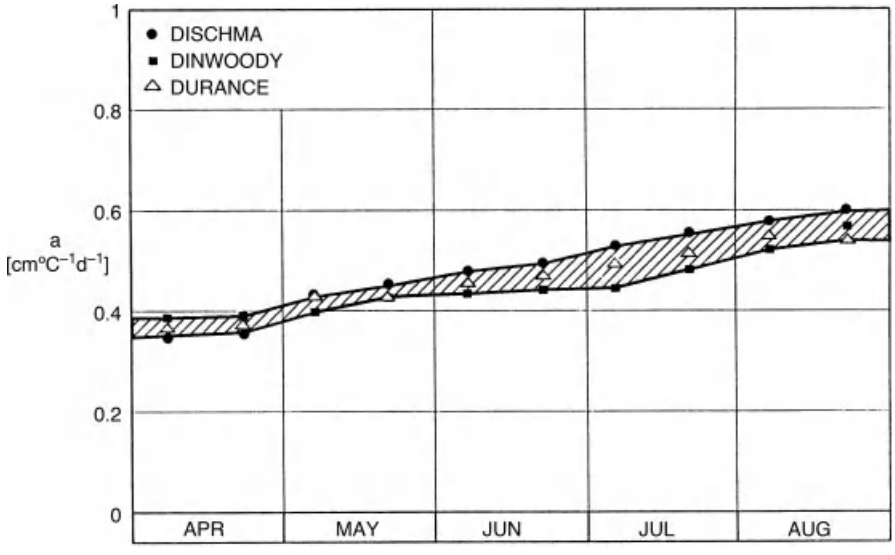


Figure B3.3.2 *Variation in average degree-day factor, F, over the melt season used in discharge predictions in three large basins: the Dischma in Switzerland (43.3 km², 1668–3146 m elevation range); the Dinwoody in Wyoming, USA (228 km², 1981–4202 m elevation range); and the Durance in France (2170 km², 786–4105 m elevation range) (after Rango, 1995, with kind permission of Water Resource Publications).*

Where a catchment covers a wide range of elevations, then it is usual to divide the area up into a number of elevation bands. Rango (1995) reports using bands of 500 m in the Snowmelt Runoff Model (SRM). Air temperatures may then be adjusted for elevation using a simple lapse rate multiplied by the elevation difference from the temperature station so that for band k

$$T_k^a = T_s^a + \gamma(E_k - E_s)/100 \tag{B3.3.4}$$

where γ is the lapse rate [K per 1000 m], T_s^a is the temperature at the recording station, E_k is the mean elevation of band k, and E_s is the elevation of the recording station. A typical lapse rate used would be of the order of 6.5 K/1000 m. A further important adjustment is to account for the change in snow-covered area within each elevation band. This can be done either by using standardised depletion curves for the changing average water equivalent of the pack in each elevation band (Figure B3.3.3) or by updating the current snow-covered area using remote sensing (Rango, 1995).

The SRM Snowmelt Runoff Model is a particularly interesting implementation of the degree-day method because of the way in which it also makes use of satellite imaging to determine snow-covered areas in predicting snowmelt runoff in large catchments (see Rango (1995) for more detail).

In some cases, shorter than daily time steps are required for runoff prediction. Rango and Martinec (1995) report a test of a modification to the degree-day method for use in the SRM model that incorporates a radiation component such that:

$$M = F^* \max(0, \bar{T} - T_F) + R_n/\lambda_m \tag{B3.3.5}$$

where M is now an hourly melt water equivalent, F^* is now a new degree-day coefficient, R_n is the net radiation and λ_m is the latent heat of melting. This modification is clearly more demanding in terms of data but can significantly improve predictions of snowmelt rates (see also Hock, 2003). Net radiation, in particular, may not be measured directly and may need

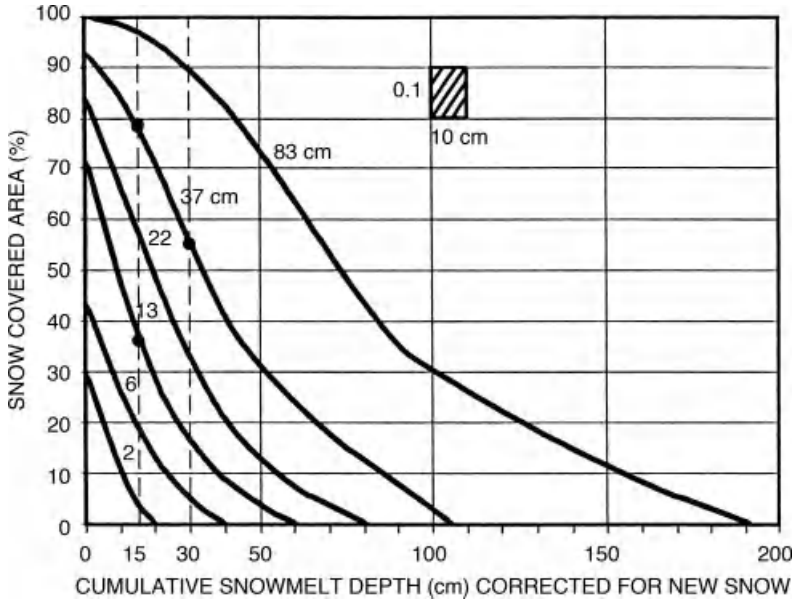


Figure B3.3.3 *Depletion curves of snow-covered area for different mean snowpack water equivalent in a single elevation zone (2926–3353 m elevation range, 1284 km^2) of the Rio Grande basin (after Rango, 1995, with kind permission of Water Resource Publications).*

the albedo and long wave outgoing radiation of the pack to be estimated. Differences in the radiation regime for different slope aspects can be accommodated within this modification. It does, however, take account of the changing radiation regime through the melt season so that the coefficient F^* shows much less variation than F over the season.

Kustas *et al.* (1994) have extended this further to a more complete energy balance melt model for use within SRM. Rango and Martinec (1995) point out that the degree-day method will continue to be used "not so much because of its simplicity but because of its modest data requirements". They suggest that the results, averaged over periods of a few days or routed through a runoff model on a large catchment, are comparable in accuracy to the more complex energy budget models available.

The assumptions of the degree-day method tend to vary with the implementation but the most important may be summarised as follows:

A1 Predicted snowmelt from a ripe snowpack is a linear function of the difference between local mean daily temperature and a threshold temperature.
A2 The degree-day factor will tend to increase as the melt season progresses.
A3 Variation in snow-covered areas can be taken into account by using local depletion curves or remote sensing.

4

Predicting Hydrographs Using Models Based on Data

Some experts are fond of saying that the simplest methods are the best. This really confuses the issue. All other things being equal, this is clearly true. Unfortunately, all other things are not usually equal. The prime criterion should be accuracy, and until equivalent accuracy is demonstrated, simplicity should be a second-order criterion.

Ray K. Linsley, 1986

Finally, a most important aspect of DBM modelling is the identification of parametrically efficient (parsimonious) models, as a way of avoiding the identifiability problems that are associated with the estimation of over-parameterized models from limited time series data .... So, a parsimonious model, in this important dynamic sense, is one that has a lowest dynamic order that is consistent with the information content in the data and whose parameters are statistically significant.

Peter Young, 2003

4.1 Data Availability and Empirical Modelling

There are two very different, but both widely held, views of modelling. The first holds that all models, however physically based their theory, are essentially tools for the extrapolation of available data in time (to different periods) and space (to different catchments). This view of modelling as induction is the subject of this chapter. The second view holds that models should as far as possible reflect our physical understanding of the processes involved. Only in this way, it is suggested, can we have faith in predictions that lie outside of the range of data available in time (e.g. in the future) and space (in different catchments). This view, modelling **almost** as deduction, is the subject of Chapter 5. It is modelling **almost** as deduction because, unfortunately, we cannot yet get away without some empiricism in the description of hydrological processes and in estimating model parameters, and may, in fact, never be able to do so (see, for example, Beven, 2006b; the discussion in Chapter 9).

Rainfall–Runoff Modelling: The Primer, Second Edition. Keith Beven.

In the first view, the approach is unashamedly empirical. The modelling problem becomes one of trying to make the most of the data that are available; indeed, to learn from the data about how the system works by trying to relate a series of inputs to a series of outputs. This is "data-based modelling", usually lumped at the catchment scale, without making much physical argument or theory about process. Another name that is used quite often is *black box* modelling. If we can successfully relate the inputs to the outputs, why worry about what is going on inside the catchment box, especially when the data available might not justify or support the calibration of a complex model (e.g. Kirkby, 1975; Jakeman and Hornberger, 1993)? The black box analogy, however, is not necessarily a good one. An approach based on an input–output analysis may, in some circumstances, lead to very different conclusions about the operation of the system than those that an accepted theoretical analysis would suggest. The data is, then, suggesting that the theoretical analysis may be wrong in some respect.

Consistency with the available observations is therefore given prominence in this approach, so it is worth remembering the discussion in Chapter 3 about the potential for error and inconsistency in hydrological observations. As in any scientific study based on observations, it is essential to critically evaluate the observations themselves before proceeding to base models or predictions on them. It is also worth noting that the inductive or empirical approach is very old; indeed it is the oldest approach to hydrological modelling. The Mulvaney *rational method* described in Section 2.1 and the early coaxial correlation or graphical technique of Figure 2.1 are essentially examples of empirical modelling based on data; attempts to find some rational similarity in behaviour between different storms and different catchments. The continuing use of the unit hydrograph approach to rainfall–runoff modelling is also an indication of the value of this type of inductive approach, where the data are available to allow it. This reflects, at least in part, a recognition of the value of data in constraining the representation of catchment responses, as the limitations of more theoretical approaches when applied to catchments with their own unique characteristics become increasingly appreciated (e.g. Sivapalan *et al.*, 2003).

Young and Beven (1994) have suggested an empirical approach that they call *data-based mechanistic (DBM) modelling* (see also Young, 1998, 2001, 2003) . The method is based on earlier work in systems analysis by Peter Young and is inductive in letting the data suggest an appropriate model structure. They suggest, however, that the resulting model should also be evaluated to see if there is a mechanistic interpretation that might lead to insights not otherwise gained from modelling based on theoretical reasoning. They give examples from rainfall–runoff modelling based on transfer functions that are discussed Section 4.5. This type of empirical data-based modelling necessarily depends on the availability of data. In the rainfall–runoff case, it is not possible to use such models on an ungauged catchment **unless** the parameters for that catchment can be estimated *a priori* (see Chapter 10). Similar arguments have been made, for example, by Jothityangkoon *et al.* (2001) and Farmer *et al.* (2003) as a *top-down modelling* methodology. Young (2003) discusses the DBM approach in relation to other top-down approaches.

4.2 Doing Hydrology Backwards

An extreme form of data-based approach to developing a hydrological model has recently been suggested by Jim Kirchner (2009). In this approach, the storage characteristics of a catchment are inferred from measured fluctuations in discharge, particularly during winter recession periods in which evapotranspiration rates are expected to be small. This is, in fact, very similar to the input–storage–output (ISO) model used as a forecasting model by Alan Lambert (1969; see Section 8.4.1). The starting point for this form of analysis is the water balance equation:

$$\frac{dS}{dt} = R_t - E_t - Q_t \tag{4.1}$$

where S is bulk catchment storage, t is time, R_t is rainfall input at time t, E is actual evapotranspiration, and Q is discharge from the catchment. Under the assumption that Q can be specified as an arbitrary function of S, then this can also be expressed in terms of the change of discharge through time as

$$\frac{dQ}{dt} = \frac{dQ}{dS}\frac{dS}{dt} = \frac{dQ}{dS}(R_t - E_t - Q_t) = g(Q)(R_t - E_t - Q_t) \tag{4.2}$$

The function $g(Q) = \frac{dQ}{dS}$ can then be defined by

$$g(Q) = \frac{\frac{dQ}{dt}}{(R_t - E_t - Q_t)} \tag{4.3}$$

It is then clear that this function is most easily determined from discharge information alone when $R << Q$ and $E << Q$, i.e. winter recession curves during dry periods (see also Brutsaert and Nieber, 1977; Lamb and Beven, 1997). Plots of $\frac{dQ}{dt}$ against Q during such periods should show the regularity of this function in a particular catchment under different conditions. Figure 4.1 shows the Kirchner example at Plynlimon in Wales.

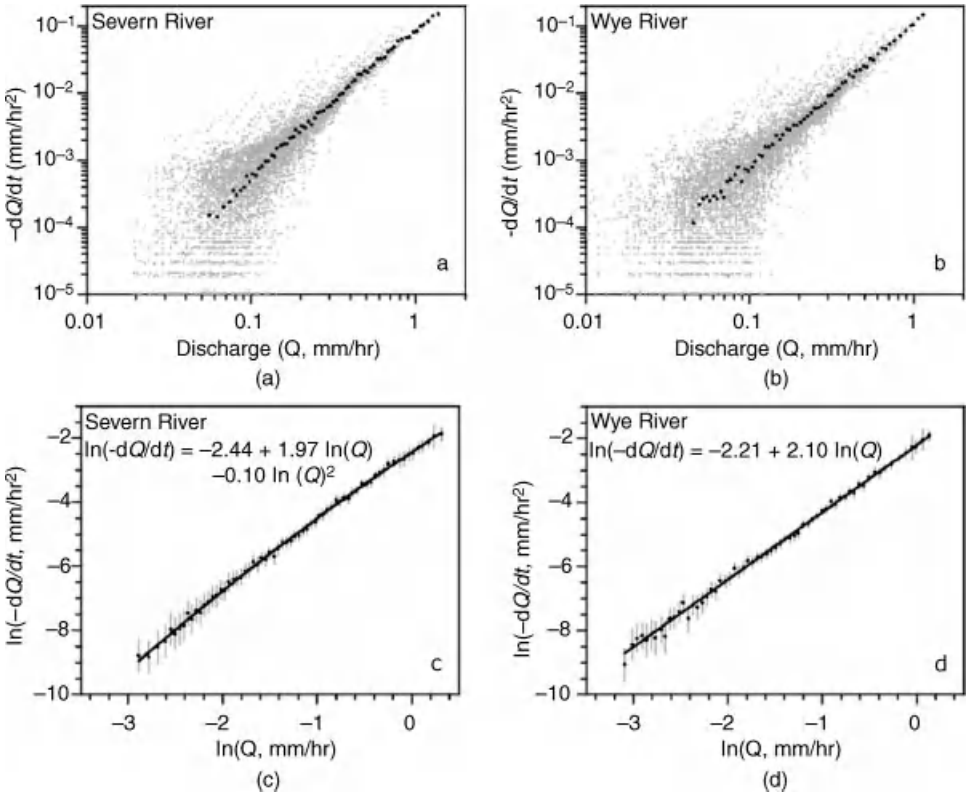


Figure 4.1 *Plots of the function g(Q) for the Severn and Wye catchments at Plynlimon: (a) and (b) time step values of $\frac{dQ}{dt}$ against Q; (c) and (d) functions fitted to mean values for increments of Q (after Kirchner, 2009, with kind permission of the American Geophysical Union).*

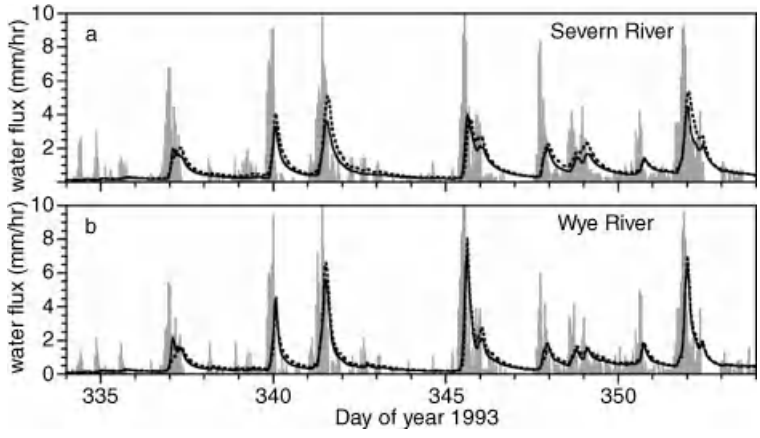


Figure 4.2 *Predicted hydrographs for the Severn and Wye catchments at Plynlimon (after Kirchner, 2009, with kind permission of the American Geophysical Union).*

Once the function $g(Q)$ has been determined, it can be used to predict discharges, given estimates of rainfall and actual evapotranspiration and a starting value of Q, from:

$$\frac{dQ}{dt} = g(Q)\,(R_t - E_t - Q_t) \tag{4.4}$$

This eliminates the need to estimate absolute values of storage (also an attraction of the ISO model in forecasting). Some care needs to be taken over the integration of Equation(4.4) because of the nonlinearity of $g(Q)$. A simple first-order integration scheme will often result in instabilities in the predictions. Kirchner suggests using a log transform, since $ln(g(Q))$ will be smoother than $g(Q)$, but a higher order scheme can also be easily be implemented.

It is worth noting that the function $g(Q)$ is necessarily an approximation of the complex and dynamic storage relationships in a catchment. It also takes no explicit account of the effects of routing delays in the catchment, resulting from wetting the soil profile and flow in the channel network. Channel routing velocities in these catchments are of the order of 1 ms^{-1}, giving a travel time for runoff in the channel network of the order of 1 hour (Beven, 1979b). This will lead to some timing errors in predicted hydrographs, although these should be small for the Plynlimon catchments (as shown in Figure 4.2).

A more interesting application of this approach is to infer effective rainfall inputs and evapotranspiration rates from the fluctuations in discharge, what Kirchner calls "doing hydrology backwards". Thus, inverting Equation (4.4) we obtain:

$$R_t - E_t = \frac{\frac{dQ}{dt}}{g(Q)} + Q_t \tag{4.5}$$

Thus, given the function $g(Q)$ and a discharge time series Q_t, the time series of the effective input $(R_t - E_t)$ can be derived. When it is raining, $(R_t - E_t)$ should be positive and R_t should be much greater than E_t. When it is dry, $(R_t - E_t)$ should be zero or negative and we would expect E_t to be greater than R_t. Thus, with some approximation, rainfalls and actual evapotranspiration can be estimated as averages over the catchment area. In some circumstances, such estimates might be more accurate than measuring rainfalls directly, especially when there are strong patterns of rainfalls (see Newson (1980) for an example at Plynlimon). Figure 4.3 shows a comparison of inferred and point-measured rainfalls for the Plynlimon catchments.

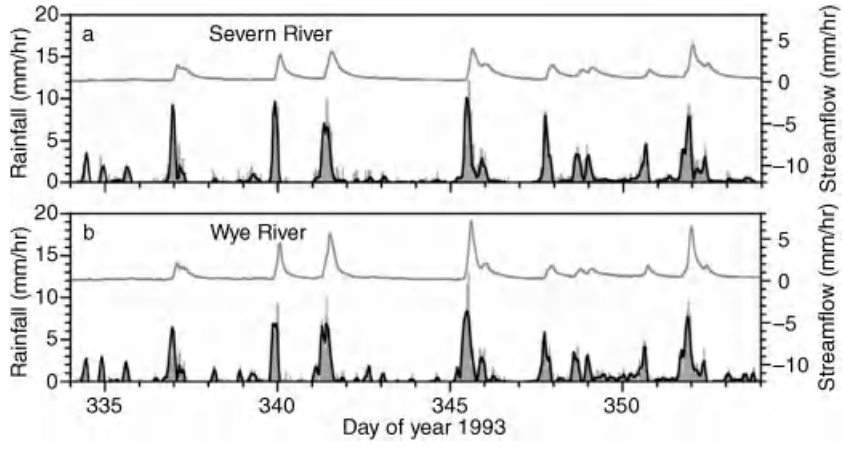


Figure 4.3 *A comparison of inferred and measured rainfalls at Plynlimon (after Kirchner, 2009, with kind permission of the American Geophysical Union).*

This approach to determining a catchment model directly from data has also been tested in Switzerland (Teuling *et al.*, 2010). Its simplicity and elimination of the need to use storage calculations are attractive and, as used in the ISO form for forecasting, it has data assimilation properties (see Section 8.4.1). It is clearly effective in small wet catchments, such as those at Plynlimon, but it is likely that it will have a limited range of applicability because of the assumption that the function $g(Q) = \frac{dQ}{dS}$ is single valued. In many catchments, with larger or multiple storage elements, slower responses, or extended wetting up periods or routing, the hysteresis in the storage–discharge relationships would be expected to be much more pronounced (e.g. Beven, 2006b; Martina *et al.*, 2011). One way of reflecting this hysteresis in both storage responses and routing is to incorprate a transfer function in relating the input to the output. One form of transfer function, traditionally used in hydrology, is the unit hydrograph.

4.3 Transfer Function Models

In this section, we look at a modern variant of the unit hydrograph approach when there are some rainfall–runoff time series data available for a catchment and it is possible to use that information to infer a representation of the response of the catchment. The unit hydrograph is a form of linear transfer functions relating input to output, as are the triangular and Nash cascade representations of the unit hydrograph shown in Figure 2.6. Modern approaches stem from work in linear systems analysis in which a general linear model is used to suggest an appropriate model structure compatible with the input–output data available (see Box 4.1). These models will also have a useful mechanistic interpretation in terms of cascades of one or more linear stores, perhaps with feedback loops. If more than one store is inferred (a second or higher order model) then the stores may be configured in series or in parallel. For example, evaluation of catchment transfer functions of the type described in Box 4.1 has frequently suggested that a parallel model structure is appropriate with a proportion of the runoff being routed through a fast pathway and the remainder through a slow pathway (Figure 4.4). This does not allow any firm conclusions to be drawn about whether surface or subsurface flow processes are involved; it does allow some characteristic time constants for the catchment to be defined in terms of the mean response times for the fast and slow flow pathways (see Box 4.1).

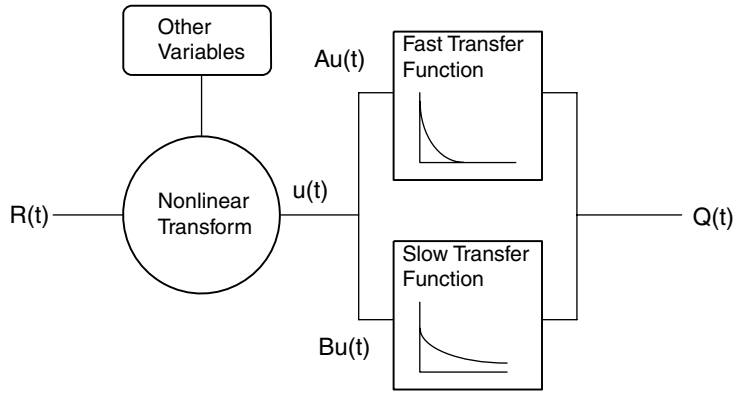


Figure 4.4 *A parallel transfer function structure and separation of a predicted hydrograph into fast and slow responses.*

Note that these time constants are often referred to as mean residence times for the storage elements inferred in this type of representation. As is much too often the case in hydrology, this nomenclature is somewhat confusing since the transfer function as used here refers to the hydrograph response, not to the actual residence times of water in the system. The response time of the hydrograph and the residence times of the water in the system are, nearly always, different. Mean residence times for the water in the system are generally much longer and also tend to vary more as the system wets and dries (see the discussions of the difference between celerities and velocities in Sections 1.4, 5.5.3 and 11.6). It is worth noting, however, that this type of transfer function approach has also been used very effectively in modelling the transport of solutes and pollutants in both surface and subsurface flows. The transfer function, then, represents the distribution of flow velocities directly. In passing, it is also worth noting this pollution transport prediction problem is a case where the transfer function model suggests that a classical theory (here the *advection–dispersion* equation) needs some modification before it can reproduce the long tails seen in tracer experiments and pollution concentration curves (see, for example, the work of Wallis *et al.*, 1989, and Green *et al.*, 1994).

The problem in applying transfer function methods to the rainfall–runoff system is that rainfall is related to stream discharge in a very nonlinear way. Many years of experience with the use of the unit hydrograph method in hydrological prediction have shown that storm runoff may be more linearly related to an "effective" rainfall, but here we wish to avoid any need to carry out any prior separations of the rainfall and runoff time series since, as discussed in Section 2.2, hydrograph separation is a pretty desperate analysis technique. However, we can interpret this experience to suggest that it may be possible to use a linear transfer function model for calculating the time distribution of the total runoff if we can find an appropriate nonlinear filter on the rainfalls to represent the runoff generation processes. The question then is how to find the appropriate form of filter.

One way is to simply assume that a certain form is physically reasonable and that constant parameter values can be found that give a good fit to the data throughout the calibration period. Early work on linear transfer functions of this type was reported by Jim Dooge (1959) and Eamonn Nash (1960). If there is truly a linear relationship between the transformed inputs and the measured output data then this should, in fact, be the case. It is the traditional approach used with unit hydrograph theory where the transformation from total rainfall to effective rainfall is based on an infiltration equation or a method such as the Φ-index method (see Section 2.2). The transfer function based IHACRES model, described below, also adopts this strategy. However, if an estimate of a transfer function is available for a catchment, we

can take advantage of the linearity assumption to apply it in an inverse sense to estimate a pattern of effective rainfalls given a (previously separated) storm runoff hydrograph (see Box 4.2). The aim is then to interpret this pattern to understand the nonlinearity of the catchment response to rainfall.

There have been earlier attempts to deal with the nonlinearity of the catchment response using transfer function approaches. Early concerns about the linearity of hydrological responses resulted in a number of attempts to formulate a nonlinear transfer function (e.g. Amorocho and Brandstetter, 1971; Diskin and Boneh, 1973). These early attempts were based on the use of Volterra series (see also the more recent studies of Ahsan and O'Connor, 1994, and Liang *et al.*, 1994). More recently, a new methodology based on an extension of the class of generalised linear models to the nonlinear case using nonlinear autoregressive moving average with exogenous inputs (NARMAX) models has been applied to hydrological problems. Tabrizi *et al.* (1998) demonstrate how NARMAX models can be used for both single input and multiple input cases.

Another ongoing issue in TFM modelling is the use of multiple rainfall input series in modelling discharges, derived either from multiple raingauges in a catchment area or from radar rainfall data. Multiple input transfer functions have been proposed, for example, by Liang *et al.* (1994), Tabrizi *et al.* (1998) and Kothyari and Singh (1999). A major problem with these approaches is the correlation to be expected amongst the multiple inputs. In the general case, there may be no unique solution to the multiple-input, single-output problem (Cooper and Wood, 1982) and a robust identification of the parameter values may be difficult. One solution to this problem has been proposed by Cooper and Wood (1982) using canonical correlation to identify an appropriate model structure and maximum likelihood estimation to identify the required parameters. For a modern account of multiple-input transfer function models, see the work of Young (2011a).

4.3.1 The IHACRES Model

The identification of unit hydrographs and component flows from rainfall, evaporation and streamflow data (IHACRES) model of Jakeman *et al.* (1990) derives from the work of Young (1975) and Whitehead *et al.* (1979) , which attempted to avoid the problem of hydrograph separation in classical unit hydrograph models by relating total rainfall to total discharge. Recent developments have been the result of a collaboration between the UK Institute of Hydrology (IH) at Wallingford and the Centre for Resource and Environmental Studies (CRES) in Canberra, Australia, resulting in an IHACRES package for PC computers. The model uses a particular set of functions to filter the rainfall to produce an effective rainfall that is then related to total discharge using a generalised linear transfer function. The rainfall filter introduces a soil storage variable and, for longer period simulations, uses temperature as an index of evapotranspiration. A number of different forms of rainfall filter have been used in different IHACRES applications. One form is as follows (see also Croke and Jakeman, 2004). If the rainfall input at time step t is denoted as R_t, while the effective rainfall is denoted as u_t, then

$$u_t = R_t \left(S_t + S_{t-1}\right)/2 \tag{4.6}$$

$$S_t = R_t + \left[1 - \frac{1}{\tau(T_i)}\right] S_{t-1} \tag{4.7}$$

$$\tau(T_i) = \tau_w \exp(10f - T_i f) \tag{4.8}$$

where S_t is the storage variable at time t, $\tau(T_i)$ is a mean residence time for the soil storage depending on mean daily temperature T_i, c controls the proportion of rainfall contributing to catchment storage, τ_w is the mean residence time for soil storage at 10^oC and f is a scaling parameter to allow for the relationship of evapotranspiration effects to this temperature difference. In many respects, this part of the IHACRES model represents a simplified form of explicit soil moisture accounting (ESMA) model (see Section 2.4). The effective rainfall, u_t then forms the input to a transfer function analysis based on the generalised

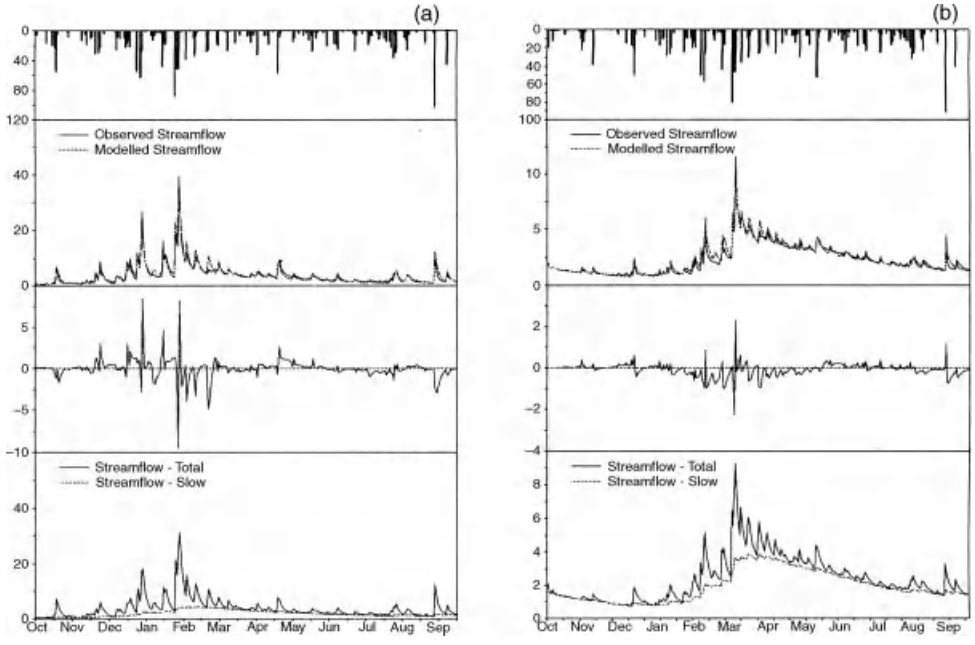


Figure 4.5 *Observed and predicted discharges using the IHACRES model for (a) Coweeta Watershed 36 and (b) Coweeta Watershed 34: Top panel: observed and predicted flows; middle panel: model residual series; lower panel: predicted total flow and model identified slow flow component (after Jakeman and Hornberger, 1993, with kind permission of the American Geophysical Union).*

linear models of Box 4.1 with the total discharge as the output. The parameters of the complete model are calibrated by fitting transfer functions to different values of c, τ_w and f until the best results are achieved. Standard errors and covariances for the transfer function model parameters can be estimated, but uncertainty in the c, τ_w and f parameters has not generally been considered.

The IHACRES model has now been applied to a wide variety of catchments (Jakeman *et al.*, 1990, 1993a; Jakeman and Hornberger, 1993; Post and Jakeman, 1996, 1999; Post *et al.*, 1998; Sefton and Howarth, 1998; Littlewood, 2002; Kokkonen *et al.*, 2003; and see, for example, Figure 4.5), including catchments subject to significant snowmelt inputs (Schreider *et al.*, 1997; Steel *et al.*, 1999) and in predictions of the impacts of climate change on catchment hydrology (Jakeman *et al.*, 1993b; Schreider *et al.*, 1996). The model has also been linked to erosion and water quality components (Jakeman *et al.* 1999). The results generally show that the parallel transfer function of Figure 4.4 is a suitable structure for rainfall–runoff simulation at the catchment scale, with one fast flow pathway and one slower flow pathway (see Box 4.1). The fast flow pathway provides the major part of the predicted storm hydrograph; the slower pathway, the major part of the recession discharge between storm periods. Note again that this does not imply anything about whether surface or subsurface flow processes, old or new water, direct flow or displacement are involved in the fast and slow flow responses. The fast flow pathway should not be interpreted as a surface flow pathway; it could equally be a response controlled by the time scale of displacement of old water from subsurface storage. This type of parallel pathway linear routing is also used in many other models (including the Xinanjiang, Arno, VIC model described in Box 2.2; HyMOD, which has been widely used in studies on the calibration of hydrological models (e.g. Boyle *et al.*, 2000;

Vrugt *et al.*, 2003; Smith *et al.*, 2008b); and the Probability Distributed Moisture (PDM) and Grid to Grid (G2G) models described in Section 6.2). The advantage of the IHACRES approach is that the data is allowed to suggest the form of the transfer function used rather than specifying a fixed structure beforehand. A recent software implementation of IHACRES (Classic Plus) is described by Croke *et al.* (2006) and a freeware version can be downloaded from the Centre for Ecology and Hydrology (see Appendix A).

The IHACRES model has the right sort of functionality to reproduce hydrological responses at the catchment scale with about the right number of parameters for those parameters to be identifiable given a period of calibration data, at least for some environments. The parameters required to apply the model are essentially the two time constants of the fast and slow pathways in the parallel transfer function, the proportion of the effective rainfall following each pathway, and the c, τ_w and f parameters of the effective rainfall filter. Jakeman and Hornberger (1993) suggest that these parameters can be considered as "dynamic response characteristics" (DRCs) of a catchment and that it might be possible to relate these DRCs to physical catchment descriptors to allow the prediction of the hydrological response of ungauged catchments. This is discussed further in Chapter 10.

4.3.2 Data-Based Mechanistic Models Using Transfer Functions

The data-based mechanistic (DBM) approach of Young and Beven (1994), as far as possible, makes no prior assumptions about the form of the model other than that a general linear transfer function approach can be used to relate an effective rainfall input to the total discharge. In the spirit of letting the data determine what the structure of the model should be, rather than making ad hoc prior decisions about model structure, they make use of time variable parameter estimation to determine the form of the effective rainfall nonlinearity. Their results suggest that the nonlinearity can often be approximated by the form

$$u_t \propto Q_t^n R_t \tag{4.9}$$

where u_t is the effective rainfall, R_t is the rainfall input, Q_t is discharge, n is a parameter and t is time. Here, discharge is being used very much as a surrogate variable for the antecedent moisture status of the catchment. In general, the measured discharge is the best readily available index of antecedent conditions in a catchment but its use in this way does mean that discharge is being used in the prediction of discharge. This is not a problem in model calibration at gauged sites; it is a problem in prediction or *real-time forecasting* but does not, in fact, turn out to be a difficult problem to overcome. In the original application of Young and Beven (1991), the bilinear model ($n = 1$) was used, but more recent results using time variable parameter estimation (see Box 4.3) have suggested values for n of between 0 and 1 (a bilinear power law model). The application of the DBM approach in this form is demonstrated in the case study discussed in Section 4.4.

Young and Beven (1994) extended this idea to the case of a time variable analysis of the transfer function gain to allow the data to suggest what the form of this nonlinear function should be. At that time, this was achieved by tracking the gain on the transfer function over the simulation and then showing that, when the system was forced by rainfall inputs, the gain showed a strong relationship with catchment wetness (in that case, a power law function with discharge, see Section 4.5 and Box 4.3). Peter Young later suggested filtering these time variable gains on the inputs (which define the "effective" rainfall) directly in terms of the ordering of some indicator variable, rather than in time, in a state dependent parameter (SDP) analysis (see Box 4.3). In this case, the indicator might be some index of wetness (such as the discharge) and this approach has been used to define a nonlinear filtering in a number of applications of the DBM approach (e.g. Young, 2000, 2002, 2003; McIntyre *et al.*, 2011), including flood forecasting (see Section 8.3). Young (2011a) provides an extensive introduction to the theory and practice of transfer function modelling techniques in both continuous and discrete time.

In a further application of these techniques to a longer period of data for a catchment in the eastern USA, Young (2000) has demonstrated that there is a seasonal pattern in the time variable

estimates of the transfer function *a* parameter (and hence in the mean residence time) as well as a similar correlation of the time variable estimates of the gain parameter *b* with discharge. It was found that the seasonal pattern in *a* could be correlated with daily mean temperatures. The correlation for this site is, however, interesting in that it suggests that the *a* parameter changes in an inverse relationship with temperature. This implies that the higher the temperature, the longer the mean residence time. This is also physically reasonable if higher temperatures are an indication of greater summertime evapotranspiration and therefore lower levels of moisture storage and consequently slower response times. In another study using temperature as an input, Young *et al.* (2007) have shown how a *hysteretic* input filter can be derived in predicting periods of snowmelt runoff (see Box 4.3). The filter is different if the temperature is rising or falling. The result is a form of degree-day method for predicting snowmelt, but one in which the coefficient is determined from the observations and is variable with time. The DBM approach has also been applied to predict surface and subsurface runoff in tropical forest catchments (Chappell *et al.*, 2006).

The DBM approach results in a minimal model relating inputs to outputs involving a nonlinear filter (identified from the observations) and normally only one or two time constants for the transfer function. As will be apparent from the use of a nonlinear filter and temperature to account for "effective rainfall", seasonal and snowmelt effects, the approach does not necessarily require that the variables are dimensionally consistent or that water or energy balances are conserved. The DBM approach has been used, for example, to provide models directly from rainfall inputs to river stage (water level) and from upstream stage to downstream stage in routing flood waves (e.g. Leedal *et al.*, 2008; Romanowicz *et al.*, 2008). Neither would be expected to maintain a water mass balance, even though this would normally seem to be a requirement for a hydrological or hydraulic model. In fact, modelling water levels directly can be advantageous in that level is easily measured and does not require the use of a rating curve to estimate discharge, which might be very uncertain for flood flows. Also, for flood forecasting and incident management, it is very often water level that is the variable of interest, not discharge. Not maintaining mass balance can also be an advantage under flood conditions when knowledge of the pattern of rainfall inputs might be poor. This makes the approach very suitable for use with data assimilation techniques in forecasting (see Chapter 8).

The parameters of the DBM model may be considered to be bulk "physical" characteristics at the catchment scale but, as in the regionalisation studies of the dynamic response characteristics of the IHACRES model noted above, it is not yet clear how well these parameters might be related to catchment characteristics nor the range of catchments for which such a simple model might be appropriate. However, what is clear is that good estimates of the parameters can be obtained from only short periods of rainfall–runoff data so that a period of field measurement of rainfalls and discharges at a site of interest might be the best way of calibrating the parameters (see Seibert and Beven, 2009).

It is worth noting that one of the features of the linear time series analysis used to derive the transfer function is that standard errors can be estimated for the transfer function parameters (see, for example, Taylor *et al.*, 2007; Young, 2011a). These standard errors can be used to evaluate the physical interpretation of the model. Young (1992; see also Young *et al.*, 2007), for example, has investigated the sensitivity of the proportion of effective rainfall going through the fast and slow flow pathways to error in the estimated parameters. The results suggest that there will be considerable uncertainty in these proportions, limiting any interpretation in terms of fast runoff contributing areas. At least in this type of model, these uncertainties can be made explicit; there are similar implications for the more complex models that are discussed in Chapters 5 and 6, but their sensitivities are seldom evaluated.

One interesting use of the DBM methodologies has been to emulate the functionality of much more complex models. Beven *et al.* (2008b) have shown how the outputs from a hydrodynamic flood routing model can be reproduced with great accuracy, including the hysteresis in the depth–discharge relationships at a site. Tych and Young (2011) have similarly applied the DBM approach to emulating the outputs from the OTIS transient storage model for pollutant transport in rivers. This could be useful in forecasting either flood wave propagation or pollutant transport to a point because the run times of the DBM emulator

would be microseconds and much faster than the full model. Even complex global carbon models can be emulated in this way and used to predict control strategies for carbon emissions (Young and Ratto, 2009; Jarvis *et al.* 2007).

4.4 Case Study: DBM Modelling of the CI6 Catchment at Llyn Briane, Wales

Thia section describes the data-based mechanistic modelling approach in relation to a specific application to the CI6 catchment at Llyn Briane in Wales (Young and Beven, 1994). The steps are as follows:

1. Fit a preliminary transfer function to the input–output data using the techniques described in Box 4.1. One or more rainfall filtering models might be tried at this stage; the important thing is to derive an estimate of the parameters of a low-order transfer function.
2. Examine the residuals, including tests for nonlinearity. If there is no nonlinear behaviour apparent then the fitted model can be accepted, otherwise go on to Step 1.
3. Calibrate a low-order transfer function (first or second order is normally adequate for rainfall–runoff modelling) using a time variable parameter estimation technique, such as the fixed interval smoothing (FIS) described in Box 4.3. Examine the variation in the parameter values, including the standard errors, through the period of the calibration (e.g. Figure B4.3.1 for the CI6 catchment).
4. Examine the nature of the time variation in the parameters, taking account of the accuracy with which the parameters are estimated. For the CI6 application, it is clear that there is only information added in estimating the b parameter when there is rainfall. The standard error of estimation on this parameter increases rapidly during recession periods.
5. Evaluate the nature of the variation in relation to other variables. In this case, a relation between this parameter and discharge can be found if only the most significant values of the b parameter estimates are used (Figure 4.6). This is physically reasonable, as with the simple bilinear rainfall filter previously used by Young and Beven (1991), in that we would expect that higher discharges suggest wetter antecedent conditions resulting in a greater proportion of the rainfall becoming discharge. This analysis suggests, however, that the rainfall filter should not be the simple multiplication by discharge used in Young and Beven (1991), but a multiplication by discharge raised to a power, here of the order of 0.65. A block diagram of the resulting model is shown in Figure 4.7.
6. Optimise the resulting model. The results are presented in Figure 4.8 as output from the TFM program (see Appendix A). The best fit for this catchment is obtained with a power of about 0.63. This is a slightly better model with different time constants (3.95 h and 80.2 h) to those of the earlier bilinear (power = 1.0) model fit for this catchment. A further example of fitting the nonlinearity using Peter Young's state dependent parameter methodology is given in Box 4.3.
7. Examine the residuals for evidence of further structure or nonlinearity. Here, Young and Beven (1994) show that, although the resulting nonlinear filter and transfer function explain more than 98% of the variance for this period of observed discharges, a third-order autoregressive model of the residuals can be used to explain more than half of the remaining error variance, but there is no evidence of further nonlinearity. Such a model of the residuals might be useful in flood forecasting (see Box 8.1).

The bilinear power law filter with parallel transfer function model has an interesting interpretation. With this model, the proportion of the rainfall that goes through the fast flow pathway is $\propto Q^n$. Without implying too much about the processes involved, this could be considered as a representation of a contributing area function for the fast responses in the catchment, whether the fast responses be due to surface or subsurface flow processes. The contributing area increases with discharge for all $n > 0$.

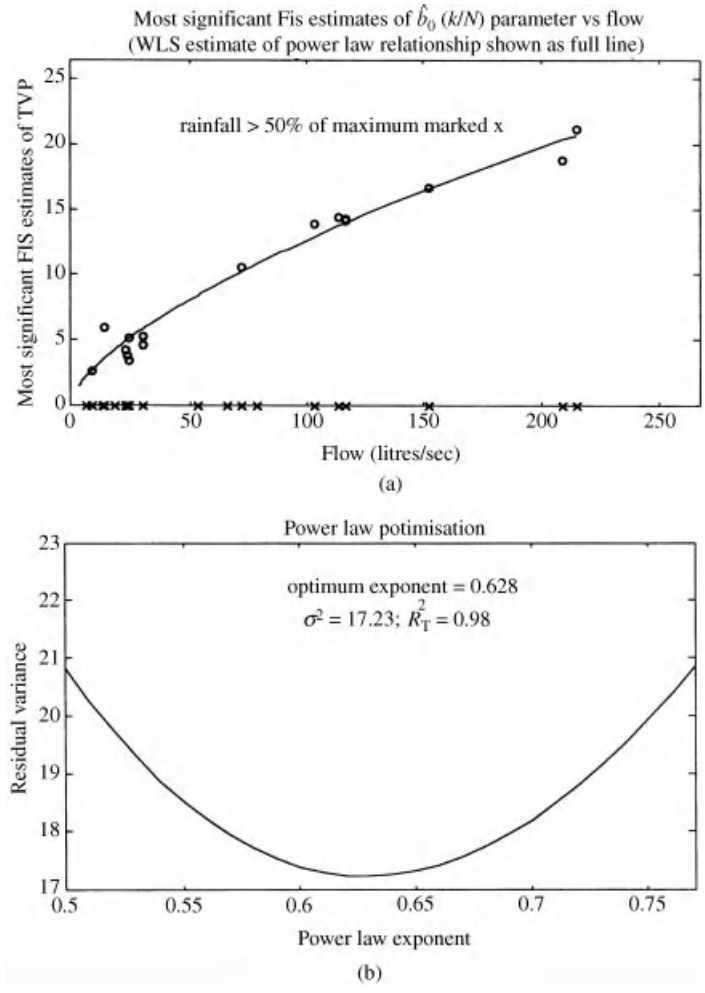


Figure 4.6 *(a) Time variable estimates of the gain coefficient in the bilinear model for the CI6 catchment plotted against the discharge at the same time step; (b) optimisation of the power law coefficient in fitting the observed discharges (after Young and Beven, 1994, with kind permission of Elsevier).*

In fact, a simple link can be made with the prediction of saturated contributing areas in TOPMODEL which depends on the distribution of an index derived from the topography in the catchment. The theory underpinning this link can be found at the end of Box 6.1.

Although of some interest, this representation of the rainfall nonlinearity still requires further evaluation since it takes no explicit account of seasonality of responses except in so far as seasonality is reflected in the contributing area filtering based on discharge. In another application of the bilinear power law model to the Canning River in Australia, Young *et al.* (1997) have shown that this approach can explain 95.8% of the variance of daily discharges over a two-year period, including dry summer periods. A split record test of the fitted model over two further years of data produced almost as good fits (88.9% and 92.4%). The IHACRES model does take account of seasonal wetting and drying by including a temperature

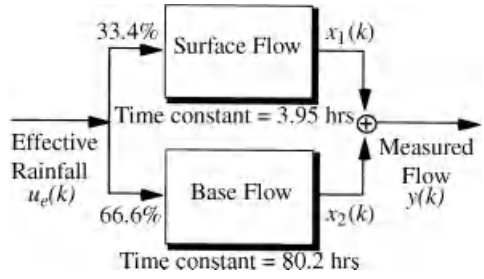


Figure 4.7 *Final block diagram of the CI6 bilinear power law model used in the predictions of Figure 4.6 (after Young and Beven, 1994, with kind permission of Elsevier).*

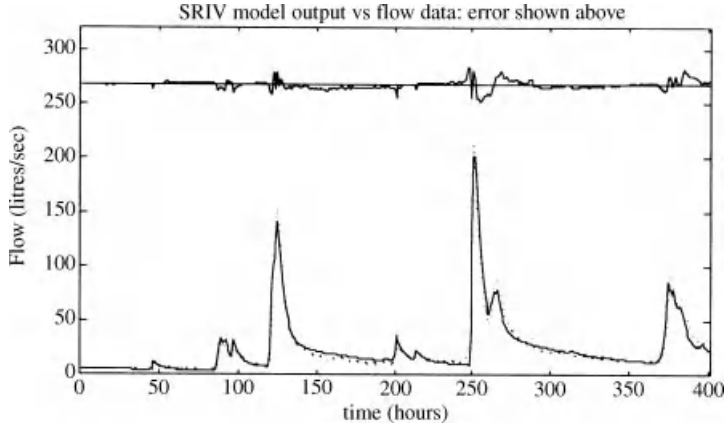


Figure 4.8 *Observed and predicted discharges for the CI6 catchment at Llyn Briane, Wales, using the bilinear power law model with $n = 0.628$ (after Young and Beven, 1994, with kind permission of Elsevier).*

variable as an input to the nonlinear effective rainfall filter component. As noted earlier, Young (2000), in a second example of modelling daily discharges for one of the Coweeta catchments, also show that temperature can be used as a surrogate variable to improve the accuracy of longer period simulations.

4.5 Physical Derivation of Transfer Functions

In Sections 4.3 and 4.4, the transfer functions have been fitted to the data using the generalised linear model developed in Box 4.1. It is possible, however, to develop a transfer function based on the form of a catchment, in a way similar to the Imbeaux/Ross time-area diagram interpretation of the unit hydrograph introduced in Section 2.2. We consider two more recent types of transfer function based on catchment form, one based on the network width function, the other on the geomorphological unit hydrograph (GUH). Note, however, that both of these approaches address only the routing problem and not how much of the rainfall to route. Thus, both require prior estimation of effective rainfalls but they can be used with a variety of effective rainfall models, including those of Chapter 2 and the nonlinear filters described earlier in this chapter. In this respect, they are in the classical unit hydrograph tradition.

4.5.1 Using the Network Width Function

The routing of runoff in catchments is a function both of the hillslope and the channel responses. Work by Kirkby (1976) and Beven and Wood (1993) has demonstrated how the time delays in small catchments tend to be dominated by the routing of surface and subsurface flows on the hillslopes; in large catchments, routing in the channel network plays the dominant role in shaping the hydrograph, especially under overbank flow conditions. If the hillslope runoff inputs to the channel network are distributed along the reaches of the network then, at least in large catchments, the shape of the hydrograph should reflect the form of the network. This is the idea behind using the network width function to derive a transfer function for runoff in the network. The width function is formed by counting the number of channel reaches at a given distance away from the outlet (see Figure 4.9). Different network shapes give different width functions. Under assumptions of a constant wave velocity in the network (which does not imply that the flow velocity must be everywhere constant (see Beven, 1979b, and the section on kinematic wave models

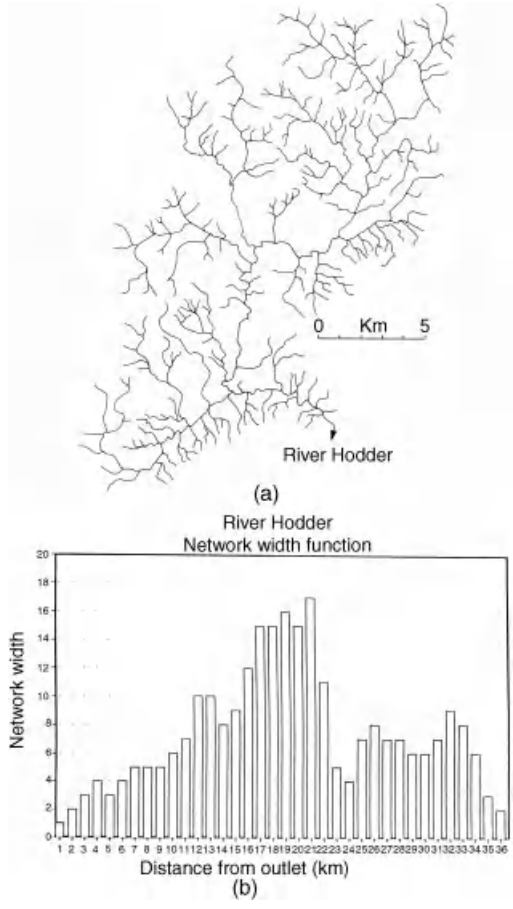


Figure 4.9 *(a) Network and (b) network width function for River Hodder catchment (261 km^2), UK (after Beven and Wood, 1993, with kind permission of Wiley-Blackwell).*

of surface flows in Chapter 5), then the width function can be used directly as a transfer function for routing runoff inputs into the channel.

This type of routing algorithm has been used, for example, in the TOPMODEL software of Chapter 6. It has the advantage that it requires only the network width function, which can be derived directly from maps or digital terrain data, and a single parameter, the wave speed in the channels. It has the disadvantage that it does not deal explicitly with routing on hillslopes and that there is no dispersion of the form of the width function in routing the flow to the outlet since, for a unit input everywhere along the channel, the shape of the resulting hydrograph will directly reflect the width function. This would be obscured, of course, if different patterns of runoff were being produced in different parts of the catchment. Both of these limitations can be relaxed. It is not difficult in an analysis of digital terrain data to derive a distance to the nearest channel for every point on the hillslopes of a catchment, so that this type of approach can be extended to routing surface runoff (at least) to the channel, perhaps using a different velocity. Secondly, Mesa and Mifflin (1986) and Naden (1992) have shown that a diffusive routing algorithm in each reach can be implemented relatively easily at the expense of introducing an additional parameter.

Diffusive network width function algorithms have also gained some recent popularity within the macroscale hydrology field to allow the routing of runoff at the continental scales in a computationally efficient way (see, for example, Gong *et al.*, 2009). The model of Naden (1992) was applied to the Thames and Severn catchments by Naden (1993) and to the Amazon and Arkansas-Red River catchments using GCM-generated rainfall inputs by Naden *et al.* (1999). It is also included as a component of the UP macroscale model discussed in Chapter 8. At this scale, it is generally necessary to estimate the routing parameters required. For a uniform channel, the effective wave velocity, c, and the dispersion parameter, D, used in this model may be approximately related to the characteristics of the flow at a site as:

$$c = \frac{3}{2} v_o \tag{4.10}$$

$$D = \frac{q_o}{2S_o}\left(1 - \frac{F_o^{\,2}}{4}\right) \tag{4.11}$$

where v_o is the mean velocity at a reference discharge q_o in a channel of bed slope S_o and Froude number F_o. In a large basin, the discharge characteristics and channel dimensions will change downstream. Work by Snell and Sivapalan (1995) and Robinson *et al.* (1995) showed how effective values for these parameters over the whole network could be related to the at-a-site and downstream hydraulic geometry of the channels. The parameter estimates have to be considered very approximate and further work needs to be done to both evaluate the routing procedures and improve them by taking full account of the variation in the parameters within these large networks. A recent application of this approach in the context of macroscale rainfall–runoff modelling is provided by Gong *et al.* (2009).

4.5.2 The Geomorphological Unit Hydrograph (GUH)

Related to the use of the network width function as a transfer function for routing runoff is the geomorphological unit hydrograph concept. This idea was initiated by Ignacio Rodriguez-Iturbe in a series of papers (summarised in Rodriguez-Iturbe, 1993) that explored the cause and effect linkages between hillslope form, runoff production, channel growth and network development. The network is a reflection of the runoff-producing mechanisms of the hillslopes operating over a long period of time, but development of the network has a feedback effect on the form of hillslopes and consequent runoff production. These geomorphological linkages result in structural regularities in the form of catchments and it should be possible to take advantage of these regularities in making hydrological predictions. The regularities have been studied by geomorphologists for a long time and are summarised in Horton's laws that express the expected relationships between channel numbers, upstream areas, lengths and slopes for different orders

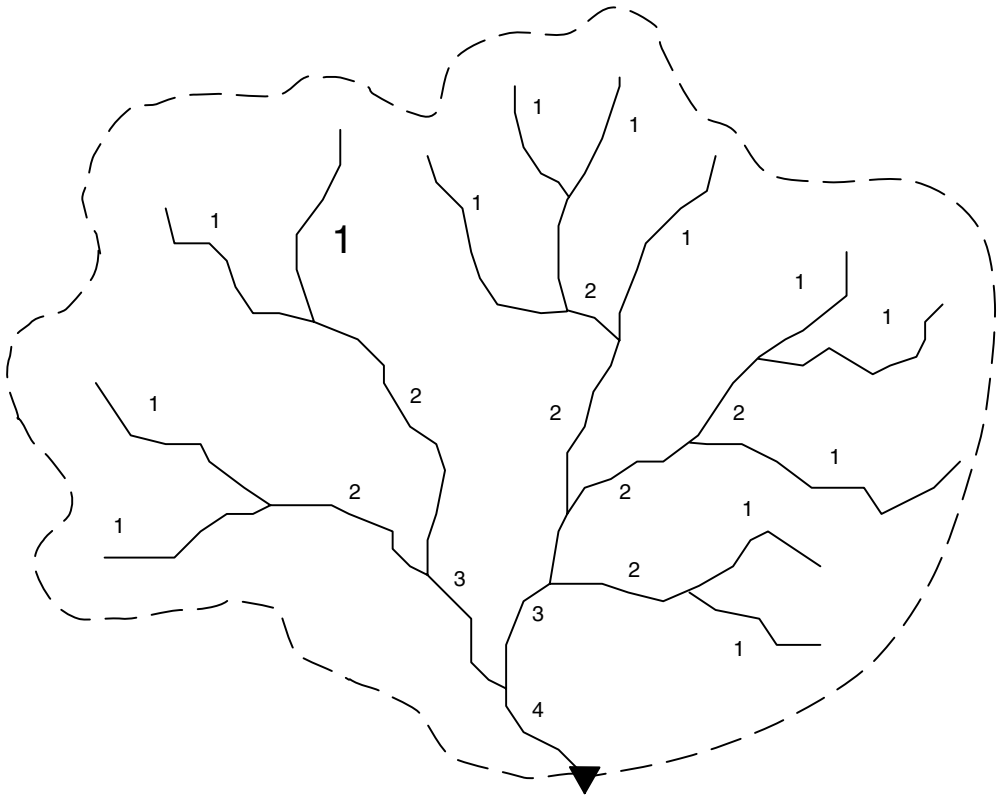


Figure 4.10 *Strahler ordering of a river network as used in the derivation of the geomorphological unit hydrograph.*

of channel where, in the Strahler ordering system, a first-order stream is a stream with no upstream junctions (an exterior link on the network), a second-order stream is formed by the junction of two first-order streams (creating an interior network link), and so on (Figure 4.10).

The geomorphological unit hydrograph is then developed by considering the probability of a "raindrop" contributing to a stream of a given order and a "holding time distribution" for both the hillslopes and the streams of each order. The important part of the theory is the way in which it uses Horton's laws to determine the probability of a drop contributing to each order of stream and the links between the stream of different orders. A third-order stream, for example, will have some direct contribution from local hillslopes, a contribution from the second-order streams feeding it from upstream, and the possibility of a contribution from additional first-order streams. The complete theory is complex and depends on making some simplifying assumptions about the holding time distributions for mathematical tractability. Thus, there have been attempts to simplify the resulting predictions in terms of functional forms of the unit hydrograph. Rodriguez-Iturbe and Valdes (1979) derived expressions for the time to peak and peak discharge of a triangular unit hydrograph; Rosso (1984) did the same for the gamma probability density function that is the mathematical form of the Nash cascade (discussed in Section 2.3). It was later shown by Chutha and Dooge (1990) that all forms of GUH that assume an exponential holding time in each reach of the network must, under reasonable geomorphic assumptions, produce a GUH that is close to a

gamma distribution in shape. Recall that the gamma probability distribution unit hydrograph is defined by two parameters, N and K (Equation 2.2), where in the Nash cascade, N is the number of linear stores in series (which does not have to be an integer number), each of time constant K. Rosso (1984) showed that the two parameters N and K could be related to the geomorphological structure of the channel network as

$$N = 3.29 \left[\frac{R_A}{R_B}\right]^{0.78} R_L^{0.07} \tag{4.12}$$

and

$$K = 0.70 \left[\frac{R_A}{R_B R_L}\right]^{0.48} L_\Omega v^{-1} \tag{4.13}$$

where R_B is called the bifurcation ratio of the network and is equal to the ratio of the number of streams of order ϖ to that of order $\varpi + 1$; R_A is the area ratio of the network and is equal to the ratio of the average catchment area of streams of order ϖ to that of order $\varpi + 1$; R_L is the length ratio of the network and is equal to the ratio of the mean length of streams of order ϖ to that of order $\varpi + 1$; L_Ω is the length of the highest order stream in the network and v is an average stream velocity.

There have now been a large number of studies that have used the GUH in various predictive contexts, from hydrograph prediction, flood frequency prediction, solute transport prediction and predicting the impacts of climate change and catchment sediment yields (Rodriguez-Iturbe (1993) and Rinaldo and Rodriguez-Iturbe (1996) provide useful summaries). Note that, as well as the bifurcation, area and length ratios of the network, there is still a velocity parameter to be calibrated, in the same way that the constant velocity network width function approach also requires a velocity parameter to be specified. This is commonly calibrated using effective rainfall and storm hydrograph data, but the values derived will depend on the formulation of the GUH used. Al-Wagdany and Rao (1998), for example, have shown that calibration of three different formulations of the GUH results in different, but correlated, velocity values. In fact, given the ease with which a full channel network can now be derived from GIS or digital terrain data, there does not seem to be all that much advantage in representing the network in terms of its geomorphological ratios over using the full network structure. Some of the detailed information about the network will be lost in the GUH ratios since Horton's laws are only approximate relationships. Beven (1986b) has discussed the use of the network width function approach as an alternative to the geomorphological unit hydrograph (see also Naden, 1992; Naden *et al.*, 1999) while Gandolfi *et al.* (1999) have used a triangular approximation to the distribution of contributing area to the channel network as the basis for a routing model. Nash and Shamseldin (1998) have suggested that the holding time assumptions of the GUH approach may be overly restrictive in shaping the form of the unit hydrograph and that little is added by the geomorphological ratios to the original Nash cascade unit hydrograph. They suggest that the GUH theory must still be considered as an untested hypothesis (even if it provides a reasonable function form for dealing with the routing process in a rainfall–runoff model).

4.6 Other Methods of Developing Inductive Rainfall–Runoff Models from Observations

4.6.1 Artificial Neural Network Concepts

An alternative approach to inductive rainfall–runoff modelling is the use of artificial neural networks (ANN). Neural networks aim to develop a predictive structure directly from the observational data. They stem from research in artificial intelligence as a simple attempt to mimic the workings of the brain in terms of nodes connected by neurones (ASCE, 2000a, 2000b). The simplest neural networks relate an

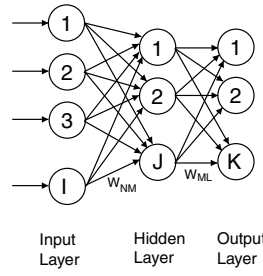


Figure 4.11 *The structure of a neural network showing input nodes, output nodes and a single layer of hidden nodes; each link is associated with at least one coefficient (which may be zero).*

input signal (here we are interested in inputs comprising rainfall and past discharge measurements) to an output signal (current or future discharge) by means of a series of weighting functions that may involve a number of layers of interconnected nodes, including intermediate "hidden layers" (Figure 4.11). Some applications have used additional filtering functions (essentially simple transfer functions) for each node in a hidden layer, so that the outputs also depend on the form and parameterisation of these functions. A variety of techniques are available for determining the appropriate model structure and weights given a *learning set* of input and output data.

The advantage of the ANN approach is that it provides a framework for reflecting complexities and nonlinearities in the data (in so far as that is possible given the available input variables). The structure of the resulting neural network should, therefore, reflect some characteristic processes for a particular catchment (in the same way as the nonlinear input filter and transfer function of the DBM approach). There is, then, an interesting question as to how far the calibrated neural network can be interpreted to provide information about the dominant processes in a catchment. Jain *et al.* (2004) related weights on the hidden nodes of an ANN runoff model to process interpretations, albeit for a network with only a small number of links.

There are clear analogies between the neural network weights and the parameters of other modelling approaches, and between the learning set and what we have before called a period of calibration data. Work in neural networks often does not draw upon this analogy but it is a useful one in that, just as an increase in the number of parameters gives a model more degrees of freedom in calibration but may result in overparameterisation with respect to information in the data set, so in a neural network an increase in the number of layers, nodes and interconnections also results in more degrees of freedom in fitting the learning set, also with the possibility of overparameterisation.

Many studies of the rainfall–runoff problem using neural network techniques have now been published (e.g. Lek *et al.*, 1996; Minns and Hall, 1996; Dawson and Wilby, 1998; Fernando and Jayawardena, 1998; Tokar and Johnson, 1999; Campolo *et al.*, 1999; Chibanga *et al.*, 2003). For the most part, these models have been produced for the purposes of N-step ahead forecasting rather than simulation over long periods. The availability of previous water level or discharge data as an input to the neural net is generally important to the success of such modelling since it allows some of the nonlinearity of the rainfall–runoff process to be reflected in the net for short-term forecasting. An interesting application is that of Campolo *et al.* (2003), using a neural net based on multiple rainfall, discharge and power production inputs to demonstrate good performance in up to six-hour-ahead forecasts for the 4000 km^2 Arno catchment in Italy (Figure 4.12). In forecasting, neural network models can also be used in the same way as transfer functions to forecast downstream river levels or discharges given some upstream data (e.g. Thirumalaiah and Deo, 1998). One study has used neural networks to predict the two-year

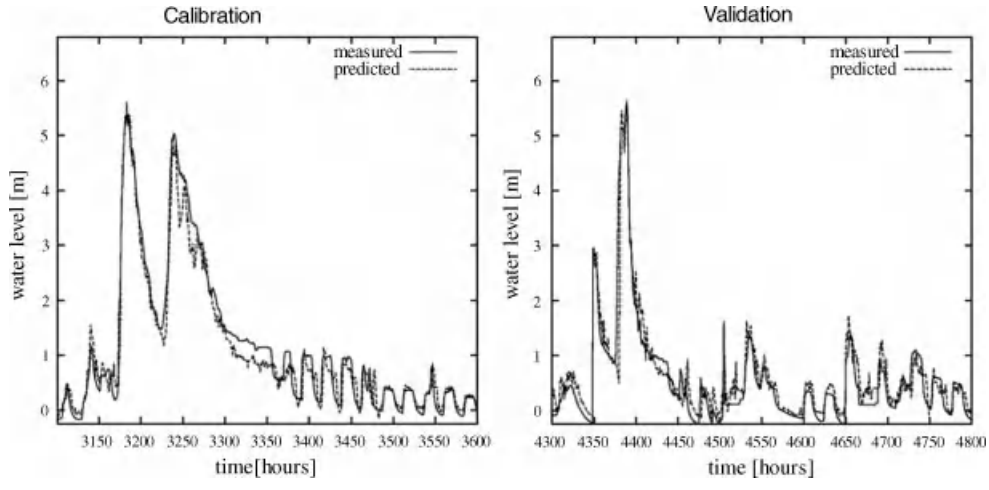


Figure 4.12 *Application of a neural network to forecasting flows in the River Arno catchment, northern Italy: one- and six-hour-ahead forecasts are based on input data of lagged rainfalls, past discharges, and power production information; the influence of power production on the flow is evident in the recession periods (after Campolo et al., 2003, with kind permission of Taylor and Francis).*

return period peak discharge for ungauged basins as an alternative to a multiple regression approach (Muttiah *et al.*, 1997).

Overparameterisation of neural net models relative to the information in the learning set is an issue with this type of model (as in any empirical model). The danger of overparameterisation is that, in general, it will lead to greater uncertainty in prediction or extrapolation, particularly in prediction or extrapolation beyond the range of the learning or calibration set. A good performance in fitting the learning set does not guarantee a good performance in prediction when the conditions go outside the range seen in the learning set (e.g. Cameron *et al.*, 2002; Gaume and Gosset, 2003; Han *et al.* 2007). As yet, there has been little work in the hydrological literature on estimating the uncertainties associated with the predictions of ANN models (but see, for example, Khan and Coulilaby, 2006). Neural networks have also been used as emulators of more complex hydrological and hydrodynamic models (Dibike *et al.*, 1999; Abrahart and See, 2007).

4.6.2 Support Vector Machines (SVM)

An alternative approach to determining a predictive model from data is the use of support vector machines (SVM; Vapnik, 1995; Cristianini and Shawe-Taylor, 2000). This approach was also developed in the field of artificial intelligence as a way of interpreting sensor data and involves a two-layer structure. The first layer is a nonlinear kernel weighting on the input variable series (the support vectors) and the second is a weighted sum of the kernel outputs. It has automatic procedures for deciding on a minimal set of support vectors and for determining the kernel weighting coefficients. Its advantages over ANN methods are that the fitting of the model, once the support vectors and appropriate kernel filters have been determined, is less likely to be over-parameterised (Bray and Han, 2004). Performance has generally been reported to be equal to or better than the results of neural network models. Disadvantages of the method include handling the large matrices of support vectors (which include longer time series than

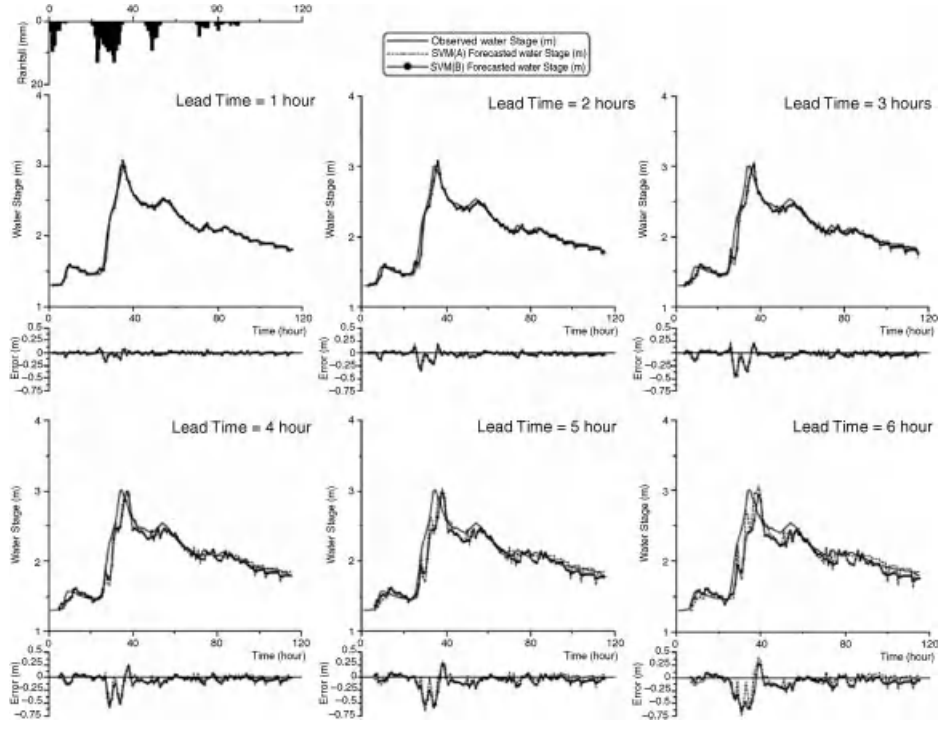


Figure 4.13 *Application of an SVM method to predict flood water levels in real time, with lead times of one to six hours (after Yu et al., 2006, with kind permission of Elsevier).*

normally included in ANN), and testing out the different possibilities for the kernel filters on the different support vectors (similar to the choice of kernel in the hidden layer of the ANN method). As with the ANN approach, there is no real mechanistic check on the resulting model, so that it is likely to perform best in prediction for events that are within the range of the training data, and should be used with care when extrapolating beyond that range to more extreme events since unreasonable predictions might then occur.

An example result for a flood forecasting problem, with predictions for different lead times, is shown in Figure 4.13 (taken from Yu *et al.*, 2006).

4.6.3 Classification and Regression Trees (CART)

The ANN and SVM methods of prediction are flexible methods of handling nonlinearities in the rainfall–runoff relationships for a catchment but are essentially parametric methods, in that they require the calibration of many different weighting coefficients given a training data set. There are some empirical methods that can be considered non-parametric strategies for model development. These generally involve classificatory algorithms in which the training data for the variables to be predicted are subdivided into classes based on values of some input control variables. Those classes are expected to reflect the different conditions that lead to different responses in the catchment. Thus, the most important input control

variables in predicting the discharge response of a catchment might be some measure or measures of rainfalls in the past and some way of characterising the antecedent state of the catchment.

Unfortunately, it might not be clear how best to define these control variables given the data available in the training set. However, we can supply a large number of potential control variables to such a classification algorithm and let it decide which are the most effective in classifying the responses. This is the approach taken, for example, by Iorgulescu and Beven (2004) using a classification and regression tree (CART) algorithm proposed by Breiman *et al.* (1984) and also orginally derived from work in artificial intelligence. This is one possible classificatory algorithm; different algorithms have different ways of finding the most important control variable and forming the resulting classification tree. When applied to this type of dynamic system, it is important that the control variables should reflect the dynamic history of the catchment in setting up the response at a particular time step. The idea is to turn the dynamic response into a static prediction problem at each time step. Thus, for a rainfall–runoff modelling problem, the dynamics depend on rainfalls in the immediate past, rainfalls in an event over the time of concentration of the catchment and longer term rainfalls and evapotranspiration in setting up the antecedent conditions. Discharge at the current time and in the immediate past might also be useful in defining the current state of the system.

Iorgulescu and Beven set up more than 1000 potential control variables, based on integrating rainfalls (and rainfall-potential evapotranspiration) over longer and longer time steps. The algorithm proceeds in a top-down order, i.e. the first split at the root node divides the data into two sets based on a threshold value of one of the control variables, these are then each split into two sets (perhaps using a different control variable threshold), and so on until a set of terminal nodes is defined. Various stopping criteria can be used in the branches of the tree, including a minimum number of data values assigned to the node (usually no fewer than six) or if all the values assigned to the node are identical. At each split, the threshold value of the control variable is chosen to give the greatest explanation of the variability in the data being split. This can also be done in different ways: least square deviation splits are commonly used; the minimum sum of the absolute deviations in the two sets can also be used. The latter is more robust with respect to outliers (Breiman *et al.*, 1984).

In series, such as discharges, the values being split are not independent but exhibit a strong autocorrelation, especially during recession periods. This affects the nature of the splits. Iorgulescu and Beven (2004) tried to allow for this by considering an additional criterion at each split that maximised the diversity of the descendent sets in the sense of containing values from different periods of the original time series. Once each branch of the tree has reached a terminal node, it is normal to consider "pruning" the tree back to try to reduce the impact of overfitting on the resulting predictions. The decision to prune is based on the prediction errors. If removing a branch of the tree improves the prediction errors then that branch can be pruned.

Note that the tree serves to directly map the input control variables (as indicative of the hydrological system) to a set of output values in the terminal nodes of the final (pruned) tree. There are no parameters or coefficients to be determined, only the threshold values for the control variables that determine the branches of the tree. The sets of output values can be used to provide predictions in different ways. Most commonly the median is used to provide a deterministic prediction (e.g. Figure 4.14), but the distribution of values in that node can be used to provide some measure of the uncertainty associated with the outputs under those conditions. The approach can also be extended to allow fitting of a model of the data values in each node, which might help in improving predictions beyond the range of the training data (e.g. Solomatine and Dulal, 2003).

It is known that the partitioning of the tree can be affected by errors in the control variables and output variable or by small changes to the training period. To try to mitigate this effect, a technique called regression forests can be used either by bootstrapping the data or by randomising the choice of splits (Breiman, 2001). Iorgulescu and Beven (2004) have applied the regression tree technique to predict the

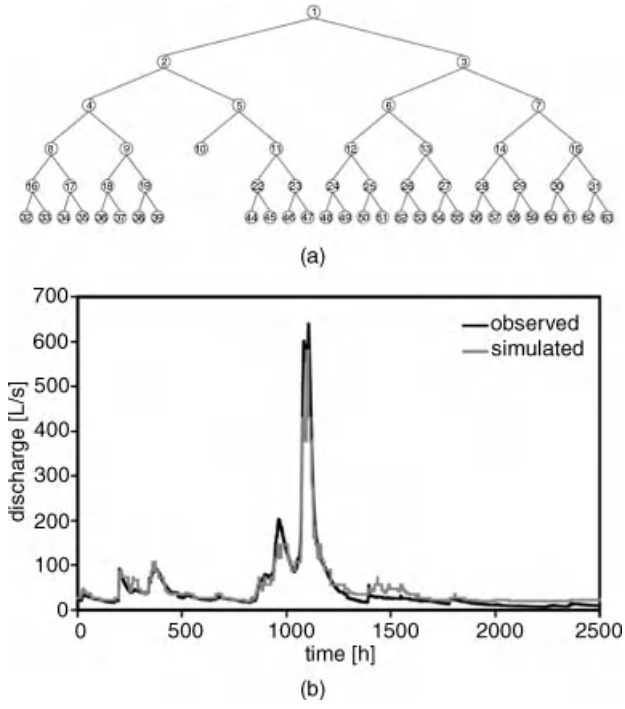


Figure 4.14 *WS2 catchment, H J Andrews Forest, Oregon: (a) depth 5 regression tree and (b) discharge prediction using a regression tree with 64 terminal nodes (after Iorgulescu and Beven, 2004, with kind permission of the American Geophysical Union).*

response of two small catchments in the H J Andrews Research Forest in Oregon, one of which had been harvested of trees and one of which had been left as a control. This method of assessing the impacts of catchment change of time is presented as a case study in Chapter 8.

4.6.4 Fuzzy Inference

When the first edition of this book appeared, a decade ago, models based on fuzzy inference were only in the initial stages of being applied to hydrological models but I thought that it might be a promising technique. There have been studies that extend early work on modelling soil water flows (Bárdossy *et al.*, 1995; Schulz and Huwe, 1997) into rainfall–runoff modelling (Özelkan and Duckstein, 2001; Hundecha *et al.*, 2001; Vernieuwe *et al.* 2005), flow forecasting (Nayak *et al.*, 2005; Alvisi *et al.*, 2006) and estimation of missing rainfall records (Abede *et al.*, 2004), but it cannot be said that fuzzy inference has had a major impact on rainfall–runoff modelling practice. It may still prove to be promising in the future, as the hydrological modelling community starts to address issues of epistemic uncertainties in simulating catchments and the limitations of available hydrological data. The related methodologies of fuzzy rules and fuzzy trees can both be used to represent rainfall–runoff information, with the results expressed as either defuzzified (crisp) numbers or fuzzy possibilistic quantities. The latter, in particular, could be useful as a way of representing some difficult uncertainties. Alvisi *et al.* (2006) raise an issue in prediction, that combinations of circumstances that have not been seen in training the fuzzy inference

will not actually produce a prediction, which might be important in a forecasting situation. If off-line simulation is treated as a learning processes, however, the identification of such periods as a source of new types of information on the catchment response might be valuable. A good introduction to fuzzy methods in rainfall–runoff modelling is provided by Bárdossy (2005).

4.6.5 Flexible Explicit Soil Moisture Accounting Models

In Section1.3, in discussing the modelling process, it was suggested that all hydrological models are conceptual approximations to the complexities of the perceptual model of how catchments work. The phrase "conceptual model" is also used in the literature to indicate a model based on a collection of conceptual storage elements. This type of model dates back to the very first hydrological models implemented as programs on digital computers, such as the Stanford Watershed Model (Crawford and Linsley, 1966). The Stanford Watershed Model established the practice of giving the various storage elements process-related names (overland flow routing store, soil water store, groundwater store, etc.) that led O'Connell (1991) to call these models "explicit soil moisture accounting" (ESMA) models. While I was seeking opinions about improvements to this book for the second edition, one or two people suggested that there should be a chapter on this type of conceptual model, since they are still very widely used in hydrological practice. I have largely resisted this suggestion, having suggested in the Preface to the first edition that they represent the past. (However, the careful reader will find this type of conceptual storage element structure underlying some of the models that are discussed elsewhere in this book, such as the Xinangjiang, Arno, VIC model of Box 2.2 and in many of the semi-distributed models in Chapter 6.)

My resistance derives from my very early experience as a debutant research student in the early 1970s when I tried to make a list of the hydrological models at that time. I stopped counting when the list went over 100. Forty years later, there are many many more, albeit that there has been also been some concentration on a smaller number of such models (particularly those that are applied on the basis of distributed hydrological response units because of the availability of supporting, GIS-linked, software systems). The sheer variety of models of this type, however, means that it is very difficult to decide which model structure (if any) might be appropriate for a particular application. Most might be able to demonstrate some success in representing catchment response, if some observational data are available with which to calibrate the parameters of the model (although see the discussion of calibration methods and issues in Chapter 7). While this might be sufficient as a basis for making predictions of catchment response for a particular purpose, it is very difficult to assess whether such models are getting the right results (in the sense of reproducing calibration data with reasonable accuracy) for the right reasons (see Klemeš, 1986; Kirchner, 2003; Beven, 2010).

This variety has, however, now been used as a form of inductive modelling approach, particularly in the Framework for Understanding Structural Errors (FUSE, Clark *et al.*, 2008). There have been a number of modelling systems in the past that have allowed different types of model components to be put together with a view to improving the representation of particular catchments (e.g. Leavesley *et al.*, 2002) but FUSE has extended this to the evaluation of tens of model structures within a calibration framework for a number of catchments in the Model Parameter Estimation Experiment (MOPEX; Duan *et al.*, 2006). Alternative approaches to evaluating multiple model structures have been taken by Fenicia *et al.* (2007b, 2008a, 2008b) and Krueger *et al.* (2010). In a combination of empirical approaches, Xiong *et al.* (2001) and Fenicia *et al.* (2007a) have reported on using fuzzy inference to combine the outputs of multiple model structures. These studies have shown that a combination of different models might give better predictions than any single model.

Some of the issues with all of these approaches were raised by Clark *et al.* (2008). How far can different model structures be differentiated given the types of performance measures and limitations of calibration data available? How far is the relative performance of particular model structures consistent between different applications? How can model structures be designed to maximise the information content in

prediction by a single model or by an ensemble of model structures? How can we evaluate whether particular model structures are given good results for the right reasons? These questions are much the same as those posed by the availability of multiple models 40 years ago, but now we have much more computing power that allows the performance of many different model structures to be evaluated. If each of these models is considered as a potential hypothesis about how a catchment works, it will continue to be frustratingly difficult to differentiate between them, given the information content of most catchment data sets with an expectation of many different models providing equally good performance (Beven, 2006a, 2010). Chapter 7 goes into much more detail about model calibration, hypothesis testing, and uncertainty estimation.

4.7 Key Points from Chapter 4

- This chapter has dealt with models for the rainfall–runoff process derived directly from data, without explicit consideration of the processes involved. These empirical or inductive models take a variety of forms from transfer function methods to data mining techniques from the field of artificial intelligence (neural networks, support vector machines, classification and regression trees, and fuzzy inference).
- For some catchments where the hysteresis between storage and discharge is not greatly different between wetting and drying, a very simple model of changes in discharge can be derived that depends on deriving a function for the rate of change of storage with discharge, but which does not require the estimation of absolute values of storage. This can also be used in "doing hydrology backwards" to derive effective rainfalls and rates of actual evapotranspiration from changes in discharge (see also Box 4.2).
- Modern transfer function techniques, an extension of the unit hydrograph approach, can be used to derive catchment scale parameters based directly on the analysis of the observations. Transfer function techniques require a nonlinear transformation of the rainfall and the results depend on the form of transformation adopted. In the data-based mechanistic approach of Young and Beven (1994), a flexible approach to model structure is adopted in which time variable parameter estimation is used to suggest a form for the nonlinear transformation.
- The resulting transfer functions are often parallel in form, with one fast flow pathway and one slow pathway. This does not directly imply any interpretation in terms of flow processes, and sensitivity analysis shows that estimates of the proportion of effective rainfall following each pathway may be subject to significant uncertainty.
- Transfer functions may also be derived directly from the structure of the channel network in the catchment. The use of the network width function and the geomorphological unit hydrograph were discussed. Both only address the problem of routing an estimate of effective rainfall.
- Neural network models can sometimes be interpreted in terms of a transfer function, but because of the possibility of large numbers of weights (parameters) to be identified for all the linkages in the network, such methods may be overparameterised and may not make accurate predictions when predicting outside the range of conditions for which they have been calibrated.
- Non-parametric methods (such as classsification and regression trees and some forms of fuzzy inference) do not generally require the estimation of parameters, but are very dependent on the range of behaviour in the training data. They also may not predict well beyond the range of that data.
- A recent method has been to test very many model structures against the available data as a form of hypothesis testing exercise. Given the limitations of readily available hydrological data, it will often be rather difficult to decide whether one structure is really better than another. It has, however, been shown that a combination of different models might give better predictions than any single model.

Box 4.1 Linear Transfer Function Models

B4.1.1 The Building Block: The First-Order Linear Store

A *linear store* is a model element for which the predicted output, Q [L^3T^{-1}], is directly proportional to the storage, S [L^3] (see Figure B4.1.1). Thus, we assume:

$$Q = S/T \tag{B4.1.1}$$

where T [T] is a parameter equivalent to the mean residence time of the store. For water, the linear store is physically equivalent to a straight-sided bucket with a hole in the bottom containing a porous material such that the outflow is laminar and proportional to the difference in head (note that this is not the case for a simple open hole when application of the Bernoulli equation shows that the velocity and discharge through the hole would be proprtional to $S^{0.5}$).

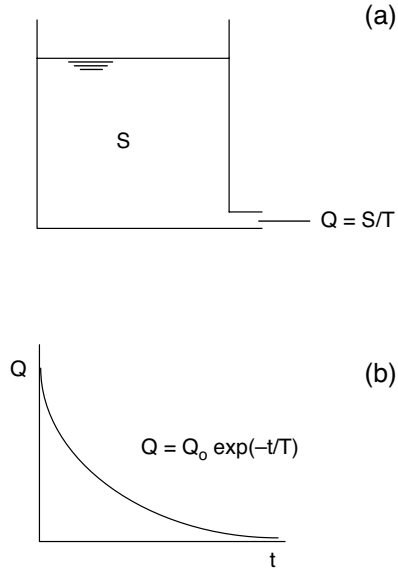


Figure B4.1.1 *The linear store.*

The mass balance equation for the linear store (or bucket) can be written as

$$\frac{dS}{dt} = u - Q \tag{B4.1.2}$$

where the differential $\frac{dS}{dt}$ is the rate of change of storage with time and u [L^3T^{-1}] is an input rate (here, an effective rainfall). To obtain an equation in the outflow Q, since $\frac{dS}{dQ} = T$, we can modify this equation to:

$$T\frac{dQ}{dt} = u - Q \tag{B4.1.3}$$

We assume here that the input sequence has already been suitably transformed to an effective input that can be related linearly to the outputs.

For simple patterns of the effective input, this equation can be solved analytically. For example, for a sudden input of effective rainfall u^* into an initially dry store at time t_o

$$Q_t = \frac{u^*}{T} \exp\{-(t - t_o)/T\} \tag{B4.1.4}$$

This is the impulse response or transfer function of the linear store expressed in continuous time. It has the form of an initial step rise followed by an exponential decline in the outflow.

In hydrology and many other modelling applications, it is often usual to have measurements of inputs and outputs at discrete time increments (e.g. every hour) rather than in continuous time. Thus, using a simple explicit finite difference form of the mass balance equation of the linear store over a discrete time step of length Δt

$$\frac{Q_t - Q_{t-\Delta t}}{\Delta t} = \frac{u_t - Q_{t-\Delta t}}{T} \tag{B4.1.5}$$

or

$$Q_t = \frac{\Delta t}{T} u_t + \left\{1 - \frac{\Delta t}{T}\right\} Q_{t-\Delta t} \tag{B4.1.6}$$

or

$$Q_t = aQ_{t-\Delta t} + bu_t \tag{B4.1.7}$$

where $a = 1 - \frac{\Delta t}{T}$; $b = \frac{\Delta t}{T}$; and to ensure mass balance between total inputs and total outputs, $a + b = 1$.

In hydrological systems, there is sometimes a delay between the start of a rainfall and the discharge starting to rise. Assuming, for the moment, that this delay can be considered a characteristic of the system under study, it can be introduced as

$$Q_t = aQ_{t-\Delta t} + bu_{t-\delta} \tag{B4.1.8}$$

where, in this discrete time equation, the delay δ must be expressed as a number of time steps.

In considering more complex linear transfer function models, it is convenient to introduce the "backward difference operator", z. This is defined as

$$z^{-1} u_t = u_{t-1} \tag{B4.1.9}$$

so that an input delayed by δ time steps may be written as

$$u_{t-\delta} = z^{-\delta} u_t \tag{B4.1.10}$$

and the discrete time mass balance equation can be written as

$$Q_t = az^{-1} Q_t + bz^{-\delta} u_t \tag{B4.1.11}$$

or rearranging this

$$Q_t = \frac{bz^{-\delta}}{1 - az^{-1}} u_t \tag{B4.1.12}$$

B4.1.2 Higher Order Transfer Function Models

Higher order transfer function models may then be constructed easily from the basic linear store building block components in either series or parallel structures (feedback structures are

also possible but are less amenable to a hydrological interpretation). For two stores in series, the individual components are multiplied together so that

$$Q_t = \left(\frac{b' z^{-\delta'}}{1 - a' z^{-1}} \right) \left(\frac{b'' z^{-\delta''}}{1 - a'' z^{-1}} \right) u_t \tag{B4.1.13}$$

or

$$Q_t = \frac{b_o}{1 - a_1 z^{-1} - a_2 z^{-2}} u_{t-\delta} \tag{B4.1.14}$$

where $b_o = b'b''$; $a_1 = a' + a''$; $a_2 = -a'a''$; and $\delta = \delta' + \delta''$

For two stores in parallel, two first-order components are added such that

$$Q_t = \left(\frac{b' z^{-\delta'}}{1 - a' z^{-1}} + \frac{b'' z^{-\delta''}}{1 - a'' z^{-1}} \right) u_t \tag{B4.1.15}$$

or, in the case where both components have the same time delay δ

$$Q_t = \frac{b_o + b_1 z^{-1}}{1 - a_1 z^{-1} - a_2 z^{-2}} u_{t-\delta} \tag{B4.1.16}$$

where $b_o = b' + b''$; $b_1 = -\left(b'a'' + b''a'\right)$; $a_1 = a' + a''$; $a_2 = -a'a''$.

By extension, a general higher order linear transfer function model may be written

$$Q_t = \frac{b_o + b_1 z^{-1} + \cdots + b_m z^{-m}}{1 - a_1 z^{-1} - a_2 z^{-2} + \cdots + a_n z^{-n}} u_{t-\delta} \tag{B4.1.17}$$

This is the general form of linear transfer function that forms the basis of the TFM package described in Appendix A. The Nash cascade discussed in Section 2.3 is one specific form of the general linear model. Combinations of linear storage elements were also investigated for rainfall–runoff modelling in the 1960s by Diskin and Kulandaiswamy (see Chow, 1964, pages 14.31–14.33; Chow and Kulandaiswamy, 1971,; Diskin and Boneh, 1973). The specific case of a second-order parallel model in hydrological applications has been discussed by Young (1992).

B4.1.3 Model Structure Identification

Given an input–output data set, there is then the problem of identifying an appropriate transfer function model structure, that is finding the best values of (m, n, δ) and the corresponding coefficients. There may be no unique answer to this problem, partly because the model will always be an approximation to what may really be a complex nonlinear relationship between the inputs and outputs, partly because there will always be some degree of error associated with the input and output data sets, and partly because the real-time delay might not be a constant or might not be precisely an integer number of time steps. It is often found that there are several model structures that give acceptably accurate simulations after calibration of the a and b coefficients, so some way of evaluating these structures is required.

There are two considerations in making such an evaluation. One is how well the model fits the data, i.e. goodness of fit. A commonly used goodness-of-fit index is the coefficient of determination or proportion of the variance in the data explained by the model (introduced

as the efficiency measure for rainfall–runoff modelling by Nash and Sutcliffe, 1970). This is defined as

$$R_t^2 = 1 - \frac{\sigma_e^2}{\sigma_o^2} \tag{B4.1.18}$$

where σ_o^2 is the variance of the observed output data calculated over all time steps used in fitting the model, and σ_e^2 is the variance of the residual differences between observed and simulated outputs at each time step assuming the residuals have zero mean. As the model fit improves, the value of R_t^2 approaches one. If the model is no better than fitting the mean of the observed outputs (when $\sigma_e^2 = \sigma_o^2$), the value of R_t^2 will be zero or less. In general the higher the model order, the better the fit will be, because there are more coefficients or degrees of freedom in the model than can be adjusted to fit the data. The aim is to find a model structure that gives a good fit but is *parsimonious* in having a small number of coefficients.

There is a danger in allowing too many coefficients or too high a model order. This might result in good values of R_t^2, but the model may then be over-fitted or over-parameterised with the result that the transfer function may not be physically realistic (perhaps with negative ordinates or oscillations) or, when used in simulation, its predictions may be very sensitive to the input sequence used. One way of checking for over-parameterisation is by looking at how well the coefficients of the model are estimated. For example, in the TFM package (see Appendix A), the transfer function model coefficients are fitted using the Simplified Recursive Instrumental Variable algorithm (Young, 2011a) which is available in the Matlab CAPTAIN Toolbox (Taylor *et al.*, 2007). This algorithm allows the standard errors associated with each of the *a* and *b* coefficients to be estimated every time a model structure is calibrated to a data set. The higher the model order (i.e. the greater the number of *a* and *b* coefficients), the larger the standard errors will tend to be. Very large standard errors on the coefficients are an indication that a model is over-parameterised. Young has suggested a criterion for model structure evaluation that combines elements of goodness of fit and standard errors on the coefficients. This "Young Information Criterion" (YIC) is defined as

$$YIC = \ln\left(\frac{\sigma_e^2}{\sigma_o^2}\right) + \ln\left(\frac{1}{N}\sum_N \sigma_e^2 \frac{P_{ii}}{\alpha_i}\right) \tag{B4.1.19}$$

where $\alpha_i, i = 1 \ldots N$ are the model coefficients and P_{ii} is the i^{th} diagonal element of a scaled parameter covariance matrix. The YIC value should be as negative as possible, indicating on its logarithmic scale that the variances of the model residuals and coefficient values are as small as possible. Over-parameterisation leads to a very rapid increase in the value of YIC. However, before accepting a transfer function model of this type, the user should always check the shape of the transfer function to ensure that it is physically reasonable for the system being studied. In the case of rainfall–runoff model or flow routing for example, a transfer function that has negative ordinates would not normally be considered acceptable.

B4.1.4 Factorising a Second-Order Model

If a model is identified as second order ($n = 2$) and has one or two *b* parameters ($m = 1, 2$), then it is possible to factorise the model into two first-order components by treating the denominator as follows. From the equations above, for both a serial and a parallel model:

$$a_1 = a' + a''$$
$$a_2 = -a'a''$$

Combining these two equations, having fitted values of the two coefficients $a_1$and a_2, gives

$$a'^2 - a_1 a' - a_2 = 0 \tag{B4.1.20}$$

that is, a quadratic equation in a'. For cases with real roots, this may be solved using the standard quadratic formula ($x = \left(-B \pm \sqrt{B^2 - 4AC}\right)/2A$ with $A = 1$, $B = -a_1$, $C = -a_2$). The two solutions will be the two coefficients, a' and a''.

Once a' and $a''$are known, for a parallel model ($m = 2$) the relative amounts of flow passing through each component can also then be determined since it is easy to show from the expressions $b_o = b' + b''$ and $b_1 = -\left(b'a'' + b''a'\right)$ that

$$b' = -\left(b_1 + a'b_o\right) / \left(a' + a''\right)$$
$$b'' = b_o - b'$$

Factorisation is also sometimes possible for higher order models with real roots, but the number of possible combinations of serial and parallel connections increases rapidly with model order.

B4.1.5 Determination of the Time Constant for a First-Order Model Component

By setting the impulse response of the continuous time and discrete time models to be equivalent, it can be shown that the a coefficient of a first-order model is related to the mean residence time of the continuous time equivalent approximately as

$$T = -\Delta t / \ln(a) \qquad \text{(B4.1.21)}$$

where Δt is the time step of the discrete time model. The mean residence time, with units of time, has a direct physical interpretation. Young and Beven (1991), for example, demonstrate that for a discrete time model fitted to hourly data for the CI6 catchment at Llyn Briane, Wales, a transfer function model suggests two parallel pathways (see Section 4.3), with mean residence times of 3.1 hours and 66.4 hours, though it must be remembered that there will be some uncertainty associated with these values. Determination of the b coefficients, b' and b'' for the two components suggests that some 28% of the effective rainfall takes the faster pathway in the parallel model and the remaining 72% the slower pathway. The fit of the model for this case is shown in Figure 4.8 ($R_t^2 = 0.991$). This application used the bilinear power filter applied to the total rainfall to create an effective rainfall series for use as an input to the transfer function modelling.

This box has described the discrete time step form of the general linear transfer function since in hydrology both input and output data are generally specified for discrete time increments. Where the sampling time step is short relative to the response time of the catchment, it might be more appropriate to use the continuous time form. Young (2011b) discusses the background to continuous time transfer functions. Routines for fitting both discrete and continuous forms are included in the MATLAB CAPTAIN Toolbox (Taylor *et al.*, 2007; see Appendix A).

B4.1.6 Summary of Model Assumptions

The assumptions made in applying the generalised linear transfer function model are as follows:

A1 The transfer function uses an effective input (such as a suitably transformed *effective rainfall*) to predict an output such as stream discharge, although differences in total volumes of input and output can be scaled by the calibrated coefficients of the numerator (b_o, b_1, ...).

A2 A transfer function is made up of one or more linear storage elements with different time constants combined in series or parallel, choosing the simplest structure that is consistent with the observations.

Box 4.2 Use of Transfer Functions to Infer Effective Rainfalls

At the start of Chapter 4, it was shown how a function for the rate of change of storage with discharge derived from discharge measurements in a catchment could be used to "do hydrology backwards" and derive estimates for effective rainfalls and actual evapotranspiration rates (Kirchner, 2009; Teulen *et al.*, 2010). This box shows how advantage can be taken of the linearity of transfer functions to similarly estimate effective rainfalls, once the transfer function is defined (see also Box 4.3). This will be applicable to a wider range of catchments since it implicitly allows for the hysteresis in the storage–discharge relationship that is reflected in the asymmetric form of the transfer function. It does, however, require that the transfer function be assumed to be linear and stationary in time.

The normal use of transfer function models in rainfall–runoff modelling is to allow the transformation of a time series of effective rainfalls into a time series of predicted discharges. Thus, assuming for demonstration purposes that the simple first-order linear transfer function model of Box 4.1 is appropriate:

$$Q_t = \frac{b_o}{1 - a_1 z^{-1}} R_{t-\delta} \tag{B4.2.1}$$

where $R_{t-\delta}$ is an effective rainfall input at time $t - \delta$, δ is a time delay, the a and b values are coefficients and z is the backward difference operator.

Because the transfer function is linear, however, once the coefficients and time delay for a particular catchment have been identified, this equation may be rewritten as:

$$R_{t-\delta} = \frac{1 - a_1 z^{-1}}{b_o} Q_t \tag{B4.2.2}$$

Thus, given the observed discharges Q_t and an estimate of the transfer function, a time series of effective rainfalls may be derived. There is, however, a problem in that a time series of effective rainfalls is needed to calibrate the transfer function coefficients. Thus, the identification of effective rainfalls in this way tends to be implemented as an iterative deconvolution procedure as follows:

1. Use a simple model (such as the Φ-index approach of Section 2.2) to transform the rainfall inputs into a first estimate of the series of effective rainfalls.
2. Use this series of effective rainfalls to calibrate a transfer function model.
3. Invert the transfer function model to derive a series of calculated effective rainfalls.
4. Repeat Steps 2 and 3 until the series of effective rainfalls and the calibrated transfer function converge to stable values over successive iterations of the process.

There have been several studies of this type of procedure within the context of unit hydrograph theory. The success of the iterative procedure may depend on the initial model of effective rainfalls used and the form of transfer function assumed. Where the transfer function is represented as a series of ordinate values, rather than the functional form of the equations above, the inverse problem is not mathematically well posed. However, Olivera and Maidment (1999) have implemented a deconvolution procedure to determine spatially varied inputs to a channel network transfer function, by initially identifying a mean catchment effective rainfall and then redistributing this to source areas on the basis of a relative runoff coefficient taken from the rational method.

Duband *et al.* (1993) have demonstrated some success in achieving convergence using this type of deconvolution procedure. The success of their "first differenced transfer function - excess rainfall and unit hydrograph by a deconvolution iterative identification technique" (FDTF-ERUHDIT) algorithm is almost certainly helped by two features. One is that they apply

the process to a whole series of rainfall events, rather than separately to individual rainfall events. The second is the fact that they write the equations in terms of the changes in discharge at each time step as the dependent variable. This neatly avoids the need for any initial hydrograph separation and helps reduce the effects of correlation in the errors in calibrating the transfer function.

Other studies of this type include one by Chapman (1996) who developed a method for the estimation of unit hydrographs directly from a collection of streamflow hydrographs that can be used without knowledge of the original rainfall records. This involves the inference of a pattern of effective rainfalls for each individual event by an iterative procedure. The method still involves an initial "baseflow separation" for each event and the resulting unit hydrographs are sensitive to the method of separation chosen. He shows that the resulting average unit hydrographs tend to have higher peaks, shorter times to peak and somewhat shorter durations than those derived by conventional methods (Chapman, 1996). The implied effective rainfalls tend to lag behind corresponding peaks in the measured rainfalls and continues after rainfall has stopped. He suggests that effective rainfalls might best be viewed as the output from a highly nonlinear storage process rather than as the result of a "loss function" as in the more conventional methods for calculation of effective rainfalls described in Box 2.2.

There is one further method of inferring a pattern of effective rainfalls that can be used without the need for an initial hydrograph separation. This is based on using a preliminary identification of a transfer function for a series of hydrographs followed by the estimation of a time variable gain parameter to track the nonlinear way in which total rainfall is related to total discharge. The time variation in the gain identified in this way may be useful in suggesting a nonlinear function for estimating an effective rainfall for prediction of further storm responses (see Young and Beven, 1994). More details of this technique are given in Box 4.3.

Box 4.3 Time Variable Estimation of Transfer Function Parameters and Derivation of Catchment Nonlinearity

There are a number of situations in which allowing the parameters of a transfer function to vary in time can be a useful analysis tool. The two main ones are in investigating the nature of the rainfall-flow nonlinearity (as in the data-based mechanistic modelling approach discussed in Section 4.2) and in adaptive real-time flow forecasting (see Section 8.4). We consider here only the simplest (but very useful) case of time variable transfer function parameter estimation: that of estimating a time variable gain.

Consider the simple first-order, discrete time transfer function of Box 4.1:

$$Q_t = \frac{b_o}{1 - a_1 z^{-1}} u_{t-\delta} \tag{B4.3.1}$$

where Q_t is the discharge output, u_t is the time series of effective rainfall inputs, a_1 and b_o are parameters, δ is a time delay and z^{-1} is the backward difference operator. To allow that the b_o or a_1 parameters might be time variable, it is necessary to make some assumptions about the nature of the time variability. Since we are primarily interested in the broad range and trends of changes in the parameters, we do not wish to allow them to vary too rapidly. One set of assumptions is based on the idea that any time variable parameter, such as the gain, can be modelled as a stochastic generalised random walk (GRW) process, which effectively acts to smooth the changes in the time step to time step values of the parameter. Changes in the parameter values can then be analysed within a framework called "fixed interval smoothing".

Fixed interval smoothing (FIS) (see, for example, Young, 1993, 2000, 2011a; Young and Beven, 1994; Young *et al.*, 1997) is based on a linearisation of the general stochastic model:

$$Q_t = \Im(\Omega_t) + \xi_t \tag{B4.3.2}$$

Here $\Im()$ is a nonlinear function of the variables

$$\Omega_t = \left[Q_{t-1}, Q_{t-2}, \ldots, u_{t-\delta}, u_{t-\delta-1}, \ldots, U_t, U_{t-1}, \ldots\right] \tag{B4.3.3}$$

where the u values are past values of effective rainfall inputs and, for generality, we include values of other exogenous variables, U, at this and previous time steps. The variable ξ_t represents the stochastic part of the model and is assumed to be a zero mean stochastic variable, independent of the variables u and U.

The first-order transfer function model above may then be written as

$$Q_t = x_t + \xi_t \tag{B4.3.4}$$

where x_t is the noise-free output of the model defined as

$$x_t = \frac{b_o}{1 - a_1 z^{-1}} u_{t-\delta} \tag{B4.3.5}$$

Let $\mathbf{p}$ represent the vector of parameter values $\left[b_o, a_1, \ldots\right]$. In the time variable parameter case, changes in any individual parameter, p_i, are described by the generalised random walk of the form

$$\mathbf{p}_{it} = \mathbf{F}_i \mathbf{p}_{it-1} + \mathbf{G}_i \eta_t \tag{B4.3.6}$$

where the elements of $\mathbf{p}_i$ are made up of two components for each parameter value, a changing level and a changing slope, and η_t is a vector of zero mean white noise inputs.

$$\mathbf{F}_i = \begin{bmatrix} \alpha & \beta \\ 0 & \gamma \end{bmatrix}, \qquad G_i \begin{bmatrix} \zeta \\ 1 \end{bmatrix} \tag{B4.3.7}$$

Some specific examples of this general model are the random walk (RW: $\alpha = 0$; $\beta = \gamma = 1$; $\zeta = 0$), the smoothed random walk (SRW: $0 < \alpha < 1$; $\beta = \gamma = 1$; $\eta_t = 0$) and the integrated random walk (IRW: $\alpha = \beta = \gamma = 1$; $\zeta = 0$).

Then, given some measurements with which to compare the model outputs, the transfer function model may be written in what is called a state space form as the two equations:

Observation equation: $\quad Q_t = \mathbf{H}\mathbf{p}_t + \xi_t$

State equation: $\quad \mathbf{p}_t = \mathbf{F}\mathbf{p}_{t-1} + \mathbf{G}\eta_t$

where, for the simple first-order transfer function above, the parameter vector $\mathbf{p}$ contains the time-varying values and slopes for the changing transfer function parameters a_1 and b_o.

The time variable state space equations are readily solved using a form of Kalman filtering (see Young, 2011a) that involves both forward filtering and backward smoothing passes through the data. The algorithm is recursive, in that the calculations are carried out time step by time step so that estimates of the updated parameters are available at every time step. In matrix notation, the time variable parameter estimation algorithm has the following form:

1. Forward pass filtering
   a. Prediction:

$$\widehat{\mathbf{p}}_{t|t-1} = \mathbf{F}\widehat{\mathbf{p}}_{t-1} \tag{B4.3.8}$$

$$\mathbf{P}_{t|t-1} = \mathbf{F}\mathbf{P}_{t-1}\mathbf{F}^T + \mathbf{G}\mathbf{N}_r\mathbf{G}^T \tag{B4.3.9}$$

   b. Correction:

$$\widehat{\mathbf{p}}_t = \widehat{\mathbf{p}}_{t|t-1} + \mathbf{P}_{t|t-1}\mathbf{H}_t^T \left[1 + \mathbf{H}_t\mathbf{P}_{t|t-1}\mathbf{H}_t^T\right]^{-1} \left\{Q_t - \mathbf{H}_t\widehat{\mathbf{p}}_{t|t-1}\right\} \tag{B4.3.10}$$

2. Backward pass smoothing

$$\widehat{\mathbf{p}}_{t|N} = \mathbf{F}^{-1}\left[\widehat{\mathbf{p}}_{t+1|N} + \mathbf{G}\mathbf{N}_r\mathbf{G}^T\mathbf{L}_t\right] \tag{B4.3.11}$$

$$\mathbf{L}_t = \left[\mathbf{I}\text{-}\mathbf{P}_{t+1}\mathbf{H}_{t+1}^T\mathbf{H}_{t+1}\right]^T\left[\mathbf{F}^T\mathbf{L}_{t+1}\text{-}\mathbf{H}_{t+1}^T\left\{Q_{t+1} - \mathbf{H}_{t+1}\widehat{\mathbf{p}}_{t+1}\right\}\right] \tag{B4.3.12}$$

$$\mathbf{P}_{t|N} = \mathbf{P}_t + \mathbf{P}_t\mathbf{F}^T\mathbf{P}_{t+1|t}\left[\mathbf{P}_{t+1|N}\text{-}\mathbf{P}_{t+1|t}\right]\mathbf{P}_{t+1|t}^{-1}\mathbf{F}\mathbf{P}_t \tag{B4.3.13}$$

with the starting condition $L_N = 0$.

In these equations, the matrices **F** and **G** are as defined in Equation (), **I** is the identity matrix, $\mathbf{N}_r$ is the noise variance ratio matrix and **P** is defined as

$$\mathbf{P}_t = \frac{\mathbf{P}_t^*}{\sigma_\xi^2} \tag{B4.3.14}$$

where σ_ξ^2 is the variance of the observation noise and $\mathbf{P}_t^*$ is the error covariance matrix of the state estimates (the states here being the parameter values). The estimation of a covariance matrix for the parameters, updated recursively at each time step, is an important part of this algorithm since the user can then follow how well the parameter values are estimated in

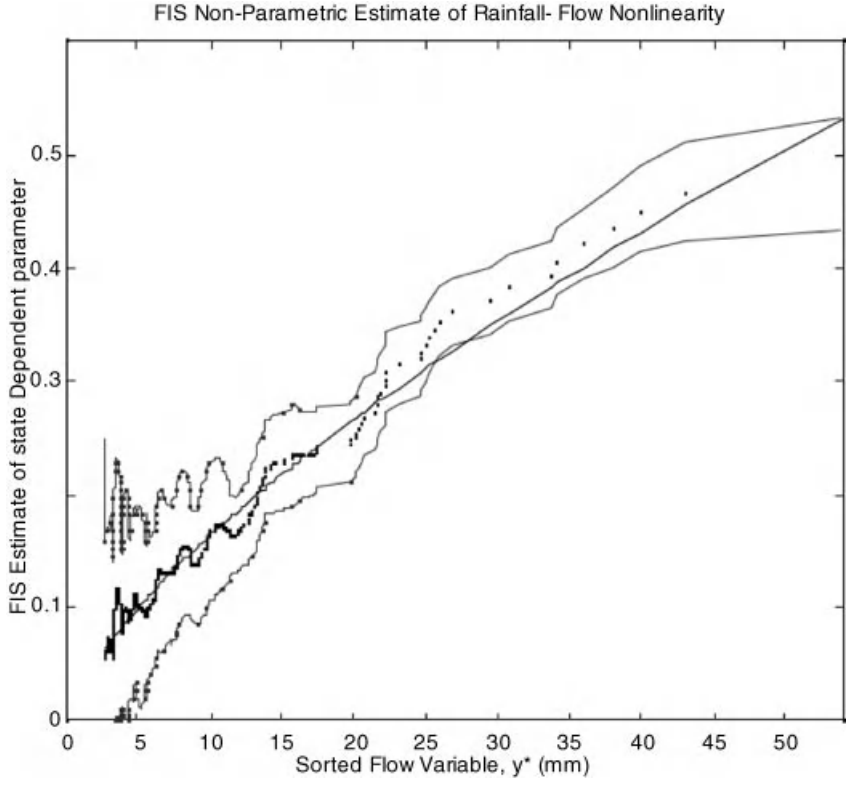


Figure B4.3.1 *Nonparametric state dependent parameter estimation of the gain coefficient in the identification of a data-based mechanistic (DBM) model using daily data at Coweeta (after Young, 2000, with kind permission of Wiley-Blackwell).*

different parts of the study period, as well as the changes in the parameters themselves. The noise variance ratio (NVR) matrix, $\mathbf{N}_r$, is normally assumed to be a diagonal matrix with elements for each of the parameters to be estimated. The NVR controls the effective memory of the algorithm. Large values allow rapid changes in the parameter values at each time step; small values mean that those changes are damped over a larger number of time steps. The NVR may be optimised to achieve the best overall forecasting performance.

In the study of Young and Beven (1994) that is described in the case study of Section 4.4, this type of FIS algorithm was used to define the time variable gain values that were used to define the nonlinear power law filter on the rainfall inputs that is shown in Figure 4.6a. In later work, this was extended by Peter Young to a more general technique that he called "state dependent parameter" (SDP) estimation. He realised that, when it was expected that the nonlinearity might be related to some index of the state of the system (wetness, in the case of a catchment system), the time variable gain values could be filtered in the dimension of that index rather than in time. This is quite easy to achieve by ordering the gain estimates by the value of the index and using FIS on the ranked values (Young, 2000). This SDP methodology actually provides a form of non-parametric function between the gain and the index variable (as shown, for example, in Figure B4.3.1). In prediction, this can either be used directly (in the form of a look-up table) or it can be represented by a simple parametric form (the power law

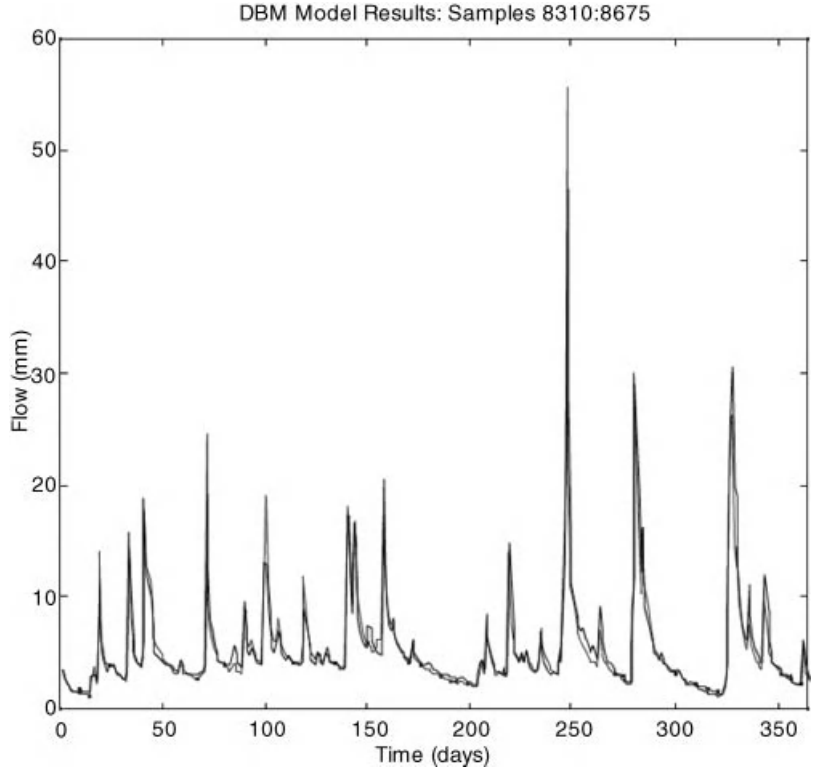


Figure B4.3.2 *Predicted discharge from a DBM model for Coweeta using the input nonlinearity of Figure B4.3.1 (after Young, 2000, with kind permission of Wiley-Blackwell). Peter Young also shows in this paper how this model can be improved even further by a stochastic model of a seasonal function of temperature, representing a small effect of evapotranspiration on the dynamics of runoff production.*

form of Figure 4.6a, with discharge as the wetness index variable, has been found to be quite useful in a number of studies). Figure B4.3.2 shows an example of the predictions of the DBM catchment model with the nonlinearities estimated in this way.

SDP estimation of the nonlinearity has now been used in a number of different catchment modelling studies (e.g. Young, 2000, 2001, 2003), in developing DBM flood forecasting models (Young, 2002, 2011b; Romanowicz *et al.*, 2006, 2008; Leedal *et al.*, 2008; see also Section 8.4) and in the emulation of hydraulic models using the upstream water level as the index variable (Beven *et al.*, 2008b; Young *et al.*, 2009). Once the **form** of the state dependency is identified, then the complete model of nonlinearity and transfer function is generally re-estimated in a final step. For the interested reader, Young (2011b) gives a good explanation of state dependent parameter estimation with a variety of examples to show the power of the methodology.

A further important use of time variable parameter estimation is in adaptive real-time flood forecasting (see Section 8.4). In this case, errors due to the nonlinearities in the runoff generation processes and limitations in the measurements of both rainfall and flow can be allowed for by using an adaptive gain parameter (Lees *et al.*, 1994; Young, 2002, 2011b). In this case, only the forward pass Kalman filtering step is normally used, since observations are only available up to time t. The parameter estimates $\widehat{\mathbf{p}}_t$ are then used to make predictions into the future, before being updated as new data is received at the next time step. The NVR is normally chosen so as to give a relatively long estimator memory, so that the parameters (and consequent forecasts) do not change too rapidly from time step to time step.

5

Predicting Hydrographs Using Distributed Models Based on Process Descriptions

Until the knowledge gained through research in the sciences concerned with hydrologic phenomena becomes adequate enough to permit good quantitative descriptions of these phenomena and their functional relationships, a very strong subjective influence will pervade the entire endeavor. As long as this is true, the efforts of practising hydrologists will remain largely in the realm of the arts. It is natural to expect that under such conditions there may arise numerous emotional controversies regarding the relative merits of apparently conflicting lines of attack to specific problems.

J. Amorocho and W. E. Hart, 1964

Models are thought experiments which help refine our understanding of the dominant processes, . . . testing whether we have a sufficient and consistent theoretical explanation of physical processes. The best model can only provide a possible explanation which is more consistent with known data than its current rivals. Every field observation, and especially the more qualitative or anecdotal ones, provides an opportunity to refute, or in some cases overturn, existing models and the theories which lie behind them.

Mike Kirkby, 1996

5.1 The Physical Basis of Distributed Models

A general model of rainfall–runoff processes requires representations of the interacting surface and subsurface processes. As noted in Section 2.5, an outline of the physics underlying such a description was first published by Freeze and Harlan (1969), although the individual process descriptions had all been established well before then. Most physically based models today are still based on the Freeze and

Rainfall–Runoff Modelling: The Primer, Second Edition. Keith Beven.

Harlan "blueprint", and many are, in fact, simplifications of that blueprint. Even so, their blueprint is not complete. Relative to the perceptual model of Section 1.4, several elements are missing, including the effects of macropores and other heterogeneities of the flow processes (although attempts to introduce the effects of macropores into hillslope and catchment scale models have been made by Zuidema, 1985; and, since then, by Bronstert and Plate, 1997; Faeh *et al.*, 1997; and Zehe *et al.*, 2005).

As noted previously, the boundaries between different types of hydrological model have become blurred in the last decade and there have also been many developments in distributed models that are not based on the continuum differential equations of the Freeze and Harlan blueprint. An example is the widely used Soil Water Assessment Tool (SWAT) model. Since these models are more explicitly semi-distributed in nature, they are considered in Chapter 6; this chapter considers only models based on solution of partial differential equations describing surface and subsurface flows (even then, many such models involve conceptual representations of other processes). We concentrate on the assumptions made in formulating the differential equations, so that other distributed modelling strategies can later be evaluated in the light of these most "physically based" models. Initially, subsurface (soil water and groundwater) and surface (overland and channel flow) flow components are considered separately, followed by their interactions. In Chapter 9, these ideas are critically examined in the context of a strategy for hydrological modelling for the future.

The last decade has seen the development of a number of models of this type, in addition to those presented in the first edition of this book (SHE, IHDM, TOPOG, DVHSM, HILLFLOW, KINEROS, CASC2D). New models that have appeared include the integrated hydrological model (InHM; VanderKwaak and Loague, 2001); the TIN-based real-time integrated basin simulator (tRIBS; Ivanov *et al.*, 2004, 2008); the Pennsylvania integrated hydrological model (PIHM; Qu and Duffy, 2007); TOUGH2 (James *et al.*, 2010); the gridded surface/subsurface hydrologic analysis (GSSHA; Downer and Ogden, 2004; Downer *et al.* 2005); CATFLOW (Zehe *et al.*, 2005; Klaus and Zehe, 2010); and the Waterloo University earth system simulator HydroGeoSphere (Therrien *et al.*, 2006). There have also been some reviews of the experience of using such models (Beven, 2001; Loague and VanderKwaak, 2004; Refsgaard *et al.*, 2010) and important tests against detailed field information (Loague *et al.*, 2000; Bathurst *et al.*, 2004; Ebel *et al.* 2007; James *et al.*, 2010). A lot has been learned from this decade's applications about the practice and limitations of this type of hydrological model and the case studies in this chapter have been changed to reflect this.

5.1.1 Subsurface Flows

The basis of all descriptions of subsurface flow used in distributed models is Darcy's law, which assumes that there is a linear relationship between flow velocity and hydraulic gradient, with a coefficient of proportionality which is now called the "hydraulic conductivity". Thus,

$$v_x = -K\frac{d\Phi}{dx} \tag{5.1}$$

where v_x is velocity in the x direction [LT^{-1}], Φ is the total hydraulic *head* [L] and K is the hydraulic conductivity [LT^{-1}]. Originally established empirically by Henri Darcy (1856) for flow through saturated sands, Richards (1931) generalised the use of Darcy's law for the case of unsaturated flow by making the assumption that the same linear relationship holds but that the constant of proportionality should be allowed to vary with soil moisture content or capillary potential. Thus,

$$v_x = -K(\theta)\frac{d\Phi}{dx} \tag{5.2}$$

where θ is volumetric soil moisture. The notation $K(\theta)$ is used to indicate that K is now a function of θ. Darcy's law can be derived from more fundamental equations of flow, called the Navier–Stokes

equations, if some assumptions are made about the nature of the pore space through which the flow is taking place, and if the flow is slow enough to stay in the laminar regime (e.g. Hassanizadeh, 1986). This is usually a good assumption for flow in a porous matrix but may break down for flow in soils with heterogeneous characteristics (where it may be difficult to define a gradient of potential except at very small scales) and macropores (where flow in large pores and flow in the porous matrix may be responding to different local gradients). Thus Darcy's law is strictly valid only over a limited range of scales (perhaps, at least in the unsaturated zone, all smaller than the normal element scale of a rainfall–runoff model, Beven, 1989, 2006b). There has been an initiative towards the development of appropriate flux equations to replace Darcy's law for application at larger hillslope element scales (e.g. Reggiani *et al.*, 1999; see Chapter 9). However, it should be noted that the use of Darcy's law to represent fluxes in large elements of a distributed model is a gross approximation and may mean that the effective value of the hydraulic conductivity required may be different from anything that could be measured in the field (see Section 5.2).

It might even not be the most appropriate description locally. There is an argument that the original laboratory experiments of Richards were not realistic in respect of flows in field soils. Richards used air pressure to create unsaturated flow conditions (a procedure that became a standard method in laboratory measurements of the relationships between soil water content and capillary pressure). This has the effect of precluding flow in larger pores that might occur in field conditions. Such flow, which can be represented as films of water around particles largely independent of the capillary potentials within the fine pores, has been described at least since Burdine (1953) using Stokes equation (Gerke *et al.*, 2010). Stokes flow does not need large continuous macropores to be a feasible description of such **bypassing** flow, and can be consistent with the type of fingering seen in Figure 1.1. It might well be that we need to revisit the Richards equation as a basis for physically based rainfall–runoff models in future.

The other important equation in this description of subsurface flow is the continuity or mass balance equation (see Box 2.3). The combination of Darcy's law with the continuity or mass balance equation results in a flow equation (generally called the Richards equation) which may be written with capillary potential ψ as the dependent variable as:

$$C(\psi)\frac{\partial \psi}{\partial t} = \nabla\left[K(\psi)\nabla\psi\right] + \frac{\partial K(\psi)}{\partial z} - E_T(x, y, z, t) \tag{5.3}$$

where ψ is the local capillary potential, $K(\psi)$ is unsaturated hydraulic conductivity, which is now expressed as a function of ψ rather than θ, $C(\psi)$ is a function of ψ defined as the rate of change of moisture content θ with change in ψ, called the *specific moisture capacity*, and $E_T(x, y, z, t)$ is a local uptake of water by plant roots to satisfy evapotranspiration (see Section 5.1.3). An equivalent equation with soil moisture content θ as the dependent variable can also be written. The derivation of both forms of the Richards equation is given in Box 5.1. The Richards equation is a partial differential equation (see Box 2.3) and, because of the nonlinear change of hydraulic conductivity with moisture content, it is a nonlinear partial differential equation. Such equations tend to be very difficult to solve analytically, except for some very simple cases of *initial conditions* and *boundary conditions* and simple forms for the nonlinear relationships relating moisture content, capillary potential and hydraulic conductivity, known as the *soil moisture characteristic curves*. Some solutions for the special case of infiltration at the soil surface are given in Box 5.2.

For most cases of interest to hydrologists, it is necessary to use an approximate numerical solution of the equation. A number of different techniques are available, including finite difference (e.g. the SHE model), finite element (e.g. IHDM, HYDRUS, InHM, HydroGeoSphere), boundary element, integrated finite difference (e.g. TOUGH2) and finite volume (e.g. PIHM) techniques. The finite volume method has become increasingly important, in part because it simplifies the coupling of different process equations at the discrete boundaries of the finite volumes and is inherently mass conservative. All these methods involve discretising the flow domain into a network of grids or elements (as shown for a finite element discretisation in Figure 5.1) and solving for values of moisture content θ or capillary potential ψ at a

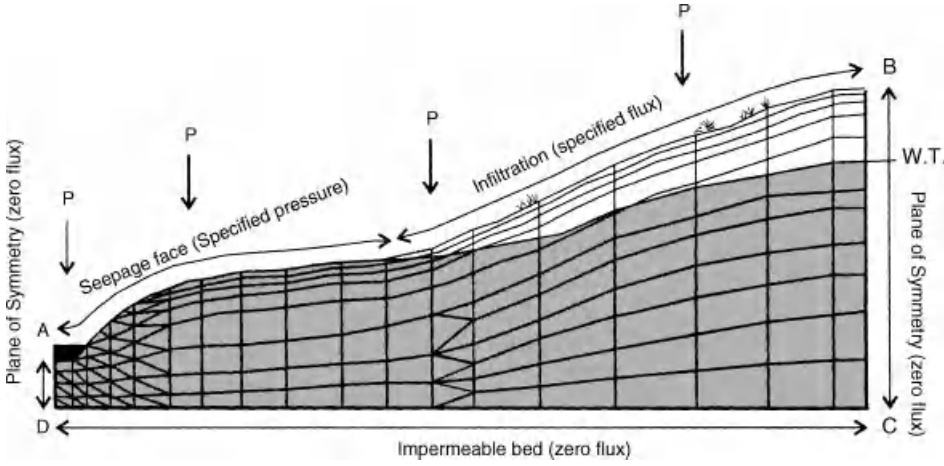


Figure 5.1 *Finite element discretisation of a vertical slice through a hillslope using a mixed grid of triangular and quadrilateral elements with a typical specification of boundary conditions for the flow domain; the shaded area represents the saturated zone which has risen to intersect the soil surface on the lower part of the slope.*

large number of nodes, either on the edges or at the centre of the elements. More details of these solution techniques are given in Box 5.3. Note that it is important to make a distinction between these solutions of the continuum partial differential equations and the semi-distributed models discussed in Chapter 6 where the fluxes between elements are based on the current states of storages in the element. In general, the latter are based on solutions of ordinary differential equations and have often been less careful about solution methodologies (see Clark and Kavetski, 2010).

Distributed models of this type are very demanding in their data requirements. Model parameters must be provided for every grid element in the flow domain and boundary conditions must be specified for every discrete length or area of the domain. Figure 5.1 shows a two-dimensional section through a hillslope with a discretisation into a finite element grid and an indication of the boundary conditions that might be applied. Specified flux boundaries are called "Cauchy-type" boundary conditions; zero flux (impermeable) boundaries are called "Neumann-type" boundary conditions; specified pressure boundaries are called "Dirichlet-type" boundary conditions.

Along the boundaries AD and BC, a symmetry boundary is usually implemented under the assumption that flow conditions are identical for the hillslope at the other side of the divide at A or the channel at D. A symmetry boundary is equivalent to having a no-flow condition along a direction normal to the boundary. Along boundary CD, a no-flow boundary is also generally assumed on the basis that the hillslope is underlain by an impermeable layer or *aquiclude*. The boundary condition along AB may be time variable. While it is raining and the soil surface has not ponded, there will be a specified flow rate equal to the net rainfall rate at the ground surface. If the soil surface reaches saturation and infiltration into the soil starts to fall below the rainfall rate then part of this boundary may be controlled as a fixed head boundary, with a potential equal to the depth of ponded water. Under dry conditions, again there may be a specified flow rate equal to the estimated loss of water from the surface as evaporation. Time-varying solutions also require the initial conditions at the start of the simulation period to be specified. The initial conditions are values of θ or ψ for every node in the flow domain at the start of the simulation.

One major problem in applying the Richards equation is specifying the nonlinear soil moisture characteristics curves for a particular location or solution grid element. Most models of this type use functional relationships between the moisture content, capillary potential and hydraulic conductivity variables

(see Box 5.4). All such relationships are defined by a number of parameter values. Parameter values need to be specified for every element in the mesh. A variety of functional forms for describing the soil moisture characteristics have been proposed. All require a number of different parameters to be specified. Most are "singled-valued" relationships, i.e. each value of θ is associated with a unique value of ψ, $K(\theta)$ and $C(\psi)$. Not all soils show such single-valued relationships, however. The appropriate curves for a soil that is wetting may be different to the appropriate curves for a soil that is drying. This is known as soil moisture hysteresis. The appropriate values of ψ, $K(\theta)$ and $C(\psi)$ then depend on the changes in θ over time. There are some models of hysteretic soil moisture characteristics available (see, for example, the review by Jaynes, 1990), but they tend to rely on idealised representations of the soil to simplify the problem of keeping track of the wetting and drying history of each node. Most models of this type have tended to neglect hysteresis of the soil moisture characteristics in actual applications.

Measuring the soil moisture characteristics, whether in the field or on samples brought back to the laboratory, is time-consuming and expensive and, in heterogeneous soils, the values obtained on one sample may not be representative of the effective grid element values needed in the model. One technique that has been developed to solve this problem is the use of what have been called *pedotransfer functions*, which attempt to provide estimates of the parameter values in mathematical descriptions of the soil moisture characteristics curves in terms of variables, such as soil texture variables, that are more easily measured (see Box 5.5). The idea is fine in principle but may need to be applied with some circumspection in practice since the pedotransfer functions currently available have generally been developed from data obtained from small sample experiments. The parameter values estimated in this way may, then, not necessarily be appropriate at the model grid scale. Most pedotransfer functions are based on regression analysis of soil characteristic parameters against the texture variables. The resulting estimates can therefore be associated with a standard error of estimation as a measure of the uncertainty associated with the estimates.

A further method for deriving the parameter values for soil moisture characteristic functions is to calibrate a model of the functions within a Richards equation solution algorithm, so as to best simulate a set of soil moisture and capillary potential data. Where this method is applied to a laboratory soil column, discharge from the column can also be used in calibration. This is also called the *inverse method*. There is a huge literature on inverse methods for groundwater problems involving only saturated flows (see, for example, Hill and Tiedeman, 2007) and one of the most widely used groundwater flow packages, MODFLOW2000, is now available from the USGS with a parameter optimisation routine as MODFLOWP (Poeter and Hill, 1997). MODFLOW has also been integrated into more general catchment rainfall–runoff models with more conceptual representations of near-surface processes, such as GSFLOW (Markstrom *et al.* 2008) and the Integrated Hydrological Modelling System (IHMS; Ragab and Bromley, 2010; Ragab *et al.*, 2010).

For unsaturated flows, the inverse method has been reviewed by Kool *et al.* (1987) and an application to the determination of hysteretic soil moisture characteristics has been made by Šimůnek *et al.* (1999). In general, the calibration of subsurface flow parameters is not a well-posed inverse problem; there is often inadequate information about the flow domain and estimated parameter values may be sensitive to errors in model structure, boundary conditions and observations. Particularly in the nonlinear case of unsaturated flow, it may be difficult to obtain a clear optimum set of parameter values (e.g. Abeliuk and Wheater, 1990; Hollenbeck and Jensen, 1998). Parameter calibration is discussed more generally in Chapter 7.

A number of computer packages, such as the now commercial HYDRUS2D and HYDRUS3D finite element programs (see Appendix A) for flow and transport calculations, are available for solving the Richards equation in one, two or three dimensions under the assumption that effective parameter values for the Darcy flow law can be specified at the element scale (Šimůnek *et al.*, 1996; Šimůnek and van Genuchten, 2008). These also include models that use a two-domain approach to the prediction of preferential flow e.g. CHAIN-2D (Mohanty *et al.*, 1998); the MACRO model (Jarvis *et al.* 1991); HYDRUS

(Akay *et al.*, 2008); and CATFLOW (Zehe *et al.*, 2005; Klaus and Zehe, 2010). The complexities of obtaining accurate solutions are such that developing solution techniques is best left to the specialist numerical analyst, but the following points are worth remembering in assessing any package:

- All the solution techniques for this type of nonlinear problem are approximate (see Box 5.3) and, because of the nonlinearities involved, it is difficult to generalise about whether one method will be more accurate than another for a particular problem.
- For any solution algorithm that is consistent with the differential equation, accuracy will depend on the time and space discretisation used. The finer the space increments (or the smaller the elements used), then the shorter the time steps will have to be to ensure stability of the solution.
- It may be necessary to use a very large number of nodes to represent the flow domain, especially in three dimensions. This will then require the solution of a very large sparse matrix equation at least once at each time step, together with the calculation of the nonlinear functions at each node. The computer time required mounts rapidly with the number of nodes.
- Problems where steep hydraulic gradients are expected, such as a wetting front during infiltration or around a pumped well, will require small elements (and consequently small time steps) to represent the gradient and flow velocities adequately in that part of the flow domain. This may seem obvious but has not always been evident in published applications of distributed models.
- Solutions with soil moisture θ as the dependent variable tend to be better for dry soil conditions; solutions with capillary potential ψ as the dependent variable tend to be better for wet soil conditions.
- Some solutions, particularly those known as *explicit* solutions in which the solution at time step t depends only on values of the nonlinear functions calculated at time step $t - 1$ (see Box 5.3), produce unstable solutions if the time step is too big. Instability in the solution will normally be seen as increasingly large oscillations in the solutions at some nodes. A good implementation of an explicit solution scheme should check for stability, albeit at the expense of additional calculations, and adjust the time step accordingly.
- *Implicit solutions*, which use variable and function values at both t and $t - 1$ (see Box 5.3), are generally more stable and can use much longer time steps, but may involve a number of iterations at each time step to converge on the solution at time t.
- There are some stability problems inherent in the solution of the Richards equation, even using an implicit time stepping scheme. This is because the nonlinear soil moisture capacity function ($C(\psi)$ in Equation (5.1)) peaks at a certain value of ψ. Thus it is possible for the solution at a node to oscillate either side of that value of ψ, while still calculating an appropriate value of $C(\psi)$.
- In heterogeneous flow domains, the values of the parameters required at the grid or element scale may be dependent on the scale of the elements. A Darcian description of the flow using effective parameter values at the model element scale is not necessarily an adequate representation of the flow processes. The effects of spatial heterogeneity of soil characteristics, time variability due to crusting and other processes and preferential flow in structured soils remain topics of research with no generally accepted descriptive models.
- Always test the model, using different time and space steps, against some simple test cases. There is no absolute guarantee that an approximate solution of a nonlinear partial differential equation will be stable and accurate under all circumstances. Validation against some test cases will at best (but also at least) give a guide.

5.1.2 Surface Runoff and Channel Routing

The physical basis of models of overland and channel flow is essentially the same. In both cases, in catchment scale rainfall–runoff modelling, one-dimensional flows downslope or downstream are usually

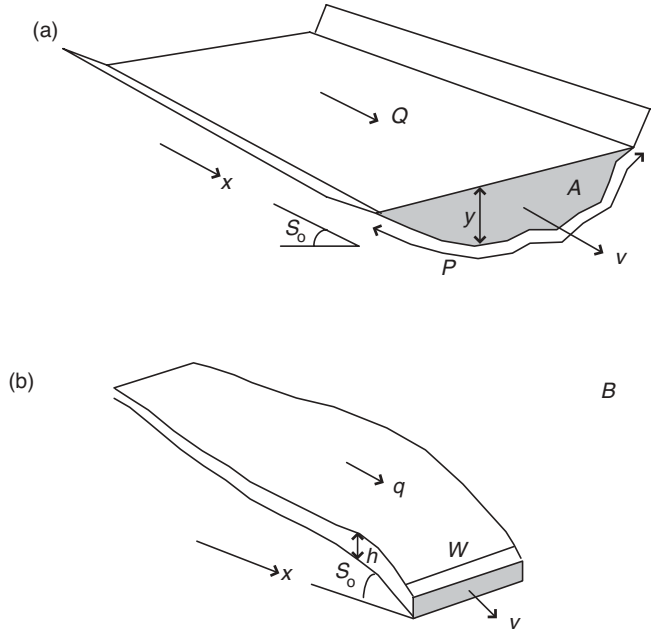


Figure 5.2 *Schematic diagram for surface flows with slope S_o and distance x measured along the slope: (a) one-dimensional representation of open channel flow with discharge Q, cross-sectional area A, wetted perimeter P, average velocity v and average depth y; (b) one-dimensional representation of overland flow as a sheet flow with specific discharge q, width W, average velocity v and average depth h.*

considered as a convenient approximation to the full three-dimensional flows. One-dimensional solutions must use average cross-sectional velocities as solution variables, even for the case of overbank flow for the channel and variable depth overland flows (Figure 5.2). The one-dimensional case may be described by the equations developed by the Barré de St. Venant (1797–1886). These equations assume that the flow may be described in terms of average cross-sectional velocities and depths and are developed from the balances of both mass and momentum in the flow. Thus for a flow of average velocity v, of average depth h, in a cross-sectional area, A, and wetted perimeter P, on a bed slope of S_o, and lateral inflow per unit length of slope or channel of q, the mass balance equation may be written as:

$$\frac{\partial A}{\partial t} = -A\frac{\partial v}{\partial x} - v\frac{\partial A}{\partial x} + q \tag{5.4}$$

and the momentum balance equation, assuming that water is incompressible, as:

$$\frac{\partial Av}{\partial t} + \frac{\partial Av^2}{\partial x} + \frac{\partial Agh}{\partial x} = gAS_o - gP\frac{f}{2g}v^2 \tag{5.5}$$

where f is the Darcy–Weisbach uniform roughness coefficient.

The St. Venant equations are a fully dynamic or dynamic wave description of the flow that can be used in routing flood waves and hydrographs down a channel or in a reach of a channel network. The derivation of these equations is given in Box 5.6, together with explanations of simplified versions known as the diffusion wave and kinematic wave approximations, resulting from neglecting different terms in Equation (5.5). As in the case of the Richards equation, solution of the St. Venant equations for cases

of practical interest generally requires an approximate numerical solution algorithm. The first attempts at an explicit finite difference solution of the St. Venant equations date back to Stoker (1957). There are now well-established finite difference schemes such as the four-point implicit method described by Fread (1973; see Box 5.6), that has been used even under the extreme conditions of routing the flood wave due to a dam break (e.g. Fread, 1985). Routing for rainfall–runoff modelling is generally not so extreme.

The other requirements to apply such a model are information about the geometry of the channel, a specification of the initial velocities and depths of flow at the start of a simulation and boundary conditions at both the upstream and downstream boundaries of a channel reach. Geometry, roughness coefficients, and initial and boundary conditions are only ever known imprecisely and some simplifying assumptions are usually necessary.

In respect of the channel geometry, it has usually been assumed that the shape of the channel can be interpolated between surveyed cross-sectional profiles measured at different distances along the channel. At low flows, this will not give a good representation of the effects of the pool/riffle geometry of the channel; at high overbank flows, it may not take proper account of the effects of embankments, field boundaries and other obstructions to the flow. These will also affect the appropriate values of the effective roughness parameter that must reflect all the causes of momentum loss in a reach of channel. Thus, the effective values might be different from that inferred from a measured velocity profile at any single point in the channel and may also require a parametric description of how the roughness coefficient changes with depth of flow, especially for overbank flow under flood conditions (see, for example, Knight *et al.*, 2010).

The boundary conditions will have an important effect on the solution. The St. Venant equations require boundary conditions to be specified for both upstream and downstream boundaries (in contrast to the kinematic wave approximation discussed below). In fact, since there are two unknowns in the solution, velocity and depth for each cross-section, two upstream boundary conditions and two downstream boundary conditions are required for every simulated reach. Junctions between reaches require some special conditions to ensure consistency in the solutions for the upstream and downstream reaches. It is, in fact, rare for the boundary conditions to be specified directly in terms of velocity and depth at each boundary: they are not generally available. Water surface elevation or stream stage is more generally available, at least at the gauging sites that often mark the boundary points for a solution in a larger river. A measurement of stage can be used with an appropriate rating curve and cross-sectional survey to get approximate estimates of discharge and cross-sectional area, from which a mean velocity can be derived. The rating curve may be measured, may be a theoretical rating of a gauging structure or may be derived by assuming that there is a locally uniform flow at the boundary. In the last case, the relationship between velocity and stage can be described by one of the uniform flow equations such as the Manning or Darcy–Weisbach equations, given knowledge of the appropriate roughness coefficient.

A little bit of care is necessary here however. Use of such a uniform flow rating curve implies that the water surface is always parallel to the bed. The fully dynamic equations, however, imply that the water surface should be steeper than the bed slope on the rising limb of a hydrograph and less steep on the falling limb, resulting in a hysteretic or looped rating curve (see, for example, Dottori *et al.* 2009). The uniform flow assumption can, then, only provide an approximate boundary condition for the solution. Again, it is important to assess what assumptions are being made in the description and solution for each process. It is also the case that, even where a measured rating curve is available, the observations on which it is based should be checked as they can show nonstationarities over time, can show a wide scatter at some stations and may not extend to the highest flood flows of interest (for good logistical and health and safety reasons!).

It must be remembered that this description of channel flow is a one-dimensional description, with the solution variables being average velocities and depths of flow. This type of description will not be as accurate during flood conditions, when local roughness coefficients, velocities and depths of flow may vary dramatically in the cross-section. Recent developments have seen two-dimensional models

coming into more widespread use. Such models can predict a pattern of depth-averaged velocities within the channel and across a flood plain. Examples of such models are TELEMAC_2D, SFV and RMA2 (see, for example, the work of Bates *et al.*, 1992, 1995, and Horritt *et al.*, 2007). Three-dimensional models of surface flows using general computational fluid dynamics packages are just starting to be used to study flows in natural channels (see the reviews by Lane, 1998, and Lane *et al.*, 1999) but are still computationally feasible only for small scale problems. Even then, much remains to be learned about appropriate representations of turbulence and momentum losses in natural channels for such models. Some comparisons of 2D models have been provided by Horritt and Bates (2001), Horritt *et al.* (2007) and Hunter *et al.* (2008).

The increasing availability of detailed topographic data from remote sensing (Lidar, SAR) has also driven demand for two-dimensional models, not only for fluvial flooding, but also for detailed mapping of the potential for pluvial flooding in urban areas, and for rapid flood spreading following breaching or over-topping of flood defences. A number of computationally efficient models have been produced, based on combining a 1D representation of the channel with a 2D diffusion simplification of the depth averaged equations on the flood plain (see Box 5.6 for a discussion of simplifications of the St. Venant equations). These models include the LISFLOOD-FP model of Bates and De Roo (2000) and JFLOW model of Bradbrook (2006). To speed the calculations further, parallel computing versions of such models have been produced (e.g. Neal *et al.*, 2010), including implementation of the JFLOW gridded model on multiple graphics processing units (GPUs) (Lamb *et al.* 2009). The advantage of GPUs is that they can have several hundred individual processors that can each carry out the relatively simple calculations for the grid elements in the flow domain with near linear speed up of the calculations.

There are a number of numerical issues with these simplified models, including stability issues of using explicit time stepping, numerical diffusion, mass conservation, and the treatment of wetting and drying cells at the edges of the flood plain. Recent advances to control these essentially numerical problems have included the implementation of adaptive time stepping schemes (Hunter *et al.* 2006) and the re-introduction of inertial terms (Bates *et al.*, 2010). However, a number of comparative studies have suggested that, given all the uncertainties involved in setting up a flood routing model, the simplified models can reproduce the observed patterns of inundation in past floods as well as the more complex finite element and finite volume solutions (Horritt and Bates, 2002; Horritt *et al.*, 2007). The advantage of the faster run times of these simplified models is that they can be run over large areas (such as the Amazon studies of Wilson *et al.*, 2007) or many times in a given application for use in model calibration or uncertainty estimation (Aronica *et al.* 2002; Di Baldassarre *et al.* 2009; Leedal *et al.* 2010). Speed of computation is also an issue if such distributed models might be used for real-time forecasting of inundation (e.g. Schumann *et al.*, 2010).

Calibration is an important issue in the application of these flow routing models. The parameters that are required are the patterns of **effective** roughness for the channel and flood plain. It is "effective" roughness because the values required to produce good simulations of the available observations depend on which model is being used, the spatial resolution of the discretisation of the flow domain, the way in which roughness elements on the flood plain (including walls, hedges, trees and buildings) dissipate momentum, and the accuracy of the available survey information. The data available for calibration is the change in levels at gauging stations and, in some cases, patterns of maximum inundation elevations from post-event surveys of wrack marks and marks on buildings (but without any accurate information on timing). Some studies have used patterns of inundation at specific times when satellite or aircraft platform sensing images have been available. Some of those images give rather uncertain estimates of the extent of inundation because of speckle and the effects of emergent vegetation (e.g. Horritt *et al.*, 2003; Di Baldassarre *et al.*, 2009) and there are often important issues about registration of the images with the representation of flood plain topography that is used in the model (e.g. Romanowicz and Beven, 2003). Such uncertainties interact with other uncertainties in setting up the model, such as the treatment of the boundary conditions (in particular, the estimates of upstream discharge), the representation of

the geometry (in particular, the channel geometry which depends on expensive survey information) and the way in which infrastructure, such as bridges, are handled (e.g. Pappenberger *et al.*, 2006b). In the best cases, distributed inundation models can give predictions with similar root mean square error to the observations (0.4 m in the case of the Carlisle flood simulations reported by Horritt *et al.*, 2010).

In common with other distributed models, there is the possibility of specifying different values of the roughness values for every element in the discretisation of the domain, but there may not be the information available to allow this. There have been recent attempts to combine information about land use and Lidar signals to make estimates of spatial patterns of roughness (e.g. Cobby *et al.* 2003). A distributed sensitivity analysis can help indicate where the model might be most sensitive to changes in roughness (e.g. Hall *et al.*, 2005) but it is generally the case that, in practical applications, the roughness parameters are assumed to be constant over large parts of the flow domain. It may also be the case that a model cannot reproduce the observed inundation everywhere. In such cases it has been suggested that different ensembles of models might be used for making predictions for different purposes (Pappenberger *et al.*, 2007a).

As in the case of subsurface flow solutions, there are a number of points that should be borne in mind when evaluating a hydraulic model of the surface flow processes:

- All numerical solutions of the flow equations are approximate and may be subject to numerical diffusion. Again a large number of nodes may be necessary to represent the flow domain and explicit solutions,. in particular, may require very short time steps.
- An important control on accuracy in representing the real flow processes is the specification of the geometry of the flow domain. Surveys are expensive and compromises are generally necessary in applications.
- The specification of boundary conditions and roughness coefficients is also important. Effective roughness coefficients may be dependent on flow depth and may have to take account of momentum losses associated with obstructions (such as trees, walls and hedges) as well as the surface roughness of the bed and banks.
- You should always test the model, using different time and space steps, against some simple test cases to evaluate the convergence and stability properties of the solution algorithm.

5.1.3 Interception, Evapotranspiration and Snowmelt

Any physically based catchment model also requires components for interception, evapotranspiration and snowmelt. In conjunction with, primarily, the subsurface flow component, they control the simulation of the antecedent conditions prior to an event, and the inputs during an event, that are important in predicting the runoff from that event. A typical set of components, as in the SHE model discussed in Section 5.2, uses a Penman–Monteith actual evapotranspiration calculation (see Box 3.1), an interception storage model such as the Rutter model (see Box 3.2) and either a full energy balance or degree-day snowmelt model (see Box 3.3).

5.2 Physically Based Rainfall–Runoff Models at the Catchment Scale

5.2.1 Coupling the Surface and Subsurface Process Descriptions: Towards a Fully Three-Dimensional Description

The distributed model defined by Freeze and Harlan (1969) was a fully three-dimensional, saturated–unsaturated subsurface flow description coupled to a two-dimensional overland flow description and a

one-dimensional channel flow description. The coupling of the different process descriptions can be achieved through common boundary conditions. For example, the depth of ponding of water on the soil surface predicted by an overland flow solution can be used to define a local head boundary for the subsurface flow solution in simulating infiltration rates. Similarly, the depth of flow predicted in the channel might provide a local head boundary condition for the prediction of fluxes from the saturated zone through the bed of the channel. In principle, therefore, the whole system of processes could be solved in one system of equations, taking proper account of all the common boundary conditions. In practice, to apply such a description at the scale of a catchment, or even at the scale of a hillslope, requires prodigious amounts of computer time, even with today's computing power. Most distributed models have therefore attempted to reduce the amount of computing power in some way although fully three-dimensional solutions are now available (such as the HYDRUS3D, MODFLOW, InHM and TOUGH2 models mentioned earlier).

A number of different strategies have been used to reduce the computational burden. The first is to use a coarser mesh, so that there are fewer nodes, a smaller number of equations must be solved at each time step and fewer parameters need to be specified. There is clearly then a danger of having a model that is not an accurate solution to the original equations. This is a very real danger; it applies to most of the distributed models that have been used in representing the rainfall–runoff process at the catchment scale to date.

A second strategy has been to reduce the dimensionality of the problem, i.e. to break it down into smaller pieces. One way to do this has been to treat the unsaturated zone, where flows are predominantly vertical, as a one-dimensional problem and the saturated zone, where flows are predominantly lateral, as a two-dimensional problem. This is the approach adopted by the SHE model (Figure 5.3; see, for example, the work of Abbott *et al.*, 1986a) and some models based on a triangular irregular network

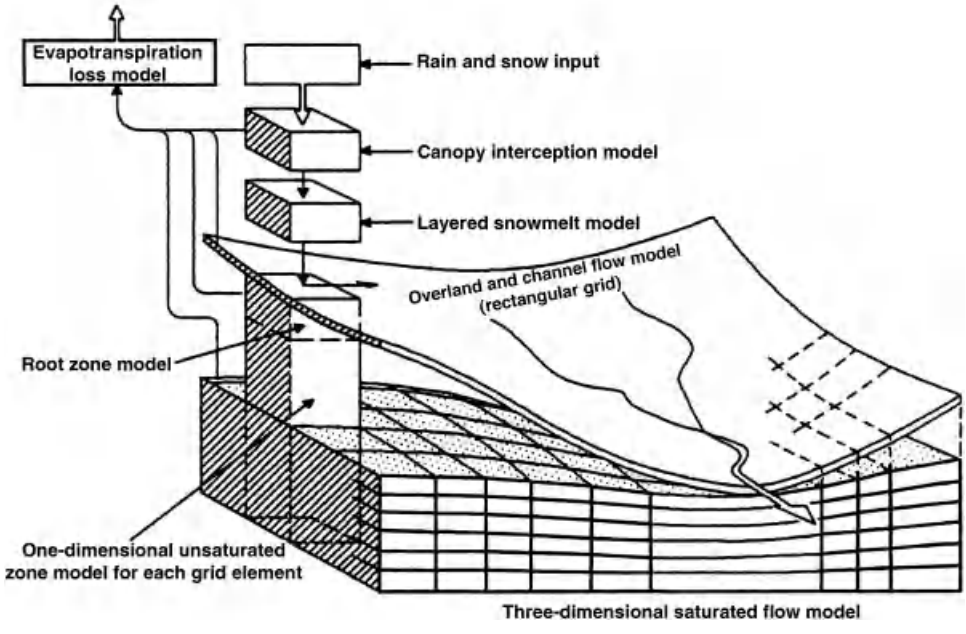


Figure 5.3 *Schematic diagram of a grid-based catchment discretisation as in the SHE model (after Refsgaard and Storm, 1995, with kind permission from Water Resource Publications).*

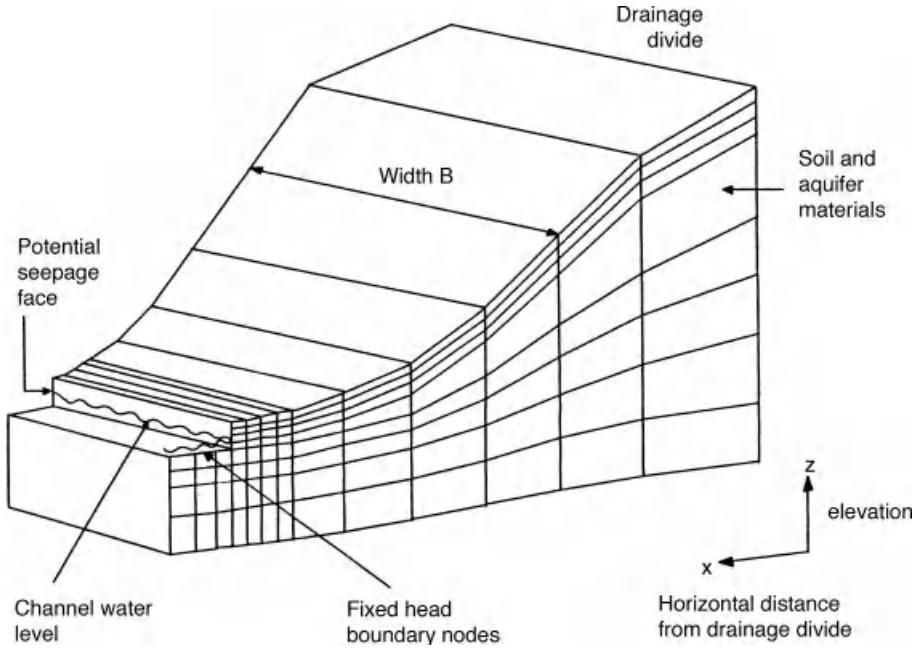


Figure 5.4 *Schematic diagram of a hillslope plane catchment discretisation as in the IHDM model (after Calver and Wood, 1995, with kind permission from Water Resource Publications).*

(TIN) discretisation of the topography such as tRIBS (Ivanov *et al.*, 2004, 2008) and PIHM (Qu and Duffy, 2007). This can give rise to numerical problems in coupling the solutions at a boundary which moves up and down with the water table as the soil wets and dries. The solution of the saturated zone problem depends on the profile of water content in the unsaturated zone of each grid element and vice versa. Normally some iterations are required to achieve convergence of the two solutions. Similar iterations may be required to achieve convergence at nodes on the soil surface, where the boundary may be changing from infiltration to ponded conditions during a storm.

An alternative strategy is to avoid decoupling the unsaturated and saturated zones and instead make the split along lines of greatest slope in the catchment to make a number of hillslope planes that are then solved separately "in parallel". The vertical section along each plane is then discretised in two dimensions, assuming that conditions across each plane can be considered uniform. This is the approach adopted by the IHDM model (Figure 5.4; see Calver and Wood, 1995) and, more recently, the CATFLOW model of Zehe *et al.* (2005). Then there are models, such as TOPOG (Vertessy *et al.*, 1993), that use variable width hillslope planes but separate the unsaturated zone and saturated zone; VSAS2 uses variable width hillslope planes but has a time variable separation of a saturated contributing area to ease the numerical problems of solving the Richards equation when part of the flow domain is fully saturated (Bernier, 1985; Prevost *et al.*, 1990; Davie, 1996); and the model of Duffy (1996), which also uses hillslope planes but solves for the moisture storage at each point, integrated over the profile of both saturated and unsaturated zones.

However, the computer power available to hydrological modellers continues to increase, and in the next decade this will result in more and more use of three-dimensional solutions of the continuum equations.

Table 5.1 *Minimal parameters required for a catchment scale, process-based model*

Parameter	Symbol	SI Units
Subsurface Flow Parameters (for each soil type/horizon element)		
Saturated hydraulic conductivity matrix	$\underline{K}_s$	ms^{-1}
Porosity	θ_s	–
Soil moisture characteristic parameters (see Box 5.4)		
Vegetation Parameters (for each vegetation type)		
Proportion direct throughfall	p	–
Interception storage capacity	C	m
Interception drainage parameter		–
Aerodynamic resistance (may vary with wind speed	r_a	sm^{-1}
Canopy resistance (may vary with other variables)	r_c	sm^{-1}
Albedo	α	–
Proportion of active roots to distribute $E_T(x, y, z, t)$ over elements		
Overland Flow Parameters (for each slope element)		
Overland flow roughness (may vary with flow depth)	f	–
Local surface slope angle	S	–
Channel Flow Parameters (for each reach element)		
Channel flow roughness (may vary with flow depth)	f	–
Overbank flow roughness (may vary with flow depth)	f	–
Local channel bedslope	S_o	–
Snow Parameters (degree-day model)		
Threshold temperature	T_o	K
Degree-day factor (may vary in time)	F	mm/day/K

This poses a question as to whether that is the best way of using the available computer resource given the uncertainties in boundary conditions, model parameters and, in fact, the continuum equation process representations themselves. This discussion is considered in Chapter 9 when we discuss the next generation of rainfall–runoff models. For the moment, we investigate the issues of defining parameter values for this type of distributed model.

Specifying all the parameters is still, however, a problem. The finer the elements, the more parameter values that must be specified. A minimal list of the parameters that might be required for a full catchment scale model is given in Table 5.1. Note that many of these parameters, while often assumed as constant for a particular model run, may in fact be dependent on other variables. Canopy resistance, for example, may require a more fundamental parameterisation to account for its variation with soil moisture, solar radiation and surface temperature (see Box 3.1); interception storage may vary with crop growth; channel flow resistance may vary with flow depth; a degree-day factor may increase during the melt season (Box 3.3). All these dependencies would need to be specified in a complete model of the processes; since a solution with thousands of elements requires many thousands of parameter values, this is only feasible by linking such models to databases for the preparation and storage of parameter values, effectively a geographical information system (GIS).

The problem is then that the normal GIS overlays of soil type, vegetation type, land use, and geology do not directly provide the information required to estimate the parameters required to run a model, especially when, as noted earlier, the **effective** values of the parameters might need to reflect the heterogeneity of the characteristics and the particular model implementation (grid element size, numerical diffusion, etc.). Thus some intepretative model is required (such as the pedotransfer functions of Box 5.5). A number of

papers have suggested that the lack of knowledge of soil heterogeneity is a fundamental constraint on the accuracy of distributed models (e.g. Beven, 2001; Hansen *et al.*, 2007; Refsgaard *et al.*, 2010) although some numerical studies have also suggested that this may not be such an issue in predicting catchment scale responses (Canfield and Goodrich, 2006).

Similar software is required for post-processing of the results. The finer the elements the more data that are produced by each simulation. The only way to assess such information easily is visually, in the form of computer graphics, and this form of distributed model has become more and more sophisticated in interacting with geographical information systems, both in setting up the model and in presenting the results (see Maidment, 2002; Refsgaard *et al.*, 2010). This is even more the case if it is necessary to convey the uncertainities in the spatial predictions of such models (e.g. Leedal *et al.*, 2010; see Figure 3.7).

5.2.2 Models Based on Grid Elements: The SHE Model

The Système Hydrologique Européen (SHE) model is the most widely known rainfall-runoff model based on grid elements. It was started in 1977 as a joint collaboration, between the UK Institute of Hydrology, the Danish Hydraulics Institute (DHI) and SOGREAH of Grenoble in France. An early description of the model was published by Beven *et al.* (1980), an explanation of the modelling philosophy was provided by Abbott *et al.* (1986a, 1986b), and the first full application (to the Institute of Hydrology River Wye experimental catchments at Plynlimon, Wales (10 km^2)) was published in a series of articles by Bathurst (1986a, 1986b). Other applications have been published, ranging from the 1.4 km^2 Rimbaud catchment in the south of France (Parkin *et al.*, 1996) to the 820 km^2 Kolar and 4955 km^2 Narmada catchments in India (Refsgaard *et al.*, 1992; Jain *et al.*, 1992). A summary of the history and application of SHE, particularly the DHI commercial version (MIKE SHE) is given by Refsgaard *et al.* (2010), who note that there are now nearly 400 installed copies of the MIKE SHE software around the world.

SHE is a grid-based model that splits the catchment into a number of square or rectangular grid elements, linked to channel reaches that run along the boundaries of the hillslope grid. The size of grid used has varied in different applications, ranging from 50 m on a side for the small 40 ha Upper Sheep Creek catchment in Idaho up to 2 km on a side for the Kolar and Narmada catchments in India. Note that in the latter case the grid size is so large that the model cannot be considered to be representing flow on the hillslopes or in the smaller channels of the catchment in any meaningful way.

Each hillslope grid element (Figure 5.3) has a specified surface elevation and model components for interception, evapotranspiration, snowmelt, and one-dimensional vertical unsaturated zone flow where appropriate. The grid elements are linked by two-dimensional surface runoff and groundwater components. Internal boundary conditions allow the coupling of surface flow and infiltration into the unsaturated zone, the unsaturated and saturated zones at the local water table, and groundwater and channel flows. Great effort has been made to ensure that the processes are properly coupled and that the numerical solutions are stable for a wide range of conditions. The model can predict a variety of runoff generation processes on each grid element, including both infiltration excess and saturation excess runoff, and the groundwater flow component can be used to simulate subsurface contributions to the hydrograph under suitable conditions. The description of the unsaturated and saturated zones are based on Darcy's law; overland and channel flows are described by a diffusion wave approximation to the St. Venant equations and various options are included for simulating interception and evapotranspiration, including the Penman–Monteith equation (see Box 3.1). Snowmelt is simulated, either using a degree-day method or a full energy balance (Bathurst and Cooley (1996) make a comparison of both implementations).

The types of parameter values required are similar to those listed in Table 5.1 and there is the potential to have different parameters for every grid element and within each grid element for different layers in

the vertical. Any application of the SHE model will therefore require the specification of thousands of parameter values. The parameter values required are effective values at the grid element scale, which may not be the same as values that might be measured locally. There is also the potential to specify fully distributed precipitation and meteorological data across the model grid elements, if the data are available. The predictions are, however, dependent on the grid scale used. A study by Refsgaard (1997), using the SHE model, is one of the few studies to have looked at the effect of the grid scale on the model predictions. This study, of the Karup catchment in Denmark, compared predictions using a finest grid of 500 m, to those with degraded grids of 1000 m, 2000 m and 4000 m. His conclusion was that it might still be possible to obtain reasonable simulations of catchment discharge above 1000 m but that it would require recalibration of parameters and possibly reformulation of some model components. Refsgaard infers that not much improvement in accuracy would be gained by using finer scale grids than 500 m but this conclusion may be conditional on the nature of the Karup catchment which is dominated by groundwater flows. Xevi *et al.* (1997) and Vázquez *et al.* (2002) have also demonstrated that the results of the SHE model are sensitive to grid size. It follows then that the effective parameter values to get good predictions of the variables of interest, such as catchment discharge, should also be expected to vary with grid size. The same effect should also be expected with other model fomulations.

The different SHE development teams have implemented impressive pre- and post-processing packages for preparing model applications and visualising the distributed predictions, including graphical animations of the predicted responses. The distributed predictions of the SHE model have also allowed other model components to be developed within the most recent versions, which are now being developed independently by the original partners. The UK version, SHETRAN, now based within the Water Resource Systems Research Unit at the University of Newcastle, has added contaminant and sediment transport components (Bathurst *et al.*, 1995, 2004; Ewen *et al.* 2000). The DHI version, MIKE SHE, has also added a contaminant transport component (Refsgaard and Storm, 1995). In both cases, the predictions of contaminant transport are based on the advection–dispersion equation. Both DHI and the University of Newcastle now have versions of SHE which make fully 3-D solutions for the unsaturated–saturated flow domain. MIKE SHE has also added options to use a simple groundwater store where a fully susubsurface solution is not justified and to predict a preferential recharge to the saturated zone as a simple proportion of the infiltration rate (Refsgaard and Storm, 1995). Such modifications undermine the way in which models purport to be "physically based".

There have been other models based on grid elements, now largely superseded by the more general modelling packages such as InHM, tRIBS and HydroGeoSphere. The fully 3-D models of Binley *et al.* (1989a, 1989b) and Paniconi and Wood (1993) use a grid-based spatial discretisation. The ANSWERS model (see for example, Beasley *et al.*, 1980; Silburn and Connolly, 1995; Connolly *et al.*, 1997), which has its origins in one of the very first fully distributed grid-based models of Huggins and Monke (1968), essentially considers only an infiltration excess runoff generation mechanism, using the Green–Ampt infiltration equation (see Box 5.2) to predict excess rainfall on each grid element. The runoff generated is then routed towards the stream channel in the direction of steepest descent from each grid element. The CASC2D model of Doe *et al.* (1996) and Downer *et al.* (2002) is similar in that it also uses a Green–Ampt infiltration equation, but it uses a 2-D diffusion wave approximation to model overland flow on the hillslopes and a 1-D diffusion wave model for the channel reaches. CASC2D was later extended to include more subsurface flow processes as the Gridded Surface/Subsurface Hydrologic Analysis (GSSHA) model (Downer and Ogden, 2004; Downer *et al.*, 2005). The 3-D version of HILLFLOW of Bronstert and Plate (1997) is a grid-based model, with the interesting option of modelling the Richards equation using the *fuzzy logic* methodology of Bárdossy *et al.* (1995). HILLFLOW also has a 2-D option for modelling individual hillslope elements in a way similar to the models discussed in Section 5.2.3 and a 1-D version for individual soil profiles. All the HILLFLOW versions have a component for modelling preferential flow in macropores, at the expense of introducing additional parameters. Bronstert (1999) provides a review of experience in using HILLFLOW in a variety of applications. Models that have been linked to

the MODFLOW groundwater code also use a gridded discretisation (e.g. IHMS (Ragab and Bromley, 2010) and GSFLOW Markstrom *et al.*, 2008).

5.2.3 Models Based on Hillslope Elements

The main alternative catchment discretisation strategy is to make a subdivision into hillslope planes (Figure 5.4). Such a subdivision is ideally made along flow lines, such that any lateral exchanges of water between adjacent hillslope elements can be neglected. Some early physically based distributed models which attempted solutions only for a single hillslope were essentially of this type (e.g. Freeze (1972) used a finite difference solution and Beven (1977) used a finite element solution). It is, of course, much easier to determine flow lines if the flow follows the form of the surface topography. Hillslope elements may then be determined on the basis of a topographic analysis of the catchment. Thus, this type of model works best where the hydrological activity is concentrated near to the soil surface and no deeper regional aquifer flows are involved. For deeper systems, a two-dimensional in plan (as in SHE) or fully three-dimensional solution of the subsurface flow domain is more appropriate.

However, there are many catchments for which the hillslope element discretisation based on surface topography provides a reasonable approximation to the flow directions. In some early catchment models of this type, the variable width, variable depth and variable slope hillslope elements were represented by "equivalent" planes of uniform width, uniform depth and uniform slope (and usually uniform soil and surface parameters). Early versions of the Institute of Hydrology Distributed Model (IHDM) were of this type, as well as some models based on Hortonian infiltration excess runoff generation that did not include a full subsurface flow solution, but treated infiltration as a "loss" (e.g. the model of Smith and Woolhiser (1971) that later developed into the KINEROS package described by Smith *et al.* (1995) – see also Section 5.5.2).

The model of Beven (1977) showed that, using finite elements, it was relatively easy to follow the actual shape of the hillslope and allow the depths of different horizons within a hillslope to vary (as in the vertical plane discretisation of Figure 5.1). This study also introduced the simple idea of including slope width in the equations so that convergent and divergent hillslopes could be represented (see also Box 5.7). This was also introduced into Version 4 of the IHDM (Beven *et al.*, 1987) and there have been further numerical improvements since (Calver and Wood, 1995). Applications are reported by Calver (1988) and Binley *et al.* (1991) for the Wye catchment at Plynlimon in Wales and by Calver and Cammeraat (1993) for an experimental hillslope in Luxembourg.

As a result of this form of discretisation, there is an implicit assumption that the soil and surface parameters are considered to be constant across the width of the hillslope (in the same way as the SHE model requires effective values for each grid element). Variations in parameter values between different soil horizons or for individual elements in a discretisation, such as that of Figure 5.1, can be represented but they must be effective values integrating over any heterogeneity across the slope. Thus, it may be difficult to measure such values in the field and Calver and Wood (1995), for example, report that their experience in using the model is that measured values of hydraulic conductivity tend to underestimate the values required to represent fast subsurface stormflow in the model.

In Australia, two similar models, THALES and TOPOG have been developed based on the TAPES-C topographic analysis package that identifies one-dimensional downslope sequences of hillslope elements from contour data without any intervening interpolation onto a raster elevation grid (see, for example, Figure 3.5). Both models use a kinematic wave approximation of downslope flows in the saturated zone and are described in more detail in Section 5.5.3. The CATFLOW model also uses hillslope segments but introduces the simulation of preferential flows in a dual porosity system (Zehe *et al.*, 2005; Klaus and Zehe, 2010).

5.3 Case Study: Modelling Flow Processes at Reynolds Creek, Idaho

Reynolds Creek is a 234 km^2 rangeland catchment in the Owyhee Mountains of Idaho, managed by the USDA North West Watershed Research Centre. It was the site of one of the very first attempts to evaluate the predictions of a distributed process based hydrological model. Stephenson and Freeze (1974) used a two-dimensional, finite difference, partially saturated Darcian subsurface flow model in an application to a vertical slice through a complex hillslope within the Reynolds Creek catchment. Model predictions were checked against field measurements made during a snowmelt season. After making initial estimates of parameter values for the soil and rock layers in the slope, they carried out a trial and error calibration of the model, adjusting the parameter values to try to improve the fit to the observations. At that time, computer constraints severely limited the number of calibration runs that could be made. The results of the best simulation are shown in Figure 5.5.

This study is interesting, even after 35 years, because it is one of the first studies to recognise that there might be limitations to the application and validation of this type of model at particular sites. Stephenson and Freeze conclude (1974, p. 293) that:

> we recognise that our calibration is less than perfect but it is probably representative of what can be attained when a fully deterministic mathematical model is applied to a field site with a fairly complete, but as always imperfect, set of field measurements.

They also noted that validating such models was a particularly difficult problem since it presupposes perfect knowledge of all the boundary conditions, parameter values and initial conditions required. Imperfect knowledge always introduces a degree of flexibility (or uncertainty) to any attempt at validation of the model.

More recently, the SHE model has been applied to the 40.4 ha Upper Sheep Creek subcatchment by Bathurst and Cooley (1996). At this site, at an elevation of over 2000 m in the headwaters of Reynolds Creek, average annual precipitation is of the order of 1016 mm, with more than 70% falling as snow. Snow accumulation is highly variable, with a deeper pack building up in the lee of a ridge where the pack may reach depths of more than 5 m.

Bathurst and Cooley simulated a single snowmelt period using both energy budget and degree-day snowmelt models within the SHE model framework. A model discretisation based on 161 square grid cells (50 m by 50 m) was used. All parameters for the model were specified on the basis of knowledge of the catchment soils and vegetation. Initial snowpack characteristics were assigned on the basis of snow course information and photographs; initial saturated zone thickness was assigned to reproduce the initial flow at the start of the simulation. It is not clear from the paper how the initial unsaturated soil moisture profiles were defined and only a 12-hour period was allowed for the model to "run in" before the predictions were compared with observations.

The study aimed to test four hypotheses about the processes in the catchment based on how well the model reproduces the stream discharge during the simulation period. These hypothesis varied in assumptions about the extent of frozen soil and depth of the effective impermeable layer. Unlike Stephenson and Freeze (1974), Bathurst and Cooley made no attempt to validate the model predictions against internal state measurements. The authors state (1996, p. 194) that

> the traditional calibration approach of adjusting the parameter values (within each hypothesis) to improve the agreement played a secondary role and was carried out under the constraint that the [parameter] values must reflect the field measurements where these exist or should otherwise lie within physically realistic limits.

The earliest run number reported in the paper is 69, the last 107. It is clear that this application, like that of Stephenson and Freeze 20 years earlier, was limited by computer run times.

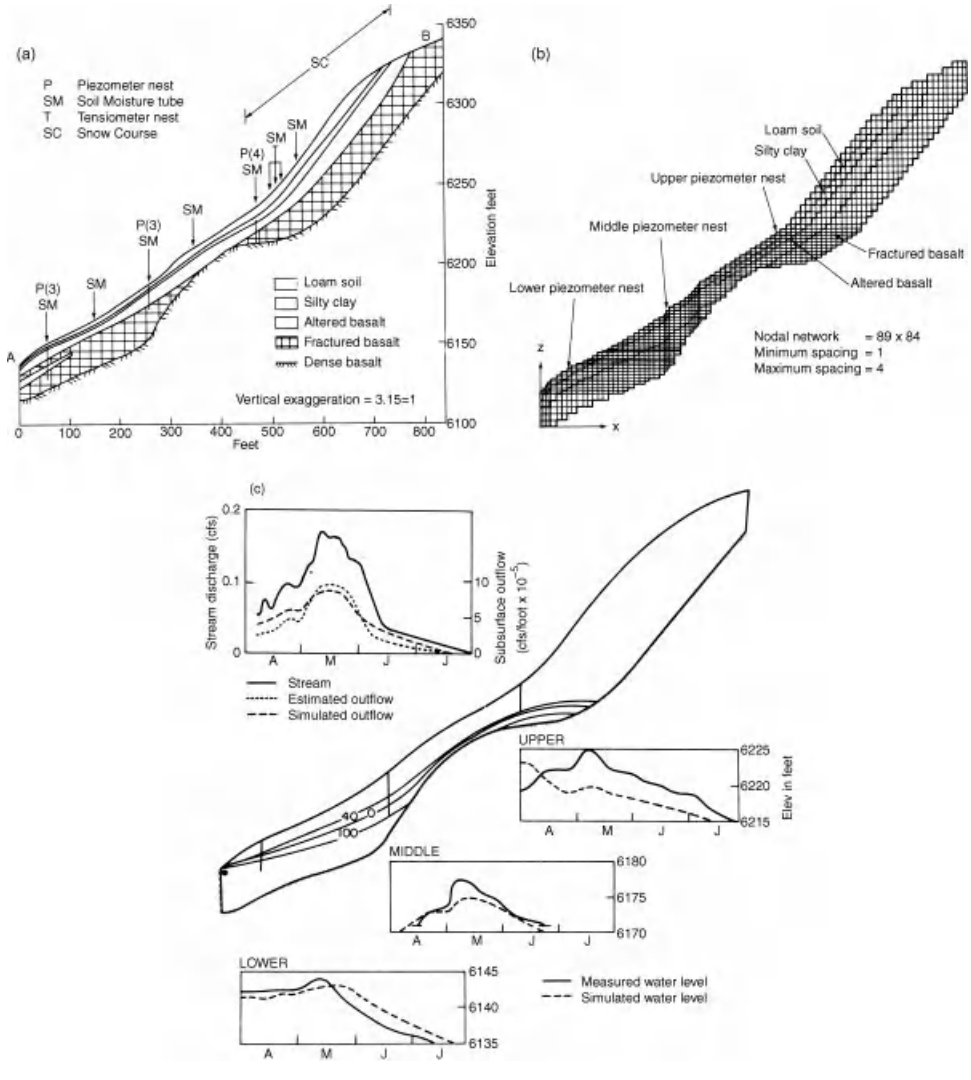


Figure 5.5 *Process-based modelling of the Reynolds Creek hillslope: (a) topography, geology and instrumentation; (b) discretisation of the hillslope for the finite difference model; (c) calibrated transient simulation results for 5 April to 13 July 1971 melt season (after Stephenson and Freeze, 1974, with kind permission of the American Geophysical Union).*

The best discharge predictions (Figure 5.6a) were found for the hypothesis that assumed that the majority of the runoff is generated by a near surface subsurface flow mechanism close to the stream while, on the rest of the slope, snowmelt infiltrates the soil surface and percolates vertically to a deep saturated zone in the porous weathered basalt. This is consistent with the earlier, much more limited, study of Stephenson and Freeze (1974), but in a sensitivity analysis it was found that different parameter sets gave equally acceptable results within the limitations of the data available for model evaluation. The

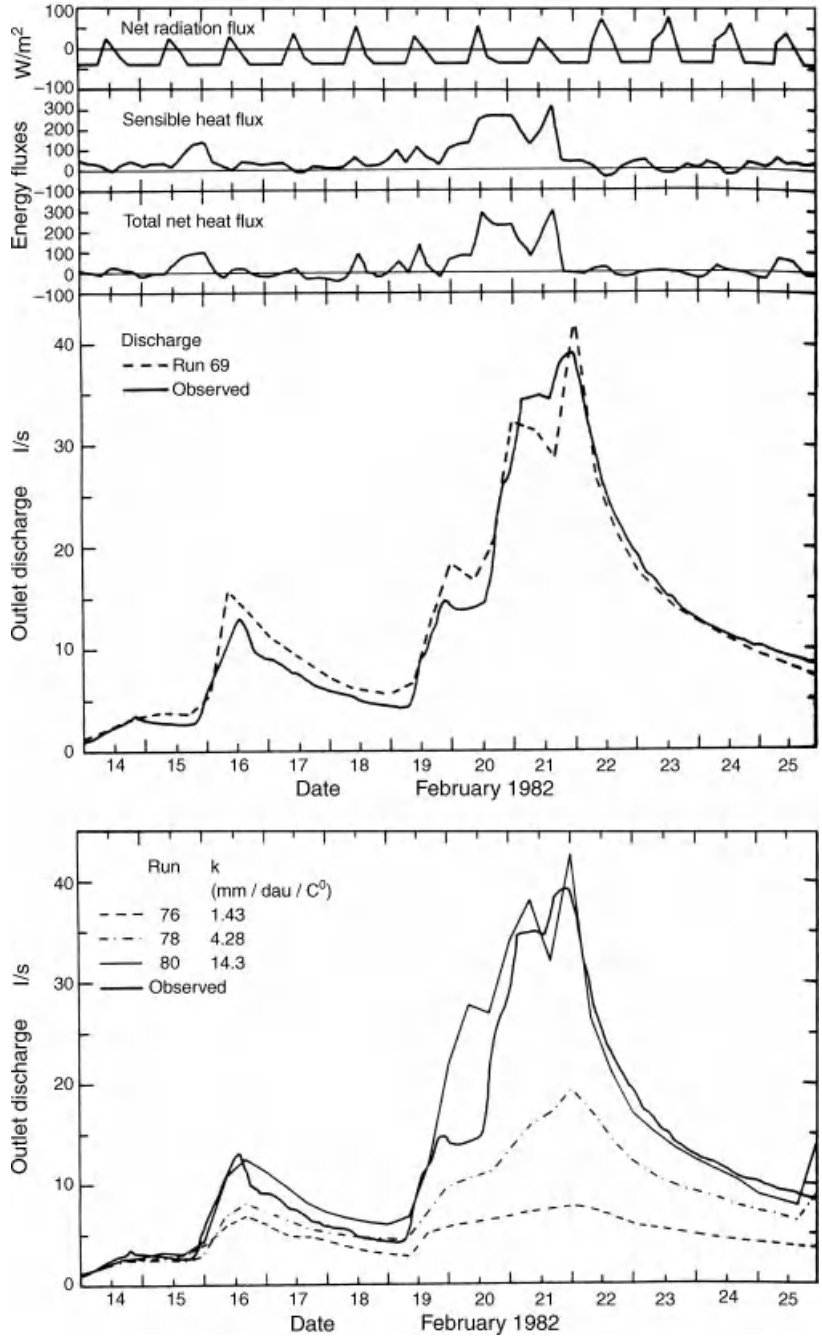


Figure 5.6 *Results of the Bathurst and Cooley (1996) SHE modelling of the Upper Sheep Creek subcatchment of Reynolds Creek: (a) Using the best-fit energy budget snowmelt model; (b) using different coefficients in a degree-day snowmelt model (with kind permission of Elsevier).*

degree-day snowmelt calculations could also produce acceptable discharge predictions (Figure 5.6b) but only after calibration of the degree-day coefficient to a value that was high relative to those reported in the literature.

5.4 Case Study: Blind Validation Test of the SHE Model on the Slapton Wood Catchment

Ewen and Parkin (1996) have outlined a methodology for the *blind validation* of a hydrological model that involves the specification of tests and criteria of success before the model simulations are compared with observed discharges or other observations. This is a form of the "proxy basin" validation test of Klemeš (1986) (see also Section 5.7). In an application of this methodology to the 1.4 km^2 Rimbaud catchment in the Maures Massif near Toulon in southern France, Parkin *et al.* (1996) tested the SHETRAN version of SHE using a grid of 100 m and only prior estimates of the parameter values based on information about the soils and vegetation. The Rimbaud catchment is one of a number of nested subcatchments in the Real Collobrier basin, managed by CEMAGREF. Uncertainty in these estimates was allowed by specifying a range for each parameter. Some general information about the runoff responses was used to set the criteria for success in the model evaluation. The evaluation was "blind" in that the modellers did not have access to the observed discharge record from the catchment before making the model runs. At that time, computer run times of the SHE model were still a significant issue and only a limited number of runs to explore the parameter space and estimate the range of model predictions were possible.

The model was evaluated on four criteria set before the start of the blind test. It was required that the predictions bounds bracket 90% of the observed discharges, 90% of the peak discharges, 11 out of 13 monthly runoff volumes and the total runoff volume. In fact the model was totally successful in only the last of these criteria. Only 78% of the discharge hydrograph, 47% of the peak flows (the model tended to generate runoff by an infiltration excess mechanism, overestimating the peaks with recessions that were too steep) and 10 of the 13 monthly flows were within the prediction bounds. No attempt was made to evaluate any internal predictions in relation to measurements. The studies of Refsgaard and Knudsen (1996) and Feyen *et al.* (2000) also demonstrated somewhat limited success in trying to validate the SHE model in this way. Somewhat greater success has been claimed by Lange *et al.* (1999) in predicting peak flows in a semi-arid environment using a simpler distributed model based on Hortonian infiltration excess concepts, but with surface runoff generation based directly on field measured plot infiltration experiments.

A more recent study of this type is the application of the SHETRAN version of SHE to the 0.94 km^2 Slapton Wood catchment in Devon, UK by Bathurst *et al.* (2004), following an intensive period of hydrological observations that included internal state measurements of water tables and soil moisture profiles. The catchment was represented by a grid of 376 squares of 50 m. Blind predictions were made of 10 features of the phreatic surface, soil water potential and surface runoff responses based on prior estimates of the parameter values. Output uncertainty bounds were determined as a function of uncertainty in the model parameter values. Again the number of runs made to estimate the uncertainty bounds was limited by available computer time.

Bathurst *et al.* (2004) note how the resulting uncertainty bounds were subjectively modified in magnitude and timing (before the modellers had sight of the evaluation data). This, it was argued, was to take account of the additional uncertainty arising from instrumentation and data processing errors (there is also the commensurability issue of what the measurements at points in the catchment mean relative to what the model is predicting on a 50 m grid). They also note how they had some prior information about the nature of the response of the catchment (initial condition information, the double peaked hydrographs in this catchment, a marshy area at the head of the stream) from previous publications about the catchment even though they were not allowed to see the observations with which the model would be compared.

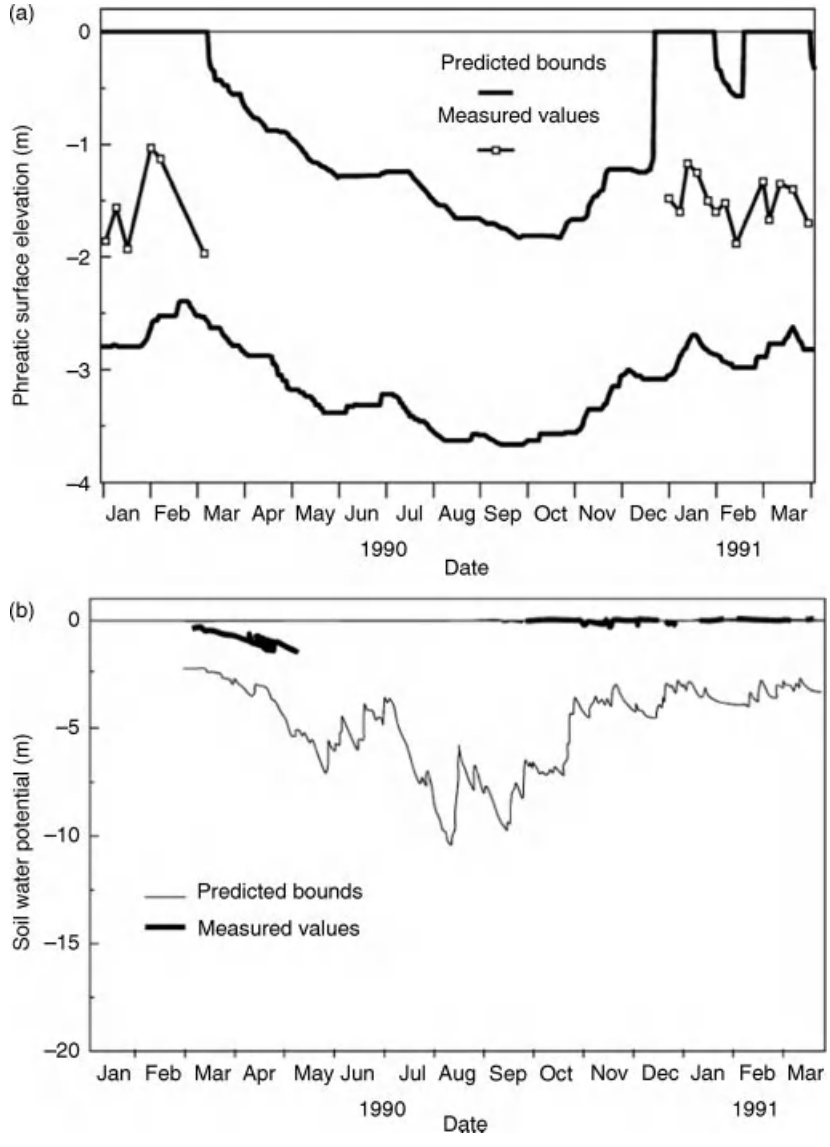


Figure 5.7 *SHE model blind evaluation tests for Slapton Wood catchment, Devon, UK (1/1/90–31/3/91): (a) comparison of the predicted phreatic surface level bounds for square (14; 20) with the measured levels for dipwell (14; 18); (b) comparison of the predicted bounds and measured weekly soil water potentials at square (10; 14) for 1.0 m depth (after Bathurst et al., 2004, with kind permission of Elsevier).*

The resulting uncertainty bounds on the model predictions were declared successful in predicting both water table levels and soil water potentials (Figure 5.7). The authors note that the bounds could have been set much wider if the only aim had been to bracket the observations. The model was also successful in predicting the total volumes of discharge outputs from the catchment (although the bounds were of the

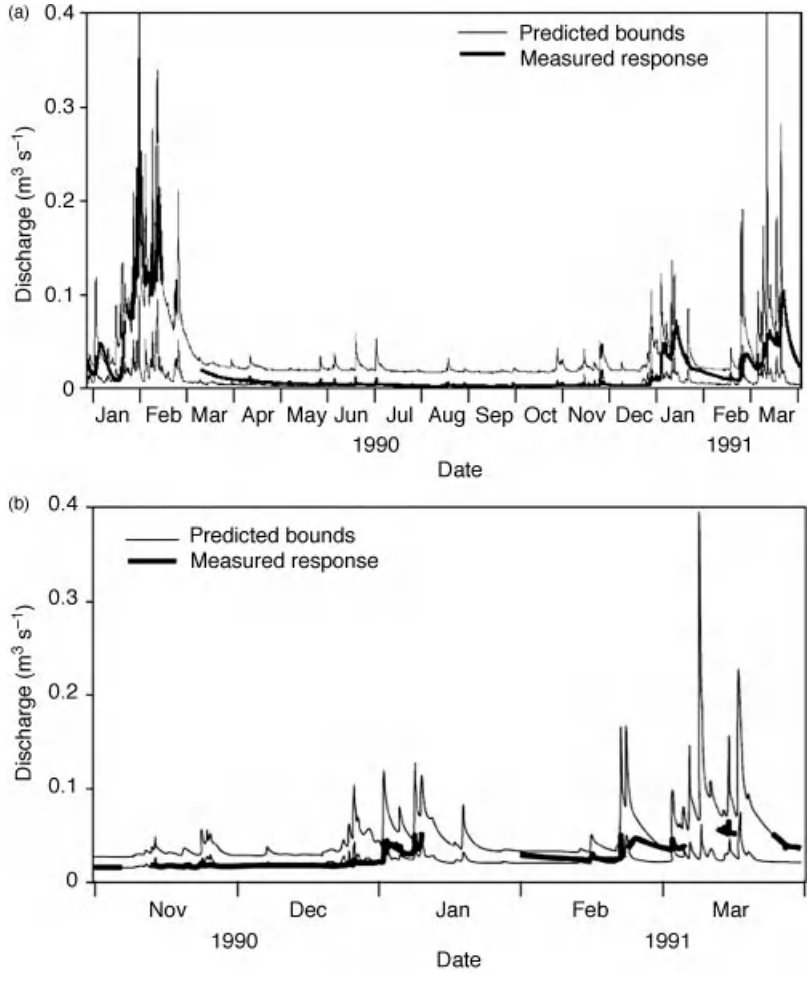


Figure 5.8 *SHE model blind evaluation tests for Slapton Wood catchment, Devon, UK (1/1/90–31/3/91): (a) comparison of the predicted discharge bounds and measured discharge for the Slapton Wood outlet weir gauging station; (b) comparison of the predicted discharge bounds and measured monthly runoff totals for the outlet weir gauging station (after Bathurst et al., 2004, with kind permission of Elsevier).*

order of 30% of the total flow) and an internal gauging site, but failed in bracketing the time series for the discharges (Figure 5.8) and the distribution of monthly flow totals at the internal site. They reasoned that this was, in part, due to the process representations in this version of SHETRAN, which could not predict a downslope subsurface flow component in the soil until it was saturated from below.

5.5 Simplified Distributed Models

The latest versions of SHE and similar physically based rainfall–runoff models represent the most complex rainfall–runoff models available. They have the advantage that they are based on physical theory though,

as we have seen, they make significant simplifying assumptions to allow a model that is computationally feasible. They also have the important advantage that they make predictions that are distributed in space so that the effects of partial changes to the catchment and the predicted spatial dynamics of the processes can be assessed. They have important disadvantages in the computational resources required and in the problems of specifying the huge numbers of parameters required over all the spatial elements of the model. These disadvantages have led to the investigation of simplified distributed models of two main types: those based on kinematic wave theory (considered in the remainder of this chapter) and probability distributed and semi-distributed models (considered in Chapter 6), in which elements with similar characteristics are grouped together to reduce the calculations required.

5.5.1 Kinematic Wave Models

Kinematic wave models are simplified versions of the surface or subsurface flow equations, resulting from making additional approximations. The first such model reported in the literature was, in fact, a grid-based model of surface runoff developed by Merrill Bernard in 1937 (see Hjelmfelt and Amerman, 1980). In another early study, Keulegan (1945) analysed the magnitude of the various terms of the St. Venant equations for shallow surface runoff over a sloping plane and concluded that a simplified equation, essentially the kinematic wave equation, was an adequate approximation. The flow routing of the Huggins and Monke (1968) grid-based model was also effectively a kinematic wave solution. There are some problems in applying kinematic wave principles in two-dimensional cases (see Section 5.5.5) and most models based on the assumptions have used a catchment discretisation based on one-dimensional hillslope planes (as in Figure 5.4). Early models used fixed width planes or planes with radial symmetry to allow analytical solutions to be made, but it is now easy to implement variable width planes in numerical solutions (see Box 5.7).

All kinematic wave models are combinations of the continuity equation with a storage-flow relationship (Box 5.7). Generally some simple mathematical function is used for the storage-flow relationship but this is not strictly necessary. Numerical solutions can use any function represented as a look-up table (even hysteretic functions, although to my knowledge hysteretic functions have never been tried in rainfall–runoff modelling based on the kinematic wave approximation). However, the resulting models are flexible and relatively easy to implement, with many analytical solutions available for simple boundary conditions. A comprehensive coverage of kinematic wave theory and its application in surface hydrology has been published by Singh (1996). The general kinematic wave equation for a variable width flow domain has the form:

$$W_x \frac{\partial h}{\partial t} = -c \frac{\partial W_x h}{\partial x} + W_x r \tag{5.6}$$

where h is a depth of flow, W_x is the width of the slope or channel, r is an inflow rate per unit area of slope or channel and c is the kinematic wave velocity or *celerity* which is, in general, a function of flow depth (but may be a constant in some special cases). The form of that function will vary with the relationship between downslope flow rate and depth of flow (see Box 5.7 for surface and subsurface flow examples).

There is one important limitation of using kinematic wave models, even in applications to one-dimensional systems. Unlike both the Richards equation for subsurface flow and the St. Venant equations and diffusion wave analogy for surface flow, the kinematic wave equation cannot reproduce the effects of a downstream boundary condition on the flow. Essentially, the effects of any disturbance to the flow generate a kinematic wave but the equation can only predict the downslope or downstream movement of these waves. Thus, a kinematic wave description cannot predict the effects of the drawdown of a water table due to an incised channel at the base of a hillslope or the backwater effects of an obstruction to the flow for a surface flow. This has led to a number of theoretical studies of the conditions under which the kinematic approximation is a valid approximation to a more complete description (see Box 5.7),

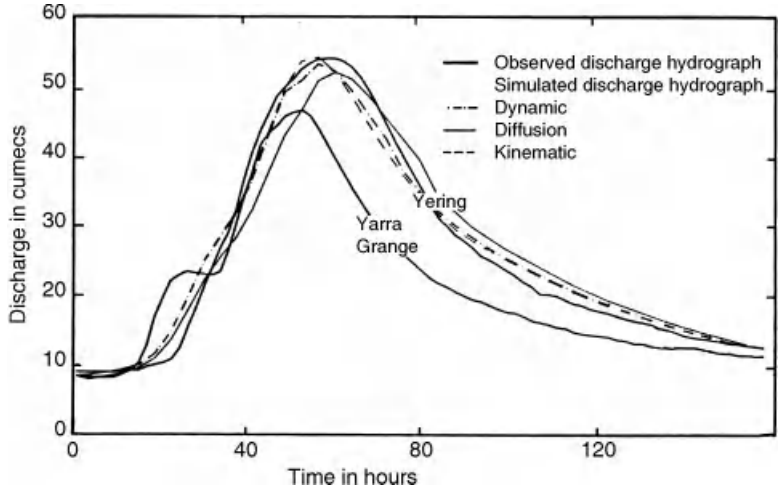


Figure 5.9 *A comparison of different routing methods applied to a reach of the River Yarra, Australia (after Zoppou and O'Neill, 1982).*

but it is worth noting that these are theoretical studies, comparing one mathematical description with another. The problems of parameter estimation and uncertain knowledge of the subsurface geometry and values of recharge or lateral inflows will often mean that these differences are not so important in real applications, and that the kinematic wave approximation may be a useful predictive model. This has been demonstrated, for example, in the study of Zoppou and O'Neill (1982) in a comparison of methods for routing flood waves on the River Yarra in Australia (Figure 5.9).

5.5.2 Kinematic Wave Models for Surface Runoff

An early exposition of the mathematics of kinematic wave theory by Lighthill and Whitham (1955) used traffic routing and flow routing in channels as example applications. This work was later developed by Eagleson (1970) for the case of overland flow routing on hillslopes to predict hydrographs. Eagleson gave analytical solutions for the case of a constant effective rainfall input. Later Li *et al.* (1975) and Smith (1980) provided simple numerical solutions that could be used for arbitrary sequences of inputs. Numerical solutions have been used in a number of catchment rainfall–runoff models based on the infiltration excess overland flow runoff mechanism, the most well-known of which are probably KINEROS (Smith *et al.*, 1995) and the US Corps of Engineers model HEC1 (Feldman, 1995). Both of these models treat a catchment area as a sequence of hillslope segments, bounded by streamlines. Flow is treated as one-dimensional in the downslope direction. Each hillslope may be represented by a single plane or by a cascade of planes of different widths and slopes. In fact, as shown in the development of Box 5.7, it is not difficult to include continuous changes in width and slope in the kinematic wave equations. This would make analytical solutions difficult, but is not a problem with numerical solutions. Goodrich *et al.* (1991) describe a finite element solution of the kinematic wave equation for overland flow for a catchment discretisation based on a triangular irregular network (TIN).

The one-dimensional description requires an appropriate function for the storage–discharge relationship. This may be different for overland and channel flows. However, it has been common in surface water hydrology to use a uniform flow relationship, such as the Manning equation, for both overland and

channel flows. The Manning equation has the form:

$$v = \frac{1}{n} S_o^{0.5} R_h^{0.67} \tag{5.7}$$

where R_h is the hydraulic radius and S_o is the local slope angle. Recall that the hydraulic radius is defined as the cross-sectional area of the flow, A, divided by the wetted perimeter, P. Thus, for flows that are wide compared with their depth $R_h = A/P \approx Wh/W \simeq h$ where W is the width of the flow and h is the local flow depth. The discharge can then be calculated as

$$Q = vhW = \frac{1}{n} W S_o^{0.5} h^{1.67} \tag{5.8}$$

This has the general form of the power law storage–discharge relationship used in Box 5.7 (where an equivalent expression is developed for the Darcy–Weisbach uniform flow equation)

$$q = bh^a \tag{5.9}$$

where the specific discharge $q = Q/W$ and, for Manning's equation, $b = S_o^{0.5}/n$ and $a = 1.67$. The kinematic wave velocity or celerity c is equal to the rate of change of discharge with storage (here $\frac{dq}{dh}$). It is an expression of the rate at which the effects of a local disturbance propagate downslope or downstream. For the power law, $c = (abh^{1-a})$ and this wave velocity always increases with discharge if $a > 1$. For $a = 1$, q is a linear function of h and both the flow velocity and the wave speed c are constant with changing discharge.

Other relationships may show different types of behaviour. Wong and Laurenson (1983) show, for a number of Australian river reaches, how the form of the relationship between wave speed and flow in a river channel may change as the flow approaches bankfull discharge and goes overbank (Figure 5.10).

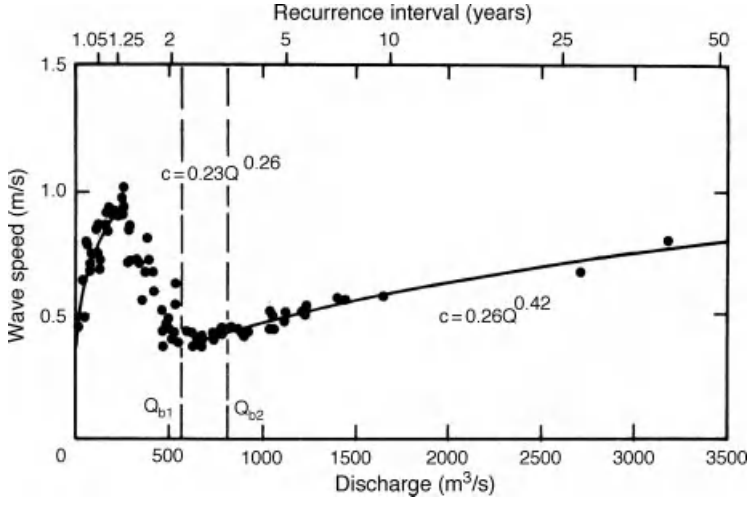


Figure 5.10 *Wave speed–discharge relation on the Murrumbidgee River over a reach of 195 km between Wagga Wagga and Narrandera (after Wong and Laurenson, 1983, with kind permission of the American Geophysical Union): Q_{b1} is the reach flood warning discharge; Q_{b2} is the reach bankfull discharge.*

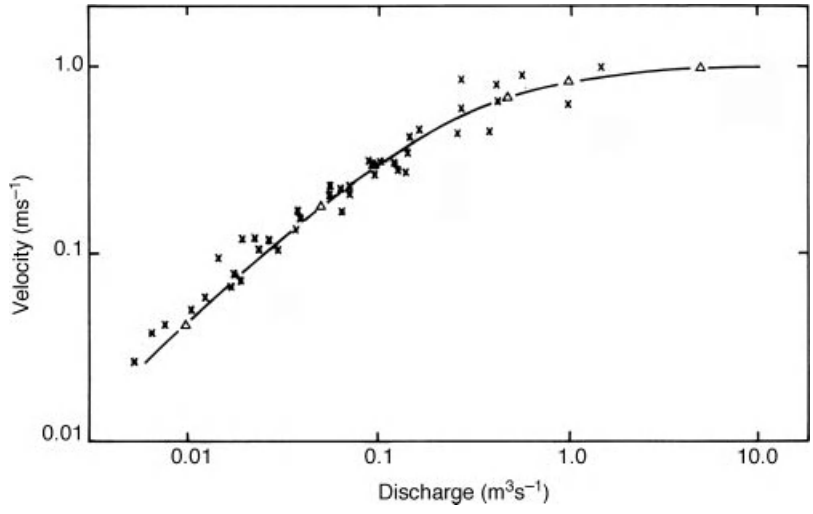


Figure 5.11 *Average velocity versus discharge measurements for several reaches in the Severn catchment at Plynlimon, Wales, together with a fitted function of the form of Equation (5.10) that suggests a constant wave speed of 1 ms^{-1} (after Beven, 1979, with kind permission of the American Geophysical Union).*

At a much smaller scale, Beven (1979b), for example, showed that field measurements in the channels of a small upland catchment suggested a velocity–discharge relationship of the form

$$v = \frac{aQ}{b + Q} \tag{5.10}$$

or, assuming an irregular channel with $Q = vA$

$$Q = a(A - b) \tag{5.11}$$

where b is interpreted as the cross-sectional area at a discharge of zero (not unreasonably in a channel with pools). This relationship gives a constant celerity of $c = a$ for all flows, even though the flow velocity of the water is always increasing with discharge (Figure 5.11). This then allows a further simplification of the routing procedure to the type of constant wave velocity network width based routing discussed in Section 4.5.1. For the particular dataset from small channels in upland Wales examined by Beven (1979b), the value of a was 1 ms^{-1}.

5.5.3 Kinematic Wave Models for Subsurface Stormflow

In the application of the kinematic wave model to saturated downslope subsurface flow, a similar representation of the catchment as one-dimensional stream tubes can be made. The important simplification made, relative to more complex descriptions of such flows, is that the hydraulic gradient can be approximated by the bed slope angle $\sin \beta$ (or, as an approximation, the surface slope). It is therefore assumed that the water table is approximately parallel to the bed (or surface). The Darcy velocity (velocity per unit cross-sectional area of the flow) is then given by

$$v_x = K_s \sin \beta \tag{5.12}$$

where v_x is the Darcian velocity (flux per unit cross-sectional area of saturated soil) measured with respect to the downslope distance x (measured along the slope), K_s is the saturated hydraulic conductivity of the soil (for the moment, assumed to be constant with depth of saturation) and $\sin\beta$ is the slope angle. The kinematic wave approximation was first applied to saturated subsurface stormflow by Henderson and Wooding (1964). Later, Beven (1981) showed that, at least for steeper slopes and high hydraulic conductivities, it could be a useful approximation to a more complete description of shallow saturated flow over an impermeable layer on a hillslope (see Box 5.7). This work was later extended to include delays associated with the propagation of a wetting front into the soil before recharge starts and different profiles of hydraulic conductivity with depth (Beven and Germann, 1982). The kinematic wave equation is a better approximation if hydraulic conductivity increases with depth of saturation, as is the case in many soils due to the increased macroporosity expected near the surface (Kirkby, 1988).

For the constant hydraulic conductivity case, an examination of the wave velocity is interesting. As shown in Box 5.7, for saturated subsurface flow, the wave velocity is given by:

$$c = K_s \sin\alpha/\varepsilon \tag{5.13}$$

where ε is a storage coefficient representing the "effective" difference between soil water content and saturation in the region above the water table. As for the $a = 1$ case for surface flow, if all the three variables controlling the wave velocity were constant, c would be constant, but in practice ε is likely to vary in magnitude both with depth of saturation and distance downslope. For wet soils, ε may be very small. In this case, comparing the expression for c with that for the Darcy velocity v_x, the wave velocity c may be very much faster than the Darcy velocity. That is to say that disturbances to the flow, such as new input of recharge, must propagate downslope faster than the Darcy velocity of the flow. The effects of recharge will also propagate downslope faster than mean pore water velocity, i.e. the average flow velocity through the part of the cross-section that is pore space rather than solids. This is given by

$$v_p = K_s \sin\alpha/\theta_s \tag{5.14}$$

where θ_s is the porosity. The wave velocity will be faster than v_p since ε will always be smaller than θ_s.

This is one explanation why storm hydrographs that have an important contribution of subsurface flow tend to show a high proportion of "old" water, even at peak flows (see Section 1.5). The effects of recharge to the saturated zone move downslope with the wave velocity, which is faster than the pore water velocity. The effect is that the discharge downslope rises more quickly than water can flow from significant distances upslope, so that some of the water which flows out of the slope must be displaced from storage in what was the unsaturated zone above the water table prior to the event. The same analysis holds for more complex descriptions of subsurface flows, but the comparison of the v_x, v_p and c velocities in the kinematic wave description illustrates the effect quite nicely.

The THALES and TOPOG models (Grayson *et al.*, 1992a, 1995; Vertessy *et al.*, 1993; Vertessy and Elsenbeer, 1999; Chirico *et al.*, 2003) both use the kinematic wave approximation on one-dimensional sequences of hillslope elements representing a catchment. Both are based on the TAPES-C digital terrain analysis package (see Section 3.7). THALES allows that each element may have different infiltration characteristics (using either the Green–Ampt or Smith–Parlange infiltration models described in Box 5.2), vertical flow in the unsaturated zone assuming a unit hydraulic gradient and using the Brooks–Corey soil moisture characteristics (Box 5.4), and downslope saturated zone flow using the one dimensional kinematic wave approximation (Box 5.7). The use of THALES in an application to the Walnut Gulch catchment is described in Section 5.6. The TOPOG-Dynamic model, developed by CSIRO in Australia, uses an analytical solution of the Richards equation to describe vertical flow in the unsaturated zone and an explicit kinematic wave solution for the lateral saturated zone fluxes (Vertessy *et al.*, 1993). Later versions include plant growth and carbon balance components for ecohydrological modelling of the effects of land use change (Dawes *et al.*, 1997).

5.5.4 Kinematic Wave Models for Snowpack Runoff

One of the earliest implementations of the kinematic wave equation in hydrology was for modelling flow through a snowpack, by Colbeck (1974). The model was later used by Dunne *et al.* (1976) and further analytical solutions have recently been outlined by Singh *et al.* (1997). All of these studies have assumed a snowpack of constant porosity and hydraulic conductivity. The earlier studies assumed that all the specified net melt would percolate downslope through the snowpack; the later study by Singh *et al.* explicitly incorporates the effects of infiltration into underlying soil. In the general case, the time-varying infiltration rate may be a function of the depth of saturation in the pack. Singh *et al.* (1997) provide analytical solutions for the case where the infiltration rate may be assumed to be constant.

5.5.5 Kinematic Shocks and Numerical Solutions

One of the problems of applying kinematic wave theory to hydrological systems is the problem of kinematic shocks (see Singh, 1996). The kinematic wave velocity can be thought of as the speed with which a particular storage or depth value is moving downslope. If the wave velocity increases with storage, then waves associated with larger depths will move downslope faster than waves associated with shallower depths. Very often, this may not be a problem. In channels, kinematic shocks are rare (Ponce, 1991). On a hillslope plane of fixed width and slope subjected to uniform rainfall, wave velocity will never decrease downslope and there will be no shock. If, however, a concave upwards hillslope is represented as a cascade of planes and subjected to a uniform rainfall, then faster flow from the steeper part of the slope will tend to accumulate as greater storage as the slope decreases, causing a steepening wave until a kinematic shock front occurs. The paths followed by kinematic waves in a plot of distance against time are known as "characteristic curves". A shock front occurs when two characteristic curves intersect on such a plot. The effect of such a shock front reaching the base of a slope would be a sudden jump in the discharge.

This is clearly not realistic. It is a product of the mathematics of the kinematic wave approximation, not of the physics of the system itself which would tend to disperse such sharp fronts. One example of a kinematic shock that can be handled analytically using kinematic wave theory is the movement and redistribution of a wetting front into an unsaturated soil or macropore system (Beven and Germann, 1982; Beven, 1984; Smith, 1983; Charbeneau, 1984; Germann, 1990). In fact, the Green and Ampt (1911) infiltration model (outlined in Box 5.2) can be interpreted as the solution of a kinematic wave description of infiltration, with the wetting front acting as a shock wave moving into the soil. Models have been developed that try to take account of such shocks by solving for the position of the characteristic curves (called the "method of characteristics" by, for example, Borah *et al.*, 1989), but most numerical solutions rely on *numerical dispersion* to take care of such shocks.

Using a finite difference form of the kinematic wave equation to obtain a solution, the approximation adds some artificial or numerical dispersion to the solution. If numerical dispersion is used to reduce the problem of shocks in a solution of the kinematic wave equation (for cases where shocks might occur), then the solution will not be a true solution of the original kinematic wave equation. There is a certain irony to this problem in that the approximate solution might then be more "realistic" since, as already noted, nature would tend to disperse such sharp fronts (Ponce, 1991). However, numerical dispersion is not well controlled; the effect will depend on the space and time increments used in the solution, together with any other solution parameters. The reader needs be aware, however, of the potential problem posed by shocks and of the fact that the approximate solution may not be consistent with the solution of the original equation in such cases. There is also a possibility that a large shock might lead to instability of the approximate numerical solution.

Shocks occur where a larger wave (a depth moving through the flow system) catches up with or meets another smaller one. This will tend to occur where flow is slowed for some reason (reduction in slope,

increase in roughness) or where there is flow convergence. A number of models that purport to use two-dimensional kinematic wave equations (in plan) have been proposed, going all the way back to the original grid-based model of Bernard in 1937. More recent examples are the catchment surface flow model of Willgoose and Kuczera (1995) and the subsurface stormflow model of Wigmosta *et al.* (1994). Mathematically, this is not really a good idea. Wherever there is flow convergence, there is the possibility of two different kinematic waves meeting and a shock front developing. This was, in fact, explicitly recognised by Bernard in the 1930s (see Hjelmfelt and Amerman, 1980) but there appears to be nothing in most of the more recent models to explicitly deal with shocks. It seems that they generally rely on numerical dispersion to smear out any effects of shock fronts. The models work and have been shown to give good results in some test cases, but if shocks are dispersed in this way they are not true solutions of the original kinematic wave equations. I am aware of this problem because I once spent several months trying to develop a finite difference two-dimensional kinematic wave model. The solution kept going unstable for some test runs on a real catchment topography. It took a long while to realise that this was not just a program bug, or a matter of getting the numerics of the approximate solution right but a problem inherent in the mathematics of a kinematic wave description. With hindsight, the reason is fairly obvious! There are, however, a number of models in the literature that claim to solve the kinematic wave equations in two dimensions, without any explicit shock-handling routines. It is not always clear how this has been achieved.

The problems of kinematic shocks can, in fact, largely be avoided in many situations by solving each node separately in the manner shown in Box 5.7, as if it was a node on a one-dimensional plane, by taking advantage of the fact that kinematic waves only move downslope. Thus, in theory, there is no dependence of the solution on downslope conditions. If the kinematic wave equation is solved for flow, q, at a node, then all the inputs from upslope can be lumped together as one input, even if they converge from more than one node. The solution is made and the resulting inputs can then be dispersed to form the inputs for one or more downslope nodes as necessary. This is the approach used in DVSHM, the subsurface kinematic model of Wigmosta *et al.* (1994), which is based on a two-dimensional raster elevation grid, solving for a depth of saturation for every grid element. Palacios-Velez *et al.* (1998) provide an algorithm for constructing such a cascade of kinematic solution elements for both TIN and raster spatial discretisations of a catchment. The only cautionary note that should be added here is that this approach still adds numerical dispersion in a way that is not well controlled and may still be subject to stability problems under rapidly changing conditions. The user should also always remember that a kinematic wave model cannot model backwater effects that arise in channels because of weirs or restrictions to the flow, or in subsurface flow as a result of groundwater ridging or the drawdown effect of an incised channel in low slope riparian areas. At least a diffusion analogy model is required to simulate such effects.

Having raised the issue of kinematic shocks and numerical dispersion, it should be added that it is not necessarily a problem that should worry us unduly. Problems of parameter calibration might well dominate any physical and theoretical approximations inherent in the numerical solution of kinematic wave equations. We can in fact use the numerical dispersion to our advantage, if the resulting model is actually more realistic. An example here is the still widely used Muskingum–Cunge channel flow routing model. The Muskingum method was originally developed as a conceptual flow routing model in the 1930s. Cunge (1969) showed that the Muskingum method was equivalent to a four-point, explicit finite difference solution of the kinematic wave approximation for surface flow. It follows that, for this interpretation, any dispersive and peak attenuation effects of the Muskingum–Cunge routing model come from numerical dispersion associated with the finite difference approximation. In applying the model, it is normal to fit its two parameters so as to match the observed peak attenuation, which allows some control over the numerical dispersion by parameter calibration. This is an interesting historical example, but the details of the Muskingum–Cunge model will not be given here as, in applications to long reaches, it suffers another serious defect of not properly allowing for the advective time delays in the channel (i.e. any change in the upstream inflow has an immediate effect on the predicted reach outflow, regardless

of the length of the reach). This means that the transfer function of the Muskingum–Cunge model often shows an initially negative response (e.g. Venetis, 1969) as a way of producing a time delay. The more general transfer function techniques, described in Chapter 4, that include the possibility of a time delay are a better approach to defining a simple flow routing model where observed hydrographs are available for model calibration. The Muskingum–Cunge model is, in fact, a specific case of the general linear transfer functions outlined in Box 4.1, mathematically equivalent to a first-order (one a coefficient) model, with two b coefficients and zero time delay, although versions have also been developed that allow the parameters to change nonlinearly with time (Todini, 2007; Price, 2009).

The simplicity of the kinematic wave equation with its straightforward combination of the continuity equation and a storage–discharge function, makes it very appealing as an approximation of the real physics. It is particularly appealing in cases where some "effective" storage–discharge function might be required to take account of limited understanding of flow over rilled surfaces or flow through structured soils that might not be well represented by the normal theory used for surface and subsurface flow (e.g Beven and Germann, 1981; Faeh *et al.*, 1997). This degree of flexibility makes it valuable as a modelling strategy. It remains necessary, however, to be aware of the limitations of the approach to essentially one-dimensional flows and the possible effects of kinematic shocks.

5.6 Case Study: Distributed Modelling of Runoff Generation at Walnut Gulch, Arizona

One of the real challenges of semi-arid hydrology is still to model the extensive data set collected by the USDA Agricultural Research Service on the well-known Walnut Gulch experimental catchment (150 km^2) in Arizona (Stone *et al.*, 2008). This has been the subject of numerous experimental and modelling studies, a selection of which are discussed in this section. Walnut Gulch is a semi-arid catchment, with 11 nested subcatchments that range from 2.3 to 150 km^2 and an additional 13 small catchment areas from 0.004 to 0.89 km^2. Spatial variability in rainfall is assessed using a network of 92 gauges. The catchment has been the subject of two intensive field campaigns combining field measurements with aircraft platform remote sensing (Kustas and Goodrich, 1994; Houser *et al.*, 1998). The perception of runoff generation in this environment is that it is almost exclusively by an infiltration excess mechanism (Goodrich *et al.*, 1994).

At the hillslope runoff plot scale in Walnut Gulch, Parsons *et al.* (1997) have compared observed and predicted discharges, together with flow depths and velocities at several cross-sections. The model used the simplified storage-based infiltration model of Box 5.2 with two-dimensional kinematic wave routing downslope. The storage–discharge relationship used was a power law with parameters that varied with the percentage cover of desert pavement in each grid cell of the model. In the first application, to a shrubland site, the model underpredicted the runoff generation but was relatively successful in predicting the shape of the experimental hydrograph (Figure 5.12). The second application, to a grassland plot, was less successful, despite a number of modifications to the model including the introduction of stochastic parameter values.

At a somewhat larger scale, Goodrich *et al.* (1994) and Faurès *et al.* (1995) have applied KINEROS (Smith *et al.*, 1995) to the 4.4 ha Lucky Hills LH-104 subcatchment to examine the importance of different antecedent soil moisture estimates and the effects of wind and rainfall pattern on the predicted discharges (Figure 5.13). KINEROS uses a Smith–Parlange infiltration equation (see Box 5.2) coupled to 1-D kinematic wave overland flow routing on hillslope planes and in channel reaches. At this scale, both studies conclude that an adequate representation of the rainfall pattern is crucial to accurate runoff prediction in this environment. Using average initial soil moisture contents from different remote sensing and modelling methods had little effect on the predictions, as did trying to take account of the effects of wind direction and

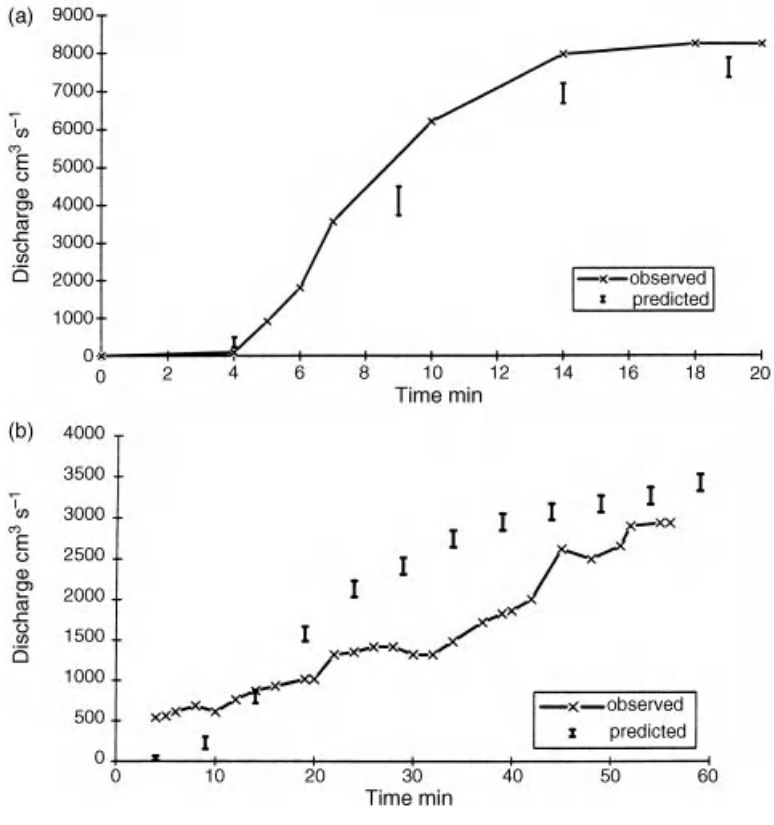


Figure 5.12 *Results of modelling runoff at the plot scale in the Walnut Gulch catchment: (a) the shrubland plot and (b) the grassland plot; the error bars on the predictions indicate the range of 10 randomly chosen sets of infiltration parameter values (after Parsons et al., 1997, with kind permission of John Wiley and Sons).*

velocity on raingauge catch. However, checking the model predictions for different numbers of raingauges they showed that combinations of four gauges (that is a density of 1 per hectare) gave a variation in predicted discharges that spanned the observed discharges and had a similar coefficient of variation to that estimated for the discharge measurements (Faurès *et al.*, 1995). Goodrich *et al.* (1994) also looked at the sensitivity of runoff production to the pattern of initial moisture content at the larger scale of the WG-11 subcatchment (6.31 km^2). They suggest that a simple basin average of initial moisture content will normally prove adequate and that, again, knowledge of the rainfall patterns is far more important. A water balance model estimate of initial moisture content did as well as the remotely sensed estimates.

Michaud and Soroochian (1994) compared three different models at the scale of the whole catchment including a lumped SCS curve number model, a simple distributed SCS curve number model and the more complex distributed KINEROS model. The modelled events were 24 severe thunderstorms (mean runoff coefficient 11%), with a raingauge density of one per 20 km^2. Their results suggested that none of the models could adequately predict peak discharges and runoff volumes, but that the distributed models did somewhat better in predicting time to runoff initiation and time to peak. The lumped model was, in this case, the least successful.

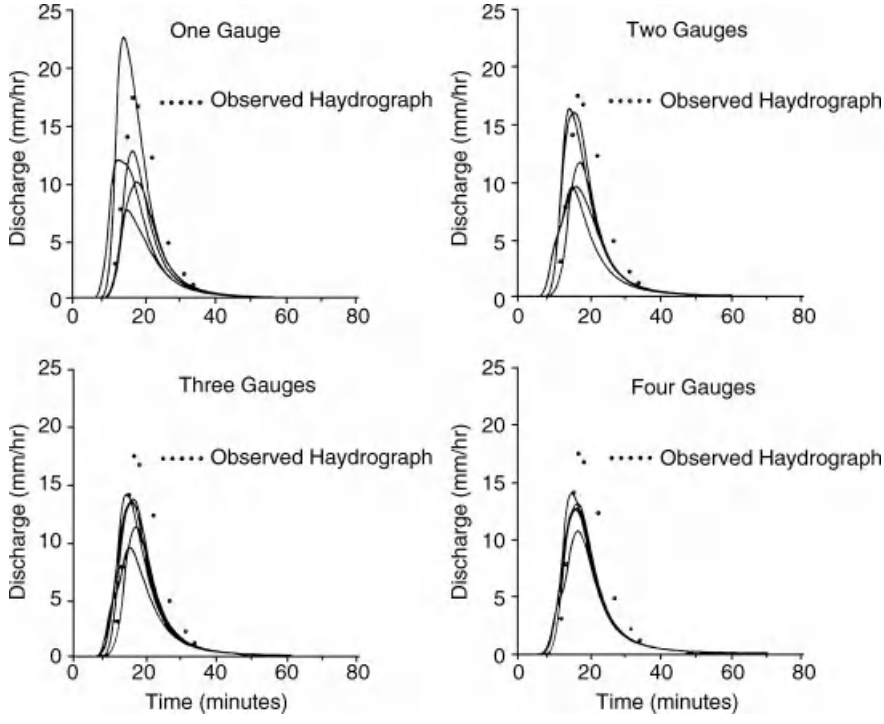


Figure 5.13 *Results of modelling the 4.4 ha Lucky Hills LH-104 catchment using KINEROS with different numbers of raingauges to determine catchment inputs (after Faurès et al., 1995, with kind permission of Elsevier).*

More recently, Goodrich *et al.* (1997) have used data from all 29 nested subcatchments within Walnut Gulch, with drainage areas ranging from 0.2 to 13 100 ha, to investigate the effects of storm area and catchment scale on runoff coefficients. They conclude that, unlike in humid areas, there is a tendency for runoff responses to become more nonlinear with increasing catchment scale in this type of semiarid catchment as a result of the loss of water into the bed of ephemeral channels and the decreasing relative size of rainstorm coverage with catchment area for any individual event. Detailed modelling studies were made using KINEROS for three of the catchments, LH-106 (0.34 ha), LH-104 (4.4 ha), and WG-11 (6.31 km^2). The model was calibrated by adjusting three multipliers applied to the distributed patterns of overland flow roughness, mean soil hydraulic conductivity and the coefficient of variation of hydraulic conductivity, which is allowed to be heterogeneous with a log normal distribution (Woolhiser and Goodrich, 1988; Smith and Goodrich, 2000), to improve the predictions for 10 calibration events. Twenty additional events were used to evaluate the calibrated models. The model was relatively successful in predicting the responses for the two small catchments, less so for WG-11 where the performance on the validation events was much worse than in calibration. The results, however, confirmed the tendency towards increasing nonlinearity at the larger scale.

KINEROS has continued to be applied at Walnut Gulch, including the modelling of sediment yield (Lopes and Canfield, 2004). One study of interest is that of Canfield and Goodrich (2006) who looked at the effect of both parameter and geometric simplifications of KINEROS in predicting both runoff and sediment yield for the LH-104 catchment. Their conclusion was that parameter simplification, by

averaging log hydraulic conductivity values at smaller scales, had little effect on runoff amounts and timing; geometric simplification, by using larger slope elements, had a somewhat greater effect but was still not significant; the effects were significant when predicting sediment yield from the catchment.

The most recent application of KINEROS to Walnut Gulch, by Yatheendradas *et al.* (2008), has been concerned with parameter estimation and uncertainty estimation by fitting observed discharge data using the generalised likelihood uncertainty estimation (GLUE) methodology (see Section 7.10). Yatheendradas *et al.* were concerned with flash flood forecasting using radar rainfall estimates as inputs. It reflects some of the difficulties of doing rainfall–runoff modelling in this kind of environment, in that the derived distributions of the most sensitive parameters were shown to vary with event. This could be for a variety of reasons, including lack of knowledge of the true pattern of antecedent conditions and patterns of rainfall for each event, as well as the limitations of representing the response as an infiltration excess phenomenon and overland flow as a sheet flow. They suggest that it might be better to use data assimilation during an event to adapt parameter values as the event proceeds, but in the case of flash floods this might reduce the time available to issue warnings to the population at risk. More details about the KINEROS model can be found in Appendix A.

Other models that have been used at Walnut Gulch include the Hortonian Infiltration and Runoff/on (HIRO_2) model, which allows for spatial heterogeneity in soil parameters (Meng *et al.*, 2008), and THALES (Grayson *et al.*, 1992a), which includes both infiltration excess and subsurface runoff components. Grayson *et al.* showed that the goodness of fit of the model and the runoff generation mechanisms simulated were both very sensitive to the parameters of the model, which were difficult to estimate from the information available. By calibrating parameters, catchment outflows could be adequately predicted using both Hortonian infiltration excess and partial area runoff generation mechanisms. They suggest that the difficulties of validating such models are acute (see also the discussion in Grayson *et al.*, 1992b).

Finally, Houser *et al.* (1998) have applied the TOPLATS variant of the simplified distributed model TOPMODEL (see Chapter 6) to Walnut Gulch. This study is included here because it is one of the few studies in rainfall–runoff modelling that has attempted to include measurements of a distributed variable within a *data assimilation* or updating framework. Data assimilation is well established in other distributed modelling fields, such as numerical weather forecasting, but has not been widely used in distributed hydrological modelling. In this study, the measured variable was surface soil moisture which was available from the Push Broom Passive Microwave Radiometer (PBMR) carried on an aircraft platform on six days during the MONSOON90 intensive field campaign, including a dry initial condition and the dry down period following a rainstorm of 50 mm. The resulting PBMR soil moisture images showed a very strong correlation with the pattern of rainfall volume interpolated for this storm, resulting in a strong spatial correlation. There are a number of difficulties in trying to make use of such data including the conversion from PBMR brightness temperature to surface soil moisture, the dependence of predicted surface soil moisture on numerous model parameters (the model has some 35 soil and vegetation parameters in all) and the choice of a method for updating the model given the PBMR images over only part of the catchment area. The study compared different methods of varying complexity (and computational burden), suggesting that there is a trade-off between the complexity of the method used and the ability to make use of all the data available. The study confirmed the importance of the rainfall forcing on the hydrologic response, not only for runoff generation but also for the spatial pattern of evapotranspiration and sensible heat fluxes back to the atmosphere.

5.7 Case Study: Modelling the R-5 Catchment at Chickasha, Oklahoma

The R-5 catchment at Chickasha, Oklahoma (9.6 ha) has been the subject of a series of modelling papers using versions of the Quasi-Physically Based Rainfall–Runoff Model (QPBRRM) originally developed by Engman (1974) and then later the Integrated Hydrology Model (InHM) (VanderKwaak

and Loague, 2001). QPBRRM has many similarities to KINEROS but uses a 1-D analytical infiltration component based on the Philip equation (see Box 5.2); a 1-D kinematic wave overland flow component applied on a constant width hillslope plane discretisation of the catchment; and a 1-D kinematic wave channel network routing algorithm. The model takes account of spatially variable runoff generation by an infiltration excess mechanism and the downslope infiltration of runoff as runon onto areas that are not yet saturated. The requirements of the model are therefore the spatially variable infiltration parameters for every point in the catchment and the effective parameters of the storage–discharge relationship for the overland and channel flow model components. InHM is a finite element solution of the coupled surface–subsurface equations in three dimensions and can incorporate heterogeneity of soil characteristics at the element scale.

The story starts with Keith Loague's PhD research, when he applied the QPBRRM model to the R-5 catchment (as well as several other catchments of different types) using data on soil charactersitics for the three main soil types as reported by Sharma *et al.* (1980) (see Loague and Freeze, 1985). Loague and Gander (1990) then added a further 247 infiltration measurements (see Figure 5.14), based on a grid of 25 m spacing (157 sites) and two transects with measurements at spacings of 2 m and 5 m. Loague and Kyriakidis (1997) used a geostatistical kriging analysis to interpolate these data to any 1 m^2 grid point in the catchment, taking account of the effects of temperature on the density and viscosity of the infiltrating water that had been ignored in a similar interpolation by Loague and Gander (1990) and the modelling

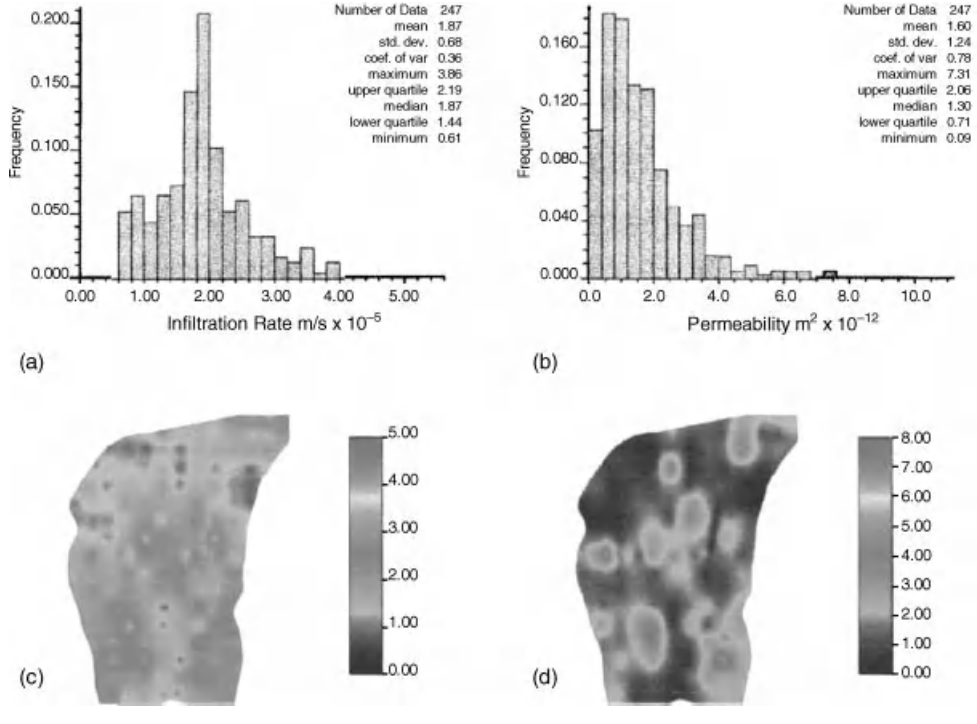


Figure 5.14 *Patterns of infiltration capacity on the R-5 catchment at Chickasha, OK: (a) distribution of 247 point measurements of infiltration rates; (b) distribution of derived values of intrinsic permeability with correction to standard temperature; (c) pattern of saturated hydraulic conductivity derived using a kriging interpolator; (d) pattern of permeability derived using kriging interpolator (after Loague and Kyriakidis, 1997, with kind permission of the American Geophysical Union).*

Table 5.2 *Results of the application of the QPBRRM model (ensemble mean of predictions from 10 stochastic permeability field realisations) to the R-5 catchment without any parameter calibration (after Loague and Kyriakidis, 1997). P_D is total storm rainfall in mm, I_{max} is peak 2-minute rainfall intensity in mm/h, Q_D is total stormflow volume in mm, Q_{PK} is hydrograph peak in l/s, and t_{PK} is time to peak in hours*

Storm	Summary Variable	Observed	Model
1	P_D	68	
	I_{max}	86	
	Q_D	27.36	0.75
	Q_{PK}	131.87	16.57
	t_{PK}	5.55	20.56
2	P_D	33	
	I_{max}	214	
	Q_D	22.43	0.49
	Q_{PK}	375.77	13.69
	t_{PK}	2.26	2.51
3	P_D	50	
	I_{max}	72	
	Q_D	10.15	1.65
	Q_{PK}	220.16	44.41
	t_{PK}	2.97	2.77

studies that had made use of their data (e.g. Loague, 1990, 1992). Loague and Kyriakidis (1997) then applied the QPBRRM model to simulate three large rainfall–runoff events (with rainfall depths between 33 and 68 mm), with a spatial grid of 2.5 m and time step of 60 s. The catchment discretisation used 959 overland flow plane segments, each with different infiltration parameters.

Given all this detailed information, it turned out that applying the model with only field-measured parameter values resulted in poorer hydrograph predictions (see Table 5.2) than in the original study by Loague and Freeze (1985). Adjusting the infiltration parameters to take account of the temperature effects in the original measurements did result in an improvement in the predictions, as did taking the ensemble average of a number of stochastically generated parameter fields relative to the mean values interpolated by kriging. The model significantly underestimates the total storm runoff for the three storms. Loague and Kyriakidis suggest that the runon process is a major limitation of the model in that even this very detailed representation does not adequately represent the nature of the overland flow pathways. In the model representation, too much overland flow infiltrates on higher conductivity elements downslope. The original model of Loague and Freeze (1985), which used much larger flow planes, had less spatial variability in infiltration rates and produced more runoff. They also suggest that the concepts that underlie the QPBRRM model may not be totally appropriate for this catchment and that subsurface processes and saturation excess runoff generation may play a much larger role in the storm response. Loague and Kyriakidis conclude (1997 p. 2895):

> One would expect that simulating rainfall–runoff events for the data-rich R-5 catchment to be somewhat more straightforward and certainly more rewarding than what is reported here. . . . Why then should we model? It is our opinion that the best use of deterministic physically based simulation models is in a concept-development mode for the design of future field experiments and optimal data collection strategies.

This history of attempts to model R-5, one of the most intensively studied in the world, was summarised as a "never ending story" by Loague *et al.* (2000). The story did not, however, end there and has continued (Loague and VanderKwaak, 2002, 2004; Loague *et al.*, 2005; Heppner and Loague, 2007; Heppner *et al.*, 2007). The role of time variant infiltration parameters, the roads and deeper subsurface flows in improving the predictions of discharge at the site have all been examined. The understanding of the hydrology at this site has changed from an assumption that the response is dominated by surface runoff, to one in which the subsurface plays a much more important role. Thus the original attempt to determine model parameter values by an extensive (and expensive) programme of infiltration measurements had only limited success. Determining parameter values by measurements deeper into the soil is, of course, much more difficult and even more expensive. This has implications for the application of detailed physically based distributed models more generally (remembering that R-5 is a catchment of only 9.6 ha).

Similar stories of using this type of model to help understanding of hillslope and catchment responses have been reported from elsewhere. InHM has also been applied to the Coos Bay site in Oregon (Ebel *et al.*, 2007, 2008; Ebel and Loague, 2008), while the TOUGH2 code has been applied to the Panola hillslope in Georgia (James *et al.*, 2010). Both of these sites have been the subject of intensive field experiments and would have made interesting case studies here. They are both worth investigating further. The Coos Bay site is particularly interesting in that it was a very steep hillslope hollow that eventually failed as a shallow landslide during a high volume storm. The failure then revealed that deeper subsurface flows through fractured bedrock, concentrating at two locations, might have contributed to the excess pressures that triggered the shallow landslide (Montgomery *et al.*, 2002). InHM has been extended to allow sediment transport simulations (Heppner *et al.*, 2007; Ran *et al.* 2007).

5.8 Good Practice in the Application of Distributed Models

Validation of distributed models is an issue that has received a great deal of recent attention in the field of groundwater modelling following a number of studies in which predictions of groundwater behaviour were not borne out by subsequent experience (see Konikow and Bredehoeft, 1992). Some of this discussion has, in fact, suggested that validation is not an appropriate term to use in this context, since no model approximation can be expected to be a valid representation of a complex reality (e.g. Oreskes *et al.*, 1994). Model evaluation has been suggested as a better term. Because distributed models make distributed predictions, there is a lot of potential for evaluating not only the predictions of discharge at a catchment outlet, but also the internal state variables, such as water table levels, soil moisture levels and channel flows at different points on the network. A decade ago, when the first edition of this book was written, there had been relatively few attempts to validate the predictions of distributed models (Bathurst, 1986; Parkin *et al.*, 1996; Refsgaard, 1997). Now there have been a few more, such as those discussed in the case studies (Bathurst *et al.*, 2004; Loague and VanderKwaak, 2002, 2004; Ebel *et al.* 2007; James *et al.*, 2010).

The lack of evaluation with respect to internal *state variables* is clearly partly due to the expense of collecting widespread measurements of such internal state variables. Hence, nearly all of these detailed evaluations have been made on small, intensively studied catchment areas for research purposes. An interesting development in this respect has been the construction in 2005 of the 6 ha artificial "Chicken Creek" catchment on an old open-cast-mine site in Germany. In constructing the catchment, significant effort was made to ensure that the base of the catchment was water tight and that the soil that was put in place was homogeneous (as far as possible given the amount of material that had to be moved), free of macropores and of known textural characteristics. The first modelling studies of this catchment have now been reported, as a competition between models using *a priori* parameter estimates (Holländer *et al.*, 2009).

The results of this comparative study have some interesting implications for good practice in the application of distributed models. It was found that the more complex models based on continuum physics (both CATFLOW and HYDRUS2D were included in the comparison) did not perform any better in predicting

discharges and water tables in the catchment than simpler models. This was interpreted as a problem of *a priori* identification of effective values of the parameters needed. Indeed, it seemed that even in this near ideal case, the uniqueness of place issue arises (Beven, 2000); there were sufficient heterogeneities and complexities in the structure of the system to make it difficult to model. It was also stressed that the interaction between modeller and model was important. The choice of modeller was as important as the choice of model, in particular in the (subjective) interpretation of information about the dominant processes at the site and expectations about the initial conditions in setting up the model runs (see also Seibert and McDonnell, 2002; Fenicia *et al.*, 2008b).

The practical applications of this type of distributed model are often at much larger catchment scales with larger calculation elements (e.g. Abbott and Refsgaard, 1996; Singh and Frevert, 2002b; Refsgaard and Henriksen, 2004). If it has proven difficult to simulate the processes in very small catchments using this type of distributed model, how should that guide good practice in practical applications of such models at larger scales? Two important consequences are apparent. The first is that there should be an expectation that the predictions of such models will be uncertain. Consequently, some effort should be made to assess and constrain that uncertainty, perhaps by a targeted measurement programme (although little information is available in the literature about the value of different types of data in this respect). The second is that the modelling process should be treated as a learning process. Holländer *et al.* reported that most of the modellers changed their qualitative view of how the catchment processes were working following a field visit. They suggested that there was a form of Bayesian learning process at work here.

For example, suppose that the parameters of a distributed model have initially been calibrated only on the basis of prior information about soil and vegetation type, with some adjustment of values being made to improve the simulation of measured discharges (although the sheer number of parameters required by distributed models makes any form of calibration difficult, see Chapter 7). Suppose that after this initial calibration a decision is made to collect more spatially distributed information about the catchment response. Measurements might be made of water table heights, soil moisture storage and some internal stream gauging sites might be installed. We would expect the predictions of the calibrated distributed model to turn out to be wrong in many places, since the calibration has taken little account of local heterogeneities in the catchment characteristics (other than the broad classification of soil and vegetation types). There is now the potential to use the new internal measurements not to evaluate the model, but to improve the local calibration, a process that will not necessarily improve the predictions of catchment discharge which was the subject of the original calibration (see also the TOPMODEL case study in Section 6.4). It will generally make a much greater improvement to predictions of the internal state variables for which measurements have now been made available. But if the new data is being used to improve the local calibration, more data will be need to make a model evaluation. In fact, in practice, there is generally little model evaluation but rather the model is adapted to take account of the new data, without necessarily any impact on the variables of greatest interest in prediction (discharge in rainfall–runoff modelling). The logical extension of this process is the "models of everywhere" concepts of Beven (2007) that are discussed in Chapter 12.

5.9 Discussion of Distributed Models Based on Continuum Differential Equations

In many respects, there has been a lot of progress in distributed hydrological models in the last decade. There is now much more computer power available to support finer grid resolutions and more sensitivity, calibration and uncertainty estimation runs. The SHE model, in particular, has been important in the development of distributed modelling technology. It is now perfectly possible to run the SHE model many times to support uncertainty estimation (Christiaens and Feyen, 2002; Vazquez *et al.*, 2009). The application of this type of model is not, however, without limitations, some of which have been discussed

by Beven (1989), Grayson *et al.* (1992b) and Beven (1996a). The latter paper summarises some of the problems as follows (Beven, 1996a, p.273):

> There is a continuing need for distributed predictions in hydrology but a primary theme of the analysis presented here is that distributed modelling should be approached with some circumspection. It has been shown that the process descriptions used in current models may not be appropriate; that the use of effective grid scale parameter values may not always be acceptable; that the appropriate effective parameter values may vary with grid scale; that techniques for parameter estimation are often at inappropriate scales; and that there is sufficient uncertainty in model structure and spatial discretisation in practical applications that these models are very difficult (if not impossible) to validate.

It is directly followed by a response from Danish Hydraulics Institute SHE modellers (Refsgaard *et al.*, 1996, p.286). Their response concludes:

> In our view the main justification for the distributed physically based codes are the demands for prediction of effects of such human intervention as land use change, groundwater abstractions, wetland management, irrigation and drainage and climate change as well as for subsequent simulations of water quality and soil erosion. For these important purposes we see no alternative to further enhancements of the distributed physically based modelling codes, and we believe that the necessary codes in this respect will be much more comprehensive and complex than the presently existing ones.

Similarly, Brian Ebel and Keith Loague have suggested that taking the distributed continuum approach to the prediction of catchment processes is one way to avoid the "fog of equifinality". While recognising the ill-posedness of the catchment modelling problem, they suggest that ensuring that the model predictions are consistent with both the physics of the processes and the internal state measurements in the catchment is the only way to progress scientific hydrology and be sure that a model is predicting the right responses for the right reasons (Ebel and Loague, 2006). How far this is possible in practical applications rather than research sites remains to be seen. The justifications that underlie the development of physically based distributed models are not in dispute. The need for prediction of the effects of distributed changes in a catchment continue to increase (see Chapter 8). However, it is somewhat difficult to see how such advances will be made unless new measurement techniques for effective parameter values or grid scale fluxes are developed. There are theoretical problems about the physics of the process descriptions (Beven, 1989, 2001, 2006b) and practical numerical problems that will need to be overcome in the future development of this type of model but the problem of parameter identification, particularly for the subsurface, will be even greater.

Refsgaard *et al.* (2010) in their review of 30 years of SHE modelling come to somewhat similar conclusions. They also point to the difficulty of defining process descriptions that reflect the complexities and heterogeneities of the real system; the problem of defining effective parameter values within a particular model structure; the potential for different model structures to produce similar predictions; the importance of uncertainty estimation; and the need for new measurement techniques for model set-up and evaluation. These issues will recur in the discussion of semi-distributed models in Chapter 6; in considering the next generation of hydrological models in Chapter 9; and in the "models of everywhere" concepts of Chapter 12. They are fundamental to the future of hydrological science. The question is whether the continuum differential equation approach that was outlined by the Freeze and Harlan blueprint in 1968 is the best or only way of doing science in hydrology. As we will see, there may be alternatives . . . .

5.10 Key Points from Chapter 5

- Fully three-dimensional models for surface and subsurface flow processes have become computationally feasible but most current physically based distributed rainfall–runoff models still discretise the catchment into lower dimension subsystems. Such a discretisation leads to only approximate representations of the processes and may lead to numerical problems in some cases.
- The effects of heterogeneities of soil properties, preferential flows in the soil and irregularities of surface flows are not generally well represented in the current generation of models. Small scale heterogeneity in the unsaturated zone suggests that new grid or element scale representations might be needed.
- The widely used SHE model is an example of a model based on grid elements, using one-dimensional finite difference solutions for channel reaches and the unsaturated zone in each grid element, and two-dimensional solutions in plan for the saturated zone and overland flow processes. It includes sediment transport and water quality components.
- The InHM model is an example of a fully 3D subsurface model, coupled to a 2D surface flow model that has been tested in applications of detailed experimental data. It has also been extended to predict sediment transport. HydroGeoSphere is another 3D modelling system that simulates flow, sediment and water quality variables.
- In some circumstances it may be possible to use simpler solutions based on the kinematic wave equation for both surface and subsurface flows.
- All distributed models require effective parameter values to be specified at the scale of the calculation elements that may be different from values measured in the field. Distributed predictions mean that distributed data can be used in model calibration but evaluation of this type of model may be difficult due to differences in scale of predictions and measurements and the fact that the initial and boundary conditions for the model cannot normally be specified sufficiently accurately.
- A small number of studies have investigated the uncertainties associated with the predictions of distributed models, based only on prior estimates of parameter values or on conditioning of prior ranges on the predictions of observed data. In both cases, the models do not do always provide acceptable simulations of all variables, particularly internal state states of the system.
- It has been argued that the use of distributed physics-based models is the best way of doing hydrological science. There are still theoretical problems that need to be overcome in dealing with heterogeneity and preferential flows in this type of model but the problem of parameter identification, particularly for the subsurface, will be even greater and significant progress will undoubtedly depend on the development of improved measurement techniques.

Box 5.1 Descriptive Equations for Subsurface Flows

The generally used description for both saturated and unsaturated subsurface flows is based on Darcy's law (Darcy, 1856) which assumes that the disharge per unit area or Darcian velocity can be represented as the product of a gradient of hydraulic potential and a scaling constant called the "hydraulic conductivity". Thus

$$v_x = -K\frac{\partial \Phi}{\partial x} \tag{B5.1.1}$$

where v_x [LT^{-1}] is the Darcian velocity in the x direction, K [LT^{-1}] is the hydraulic conductivity, and Φ [L] is the total potential ($\Phi = \psi + z$ where ψ [L] is the capillary potential and z is the elevation above some datum). In the case of unsaturated flow, the hydraulic conductivity changes in a nonlinear way with moisture content so that

$$K = K(\theta) \tag{B5.1.2}$$

where θ [-] is volumetric moisture content.

Combining Darcy's law with the three-dimensional mass balance equation gives:

$$\frac{\partial \rho\theta}{\partial t} = -\frac{\partial \rho v_x}{\partial x} - \frac{\partial \rho v_x}{\partial x} - \frac{\partial \rho v_x}{\partial x} - \rho E_T(x, y, z, t) \tag{B5.1.3}$$

where ρ is the density of water (often assumed to be constant) and $E_T(x, y, z, t)$ [T^{-1}] is a rate of evapotranspiration loss expressed as a volume of water per unit volume of soil that may vary with position and time. Combining Equation (B5.1.1) with Equation (B5.1.3) gives the nonlinear partial differential equation now known as the Richards equation after L. A. Richards (1931):

$$\frac{\partial \rho\theta}{\partial t} = \frac{\partial}{\partial x}\left[\rho K(\theta)\frac{\partial \Phi}{\partial x}\right] + \frac{\partial}{\partial y}\left[\rho K(\theta)\frac{\partial \Phi}{\partial y}\right] + \frac{\partial}{\partial z}\left[\rho K(\theta)\frac{\partial \Phi}{\partial z}\right] - \rho E_T(x, y, z, t) \tag{B5.1.4}$$

or, remembering that $\Phi = \psi + z$

$$\frac{\partial \rho\theta}{\partial t} = \frac{\partial}{\partial x}\left[\rho K(\theta)\frac{\partial \psi}{\partial x}\right] + \frac{\partial}{\partial y}\left[\rho K(\theta)\frac{\partial \psi}{\partial y}\right] + \frac{\partial}{\partial z}\left[\rho K(\theta)\left(\frac{\partial \psi}{\partial z} + 1\right)\right] - \rho E_T(x, y, z, t) \tag{B5.1.5}$$

or, in a more concise form using the differential operator ∇ (see Box 2.2),

$$\frac{\partial \rho\theta}{\partial t} = \nabla\left[\rho K(\theta)\nabla\psi\right] + \frac{\partial \rho K(\theta)}{\partial z} - \rho E_T(x, y, z, t) \tag{B5.1.6}$$

This form of the equation assumes that the hydraulic conductivity at a given moisture content is equal in all flow directions, i.e. that the soil is "isotropic", but in general the soil or aquifer may be "anisotropic", in which case, the hydraulic conductivity will have the form of a matrix of values. This equation involves two solution variables Φ and θ. It can be modified to have only one solution variable by making additional assumptions about the nature of the relationship between ψ and θ. For example, defining the *specific moisture capacity* of the soil as $C(\psi) = \frac{d\theta}{d\psi}$ and assuming that the density of water is constant, then the Richards equation may be written

$$C(\psi)\frac{\partial \psi}{\partial t} = \nabla\left[\underline{K}(\psi)\nabla\psi\right] + \frac{\partial K_z(\psi)}{\partial z} - E_T(x, y, z, t) \tag{B5.1.7}$$

where $\underline{K}(\psi)$ and $K_z(\psi)$ now indicate that hydraulic conductivity depends on direction and is treated as a function of ψ for the unsaturated case. To solve this flow equation, it is necessary to define the functions $C(\psi)$ and $\underline{K}(\psi)$ (see Box 5.4). To make things more complicated, both

$C(\psi)$ and $\underline{K}(\psi)$ may not be simple single valued functions of ψ or θ but may be subject to *hysteresis*, that is varying with the history of wetting and drying at a point. Various models of hysteretic soil characteristics have been proposed (see Jaynes, 1990) but they are not often used in rainfall–runoff modelling. For saturated soil and in the saturated zone of an aquifer, $\underline{K}(\psi)$ approaches the saturated conductivity $\underline{K}_s$ and $C(\psi)$ takes on a very small value due to the compressibility of the soil or aquifer.

An equivalent form of the Richards equation may be written with soil moisture content θ as the dependent variable as:

$$\frac{\partial \theta}{\partial t} = \nabla \left[\underline{K}(\theta) \frac{d\psi}{d\theta} \nabla \theta \right] + \frac{\partial K_z(\theta)}{\partial z} - E_T(x, y, z, t) \tag{B5.1.8}$$

The product $\underline{K}(\theta)\frac{d\psi}{d\theta}$ is known as the *diffusivity* of the soil and is often written as $\underline{D}(\theta)$.

The Richards equation applies to both saturated and unsaturated flow through a porous medium. For near surface flows, it is normally assumed that the water is of constant density and that the soil is also incompressible. For deep aquifers, such an assumption may not be valid and it may be necessary to take account of the compressibility of both water and rock. A typical parameterisation takes the form:

$$n(P) = \theta_s(1 + bP) \tag{B5.1.9}$$

where $n(P)$ is the porosity of the soil at pore pressure P, for $P > 0$, θ_s is the porosity at atmospheric pressure, and b is a compressibility coefficient that varies with the nature of the soil or rock.

The temperature dependence of hydraulic conductivity due may also be important and may be expressed in the form:

$$K(\psi) = k_s k_r(\psi) \frac{\rho g}{\mu} \tag{B5.1.10}$$

where k_s is called the intrinsic permeability of the soil at saturation which should be a characteristic only of the porous medium, $k_r(\psi)$ is a relative conductivity ($0 < k_r(\psi) < 1$) varying with capillary potential ψ, and μ is the dynamic viscosity of the soil. Both ρ and μ vary with temperature.

The Richards equation does not have an analytical solution for most cases of interest and solutions must be obtained by approximate numerical calculations (see Box 5.3).

The following assumptions are made in developing this form of the Richards equation:

A1 For both saturated and unsaturated flow, flow velocity may be assumed to be a linear function of the gradient of hydraulic potential in accordance with Darcy's law.
A2 Functional relationships can be specified for the soil moisture characteristic curves to relate moisture content, capillary potential and hydraulic conductivity of the soil.
A3 Fluxes of water in vapour form can be neglected.

Additional assumptions that are often made in applying the Richards equation are:

A4 The soil moisture characteristics are non-hysteretic.
A5 The hydraulic conductivity tensor is isotropic.
A6 The porous medium is incompressible.
A7 The water is of constant temperature and density.

Assumptions A4–A7 reduce the number of parameters that must be specified before the model can be run.

Box 5.2 Estimating Infiltration Rates at the Soil Surface

Many hydrological models have had as their basis the Horton infiltration excess concept of runoff generation. This includes some distributed models, in which the only consideration of subsurface flow processes has been the prediction of infiltration at the soil surface (see Section 5.5). Other models, including some unit hydrograph models, have also used an infiltration excess concept to calculate how much of a rainstorm to route as runoff. Thus, the estimation of infiltration at the soil surface has traditionally been an important part of hydrological theory.

Most infiltration experiments show that at the start of rainfall infiltration rates are high and then decline gradually over time (there are rare exceptions where, for example, the breakdown of a surface crust or reduction in hydrophobicity may result in increasing infiltration rates over time). The initially high rates are due to the effects of capillary potential drawing water into the dry soil in addition to the effects of gravity. If there is a time before the surface of the

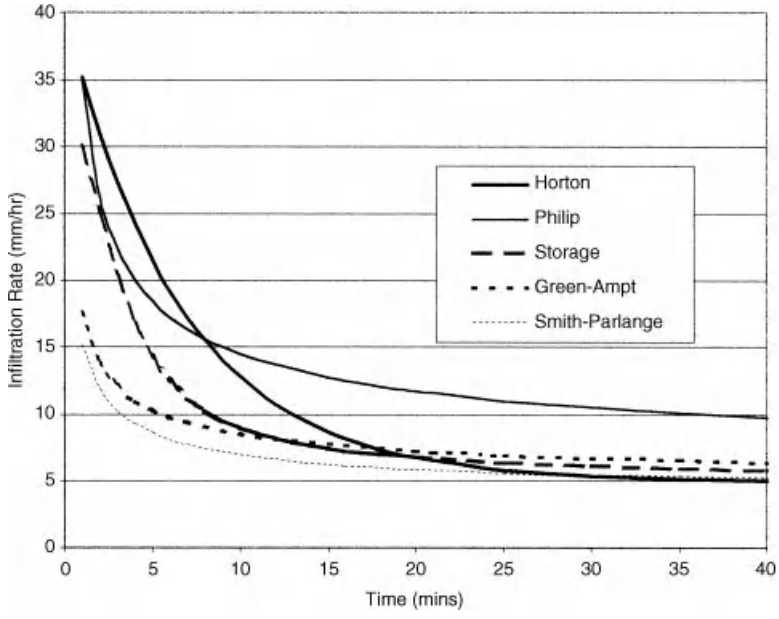


Figure B5.2.1 *Predictions of infiltration equations under conditions of surface ponding.*

soil reaches saturation then initially the infiltration may be limited by the rainfall rate, i.e. all the rainfall infiltrates into the soil. The soil may only start to limit infiltration rates once this *time to ponding* has been reached. After a long period of heavy rainfall, the wetting front will have moved some distance into the soil, the effects of capillary potential will be small, the potential gradient will be dominated by gravity and the infiltration rate will approach the effective hydraulic conductivity of the soil (see Figure B5.2.1).

This decline of infiltration capacity can be predicted for any arbitrary initial soil profile and pattern of rainfall intensities by a numerical solution of the Richards equation (see Box 5.1). Such solutions can also easily take account of arbitrary layering in the soil, including surface crusting, given some knowledge of the soil moisture characteristics for the different layers. Dual porosity solutions can make some attempt at representing the effects of macropores on infiltration (e.g. those of Jarvis *et al.*, 1991; Bronstert and Plate, 1997; and Mohanty *et al.*, 1998).

However, an accurate simulation of the time to ponding and infiltration capacity of the soil may require very small depth and time increments to resolve the rapid changes in hydraulic gradients in time and space that control the change in infiltration capacity. Computer run time therefore becomes an issue, particularly where distributed predictions of infiltration capacity may be required simulating the generation of surface runoff in a heterogeneous catchment. Thus, there remains a need for simpler analytical solutions of infiltration into the soil surface. There are many such equations reported in the literature. A selection of the most widely used are discussed in the following sections.

B5.2.1 The Horton Infiltration Equation

Horton (1933, 1940) described this type of curve by an empirical function of the form

$$f(t) = \left(f_o - f_c\right) \exp\left\{-kt\right\} + f_c \tag{B5.2.1}$$

where f_o is an initial infiltration capacity, f_c is a final infiltration capacity and k is an empirical coefficient. The three parameters, f_o, f_c and k, are a function of soil type but may also depend on the antecedent state of the soil. The final infiltration capacity, f_c, will be close to the hydraulic conductivity of the soil at field saturation. Although this form of equation was based on empirical evidence, Eagleson (1970) has shown that it is an approximate solution of the Richards equation under certain simplifying assumptions. As noted in Chapter 1, however, Horton did not think of the decline of infiltration as being controlled by the soil moisture profile. He observed that the surface of the soil could saturate without the profile being saturated and concluded that the infiltration capacity was controlled by processes at the surface, for example the redistribution of fine particles by rainsplash that would block the larger pores. He therefore assigned the decline in capacity to what he called "extinction phenomena" (Beven, 2004b). Others have also suggested that the Richards equation may not be the most appropriate description of infiltration into soils (e.g. Beven and Germann, 1981, 1982; Germann, 1985; Germann *et al.*, 2007).

B5.2.2 The Green–Ampt Infiltration Equation

There are other, more direct, ways of deriving an infiltration equation from the Richards equation. Soil physical theory suggests that infiltration can be described (at least in the absence of major macropores) by the Richards equation which is based on the nonlinear form of Darcy's law for partially saturated flow (see Box 5.1). While there are no general analytical solutions to the Richards equation, a number of different solutions are available for infiltration at the soil surface based on different simplifying assumptions. Green and Ampt (1911), for example, assumed that the infiltrating wetting front forms a sharp jump from a constant initial moisture content ahead of the front to saturation at the front. This allows a simple form of Darcy's law to be used to represent the infiltration such that infiltration rate f is calculated as

$$f(t) = \tilde{K}_s \left(\frac{h_o + \psi_f}{z_f} + 1 \right) \tag{B5.2.2}$$

where $\tilde{K}_s$ is the hydraulic conductivity of the soil at field saturation, h_o is the depth of ponded water on the soil surface, ψ_f is a parameter related to the difference in capillary potential across the wetting front and z_f is the depth of penetration of the wetting front. The $\frac{h_o+\psi_f}{z_f}$ term is due to the capillary potential gradient, here estimated from an effective difference in capillary potential across the wetting front averaged over the depth of penetration. The magnitude of this term reduces as the wetting front goes deeper leading to the decline in infiltration capacity of the soil. This term is added to the gravitational term (1). It stays at unity regardless of the depth of the wetting front and, when multiplied by the effective saturated hydraulic conductivity, gives the final infiltration capacity.

Morel-Seytoux and Khanji (1974) have shown that a parameter called the "capillary drive", with units of length and defined as $C_D = \psi_f \delta\tilde{\theta}$, is a relatively constant parameter for a range of initial moisture conditions defined as $\delta\tilde{\theta} = (\tilde{\theta} - \theta_i)$ which is the change in moisture content between the initial state of the soil, θ_i, and field saturation, $\tilde{\theta}$. The Green–Ampt equation is then better applied in the form:

$$f(t) = \frac{\tilde{K}_s}{B}\left(\frac{h_o \delta\tilde{\theta} + C_D}{z_f \delta\tilde{\theta}} + 1\right) \tag{B5.2.3}$$

where B is an additional parameter proposed by Morel-Seytoux and Khanji (1974) to allow for air pressure effects, called the "viscous resistance correction factor" ($1 < B < 1.7$).

The original Green–Ampt infiltration equation assumes constant soil characteristics with depth. An analysis of a wide variety of soil moisture characteristics data by Rawls *et al.* (1983; see also Box 5.5) has led to a classification of the Green–Ampt parameters by soil texture (see Figure B5.2.2). These type of relationship should be used with care, however, since they are based on measurements of small samples brought back to the laboratory, not field measurements at the plot scale. Beven (1984) has produced a solution with equivalent assumptions for the case where hydraulic conductivity declines exponentially with depth, often a useful approximation of real soil characteristics.

B5.2.3 The Philip Infiltration Equation

Philip (1957) obtained an analytical solution to the Richards equation by assuming a delta function change in diffusivity for the soil across the wetting front. His widely used infiltration equation has the form

$$f(t) = 0.5St^{-0.5} + A \tag{B5.2.4}$$

where S is called the "sorptivity" of the soil and is calculated from knowledge of the soil moisture characteristics of the soil and A is a final infiltration capacity equivalent to the f_c of the Horton equation or $\tilde{K}_s$ of the Green–Ampt equation. The effects of the sorptivity term gradually reduce with increasing time, eventually leaving the final infiltration capacity as a function of the effective saturated hydraulic conductivity of the soil.

B5.2.4 The Smith–Parlange Infiltration Equation

Under assumptions of a diffusivity that changes exponentially with moisture content, Smith and Parlange (1978) derived another widely used infiltration equation that takes the form:

$$f(t) = \tilde{K}_s \frac{\exp F(t)/C_D}{\exp F(t)/C_D - 1} \tag{B5.2.5}$$

where $F(t) = \int_0^t f(t)dt$, and C_D is the capillary drive, as above.

B5.2.5 An Infiltration Equation Based on Storage Capacity

Treating infiltration capacity as a function of infiltrated volume can also be used to treat the case where overland flow is produced as a result of the topsoil layer becoming saturated as a result of a limitation on vertical flow at some depth within the soil. This can occur either where a thin soil overlies an impermeable bedrock or where there is some horizon of lower permeability at some depth into the soil profile (e.g. Taha *et al.*, 1997). In this circumstances, infiltration rates might be controlled more by a saturation excess than by a surface infiltration

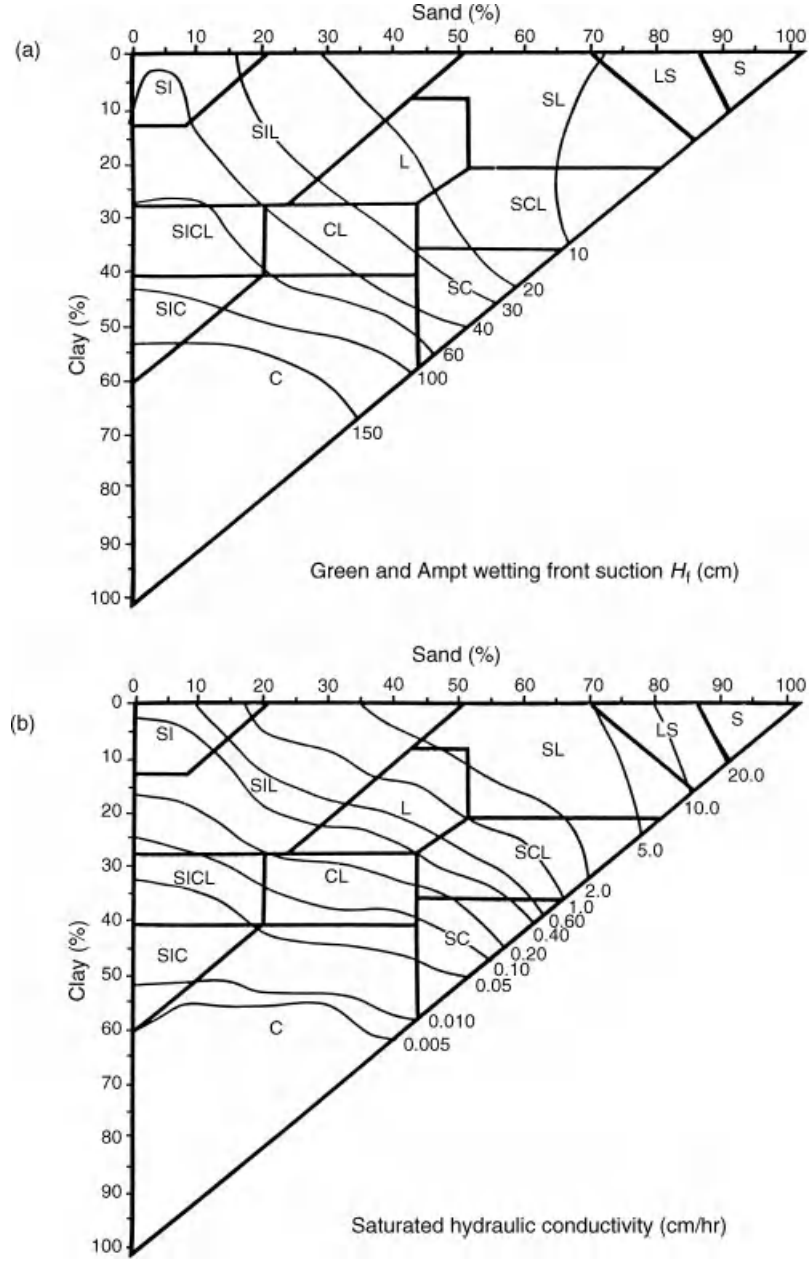


Figure B5.2.2 *Variation of Green–Ampt infiltration equation parameters with soil texture (after Rawls* et al.*, 1983, with kind permission from Springer Science+Business Media B.V).*

excess process. This situation was addressed by Kirkby (1975; see also Scoging and Thornes, 1982) using an infiltration equation of the form:

$$f(t) = B/H + A \tag{B5.2.6}$$

where H is the current depth of storage, A is the long term infiltration rate (which may now be controlled at depth) and B is a constant. Updating of H at successive time steps allows application to any irregular sequence of rainfall inputs using:

$$\text{for r(t)} > \text{f(t)}, \ \frac{dH}{dt} = r(t) - f(t)$$

and

$$\text{for r(t)} < \text{f(t)}, \ \frac{dH}{dt} = r(t) - A$$

For steady rainfall inputs, this results in infiltration capacity being an inverse function of time rather than the square root of time, as in the Philip (1957) equation.

B5.2.6 Time to Ponding and the Time Compression Assumption

Whatever type of function is used to describe infiltration, effective rainfalls are calculated as the excess of rainfall over infiltration following the time to ponding. For rainfall intensities that are irregular in time, this is often most easily done in terms of comparing the cumulative rainfall and cumulative infiltration rates, estimating the time of ponding as the point at which the cumulative infiltration satisfies one of the solutions above (this is called the *time compression assumption)*.

Thus, using the Green–Ampt equation as an example, the cumulative infiltration at any time up to the time of ponding is the integral of the sequence of rainfall intensities:

$$F(t) = \int_0^{t_p} r(t)\, dt \tag{B5.2.7}$$

Under the sharp wetting front assumption, this is also equal to

$$F(t) = z_f(\tilde{\theta} - \theta_i) \tag{B5.2.8}$$

Thus, at each time step the depth of the wetting front can be calculated and used in Equation (B5.2.3) to calculate the current infiltration capacity, f_t. If this is less than the current rainfall intensity, r_t, then the time of ponding has been reached. Similar arguments can be used for other surface control infiltration equations to determine the time of ponding (see, for example, Parlange *et al.*, 1999). For storage-based approaches, a minimum storage before surface saturation can be introduced as an additional parameter if required.

B5.2.7 Infiltration in Storms of Varying Rainfall Intensity

In storms of varying rainfall intensity, the input rate may sometimes exceed the infiltration capacity of the soil and sometimes not. When input rates are lower than the infiltration capacity of the soil, there may be the chance for redistribution of water within the soil profile to take place, leading to an increase in the infiltration capacity of the soil. This is handled easily in storage capacity approaches to predicting infiltration, such as that of Kirkby (1975) noted above, but such approaches do not have a strong basis in soil physics. Clapp *et al.* (1983) used an approach based on a Green–Ampt or kinematic approach to the infiltration and redistribution of successive wetting fronts. More recently Corradini *et al.* (1997) have produced a similar approach based on a more flexible profile shape than the piston-like wetting front of the Green–Ampt approach. It is worth noting that methods based on soil physics may also be limited in

their applicability in this respect since they take no account of the effects of macropores in the soil on infiltration and redistribution (though see the work of Beven and Clarke (1986) for one attempt at an infiltration model that takes account of a population of macropores of limited depth).

B5.2.8 Derivation of Soil Moisture Characteristics from Infiltration Measurements

Measuring the rate at which water infiltrates into the soil surface is one of the simplest ways of trying to assess the hydrological characteristics of soils. A number of techniques are used, including ponding water within single and double ring infiltrometers, porous plate infiltrometers with pressure controls that can be used to exclude infiltration into macropores, and plot scale sprinkler experiments. There is a long history of using analytical solutions to the infiltration equation (including three dimensional effects beneath an infiltration ring of limited dimension) to derive soil moisture characteristics from infiltration rates. A recent example is provided by the BEST technique of Lassabatère *et al.* (2006) that is based on scaled forms of infiltration equation presented by Braud *et al.* (2005).

B5.2.9 Estimating Infiltration at Larger Scales

All of the equations above are for the prediction of infiltration at the point scale. They have often been applied wrongly at hillslope or catchment scales, as if the same point scale equation applied for the case of heterogeneous soil properties. However, because of the nonlinearities inherent in any of the infiltration equations, this is not the case. It is therefore necessary to predict a distribution of point scale infiltration rates before averaging up to larger scales. Similar media theory (see Box 5.4) can be useful in this case, if it can be assumed that the soil can be treated as a distribution of parallel non-interacting heterogeneous columns. Clapp *et al.* (1983), for example, have based a field scale infiltration model on this type of representation. Philip (1991) has attempted to extend the range of analytical solutions to the case of infiltration over a sloping hillslope, albeit a hillslope of homogeneous soil characteristics.

There have been many hypothetical studies of the effects of random variability of soil properties and initial conditions on infiltration at the hillslope and catchment scales, some taking account of the infiltration of surface flow from upslope, others not (see, for example, the study of Corradini *et al.*, 1998). These have mostly been based on purely stochastic heterogeneity but it is worth noting that some variation may be systematic (see, for example, the dependence of infiltration on depth of overland flow and vegetation cover included in the model of Dunne *et al.*, 1991). To apply such models to a practical application will require considerable information on the stochastic variation in soil properties and, even then, will not guarantee accurate predictions of runoff generation (see the discussion of the R-5 catchment case study in Section 5.7).

Storage-based approaches need not be considered only as point infiltration models. A function such as that of Equation (B5.2.6) may equally be considered as a conceptual representation of basin wide infiltration, with the storage as a mean basin storage variable. Other storage approaches have taken a more explicit representation of a distribution of storage capacities in a catchment in predicting storm runoff. These include the Stanford Watershed model and variable infiltration capacity models (see Section 2.4 and Box 2.2), and the probability distributed model (see Section 6.2).

A further widely used method for predicting runoff generation at larger scales, often interpreted as an infiltration model, is the USDA Soil Conservation Service (SCS) curve number approach. This is discussed in more detail in Chapter 6 and Box 6.3, where it is shown that it can be interpreted in terms of both spatially heterogeneous infiltration rates and dynamic contributing areas generating saturation excess surface (and possibly subsurface) runoff.

Box 5.3 Solution of Partial Differential Equations: Some Basic Concepts

As noted in the main text, it is nearly always not possible to obtain analytical solutions to the nonlinear differential equations describing hydrological flow processes for cases of real interest in practical applications, such as rainfall–runoff modelling. The approximate numerical solution of nonlinear differential equations is, in itself, a specialism in applied mathematics and writing solution algorithms is something that is definitely best left to the specialist. A major reason for this is that it is quite possible to produce solution algorithms that are inaccurate or are not **consistent** with the original differential equation (i.e. the approximate solution does not converge to the solution of the original equation as the space and time increments become very small). It is also easy to produce solutions for nonlinear equations that are not **stable**. A stable solution means that any small errors due to the approximate nature of the solution will be damped out. An unstable solution means that those small errors become amplified, often resulting in wild oscillations of the solution variable at adjacent solution nodes or successive time steps. The aim of this box is to make the reader aware of some of the issues involved in approximate numerical solutions and what to look out for in using a model based on one of the numerical algorithms available.

The differential equations of interest to the hydrologist (such as the Richards equation of Box 5.1) generally involve one or more space dimensions and time. An approximate solution then requires a discretisation of the solution in both space and time to produce a grid of points at which a solution will be sought for the dependent variable in the equation. Figure B5.3.1 shows some ways of subdividing a cross-section through a hillslope into a grid of solution points or nodes in space and Figure B5.3.2 shows a regular discretisation into both time and space increments, Δt and Δx, for a single spatial dimension.

Starting a numerical solution always requires a complete set of nodal values of the solution variable at time $t = 0$. The algorithm then aims to step the solution through time, using time steps of length Δt, to obtain values of the variable of interest for all the nodes at each time step. The most easily understood numerical approximation method is the *finite difference* method. The original SHE model (Section 5.2.2) uses a finite difference solution of the surface and subsurface flow equations on a regular spatial grid. The differentials in the flow equation are replaced directly by differences. For example, to calculate a spatial differential for a variable ψ at node i in the interior of a grid at time step j, one possible difference approximation is:

$$\left\{\frac{\partial \psi}{\partial x}\right\}_{i,j} \approx \frac{\psi_{i+1,j} - \psi_{i-1,j}}{2\Delta x} \tag{B5.3.1}$$

This form is called a "centred difference approximation". It is a more accurate approximation than the forward difference that is often used for the time differential:

$$\left\{\frac{\partial \psi}{\partial t}\right\}_{i,j} \approx \frac{\psi_{i,j+1} - \psi_{i,j}}{\Delta t} \tag{B5.3.2}$$

The forward difference is convenient to use for any differential in time since, given the nodal values of ψ at time step j, the only unknowns are the values at time step $j + 1$. A centered approximation for a second-order spatial differential involving a spatially variable coefficient, K, may be written as

$$\begin{aligned}\left\{\frac{\partial}{\partial x}\left(K\frac{\partial \psi}{\partial x}\right)\right\}_{i,j} &\approx \frac{K_{i,j}}{\Delta x}\left(\frac{\psi_{i+1,j} - \psi_{i,j}}{\Delta x} - \frac{\psi_{i,j} - \psi_{i-1,j}}{\Delta x}\right) \\ &\approx K_{i,j}\left(\frac{\psi_{i+1,j} - 2\psi_{i,j} - \psi_{i-1,j}}{\Delta x^2}\right)\end{aligned}$$

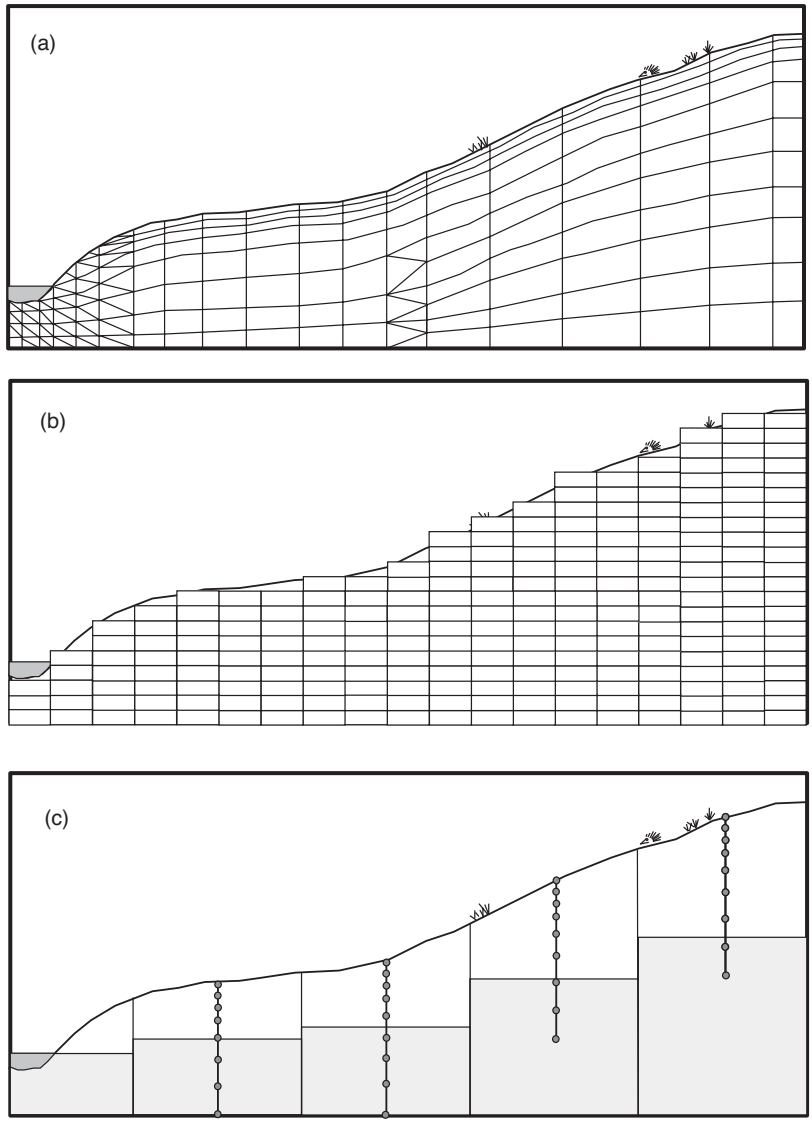


Figure B5.3.1 *Discretisations of a hillslope for approximate solutions of the subsurface flow equations: (a) Finite element discretisation (as in the IHDM); (b) Rectangular finite difference discretisation (as used by Freeze, 1972); (c) Square grid in plan for saturated zone, with one-dimensional vertical finite difference discretisation for the unsaturated zone (as used in the original SHE model).*

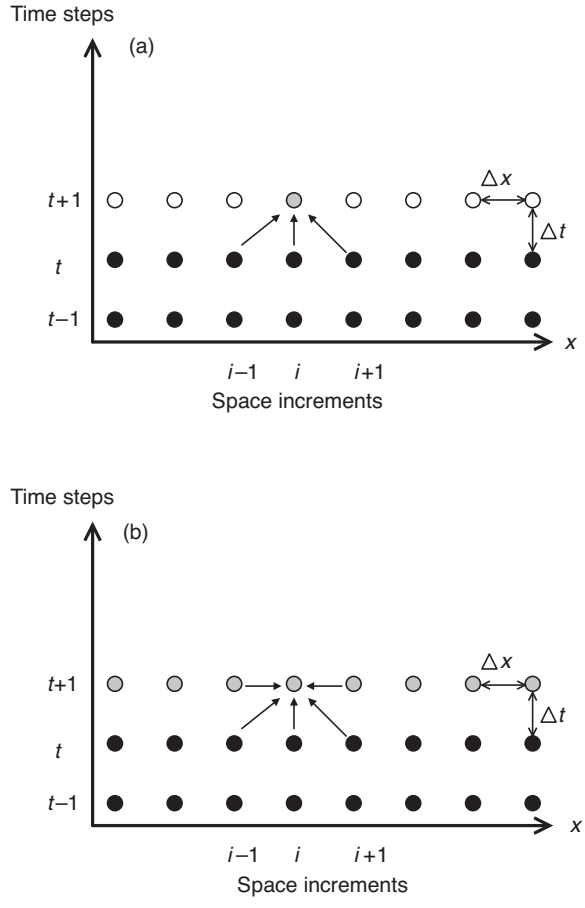


Figure B5.3.2 *Schematic diagram of (a) explicit and (b) implicit time stepping strategies in approximate numerical solutions of a partial differential equation, here in one spatial dimension, x (arrows indicate the nodal values contributing to the solution for node* $(i, t+1)$*; nodal values in black are known at time t; nodal values in grey indicate dependencies in the solution for time* $t+1$*).*

When an equation involves both space and time differentials, the simplest (but not necessarily the most accurate) solution is to evaluate all the spatial differentials at time step j and use them to solve for the dependent variable at time step $j+1$. In this strategy, the calculations for each node can be carried out independently of all the other nodal values at time $j+1$ as shown in Figure B5.3.2a. This is known as a solution that is *explicit* in time. It results in a solution algorithm that has many attractions in that it is not necessary to assemble large matrices and the solution is easily implemented on parallel computers. The disadvantage is that the solution is only conditionally stable, especially for problems that are strongly nonlinear, such as those encountered in hydrology. Thus, it is necessary to ensure that the time step is short enough that the solution remains stable. For nonlinear problems, this may mean that very small time steps are required. It is possible to calculate a stability criterion and to ensure that the time step is small enough to ensure stability. The smaller the spatial discretisation used, in general, the smaller the time step that is required to ensure stability. Such criteria take the form $\Delta t = c\Delta x$, where c is a local wave speed and Δx is the local space increment of the model discretisation.

There are some types of differential equations, known as elliptic equations, for which explicit methods are not suitable since the effective wave speed is theoretically infinite. In hydrology, flow in a saturated soil or aquifer, where the effective specific moisture capacity or storage coefficient is very small, results in a quasi-elliptic equation which is one reason why some physically based models have had significant problems with stability of solutions.

A more robust method arises if some form of averaging of the estimates of the spatial differentials at time steps j and $j+1$ is used, e.g.

$$\left\{\frac{\partial\psi}{\partial x}\right\}_{i,j} \approx \varpi\left(\frac{\psi_{i+1,j+1}-\psi_{i-1,j+1}}{2\Delta x}\right)+(1-\varpi)\left(\frac{\psi_{i+1,j}-\psi_{i-1,j}}{2\Delta x}\right) \tag{B5.3.3}$$

where ϖ is a weighting coefficient in time. An *explicit solution* has $\varpi = 0$. Any solution algorithm with $\varpi > 0$ is known as an *implicit* scheme (Figure B5.3.2b). A central difference, or Crank–Nicholson scheme, has $\varpi = 0.5$, while a backward difference or fully implicit scheme has $\varpi = 1$. For linear problems, all implicit algorithms with $\varpi > 0.5$ can be shown analytically to be unconditionally stable, but no such general theory exists for nonlinear equations. The central difference approximation is superior to the backward difference scheme in principle, since the truncation error of the approximation is smaller. However, the backward difference scheme has been found to be useful in highly nonlinear problems. Details of a four-point implicit finite difference scheme to solve the one dimensional St. Venant channel flow equations are given in Box 5.6; a simpler implicit scheme to solve the one-dimensional kinematic wave equation is given in Box 5.7.

With an implicit scheme, the solution for node i at time step $j+1$ involves values of the dependent variable at other surrounding nodes, such as at $i+1$ and $i-1$ at time step $j+1$, which are themselves required as part of the solution (as shown in Figure B5.3.2b). Thus it is necessary to solve the problem as a system of simultaneous equations at each time step. This is achieved by an iterative method, in which an initial guess of the values at time $j+1$ is used to evaluate the spatial differentials in the implicit scheme. The system of equations is then solved to get new estimates of the values at time $j+1$ which are then used to refine the estimates of the spatial differentials. This iterative procedure is continued until the solution converges towards values that change by less than some specified tolerance threshold between successive iterations. A successful algorithm results in a stable solution with rapid convergence (a small number of iterations) at each time step. At each iteration, the system of equations is assembled as a matrix equation, linearised at each iteration for nonlinear problems in the form:

$$[A]\{\Psi\} = \{B\} \tag{B5.3.4}$$

where $[A]$ is a two-dimensional matrix of coefficients known at the current iteration, $\{\Psi\}$ is a one-dimensional vector of the unknown variable and $\{B\}$ is a one-dimensional vector of known values. Explicit schemes can also be expressed in this form but are solved only once at each time step. For a large number of nodes, the $[A]$ matrix may be very large and **sparse** (i.e. having many zero coefficients) and many algorithms use special techniques, such as indexing of non-zero coefficients, to speed up the solution.

Solution methods may be direct, iterative or a combination of the two. Direct methods carry out a solution once and provide an exact solution to the limit of computer round-off errors. The problem with direct methods is that, with a large sparse $[A]$ matrix, the computer storage required may be vast. Examples of direct methods include Gaussian elimination and Cholesky factorisation.

Iterative methods attempt to find a solution by making an initial estimate and then refining that estimate over successive iterations until some error criterion is satisfied. For implicit solutions of nonlinear problems, the nonlinear coefficients can be refined at each iteration. Iterative methods may not converge for all nonlinear problems. Steady state problems, for example, often pose greater problems in obtaining a solution than transient problems and convergence may depend on making a good initial guess at the solution. This is often easier in the transient

case where a good initial guess will be available by extrapolating the behaviour of the solution variable from previous time steps. Iterative methods can, however, be fast and generally require much less computer storage than direct methods for large problems. Examples of iterative solution methods include Picard iteration, Newton iteration, Gauss–Seidal iteration, successive over-relaxation and the conjugate gradient method. Paniconi *et al.* (1991) have compared the Picard and Newton iterations in solving the nonlinear Richards equation by a *finite element method*. The Newton iteration method requires the calculation of a gradient matrix at each iteration, but will generally converge in fewer iterations. They suggested that it was more efficient in certain strongly transient or highly nonlinear cases (but also recommend the non-iterative Lees method for consideration). The pre-conditioned form of the conjugate gradient method is very popular in the solution of groundwater problems and is readily implemented on parallel computers (Binley and Beven, 1992).

Although we have discussed the considerations of explicit and implicit schemes, direct and iterative solutions, stability and convergence in the context of finite difference approximations, they apply to all solution algorithms including finite element and finite volume techniques. Note that with all these schemes, it is possible to have an algorithm that is stable and consistent but not accurate, if an injudicious choice of space and time increments is made. It may be very difficult to check accuracy in a practical application except by testing the sensitivity of the solution to reducing the space and time increments. This is something that is often overlooked in the application of numerical methods since checks on the accuracy of the solution may often be expensive to carry out.

Finite element solutions are also commonly used in hydrological problems, such as in the InHM and HydroGeoSphere (VanderKwaak and Loague, 2001; Therrien *et al.*, 2006). The finite element method has an important advantage over finite difference approximations in that flow domains with irregular external and internal boundaries are more realistically represented by elements with straight or curved sides. No-flow and specified flux boundaries are also more easily handled in the finite element method. The solution nodes in the finite element method lie (mostly at least) along the boundaries of the elements. Spatial gradients are represented by interpolating the nodal values of the variable of interest within each element using *basis functions*. The simplest form of basis function is simple linear interpolation. Higher order interpolation can be used, but requires more solution nodes within each element. Ideally, the form of interpolation should be guided by the nature of the problem being solved but this is often difficult for problems involving changes over time and in which different solution variables are nonlinearly related, implying that different basis functions should be used. A higher order finite element interpolation for time differentials can also be used.

Over the last decade, finite volume techniques have become more popular. In the finite volume method, each node in the solution grid is surrounded by a small volume of the flow domain. Since these need not be regularly spaced or a particular shape the finite volume technique is well suited for unstructured grids representing complex grids. Its other main advantage is that it is inherently mass conservative. This is because divergence terms in the partial differential equation are converted to fluxes across the boundaries of the finite volume using the divergence theorem (see LeVeque, 2002). The change in mass in each finite volume is then calculated exactly from the net balance of the surface fluxes across all the boundaries. The finite volume approach has been used in hydrological models for surface flows in the SFV model of Horritt *et al.* (2007), for saturated subsurface flows by Jenny *et al.* (2003) and for coupled surface and subsurface flows by Qu and Duffy (2007), He *et al.* (2008) and Kumar *et al.* (2009).

There is a vast literature on numerical solution algorithms for differential equations. A good detailed introduction to finite element and finite difference algorithms used in hydrology may be found in Pinder and Gray (1977). The recent developments in finite volume methods also appear to be promising. However, it is always important to remember that all these methods are approximations, especially in the case of nonlinear problems, and that approximations involve inaccuracies even if the solution is stable, convergent and mass conserving. One form of inaccuracy commonly encountered is that of numerical dispersion. It is an important

consideration in problems involving strong advection, such as the propagation of a sharp wetting front into the soil, the movement of a steep flood wave down a river or the transport of a contaminant in a flow with a steep concentration gradient. Numerical solutions using a fixed nodal grid will inevitably smear out any rapid changes in moisture content or concentration in such problems. This smearing is called "numerical dispersion" and becomes greater as the grid gets coarser, even if the solution apparently stays stable. Sharp fronts may also lead to oscillation in many solution schemes, a product of the approximate solution not the process. There have been some techniques, such as "upwind" differencing, developed to try to minimise such problems, especially in one-dimensional advection problems, but the lesson is that these types of solution must be used with care.

Box 5.4 Soil Moisture Characteristic Functions for Use in the Richards Equation

Use of the Richards equation to predict water flow in unsaturated soil (see Box 5.1) requires the specification of the nonlinear functions $C(\psi)$ and $K(\psi)$ if solving for capillary potential ψ, or $D(\theta)$ and $K(\theta)$ with moisture content θ, as the dependent variable. These functions are time consuming to measure directly, even in the laboratory, and are generally complicated by being multi-valued hysteretic functions dependent on the history of wetting and drying. For modelling purposes, it is often assumed that simpler, single-valued functions can be used. Two sets of widely used functions are presented here: those suggested by Brooks and Corey (1964) and those of van Genuchten (1980). Both specify forms for $\theta(\psi)$ and $K(\theta)$, from which the specific moisture capacity, $C(\psi)$, and diffusivity, $D(\psi)$, can be derived. Smith *et al.* (1993) proposed a form that is an extension of both the Brooks–Corey and van Genuchten forms but this has not been widely used. Some theoretical developments have involved developing functional forms based on assumptions of a fractal pore space (see, for example, Tyler and Wheatcraft, 1992; Pachepsky and Timlin, 1998). Jaynes (1990) reviews multi-valued hysteretic forms.

B5.4.1 The Brooks–Corey Functions

In the Brooks–Corey functions, moisture content and capillary potential are related as

$$\frac{\theta - \theta_r}{\theta_s - \theta_r} = \left(\frac{\psi_o}{\psi}\right)^{\lambda} \tag{B5.4.1}$$

while hydraulic conductivity and moisture content are related as

$$\frac{K(\theta)}{K_s} = \left(\frac{\theta - \theta_r}{\theta_s - \theta_r}\right)^{3+2\lambda} \tag{B5.4.2}$$

B5.4.2 The van Genuchten Functions

In the van Genuchten functions, moisture content and capillary potential are related as

$$\frac{\theta - \theta_r}{\theta_s - \theta_r} = \left(\frac{1}{1 + \{\psi/\psi_o\}^{\lambda+1}}\right)^{\frac{\lambda}{\lambda+1}} \tag{B5.4.3}$$

while hydraulic conductivity and moisture content are related as

$$\frac{K(\theta)}{K_s} = \left(\frac{\theta - \theta_r}{\theta_s - \theta_r}\right)^{\lambda+1} \left[1 - \left(1 - \left\{\frac{\theta - \theta_r}{\theta_s - \theta_r}\right\}^{\frac{\lambda+1}{\lambda}}\right)^{\frac{\lambda}{\lambda+1}}\right]^2 \tag{B5.4.4}$$

B5.4.3 Comparing the Soil Moisture Characteristics Functions

Both sets of relationships have five parameters that define the shape of the functions:

- θ_s is the saturated porosity of the soil;
- θ_r is the residual moisture content;

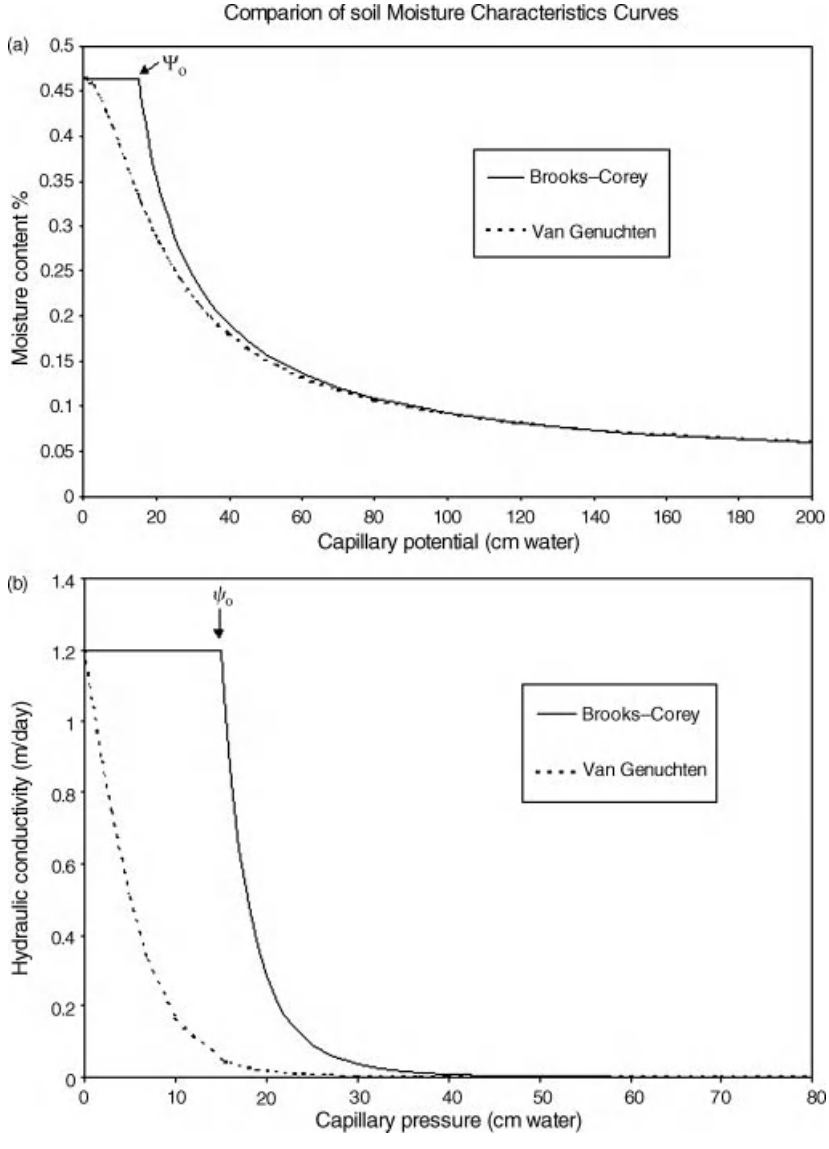


Figure B5.4.1 *Comparison of the Brooks–Corey and van Genuchten soil moisture characteristics functions for different values of capillary potential: (a) soil moisture content and (b) hydraulic conductivity.*

- K_s is the saturated hydraulic conductivity;
- ψ_o is called the "bubbling" potential;
- λ is called the "pore size" index.

Note that both ψ_o and λ are best viewed as fitting rather than physical parameters and may take different numerical values in the Brooks–Corey and van Genuchten equations when both functions are fitted to the same set of data.

The two functions, calculated with identical values of the five parameters, are plotted in Figure B5.4.1. The smoothness of the van Genuchten functions in the area close to saturation is numerically advantageous when used in modelling (and, in most cases, is probably a more realistic representation of the soil wetting and drying characteristics). A variant on the Brooks–Corey relationships has been suggested by Clapp and Hornberger (1978), using a parabolic function to smooth the sharp break at the bubbling potential, ψ_o, while Mohanty *et al.* (1997) have suggested piecewise functions that can take account of a rapid increase in hydraulic conductivity due to preferential flows close to saturation.

B5.4.4 Similar Media Scaling in Representing the Soil Moisture Characteristics Functions

An early attempt to deal with the heterogeneity of soil properties that has proven to be quite robust is the local scale scaling theory of Miller and Miller (1956). They derived relationships for the soil characteristic of different soils that were geometrically similar in terms of the shape and packing of the particles, differing only in terms of a characteristic length scale (Figure B5.4.2). In this case, the water content of the soil, θ, can be related to the capillary potential, ψ, as

$$\theta = \Omega_1 \left(\frac{\alpha}{\sigma} \gamma \psi \right) \tag{B5.4.5}$$

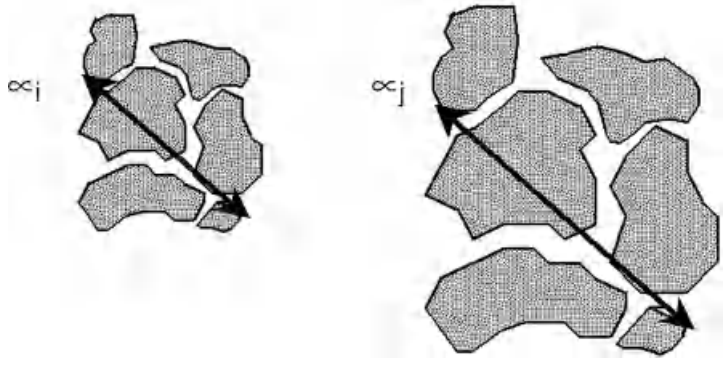


Figure B5.4.2 *Scaling of two similar media with different length scales α.*

where σ is surface tension, γ is the specific weight of water, α is the characteristic length and the function $\Omega_1()$ is the same for all soils scaled in this way.

A similar relationship can be derived for the hydraulic conductivity of similar soils as

$$K = \frac{\alpha^2}{\gamma \mu} \Omega_2(\theta) \tag{B5.4.6}$$

where μ is the dynamic viscosity of water and Ω_2 is again constant for soils that are geometrically similar.

Thus, the soil moisture characteristics of similar soils can be scaled by representing their variability in terms of the distribution of the characteristic length α. Note that a consequence of the geometric similarity is that the porosity of the soil θ_s is a constant for all similar soils. This type of scaling has been applied in many studies (e.g. Simmons *et al.*, 1979; Clausnitzer *et al.*, 1992); its relevance here is as a way of incorporating a simple description of soil variability into predictions of infiltration and runoff generation (e.g. Luxmoore and Sharma, 1980; Clapp *et al.*, 1983; Warrick and Hussen, 1993).

Using the Brooks–Corey characteristics as an example, assuming that we have knowledge of the parameters for one particular soil with α=1, then for any other similar soil:

$$\frac{\theta - \theta_r}{\theta_s - \theta_r} = \left(\alpha \frac{\psi_o}{\psi}\right)^{\lambda} \tag{B5.4.7}$$

while saturated hydraulic conductivity is given by

$$K_s(\alpha) = \alpha^2 K_s^1 \tag{B5.4.8}$$

where K_s^1 is the saturated hydraulic conductivity of the reference soil. The expression for relative hydraulic conductivity, Equation (B5.4.2), stays the same.

The analysis can be extended to other parameters dependent on the soil moisture characteristics of the soil. Thus, the Philip infiltration equation (see Box 5.2) can be scaled as:

$$f(t) = 0.5\alpha^{1.5} S^1 t^{-0.5} + \alpha^2 A^1 \tag{B5.4.9}$$

where S^1 and A^1 are the sorptivity and final infiltration capacity parameters for a reference soil with α=1. An analysis of multiple infiltration curves is one way of calculating a distribution of α values for a particular soil type (see for example, Shouse and Mohanty, 1998). An example of the scaling of infiltration curves in this way is demonstrated in Figure B5.4.3.

The Miller and Miller *similar media* concept is just one possible scaling theory that could be used to provide a simple representation of heterogeneity of soil properties. A number of other possibilities have been reviewed by Tillotson and Nielsen (1984). Mohanty (1999) has proposed a method for scaling the properties of soils with macropores, treating their soil moisture characteristics as continuous curves spanning the matrix/macropore pore sizes. This assumes that such a dual porosity medium responds as a continuum, which may not always be a good assumption.

B5.4.5 Identification of the Soil Moisture Characteristics Function Parameters

Both the Brooks–Corey and the van Genuchten soil moisture characteristics functions require the calibration of a number of parameter values before they can be used in a predictive model. Both laboratory and field measurements may be used in calibration, as well as indirect techniques based on *pedotransfer functions* (discussed in Box 5.5). The resulting parameter values are not independent of the measurement techniques and parameter identification techniques used.

In principle, the soil moisture characteristics and conductivity characteristics at different capillary pressures can be measured directly in the laboratory. Different functions can then be fitted directly to the measured curves using a nonlinear least squares regression procedure (see for example Kool *et al.*, 1987). However, such measurements are only currently possible on small samples which tend to show significant heterogeneity from sample to sample, especially close to saturation.

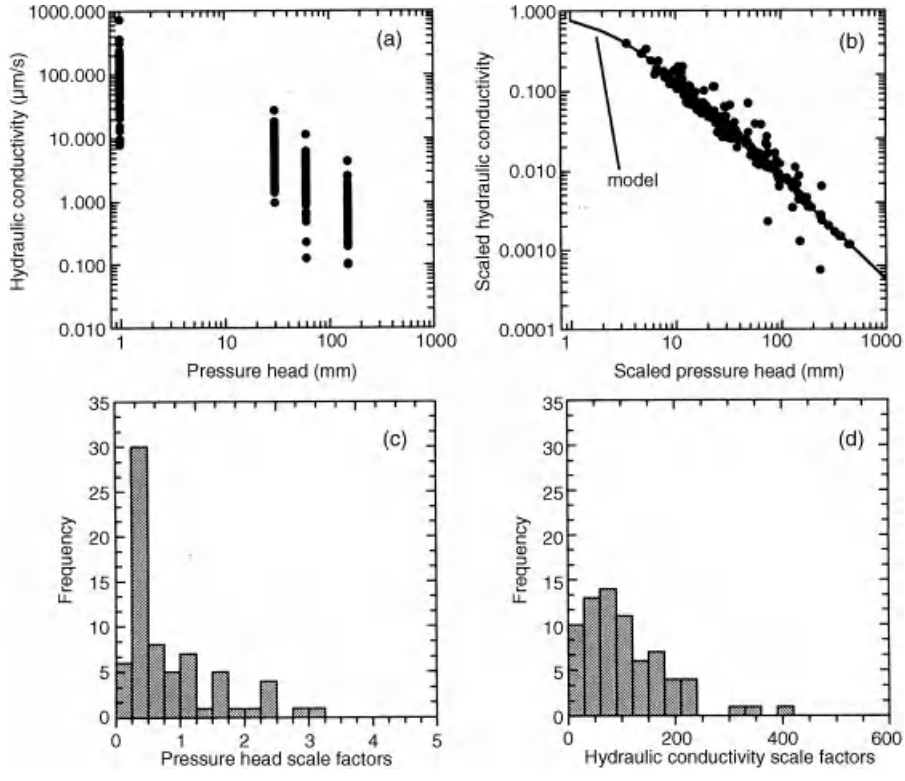


Figure B5.4.3 *Scaling of unsaturated hydraulic conductivity curves derived from field infiltration measurements at 70 sites under corn rows on Nicollet soil, near Boone, Iowa (after Shouse and Mohanty, 1998, with kind permission of the American Geophysical Union).*

Identification is also possible from field measured infiltration rates (see Box 5.2) or moisture profiles by an inverse identification using a model based on the Richards equation. The parameter values are adjusted until a good fit is obtained between observed and predicted discharges. Similar methods can be used on large undisturbed columns of soil in the laboratory (but repacked soil columns may not be representative of field conditions). Because of interactions between the parameters, physical effects (such as hysteresis) and, perhaps, preferential flows, it is often not possible to find a single parameter set that gives a best fit to the data (see Mishra and Parker, 1989; Abeliuk and Wheater, 1990).

Box 5.5 Pedotransfer Functions

Pedotransfer functions allow the estimation of the soil moisture characteristics curves on the basis of information about more easily measured variables, particularly textural variables. Pedotransfer functions are developed from experimental measurements made on a large number of samples. Two types of pedotransfer functions may be distinguished. In the first, equations are developed for values of moisture content and hydraulic conductivity for specific values

of capillary potential (e.g. Gupta and Larson, 1979). In the second, equations are developed to estimate the parameters of functional forms of the soil moisture characteristics, such as the Brooks–Corey or van Genuchten curves of Box 5.4, or parameters of application-specific functions, such as the infiltration equations of Box 5.2 (e.g Rawls and Brakensiek, 1982; Cosby *et al.*, 1984; van Genuchten *et al.*, 1989; Vereeken *et al.*, 1989, 1990; Schaap and Leij, 1998).

Rawls and Brakensiek (1982) provided regression equations for the Brooks–Corey functions as a function of soil properties based on several thousand sets of measurements collected by the US Department of Agriculture (USDA). A summary of this regression approach is given by Rawls and Brakensiek (1989). A typical equation for K_s in terms of the variables C (clay; 5% $< C <$ 60%), S (sand; 5% $< S <$ 70%), and porosity, θ_s is:

$$\begin{aligned} K_s = \exp[&19.52348\theta_s - 8.96847 - 0.028212C + 0.00018107S^2 \\ &-0.0094125C^2 - 8.395215\theta_s^2 + 0.077718S\theta_s - 0.00298S^2\theta_s^2 \\ &-0.019492C^2\theta_s^2 + 0.0000173S^2C + 0.02733C^2\theta_s \\ &+0.001434S^2\theta_s - 0.0000035C^2S] \end{aligned}$$

where K_s is in cmh^{-1} and the porosity θ_s can be estimated from measured dry bulk density ρ_d as

$$\theta_s = 1 - \frac{\rho_d}{\rho_s} \tag{B5.5.1}$$

where ρ_s is the density of the soil mineral material ($\approx$ 2650 kgm^{-3}) or, from another equation, to estimate θ_s in terms of C, S, % organic matter, and cation exchange capacity of the soil. These are all variables that are often available in soil databases. Rawls and Brakensiek also provide equations for adjusting porosity to allow for entrapped air, to correct for frozen ground and for surface crusting, to account for the effects of management practices, and for the parameters of various infiltration equations including the Green–Ampt and Philip equations of Box 5.2. Some special pedotransfer functions have also been developed, such as those of Brakensiek and Rawls (1994) to take account of the effects of soil stoniness on infiltration parameters.

It is necessary to use all these equations with some care. Equations such as that for K_s above have been developed from data generally collected on small samples (the USDA standard sample for the measurement of hydraulic conductivity was a "fist-sized fragment" (Holtan *et al.*, 1968) that would exclude any effects of macroporosity). There is also considerable variability within each textural class. The apparent precision of the coefficients in this equation is, therefore, to some extent misleading. Each coefficient will be associated with a significant standard error, leading to a high uncertainty for each estimate of K_s. In the original paper of Rawls and Brakensiek (1982), the order of magnitude of these standard errors is given. In some later papers, this is no longer the case. The estimates provided by these equations are then apparently without uncertainty. This gives plenty of potential for being wrong, especially when in the application of a catchment scale model it is the effective values of parameters at the model grid element scale that are needed. Some evaluations of the predictions of pedotransfer functions relative to field measured soil characteristics have been provided by Espino *et al.* (1995), Tietje and Tapkenhinrichs (1993), Romano and Santini (1997) and Wagner *et al.* (1998).

An alternative approach to deriving pedotransfer functions has been to use neural network methods (e.g. those by Schaap and Bouten, 1996; and Schaap *et al.*, 1998). Uncertainty in the resulting functions cannot, then, be predicted by standard regression confidence estimates but can be estimated using a bootstrap method (Shaap and Leij, 1998). Figure B5.5.1 gives

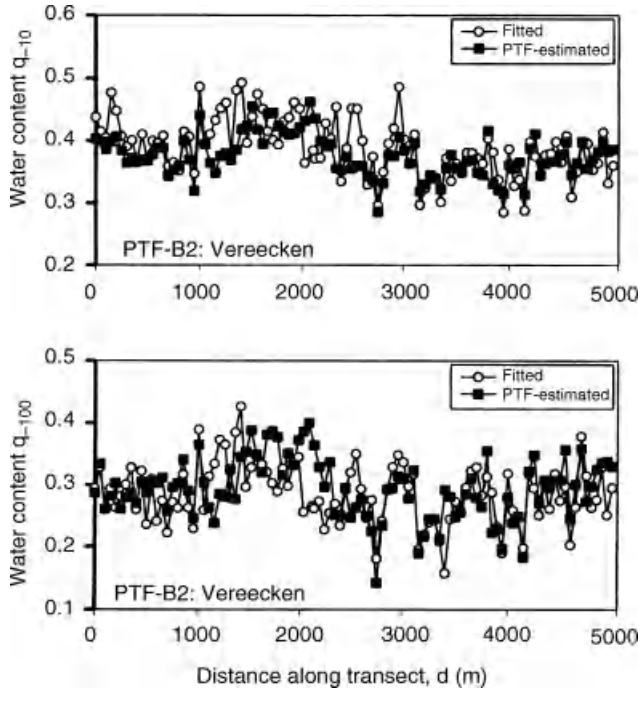


Figure B5.5.1 *A comparison of values of soil moisture at capillary potentials of -10 and -100 cm curves fitted to measured data and estimated using the pedotransfer functions of Vereeken* et al. *(1989) for different locations on a transect (after Romano and Santini, 1997, with kind permission of Elsevier).*

an indication of the significant uncertainty in the parameters and moisture characteristics for a typical sand and clay soil using this method, based on the analysis of a database of 2085 samples. The greater uncertainty for the clay soil reflects the smaller number of samples of fine textured soils in the database.

Pedotransfer functions for different soil texture classes have been included in the US STATSGO soils database (USDA SCS, 1992) and the European HYPRES soils database (Wösten, 1999). SOILPAR 2.0 (Acutis and Donatelli, 2003) includes 15 regression-based methods for estimating different soil characteristics and functional curves. ROSETTA (Schaap *et al.*, 2001) uses a neural network as a method of estimating soil parameters and SINFERS (McBratney *et al.*, 2002) is an inference system based on rules.

Box 5.6 Descriptive Equations for Surface Flows

We consider only a one-dimensional (downslope or downstream) description of surface flows here, but the principles apply also to the two-dimensional, depth integrated St. Venant equations. In one dimension, it is assumed that the flow can be adequately represented by a flow velocity, v [LT^{-1}], averaged over the local cross-sectional area, A [L^2] such that discharge $Q = vA$ [L^3T^{-1}]. Thus, since both v and A vary with discharge, there are two solution variables but for surface flow it is less usual to assume a simple functional relationship between them (although see the discussion of the kinematic wave approximation below). Thus two equations

are required to solve for the two variables. These equations were first formulated by the Barré de St. Venant in terms of a mass balance and a balance of momentum. Using the cross-sectional area of the flow to represent storage per unit length of channel, then as shown in Box 2.3, the mass balance is given by

$$\begin{aligned}\frac{\partial A}{\partial t} &= -\frac{\partial Q}{\partial x} + q \\ &= -\frac{\partial vA}{\partial x} + q \\ &= -A\frac{\partial v}{\partial x} - v\frac{\partial A}{\partial x} + q\end{aligned}$$

where x is distance downslope or downstream and q [L^2T^{-1}]is the net lateral inflow rate per unit length of channel.

A second equation linking v and A can be developed from the momentum balance of the flow. The control volume approach of Box 2.3 can also be used to derive the momentum balance which may be expressed in words as:

spatial change in hydrostatic pressure	+	loss in potential energy	−	friction loss	=	temporal change in local momentum	+	spatial change in momentum flux

or

$$-\frac{\partial Ap}{\partial x} + \rho g A S_o - \tau P = \frac{\partial \rho A v}{\partial t} + \frac{\partial \rho A v^2}{\partial x} \tag{B5.6.1}$$

where g [LT^{-2}] is the acceleration due to gravity, S_o [−] is the channel bed slope, P is the wetted perimeter of the channel [L], τ is the boundary shear stress [$ML^{-1}T^{-2}$], and p is the local hydrostatic pressure at the bed [$ML^{-1}T^{-2}$]. We can substitute for p and τ as

$$p = \rho g h \tag{B5.6.2}$$

and

$$\tau = \rho g R_h S_f \tag{B5.6.3}$$

where h [L] is an average depth of flow, S_f [−] is the friction slope which is a function of the roughness of the surface or channel, and R_h [L] is the hydraulic radius of the flow ($= A/P$).

With these substitutions, and dividing through by ρ under the assumption that the fluid is incompressible, the momentum equation may be rearranged in the form

$$\frac{\partial Av}{\partial t} + \frac{\partial Av^2}{\partial x} + \frac{\partial Agh}{\partial x} = gA(S_o - S_f) \tag{B5.6.4}$$

The friction slope is usually calculated by assuming that the rate of loss of energy is approximately the same as it would be under uniform flow conditions at the same water surface slope so that one of the *uniform flow* equations holds locally in space and time. The Manning equation (5.7) is often used, but an alternative is the Darcy–Weisbach uniform flow equation

$$v = \left[\frac{2g}{f} S_f R_h\right]^{0.5} \tag{B5.6.5}$$

where f is the Darcy–Weisbach resistance coefficient, so that an alternative form of the momentum equation can then be written as:

$$\frac{\partial Av}{\partial t} + \frac{\partial Av^2}{\partial x} + \frac{\partial Agh}{\partial x} = gAS_o - gP\frac{f}{2g}v^2 \tag{B5.6.6}$$

Again, the St. Venant equations are nonlinear partial differential equations (hyperbolic, in this case) that do not have analytical solutions except for some very special cases and approximate numerical solutions are necessary. As noted in the main text, the first attempt at formulating a numerical solution was by Stoker (1957). This used explicit time stepping which generally requires very short time steps to achieve adequate accuracy. Most solution algorithms used today are based on implicit time stepping (see the explanation in Box 5.3), such as the four-point implicit method described by Fread (1973).

B5.6.1 Summary of Model Assumptions

We can summarise the assumptions made in developing this form of the St. Venant equations as follows:

A1 The flow can be adequately represented by the average flow velocity and average flow depth at any cross-section.
A2 The amplitude of the flood wave is small relative to its wavelength so that pressure in the water column at any cross-section is approximately hydrostatic (pressure is directly proportional to depth below the water surface).
A3 The water is incompressible and of constant temperature and density.
A4 The friction slope may be estimated approximately using one of the uniform flow equations (such as the Manning or Darcy–Weisbach equations) with actual flow velocities and depths.

B5.6.2 Simplifications of the St. Venant Equations

The St. Venant equations are based on hydraulic principles but they are clearly an approximation to the fully three-dimensional flow processes in any stream channel. There are various further approximations to the St. Venant equations that are produced by assuming that one or more of the terms in the momentum equation, Equation B5.6.1, can be neglected. The two main approximate solutions are the diffusion wave approximation:

$$\frac{\partial A\rho g h}{\partial x} = \rho g A(S_o - S_f) \tag{B5.6.7}$$

and the kinematic wave approximation:

$$\rho g A(S_o - S_f) = 0 \tag{B5.6.8}$$

so that

$$S_o = S_f \tag{B5.6.9}$$

which reflects the assumption for the kinematic wave equation that the water surface is always parallel to the bed. This is also, of course, the assumption made in equations describing a uniform flow so, assuming again that the Darcy–Weisbach uniform flow equation is a good approximation for the transient flow case,

$$v = \left[\frac{2g}{f} S_o R_h\right]^{0.5} \tag{B5.6.10}$$

For a channel that is wide relative to its depth or an overland flow on a relatively smooth slope, then $R_h \approx h$ and this equation has the form of a power law relationship between velocity and storage:

$$v = bh^{0.5} \tag{B5.6.11}$$

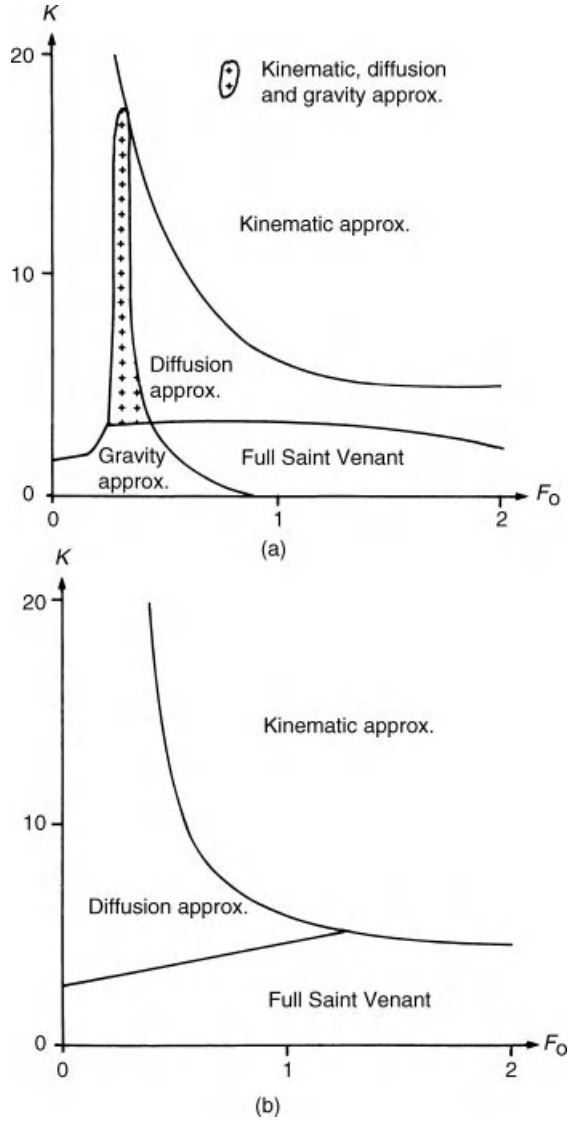


Figure B5.6.1 *Ranges of validity of approximations to the full St. Venant equations defined in terms of the dimensionless Froude and kinematic wave numbers (after Daluz Vieira, 1983, with kind permission of Elsevier).*

or, for discharge per unit width q (=vh) :

$$q = bh^{1.5} \tag{B5.6.12}$$

This type of power law has been widely used in kinematic wave descriptions of overland and channel flows (see Box 5.7). Use of a different uniform flow equation, such as the Manning equation (5.7), may result in a different power (1.67, in the case of the Manning equation).

A number of studies have examined the theoretical limits of acceptability of approximations to the full St. Venant equations for different surface flows. A study by Daluz Vieira (1983), based on extensive numerical analysis, provided ranges of validity for different approximations to the St. Venant equations, including the kinematic wave equation, in terms of two dimensionless numbers (see Figure B5.6.1): a Froude number, Equation (B5.6.13), and a kinematic wave number, Equation (B5.6.14).

$$F_o = C\left(\frac{tan\theta}{g}\right)^{0.5} \tag{B5.6.13}$$

$$\kappa = \left(\frac{g^3 L sin\theta}{C^4 q^2}\right)^{0.333} \tag{B5.6.14}$$

These studies suggest where the simplified models give a reasonable theoretical agreement with the full St. Venant equations. However, uncertainty in effective parameter values and boundary conditions may mean that the simplified models are useful under a wider range of conditions. The flood routing example of Figure 5.9 is a good illustration of this. In a real application, the shear stress and lateral inflow terms are not usually well known. Both terms may vary in both space and time as the river stage or depth of overland flow changes during an event. For river flow, in particular, there may be important changes in the effective shear stress if the flow exceeds the bankfull stage and starts to spill onto a flood plain (see, for example, the work of Knight *et al.*, 1994, 2010).

Box 5.7 Derivation of the Kinematic Wave Equation

The kinematic wave equation arises from the combination of a mass balance or continuity equation, expressed in terms of storage and flows, and a functional relationship between storage and flow that may be nonlinear but is single-valued, that is to say that there is a single value of flow at a point corresponding to any value of storage at that point. Consider a one-dimensional downslope overland flow on a slope of constant width. Let x be distance along the slope, h the depth of flow (which acts as the storage variable), and q the mean downslope velocity at any x (which is the flow variable). The mass balance equation can then be expressed as the partial differential equation

$$\frac{\partial h}{\partial t} = -\frac{\partial q}{\partial x} + r \tag{B5.7.1}$$

where r is a rate of addition or loss of water per unit length and width of slope at point x, and t is time.

The functional relation between h and v may be of many forms but a common assumption is the power law (see, for example, the kinematic approximation for surface flows in Box 5.6):

$$q = bh^a \tag{B5.7.2}$$

Thus, assuming a and b are constant and combining these two equations, to yield a single equation in h

$$\frac{\partial h}{\partial t} = -abh^{(a-1)}\frac{\partial h}{\partial x} + r \tag{B5.7.3}$$

or

$$\frac{\partial h}{\partial t} = -c\frac{\partial h}{\partial x} + r \tag{B5.7.4}$$

where $c = \frac{\delta q}{\delta h} = abh^{(a-1)}$ is the kinematic wave velocity or *celerity*. This is a kinematic wave equation. The celerity is the speed with which any disturbance to the system will be propagated downslope. It is worth noting again that kinematic wave equations can only propagate the effects of disturbances in a downslope or downstream direction. They cannot predict any backwater effects in rivers or drawdown effects due to a channel for subsurface hillslope flows. They do have the advantage, however, that for some simple cases, such as the case of constant input rates, analytical solutions exist for both surface and subsurface cases (see, for example, the work of Eagleson, 1970; or Singh, 1996).

For more complex cases, such as variable width slopes and arbitrary patterns of inputs, it may be necessary to resort to a numerical solution but a finite difference approximation is very easy to formulate and include in a hydrological model. The four-point implicit finite difference scheme of Li *et al.* (1975) has proven to be robust in a variety of applications. Li *et al.* point out that solving for the flow rate, q, rather than the storage variable, h, has some numerical advantages. The equivalent kinematic wave equation in q is:

$$\frac{\partial q}{\partial t} = -c\frac{\partial q}{\partial x} + cr \tag{B5.7.5}$$

The kinematic wave equation is easily extended to the case of a slope or channel for which the width, W_x, is varying downslope so that

$$W_x\frac{\partial q}{\partial t} = -c\frac{\partial W_x q}{\partial x} + cW_x r \tag{B5.7.6}$$

For surface runoff, both overland flow and channel flow, the kinematic wave approach is a good approximation to the full dynamic equations as the roughness of the surface or channel gets greater and the bedslope gets steeper (see Box 5.6).

The kinematic wave approach can also be adapted for the case of saturated downslope subsurface flow, in which h represents a depth of saturation above a water table and account must be taken of the effective storage deficit in the unsaturated zone above the water table which will affect the rise and fall of the water table (Beven, 1981). In the subsurface flow case, downslope flow rate per unit width of slope can often be approximated by a function such as:

$$q = K_s h \sin\alpha \tag{B5.7.7}$$

for a saturated hydraulic conductivity, K_s, that is a constant with depth of saturation, or for a soil in which hydraulic conductivity falls with increasing depth, by a function of the form:

$$q = K_o \exp(-f\left\{D - h\right\})\sin\alpha \tag{B5.7.8}$$

where K_o is the saturated hydraulic conductivity of the soil at the soil surface, D is the depth of the soil to the impermeable layer, α is the local slope angle, and f is a coefficient that controls how rapidly hydraulic conductivity declines with depth. These functions can be interpreted as a form of Darcy's law in which the effective downslope hydraulic gradient is assumed to be equal to the local slope, $\sin\alpha$.

The kinematic wave equation may then be written as

$$\varepsilon W_x\frac{\partial h}{\partial t} = -\frac{\partial W_x q}{\partial x} + W_x r \tag{B5.7.9}$$

where ε is an effective storage coefficient (here assumed constant). Substituting for h gives the same equation as for surface flow on a variable width slope:

$$W_x \frac{\partial q}{\partial t} = -c \frac{\partial W_x q}{\partial x} + c W_x r \tag{B5.7.10}$$

but here $c = \frac{K_s \sin\alpha}{\varepsilon}$ for the constant conductivity case and $c = \frac{K_s \sin\alpha}{\varepsilon} \frac{\exp(fh)}{f}$ for the exponentially declining conductivity case.

For saturated subsurface runoff on a hillslope, Beven (1981) showed that the kinematic wave description was a good approximation to a more complete Dupuit–Forchheimer equation description, if the value of a non-dimensional parameter, defined by

$$\chi = \frac{4K_s \sin\beta}{i} \tag{B5.7.11}$$

where i the effective rate of storm recharge to the slope, was greater than about 0.75. When this condition is met, any drawdown of the water table at the lower end of the slope due to an incised channel, is unlikely to have a great effect on the predicted discharges.

Given the mass conservation of Equation (B5.7.1), there is only one primary assumption underlying the derivation of the kinematic wave equation:

A1 A functional relationship between storage and discharge can be specified for the particular flow process being studied.

Several examples of such relationships for both surface and subsurface flow have been demonstrated above. The limitations of the kinematic wave approach must be appreciated, but a major advantage is that it is not restrictive in its assumptions about the nature of the flow processes, only that discharge should be a function of storage. Analytical solutions of the kinematic wave equation require that this functional relationship should be univalued (and generally of simple form). The simplicity of the kinematic wave assumptions have allowed the exploration of analytical solutions for different shapes of hillslope (e.g. Troch *et al.*, 2002; Norbiato and Borga, 2008). Numerical solutions do not have such a restriction and it is possible to envisage a kinematic wave solution that would have a hysteretic storage–discharge relationship that would more closely mimic the solution of the full surface or subsurface flow equations (in the same way that hysteretic soil moisture characteristics are sometimes used in unsaturated zone models Jaynes, 1990). It seems that no-one has tried to implement such a model in hydrology, although there has been some analysis of the storage–discharge hysteresis that arises at the hillslope and catchment scale (e.g. Ewen and Birkinshaw, 2006; Beven, 2006b; Norbiato and Borga, 2008; Martina *et al.*, 2011).

6

Hydrological Similarity, Distribution Functions and Semi-Distributed Rainfall–Runoff Models

Faced with this situation, it has been usual to incorporate what knowledge we have about the operation of the processes into some conceptual model of the system It is common that some parts of a complex conceptual model may be more rigorously based in physical theory than others. Even the most physically-based models, however, cannot reflect the true complexity and heterogeneity of the processes occurring in the field. Catchment hydrology is still very much an empirical science.

George Hornberger *et al.*, 1985

6.1 Hydrological Similarity and Hydrological Response Units

In any catchment the hydrologist is faced with a wide variety of geology, soils, vegetation and land use, and topographic characteristics that affect the relationship between rainfall and runoff. One way of taking these characteristics of any individual catchment into account is the type of fully distributed model that was discussed in Chapter 5 but, as was shown there, such models are difficult to apply because of their demands of both input data, much of which is not directly measureable, and computational resources. However, in any catchment, there may be many points that act in a hydrologically similar way with a similar water balance and similar runoff generation characteristics whether by surface or subsurface processes. If it were possible to classify points in the catchment in terms of their hydrological similarity, then a simpler form of model could be used based on a distribution of functional hydrological responses in the catchment without the need to consider every individual point separately. We should note at this point that it is also possible to consider defining hydrological similarity at the catchment

Rainfall–Runoff Modelling: The Primer, Second Edition. Keith Beven.

scale. This has been used in some approaches to the ungauged catchment problem and so discussion of similarity in this sense is discussed in Chapter 10.

There are three main approaches to attempting to use such a distribution of different responses **within** a catchment to model rainfall–runoff processes. The first is a statistical approach, based on the idea that the range of responses in a catchment area can be represented as a probability distribution of conceptual stores without any explicit consideration of the physical characteristics that control the distribution of responses. This approach therefore has much in common with the transfer function models of Chapter 4 and the example we outline in Section 6.2, the Probability Distributed Model of Moore and Clarke (1981), uses a similar parallel transfer function for routing of the generated runoff.

This purely statistical approach does not require any formal definition of similarity for different points in the catchment. The second approach is based on an attempt to define the hydrological similarity of different points in a catchment based on simple theory using topography and soils information. For catchments with moderate to steep slopes and relatively shallow soils overlying an impermeable bedrock, topography does have an important effect on runoff generation, at least under wet conditions, arising from the effects of downslope flows. This was the basis for the index of hydrological similarity introduced by Kirkby (1975) that was developed into a full catchment rainfall–runoff model, TOPMODEL, by Beven and Kirkby (1979) (see Section 6.3). The basic assumption of TOPMODEL is that all points in a catchment with the same value of the topographic index (or one of its variants) respond in a hydrologically similar way. It is then not necessary to carry out calculations for all points in the catchment, but only for representative points with different values of the index. The distribution function of the index allows the calculation of the responses at the catchment scale.

Another type of distribution function model that attempts to define similarity more explicitly is that based on the idea of *hydrological response units* or HRUs. These are parcels of the landscape differentiated by overlaying maps of different characteristics, such as soils, slope, aspect, vegetation type, etc. This type of classification of the landscape is very much easier to achieve now that maps of such characteristics can be held on the databases of geographical information systems and producing overlay maps of joint occurrences of such characteristics is a matter of a few simple mouse clicks on a personal computer. Freeware tools, such as Google Maps, also allow this information to be readily overlain onto base map or satellite image information, with zoom and other features already incorporated. These visualisation facilities are a major reason why the use of this type of model has expanded dramatically in the last decade. An example of the resulting landscape classification has already been seen in Figure 2.7. Models of this type are "distributed" in the sense of discretising a catchment into parcels of similar characteristics, but are generally lumped in the representation of the processes at the HRU scale. We will, therefore, refer to them as semi-distributed models (remembering that the distributed models of Chapter 5 are also effectively lumped at the discretisation scale, but in that case the continuum equations could, in principle, be refined to smaller and smaller discretisations in a consistent way, albeit at ever increasing computational cost). Just as there have been many different lumped catchment models with slightly different process conceptualisations, the semi-distributed models have tended to use directly analogous models for each HRU parcel (see Section 6.6). This reflects their origins; they are effectively the modern offspring of the lumped conceptual (or ESMA) catchment models.

These distribution functions and semi-distributed models are easier to implement and require much less computer time than fully distributed models, especially those based on distribution functions. They are clearly an approximation to a fully realistic distributed representation of runoff generation processes (but, as discussed earlier, so are the continuum partial different equation models of Chapter 5). It is not yet clear that the fully distributed models are distinctly advantageous in practical applications (see also Chapter 9) and the simplicity of the distribution function models has allowed them to be used to provide important insights into the modelling process.

6.2 The Probability Distributed Moisture (PDM) and Grid to Grid (G2G) Models

In many ways, the PDM model is a simple extension of some of the lumped storage models developed in the 1960s (and later) to the case of multiple storages representing a spatial distribution of different storage capacities in a catchment. It is a logical extension, in that we would expect that a distribution of storages might be a better representation of the variability in the catchment than simple lumped storage elements. However, in the original form outlined by Moore and Clarke (1981), the model makes no real attempt to relate the distribution of storages to any physical characteristics of the catchment. In fact, one of their main reasons for introducing a distribution of storages was to make the calibration problem easier since they found that they obtained smoother response surfaces for their new model formulation in comparison with models based on ESMA-type storage elements lumped at the catchment scale. A smoother response surface, in general, makes it easier for an automatic parameter optimisation routine to find the best fitting set of parameter values (but see the discussion of parameter calibration in Chapter 7).

The basic idea of the PDM model is illustrated in Figure 6.1. The multiple storage elements are allowed to fill and drain during rainstorm and interstorm periods respectively. If any storage is full then any additional rainfall is assumed to reach the channel quickly as storm runoff. A slow drainage component is allowed to deplete the storages between storms, contributing to the recession discharge in the channel and setting up the initial storages prior to the next storm. Evapotranspiration is also taken from each store during the interstorm periods.

In any storm, clearly those stores with the smallest storage capacity are filled first and start to produce rapid runoff first. Each storage capacity is assumed to represent a certain proportion of the catchment so that, as the stores fill, the proportion of the area producing fast runoff can also be calculated. This area expands during rainstorms and contracts between rainstorms so that, in essence, the distribution

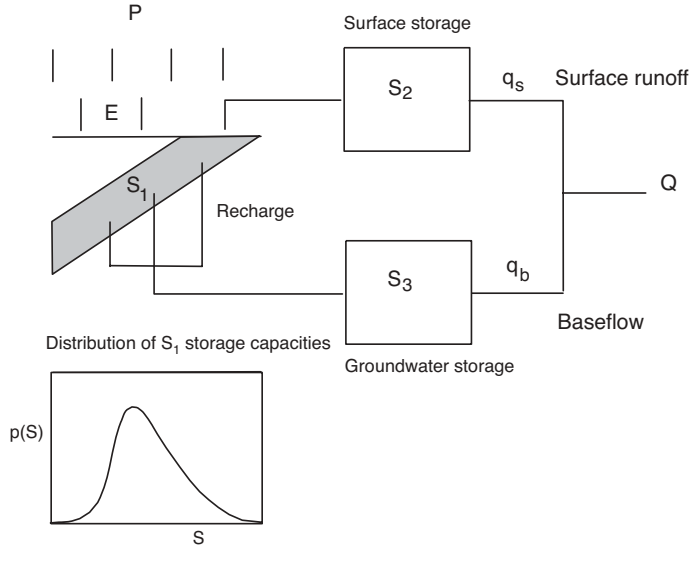


Figure 6.1 *Structure of the Probability Distributed Moisture (PDM) model.*

of stores represents a dynamic contributing area for runoff generation. In the published descriptions of PDM models, the authors distinguish between surface runoff and baseflow components. This is not, however, a necessary interpretation and, as with the transfer function models of Chapter 4, it is sufficient to recognise them as fast and slow runoff. The distribution of storages used in the model is only a distribution of conceptual storages and the question arises as to what form of distribution might be appropriate for a given catchment.

Moore and Clarke (1981) show that a variety of distributions can be easily incorporated into this type of model structure and they derive analytical equations for the responses of different distributions. Their work was extended by Hosking and Clarke (1990) who show how the model can be used to derive a relationship between the frequencies of storm rainfall and flow peak magnitudes in an analytical form. Moore (1985) examines the case where the stores lose water to deep drainage and evapotranspiration, while Moore and Clarke (1983) link the model to predicting sediment production as well as discharges. A review of PDM model concepts and equations has been provided by Clarke (1998). The simplicity of the model has allowed it to be used for long runs to derive flood frequencies (Lamb, 1999; Lamb and Kay, 2004; Calver *et al.*, 2009) and also, in a more distributed application, with radar rainfall and snowmelt inputs for flood forecasting (Moore *et al.*, 1994; Bell and Moore, 1998; Moore *et al.*, 1999). In the latter application, a separate PDM model is used for each radar rainfall pixel (an element of size 2 km by 2 km for UK radar rainfalls) so that any effect of the spatial distribution of rainfalls is preserved. Some attempt has also been made to reflect the different soil and topographic characteristics of these landscape units by varying the parameters of the distribution of stores in each element according to the soil type and average slope angle. The PDM model has also been coupled to a distributed snowmelt model (Moore *et al.*, 1999); a form of the model has been used as a macroscale hydrological model (Arnell, 1999); and an alternative method of redistributing storage between the storage elements has been suggested (Senbeta *et al.*, 1999).

The advantages of the PDM model are its analytical and computational simplicity. It has been shown to provide good simulations of observed discharges in many applications so that the distribution of conceptual storages can be interpreted as a reasonably realistic representation of the functioning of the catchment in terms of runoff generation. However, no further interpretation in terms of the pattern of responses is possible, since there is no way of assigning particular locations to the storage elements. In this sense, the PDM model remains a lumped representation at the catchment (or subcatchment element in the distributed version) scale.

In fact, an analogy can be drawn between the structure of the PDM model and some lumped catchment models, such as the VIC model, which use a functional relationship between catchment storage and the area producing rapid runoff (see Figure B2.2.1). The form of this relationship is controlled by parameters that are calibrated for a particular catchment area but then imply a certain distribution of storage capacities in the catchment in a similar way to the PDM model. Both models also use parallel transfer function routing for fast and slow runoff (surface runoff and baseflow in Figure 6.1), similar to the transfer function models discussed in Chapter 4.

The most recent formulation of the PDM model has seen it used as a semi-distributed model, with PDM elements representing grid squares feeding a grid-to-grid routing method (Figure 6.2). This Grid to Grid (G2G) model is directly descended from the distributed forms of PDM model noted above and makes use of the same way of relating the maximum storage in the runoff generation function to the local mean slope for each grid. Thus:

$$S_{max} = c_{max}\left(1 - \frac{\bar{g}}{g_{max}}\right) \tag{6.1}$$

where S_{max} is the local maximum storage capacity, c_{max} and g_{max} are maximum storage and gradient parameters common to all grid cells, and $\bar{g}$ is the local mean topographic gradient. This provides a

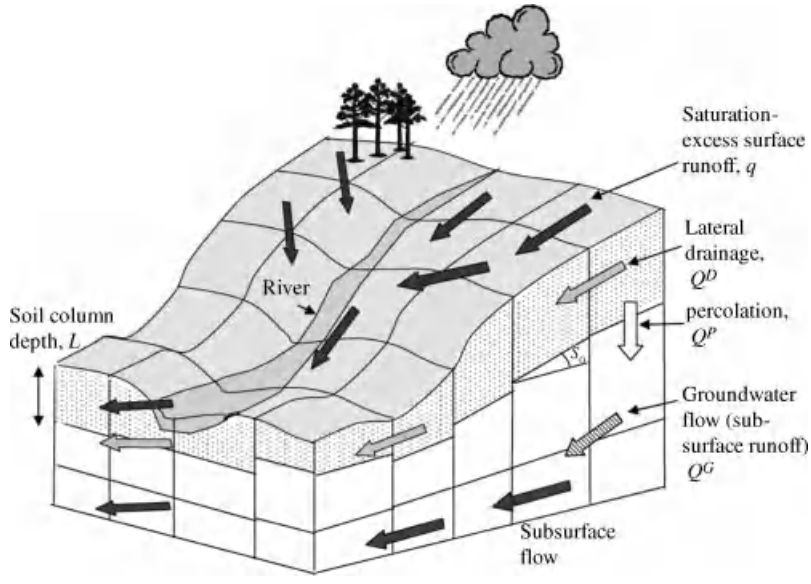


Figure 6.2 *Integration of PDM grid elements into the G2G model (after Moore et al., 2006, with kind permission of IAHS Press).*

simple way of distributing the runoff generation representation within the G2G model at the cost of only two parameters. Soil characteristics are defined using information from the hydrology of soil types (HOST) classification (Boorman *et al.*, 1995) linked to the SEISMIC soil properties database which are mapped to 1 km scale (Moore *et al.*, 2006). The model is completed by routing algorithms (hence the name "grid to grid") based on simplifications of the kinematic wave theory for surface and channel flow routing. Subsurface routing between grid elements is achieved by means of a nonlinear function of groundwater storage.

G2G has also been used for real-time forecasting driven by radar rainfall inputs (Cole and Moore, 2008) and a model-specific method for data assimilation of observed discharges by modifying the storages in individual grid elements to correct the overall prediction of discharge. Recent applications have included the use of the G2G model driven directly from inputs from a 25 km Regional Climate Model using either historical reanalysis or future climate projections (Bell *et al.*, 2007a, 2007b, 2009). This has included an assessment of reproducing flood frequency statistics based on continuous simulation (Bell *et al.* 2007a) which is discussed further in Chapter 8.

The G2G model is the first example of a hydrological model for the whole of the UK. It makes use of calibration where gauging station data are available, but can also provide predictions for ungauged subcatchments and catchments where data are not available. Any grid element in the country can be interpreted (or colour coded in a visualisation) for the current state of the flow in either absolute terms or as a frequency of occurrence. This makes the model a good tool for forecasting purposes and the G2G model is now being used operationally in the UK (by the joint Environment Agency/Met Office Flood Forecasting Centre) to make predictions of potential flooding with long lead times. The model is driven with the Met Office Global and Regional Ensemble Prediction System (MOGREPS) forecasts. The regional products of MOGREPS, which cover most of the North Atlantic and Europe, are nested within the global predictions. Each has 24 ensemble members, with perturbations to both the initial conditions and forecast model, with a grid scale for the Regional Model of 18 km. The G2G model is run on a 1 km

grid for the whole of the UK up to five days ahead, enabling a probabilistic evaluation of the potential for flooding across the country.

6.3 TOPMODEL

A simple approach to predicting spatial patterns of responses in a catchment is represented by TOPMODEL (see Beven *et al.*, 1995; Beven, 1997). TOPMODEL may be seen as a product of two objectives. One objective is to develop a pragmatic and practical forecasting and continuous simulation model. The other objective is to develop a theoretical framework within which perceived hydrological processes, issues of scale and realism, and model procedures may be researched. Parameters are intended to be physically interpretable and their number is kept to a minimum to ensure that values determined by a calibration exercise should be more easily identifiable, while still allowing a mapping of the predictions back into the catchment based on the pattern of a topographic index derived from an analysis of flow paths in the catchment. The model, in practice, represents an attempt to combine the computational and parametric efficiency of a distribution function approach with the link to physical theory and possibilities for more rigorous evaluation of spatial patterns of predictions offered by a fully distributed model. It has been seen as a reference against which other modelling concepts might be compared (Buytaert *et al.*, 2008). Some sources of TOPMODEL software are given in Appendix A.

6.3.1 The Background Theory

TOPMODEL can be considered to be a further approximation to the kinematic wave description of the subsurface flow system of Section 5.4. This link is made explicit by Kirkby (1997) and Wigmosta and Lettenmaier (1999). It is premised upon two basic assumptions: that the dynamics of the saturated zone can be approximated by successive steady state representations of the saturated zone on an area a draining to a point on a hillslope (Figure 6.3) and that the hydraulic gradient of the saturated zone can be approximated by the local surface topographic slope, tan β (Box 6.1).

These assumptions lead to simple relationships between catchment storage (or storage deficit below saturation) in which the main factor is the Kirkby topographic index ($a/\tan\beta$) (Kirkby, 1975). The Kirkby index represents the propensity of any point in the catchment to develop saturated conditions. High values

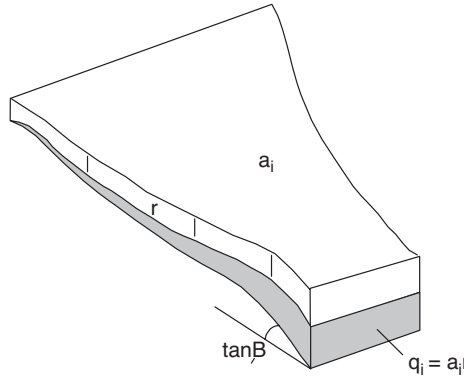


Figure 6.3 *Definition of the upslope area draining through a point within a catchment.*

are caused by either long slopes or upslope contour convergence, and low slope angles. Points with the same value of the index are predicted as having the same hydrological responses. The topographic index approach was developed into a complete rainfall–runoff model by Beven and Kirkby (1979) and has been generalised since to allow for differences in soil characteristics within the catchment (see below and Box 6.1). The assumptions are similar to those used in the development of the "wetness" index developed independently by O'Loughlin (1981, 1986) and used in the model of Moore *et al.* (1986). In passing, it is worth mentioning that Horton (1936) had already used a very similar concept, though in his case he considered only the maximum water table that might be supported by input at the infiltration capacity of the soil (see Beven, 2004c, 2006d).

TOPMODEL in its original form takes advantage of the mathematical simplifications allowed by a third assumption: that the distribution of downslope transmissivity with depth is an exponential function of storage deficit or depth to the water table:

$$T = T_o e^{-D/m} \tag{6.2}$$

where T_o is the lateral (downslope) transmissivity when the soil is just saturated [L^2/T], D is a local storage deficit below saturation expressed as a depth of water [L] and m is a model parameter controlling the rate of decline of transmissivity in the soil profile, also with dimensions of length [L]. A physical interpretation of the decay parameter m is that it controls the effective depth or active storage of the catchment soil profile. A larger value of m effectively increases the active storage of the soil profile. A small value generates a shallow effective soil, but with a pronounced transmissivity decay.

Given this exponential transmissivity assumption, it can be shown that the appropriate index of similarity is $\ln(a/\tan\beta)$ or, if the value of T_o is allowed to vary in space, $\ln(a/T_o \tan\beta)$ such that, given a mean storage deficit over a catchment area $\overline{D}$, a local deficit at any point can be calculated as (see Box 6.1):

$$D_i = \overline{D} + m\left[\gamma - \ln(a/T_o \tan\beta)\right] \tag{6.3}$$

where γ is the mean value of the index over the catchment area. Thus, every point having the same soil/topographic index value $(a/T_o \tan\beta)$ behaves functionally in an identical manner. The $(a/T_o \tan\beta)$ variable is therefore an index of hydrological similarity. Other forms of transmissivity profile assumption lead to different forms for the index and local deficit calculation (see Box 6.1). Of particular interest are points in the catchment for which the local deficit is predicted as being zero at any time step. These points, or fraction of the catchment, represent the saturated contributing area that expands and contracts with the change in $\overline{D}$ as the catchment wets and dries (Figure 6.4). This simplest form of estimate for the local deficit D_i also assumes that the theoretical negative deficits (steady-state oversaturation on the contributing area) can be included in the calculation of the mean catchment deficit. Saulnier and Datin (2004) show how this can be modified to allow for dynamic changes in the contributing area and account only for the deficit on the non-saturated area. The equations can also be derived in terms of water table depth rather than storage deficit but this introduces at least one additional effective storage parameter (Beven *et al.*, 1995). In each case, there is a relationship between the transmissivity profile assumed and the form of the recession curve at the catchment scale produced by soil drainage. For the exponential transmissivity assumption, the derived recession curve function is given by:

$$Q_b = Q_o e^{-\overline{D}/m} \tag{6.4}$$

where $Q_o = Ae^{-\gamma}$ for a catchment area of A. This equation (and equivalent forms for different transmissivity assumptions) is derived under the assumption that the effective hydraulic gradients for the subsurface flow do not change with time, as would be predicted by a more complete analysis.

The calculation of the index for every point in the catchment requires knowledge of the local slope angle, the area draining through that point and the transmissivity at saturation. The spatial distribution

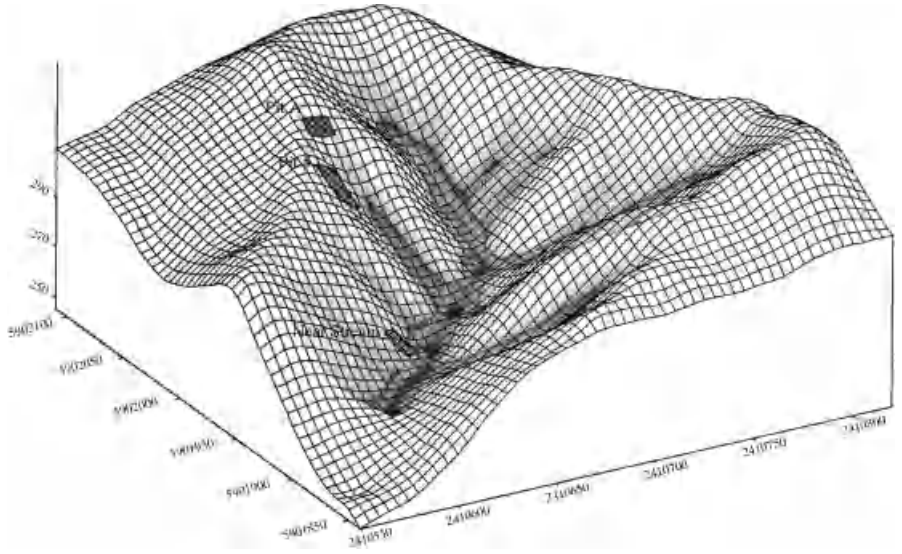


Figure 6.4 *The ln(a/ tan β) topographic index in the small Maimai M8 catchment (3.8 ha), New Zealand, calculated using a multiple flow direction downslope flow algorithm; high values of topographic index in the valley bottoms and hillslope hollows indicate that these areas are predicted as saturating first (after Freer, 1998).*

of $(a/\tan\beta)$ (see Figure 6.4) may be derived from analysis of a digital terrain model (DTM) or a digital elevation map (DEM) of the catchment (see Section 6.3.2). Specifying a spatial distribution for T_o is generally much more problematic, since there are no good measurement techniques for obtaining this parameter. In most applications, it has been assumed to be spatially homogeneous, in which case the similarity index reduces to the form $(a/\tan\beta)$.

To calculate the surface (or subsurface) contributing area, the catchment topographic index is expressed in distribution function form (Figure 6.5). Discretisation of the $(a/\tan\beta)$ distribution function brings computational advantages. Given that all points having the same value of $(a/\tan\beta)$ are assumed to behave in a hydrologically similar fashion, then the computation required to generate a spatially-distributed local water table pattern reduces to one calculation for each $(a/\tan\beta)$ class; calculations are not required for each individual location in space. This approach should be computationally more efficient than a solution scheme that must make calculations at each of a large number of spatial grid nodes, a potentially significant advantage when parameter sensitivity and uncertainty estimation procedures are carried out.

In a time step with rainfall, the model predicts that any rainfall falling upon the saturated source area will reach the stream by a surface or subsurface route as storm runoff, along with rainfall in excess of that required to fill areas where the local deficit is small. The calculated local deficits may also be used to predict the pattern of subsurface stormflow contributing areas or flow through different soil horizons (Robson *et al.*, 1992) if they can be defined by some threshold value of deficit (or water table depth).

The model is completed by a representation of the unsaturated zone and a flow routing component. Both have been kept deliberately simple to facilitate parameter estimation. It is particularly difficult to account explicitly for the effects of local soil heterogeneity and macroporosity. No entirely satisfactory mathematical description is currently available of unsaturated flow in structured soils with parameters that can be identified at a practical prediction scale and if parameter values are to be determined by calibration then minimal parameterisation is advantageous. Current versions of TOPMODEL use two

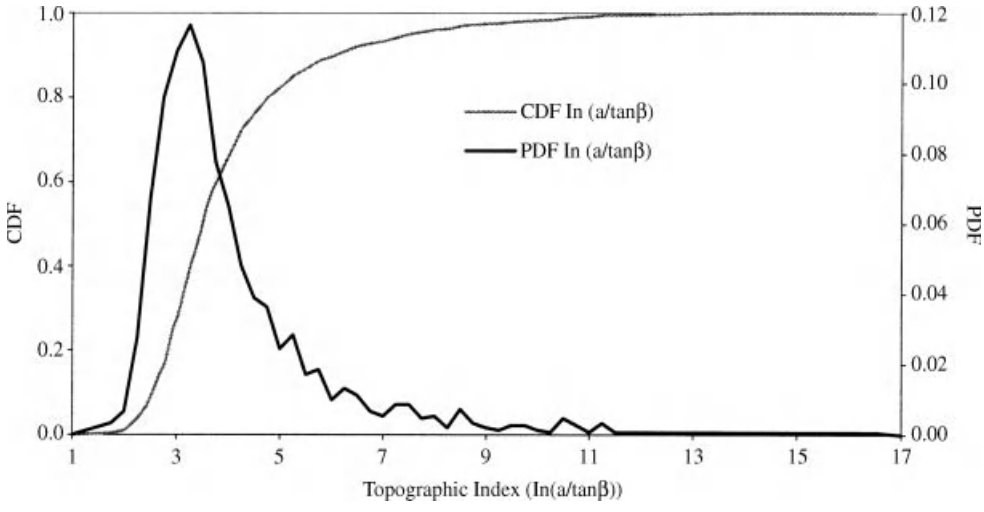


Figure 6.5 *Distribution function and cumulative distribution function of topographic index values in the Maimai M8 catchment (3.8 ha), New Zealand, as derived from the pattern of Figure 6.4.*

stores to represent the unsaturated zone: one representing interception and root zone storages for which additional deficits due to evapotranspiration are calculated; and a drainage store which controls recharge to the saturated zone. Both introduce one additional parameter.

There is no reason why this part of the model should not be made more complex if sufficient information is available to justify that complexity. Indeed, more complex representations of the soil and vegetation have been linked to TOPMODEL concepts in the TOPLATS formulation of Famiglietti *et al.* (1992); by Seuffert *et al.* (2002); in RHESSys (Band *et al.* 1991, 1993; Fagre *et al.* 1997; Hartman *et al.* 1999); in the MACAQUE model of Watson *et al.* (1999); and in the ISBA-TOPMODEL land surface parameterisation (e.g. Pellenq *et al.*, 2003; Vincendon *et al.*, 2010). These can also be used as semi-distributed models with calculations made for every pixel within a catchment (e.g. Houser *et al.*, 1998). This extension introduces further vegetation and soil parameters. Representations of infiltration excess runoff can also be included and, in some other applications of TOPMODEL, more explicitly physically-based approaches to infiltration have been adopted (Beven, 1986a, 1986b, 1987; Sivapalan *et al.*, 1990; Wood *et al.*, 1990). However, these approaches also introduce extra parameters to the model which may be difficult to calibrate, especially for the case of spatially heterogeneous soils. Some variants on the TOPMODEL concepts are discussed in Box 6.1.

6.3.2 Deriving the Topographic Index

An analysis of catchment topography is required in order to derive the $(a/\tan\beta)$ distribution function (Figure 6.5). In order to obtain discrete values of $(a/\tan\beta)$ some sampling of topography is implied. Early development of TOPMODEL relied upon the manual analysis, based on map and air photo information, of local slope angles, upslope contributing areas and cumulative areas. Beven and Kirkby (1979) outlined a computerised technique used to derive the topographic index distribution function (and the overland flow delay histogram) based on the division of the catchment into sub-basin units. Each unit was then discretised into small "local" slope elements on the basis of dominant flow paths (inferred from lines of greatest slope in a way similar to the TAPES-C flow path analysis software of Grayson *et al.*, 1995). Calculation

of $(a/\tan\beta)$ was carried out for the downslope edge of each element. Although an approximation, this method was felt to be justified by its relative efficiency and because field observations of flow paths could be used in defining the slope elements to be analysed. In particular, the effects of field drains and roads in controlling effective upslope contributing areas could be taken into account. Such human modifications of the natural hydrological flow pathways are often important to the hydrological response but are not normally included in the digital terrain maps (DTMs) of catchment areas commonly used in topographic analysis today.

However, given a DTM, more computerised methods are now available. Quinn *et al.* (1995) demonstrate the use of digital terrain analysis (DTA) programs, based on raster elevation data, in application to catchment modelling studies based on TOPMODEL. There are subjective choices to be made in any digital terrain analysis. Techniques of determining flow pathways from raster, contour and triangular irregular network DTMs are discussed in Section 3.7. Different DTA techniques result in different flow-path definitions and therefore different calculations of "upslope contributing area" for each point in the catchment. The resolution of the DTM data will also have an effect. The DTM must have a fine enough resolution to reflect the effect of topography on surface and subsurface flow pathways adequately. Coarse resolution DTM data may, for example, fail to represent some convergent slope features. However, too fine a resolution may introduce perturbations to flow directions and slope angles that may not be reflected in subsurface flow pathways which, in any case, will not always follow the directions suggested by the surface topography (see, for example, the work of Freer *et al.* (1997), who suggest that bedrock topography may be a more important control in some catchments). The appropriate resolution will depend on the scale of the hillslope features, but 50 m or better data is normally suggested. Anything much larger and, in most catchments, it will not be possible to represent the form of the hillslopes in the calculated distribution of the $(a/\tan\beta)$ index.

Experience suggests that the scale of the DTM used and the way in which river grid squares are treated in the DTA do affect the derived topographic index distribution, in particular inducing a shift in the mean value of $(a/\tan\beta)$. There is a consequent effect on the calibrated values of parameters (especially the transmissivity parameter) in particular applications. This is one very explicit example of how the form of a model definition may interact with the parameter values required to reproduce the hydrology of a catchment. In this case, two different topographic analyses of a catchment may require different effective transmissivity values for hydrograph simulation. However, in one study, Saulnier *et al.* (1997b) have suggested that excluding river grid squares from the distribution results in calibrated transmissivity parameter values that are much more stable with respect to changing DTM resolution. An alternative adjustment has been suggested in developing a block version of TOPMODEL for use at very large scales (BTOPMC) in which the a developed from topographic analysis within a block can be adjusted by a function $f(a)$ within the range 0–1 (Ao *et al.*, 2006; Takeuchi *et al.*, 2008). BTOPMC is now being used to support a global flood alert system (GFAS) for large scale catchments in sparsely gauged areas, driven by satellite-derived rainfall estimates at the UNESCO International Centre of Excellence for Water Hazard and Risk Management (ICHARM) in Japan.

The simplicity of the TOPMODEL structure has allowed this problem to be studied in some detail but similar considerations of an interaction between grid scale and model parameters must apply to even the most physically based, fully distributed models (Beven, 1989; Refsgaard, 1997; Kuo *et al.*, 1999). In the last decade, there has been an interesting analysis of this grid-scale dependence of the topographic index by Nawa Pradhan *et al.* (2006, 2008). They point out that both the effective contributing area and the steepest gradients in a raster grid are dependent on the size of the grid elements and propose ways of adjusting both, when topographic data are available at one scale and the target resolution is at a finer scale. The adjustment of the a value is simply to divide by the ratio of the grid lengths (coarse/target). The adjustment for slope is based on assuming that the fractal properties of topography imply scaling similarity, at least for a range of scales (this tends to break down at very fine scales). They make use of a relationship suggested by Zhang *et al.* (1999) to estimate a local fractal dimension D from the standard

deviation of the elevations in a local area, such as a 3-by-3 array around the point of interest. Thus:

$$D = 1.13589 + 0.08452 \ln \sigma \tag{6.5}$$

Then if the local steepest gradient $G_{steepest}$ is calculated at the coarsest resolution, the equivalent value at the target resolution can be defined by

$$G_{steepest} = \alpha d^{(1-D)} \tag{6.6}$$

where d is the distance between grid points (either in the principal or diagonal direction of the grid, depending on the direction of greatest slope) and α can be determined at the coarse grid scale by inverting this equation since $G_{steepest}$, d and D are known. Under unifractal scaling, α can be assumed constant with change in scale, so that a scale's steepest slope at the target scale $g_{steepest}$ can be calculated as

$$g_{steepest} = \alpha d*^{(1-D)} \tag{6.7}$$

where $d*$ is the equivalent distance between points in the same flow direction at the target grid scale.

Note that no additional points are being added here. It is assumed that data are only available at the coarse resolution, but that this might not adequately represent the hillslope forms so that scaling down to a finer target resolution might lead to better simulations. The importance of this scaling is demonstrated by Pradhan *et al.* (2006). They show the dependence of the topographic index distribution on grid resolutions from 50 m to 1000 m. They show how a good estimate of the distribution at the 50 m resolution can be obtained if only 1000 m data are available. They also show how, without this adjustment, different parameter values are required in TOPMODEL for the different grid resolutions, but if all the resolutions are scaled to a target resolution of 50 m and the calibrated parameter values for that scale are used to make predictions, only slightly pooorer simulations are obtained than those calibrated specifically for the 50 m resolution grid. Thus it seems, despite the gross simplifying assumptions that are made, a lot of the scale dependence effect on parameter values can be eliminated. It is also worth noting, however, that this analysis has been somewhat overtaken by events. As seen in Chapter 3, satellite-derived topographic data at 30 m resolution are now available for almost the whole globe in the ASTER global digital elevation model.

It is worth noting that a parameterisation of the $(a/\tan\beta)$ distribution may sometimes be useful. Sivapalan *et al.* (1990) introduced the use of a gamma distribution in their scaled version of TOPMODEL. Wolock (1993) also gives details of a gamma distribution version for continuous simulation. An analogy with the statistical PDM model of Section 6.2 becomes apparent in this form. The advantage of using an analysis of topography to define the index distribution beforehand is that there are then no additional parameters to estimate. This is only an advantage, however, where the topographic analysis allows a realistic representation of the similarity of hydrological responses in the catchment, which clearly depends on the validity of the simplifying assumptions that underlie the index.

6.3.3 Applications of TOPMODEL

Simulation of humid catchment responses

TOPMODEL was originally developed to simulate small catchments in the UK (Beven and Kirkby, 1979; Beven *et al.*, 1984). These studies showed that it was possible to get reasonable results with a minimum of calibration of parameter values. A summary table of applications is given by Beven (1997). More recent applications include Franks *et al.* (1998) and Saulnier *et al.* (1998) in France; Lamb *et al.* (1998a, 1998b) in Norway (see Section 6.4); Quinn *et al.* (1998) in Alaska; Cameron *et al.* (1999) in Wales; Dietterick *et al.* (1999) in the USA; Donnelly-Makowecki and Moore (1999) in Canada; and Güntner *et al.* (1999) in Germany (Figure 6.6). In most of these cases, it has been found that, after calibration

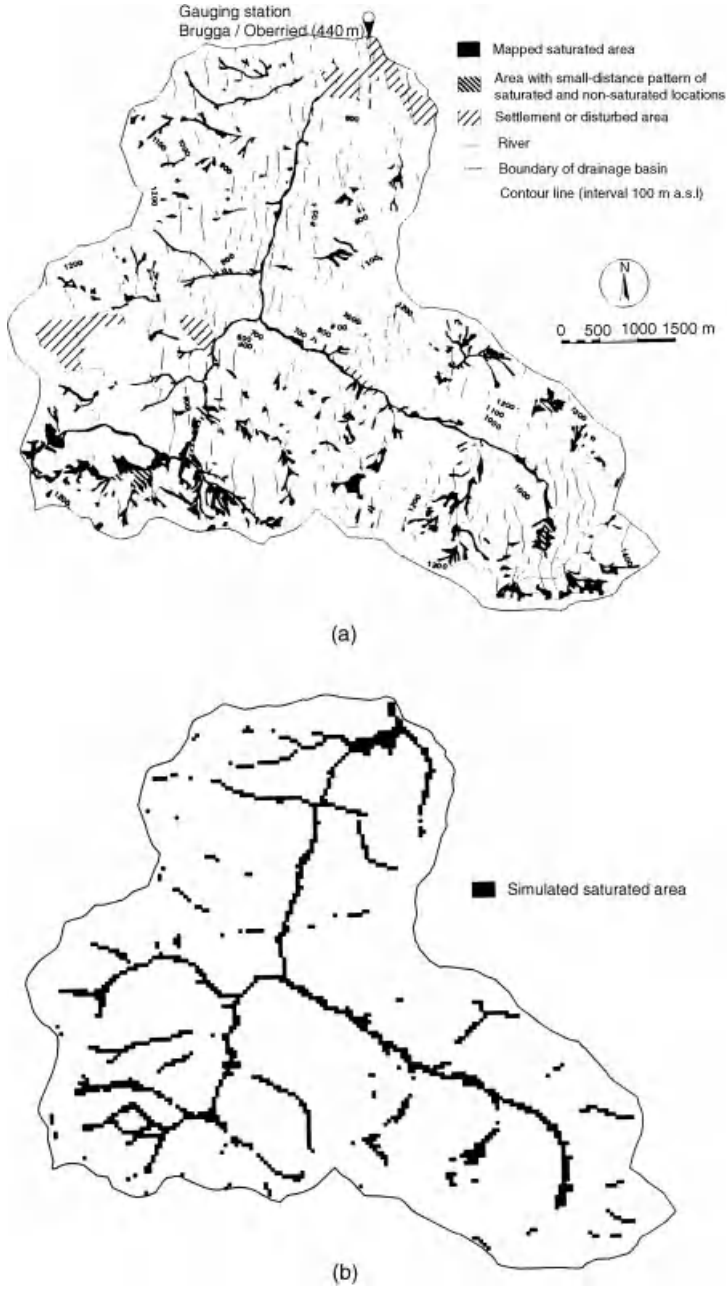


Figure 6.6 *Spatial distribution of saturated areas in the Brugga catchment (40 km^2), Germany: (a) mapped saturated areas (6.2% of catchment area); (b) topographic index predicted pattern at same fractional catchment area assuming a homogeneous soil (after Güntner et al., 1999, with kind permission of John Wiley and Sons).*

of the parameters, TOPMODEL provides good simulations of stream discharges and broadly believable simulations of variable contributing areas.

Catchments with deeper groundwater systems or locally perched saturated zones may be much more difficult to model. Such catchments tend to go through a wetting up sequence at the end of the summer period in which the controls on recharge to any saturated zone and the connectivity of local saturated zones may change with time. An example is the Slapton Wood catchment in southern England modelled by Fisher and Beven (1995).

Simulation of drier catchment responses

A model that purports to predict fast catchment responses on the basis of the dynamics of saturated contributed areas may not seem to be a likely contender to simulate the responses of catchments that are often dry, such as in Mediterranean or savannah climates. However, Durand *et al.* (1992) have shown that TOPMODEL can successfully simulate discharges in such catchments at Mont Lozère in the Cevennes, southern France, at least after the calibration of some parameters.

Experience in modelling the Booro-Borotou catchment in the Côte d'Ivoire (Quinn *et al.*, 1991), Australia (Barling *et al.*, 1994) and catchments in the Prades mountains of Catalonia, Spain (Piñol *et al.*, 1997), suggests that TOPMODEL will only provide satisfactory simulations once the catchment has wetted up. In many low precipitation catchments, of course, the soil may never reach a "wetted" state and the response may be controlled by the connectivity of any saturated downslope flows. TOPMODEL assumes that there is connected downslope saturation everywhere on the hillslope; before such connectivity is established, a dynamic index would be required (Barling *et al.*, 1994; Beven and Freer, 2001; Vincendon *et al.*, 2010). Such catchments also tend to receive precipitation in short, high-intensity storms. Such rainfalls may lead, at least locally, to the production of infiltration excess overland flow which is not usually included in TOPMODEL (but see the work of Beven (1986) and Sivapalan *et al.* (1990) for example applications including infiltration excess calculations). The underlying assumptions of the TOPMODEL concepts must always be borne in mind relative to the pertinent perceptual model for a particular catchment.

6.3.4 Testing the Hydrological Similarity Concept in TOPMODEL

TOPMODEL may be expected to perform best when tested against catchments where its assumptions are met, in particular those of an exponential saturated zone store, a quasi-parallel water table and a topographic control on water table depth. A full critique of the TOPMODEL concepts can be found in Beven (1997). There are certainly limitations on both the geographical and seasonal validity of the TOPMODEL concepts, but they do provide a basis for thinking about the distributed nature of catchment responses. It has always been stressed that TOPMODEL is not a fixed model structure but rather a set of concepts that should be modified if it is perceived that a catchment does not conform to the basic assumptions. Ways of relaxing the basic assumptions are discussed in Box 6.1.

The main limitation imposed by the model assumptions is that of the quasi-steady-state water table shape. This has been criticised by Wigmosta *et al.* (1999), who compare the results of TOPMODEL with a dynamic simulation based on a subsurface kinematic wave solution. They show that although TOPMODEL could usually be calibrated to produce a reasonable simulation of discharge hydrographs produced by the kinematic wave model, the resulting effective parameter values tended to be high and the steady state assumption did not produce reasonable predictions of the dynamic changes in the water table. It is the steady state assumption that allows TOPMODEL to make use of similarity in greatly increasing computational efficiency. This is useful for a number of purposes, not the least of which is the exploration of predictive uncertainty (considered in Chapter 7). The approach can also be modified to allow more dynamic calculations whilst retaining the concept of the index (see Section 6.3.5 and the TOPKAPI variant in Section 6.5).

6.3.5 Dynamic TOPMODEL

While TOPMODEL has been very widely used in rainfall–runoff modelling, it is clear from the above results and discussion that the assumptions on which it is based are not going to be valid except in a relatively limited range of catchments, those that are frequently wetted and have relatively thin soils and moderate topography. It might produce a good fit when calibrated to observed discharges for a wider range of catchments, but it does not then produce good results for good reasons. It is serving only to provide fast runoff generation that is a nonlinear function of storage deficit (in much the same way as other models, such as the PDM model of Section 6.2, the nonlinear filter in the DBM of Section 4.3.2 and the VIC model of Box 2.2).

A particular issue with the classic formulation is the assumption that there is always connectivity over the whole of the upslope area that underlies the representation of the water table as a succession of steady states. Barling *et al.* (1994) have shown that a better relationship could be found between saturated area and a topographic index, if the index was calculated using only an effective upslope contributing area rather than the full upslope area all the way to the divide. This effective upslope area would be expected to be small when the catchment was dry and increase as the catchment wets up. In fact, Western *et al.* (1999) show that patterns of near surface soil moisture only show the effects of a topographic control on downslope flows under relatively wet conditions in the Tarrawarra catchment in Australia. This is, in fact, another reason why calibrated values of transmissivity in TOPMODEL might be high. Since, in the soil's topographic index ($a/T_o \tan\beta$), a and T_o appear in ratio, a high value of T_o can compensate for an overestimation of the effective upslope area a.

The result is that in environments that are subject to strong drying periods, the application of TOPMODEL often has difficulties in simulating the first few events during a wetting up period (e.g. Piñol *et al.* 1997). With this in mind, Beven and Freer (2001a) proposed a modified TOPMODEL, called "Dynamic" TOPMODEL. This was more flexible in defining the area units for which calculations are made. In particular, it was possible to separate out areas of different a and $\tan\beta$ values. Other information on catchment characteristics, such as the pattern of T_o, can also be taken into account where it is available. Each elemental area then uses a local numerical implicit kinematic wave solution for the saturated zone, taking account of local slope angles. Calculated outflows from each elemental area are then distributed to appropriate downslope elements. The model can be set up without imposing a regular spatial grid (but the kinematic subsurface solution is otherwise similar to the gridded DVSHM model of Wigmosta *et al.*, 1999). In this framework, not all elements need to produce downslope flows under dry conditions, introducing more dynamic hillslope responses than the original TOPMODEL representation as a succession of quasi-steady states. The unsaturated zone is treated in a similar way to TOPMODEL. Dynamic TOPMODEL has been successfully used in a variety of applications to catchments in England (Beven and Freer, 2001; Younger *et al.*, 2008, 2009), Wales (Page *et al.*, 2007), USA (Peters *et al.*, 2003) and Luxembourg (Liu *et al.*, 2009).

6.4 Case Study: Application of TOPMODEL to the Saeternbekken Catchment, Norway

In most rainfall–runoff modelling studies, there are generally few internal state measurements with which to check any distributed model predictions. The potential of making such checks with distributed models raises some interesting questions about model calibration and validation. One study where distributed predictions have been checked is in the application of TOPMODEL to the Saeternbekken MINIFELT catchment in Norway (Lamb *et al.*, 1997, 1998a, 1998b). This small subcatchment of only 0.75 ha has a network of 105 piezometers and four recording boreholes (Figure 6.7a; Myrabø, 1997; Erichsen and Myrabø, 1990). The distribution of the $ln(a/\tan\beta$ topographic index is shown in Figure 6.7b.

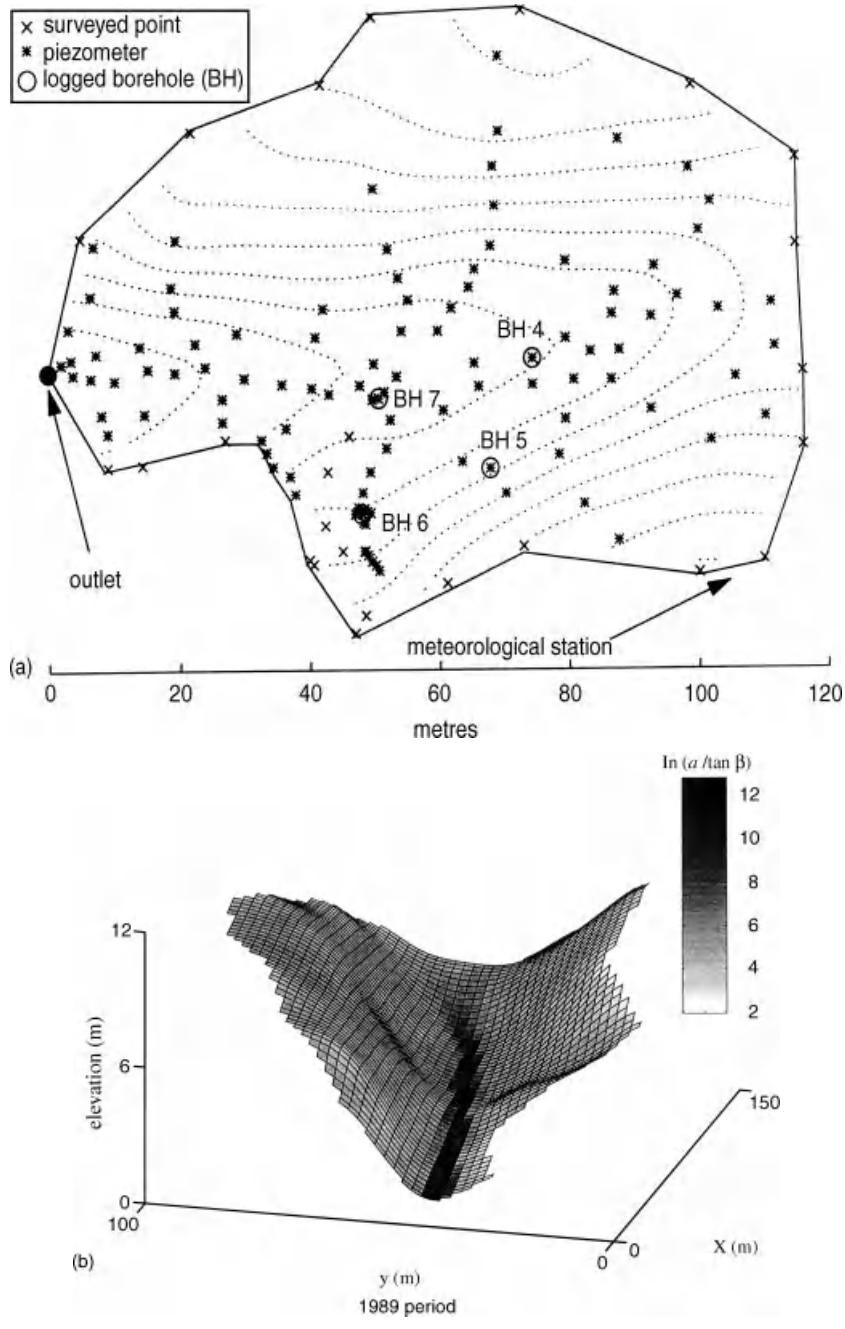


Figure 6.7 *Application of TOPMODEL to the Saeternbekken MINIFELT catchment, Norway (0.75 ha): (a) topography and network of instrumentation; (b) pattern of the ln(a/ tan β) topographic index; (c) prediction of stream discharges using both exponential (EXP) and generalised (COMP) transmissivity functions (after Lamb et al., 1997, with kind permission of John Wiley and Sons).*

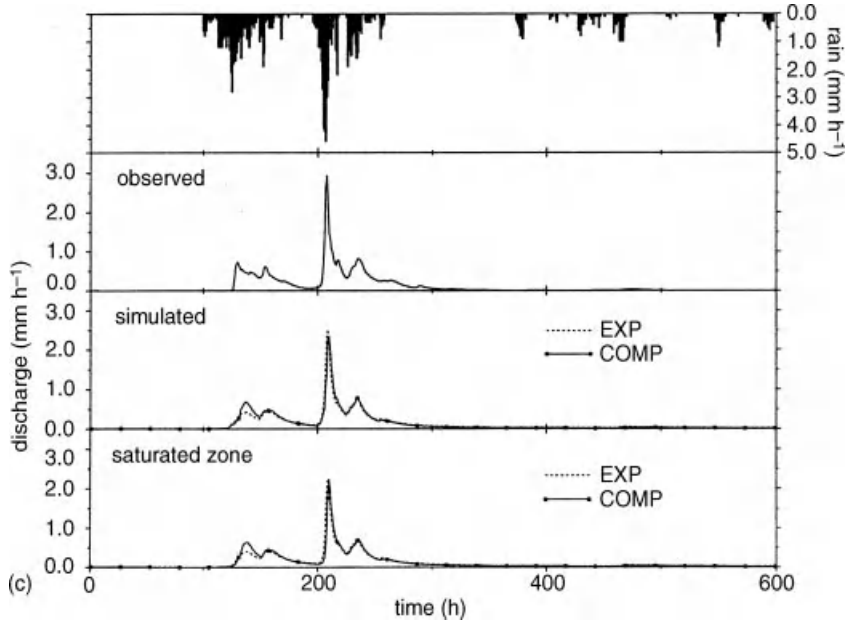


Figure 6.7 *(continued)*

Figure 6.7c shows the discharge predictions of two variants of TOPMODEL. In both cases, the models were calibrated on rainfall–runoff events in 1987 and the results shown are for a separate evaluation period in 1989.

The EXP model in Figure 6.7c is essentially the original exponential transmissivity function version of TOPMODEL that is described in detail in Box 6.1. The COMP model uses a technique used by Lamb *et al.* (1997, 1998b) in which a recession curve analysis is used to define an arbitrary discharge–storage deficit relation that can replace the exponential transmissivity assumption of the original model. In the case of the Saeternbekken MINIFELT catchment, a composite curve with exponential and linear segments was found to be suitable. Other functions can also be used (see Box 6.1). There was little to choose between the models in discharge prediction.

TOPMODEL, however, also allows the mapping of the predictions of storage deficit or water table level back into the space of the catchment. While, given the approximate assumptions of the TOPMODEL approach, the results are not expected to be accurate everywhere in the catchment, we would want to reject the model if it was shown to give very poor predictions of such internal measurements. Results for the recording boreholes are shown in Figure 6.8 and for the piezometer levels for five different discharges in Figure 6.9. For these internal data, the composite model appears to give slightly better results, but clearly both models are limited in their accuracy (notably in predicting the significant depths of ponding recorded in some of the piezometers, since neither model has any direct representation of depression storage in the hummocky terrain). In both cases, these predictions use the global parameter values calibrated for the 1987 period. This effectively assumes that the transmissivity function in both models is uniform throughout the catchment. The recording borehole data were used only in the estimation of an effective storage coefficient, for which a catchment median calibrated value was 0.06. The piezometer data were not used in calibration at this stage.

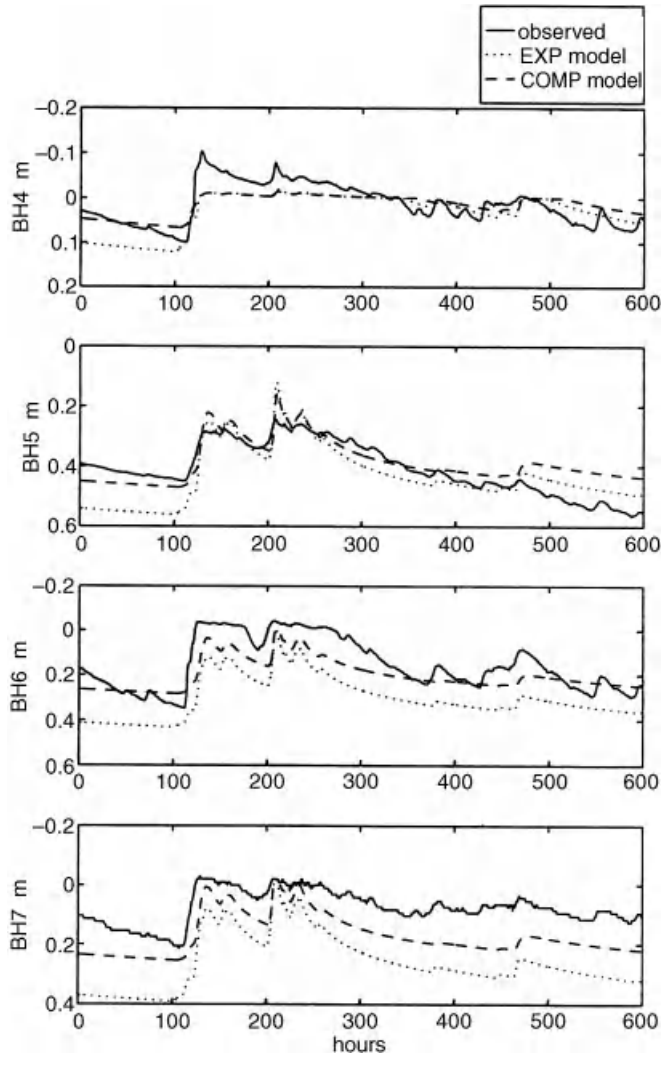


Figure 6.8 *Predicted time series of water table levels for the four recording boreholes in the Saeternbekken MINIFELT catchment, Norway, using global parameters calibrated on catchment discharge and recording borehole data from an earlier snow-free period in October–November 1987 (after Lamb et al., 1997, with kind permission of John Wiley and Sons).*

These results are reasonable but could perhaps be improved in a number of ways. The need for some surface depression storage has already been noted. The steady state assumption of the TOPMODEL index approach may not be appropriate. It is also known that soil heterogeneity can be an important source of variability in hydrological response. In the case of these water table predictions, it might be possible to improve the simulations by allowing for some local variability in the transmissivity function. This possibility was investigated for this catchment by Lamb *et al.* (1997).

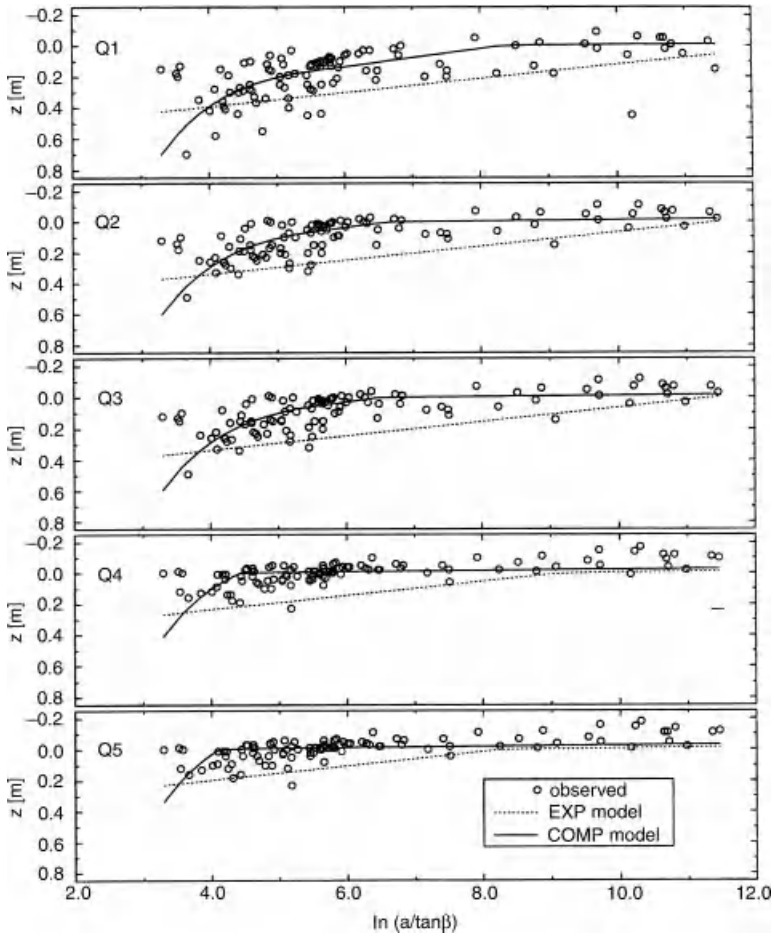


Figure 6.9 *Predicted local water table levels for five discharges (0.1 to 6.8 mm/hr) in the Saeternbekken MINIFELT catchment, Norway, using global parameters calibrated on catchment discharge and recording borehole data from October–November 1987 (after Lamb et al., 1997, with kind permission of John Wiley and Sons).*

Figure 6.9 shows the results of local calibration of saturated transmissivity, T_o, and effective porosity, $\delta\theta$, parameters for two of the recording boreholes for the 1987 calibration period. The results show a significant improvement. On the basis that the variation of T_o might be much greater than that of the effective porosity, an apparent T_o value was calculated for each piezometer to match the observed and predicted water table elevations at each point for each of the five available data sets (points showing significant ponding were excluded). A similar approach to adjusting local transmissivities has also been used by Jordan (1994) and Seibert *et al.* (1997). In this case, it was found that the adjustment necessary was positively correlated with the topographic index, implying higher apparent transmissivities at higher index values. The correlations were stronger for the composite model with coefficients varying from

0.53 to 0.92 for the five discharges. The form of this relationship then suggested a modified index of similarity as a power function of ($a \tan \beta$). For this catchment, this resulted in a much more rapid increase in predicted saturated area with discharge than in the unmodified model. A similar correlation and behaviour was found later for a small humid catchment in the Jizera Mountaints of the Czech Republic (Blazkova *et al.*, 2002a).

This study raises a number of interesting issues. Firstly, although we do expect from many different field studies that the soil characteristics should be heterogeneous in space, use of local measurements to calibrate local parameter values restricts the value of the internal data in evaluating the model. In addition, such local calibrations of, for example, a transmissivity value should be expected to be dependent on both model structure and model grid scale (e.g. Saulnier *et al.*, 1997b). Finally local calibrations can only be made for points at which measurements are available; the Saeternbekken MINIFELT catchment is very unusual, if not unique, in having so many internal measurement points. Even in this very small catchment, there is the problem of extrapolating to other points in the catchment. In the Lamb *et al.* (1997) study, this could be achieved because a suitable correlation with the ($a/\tan \beta$) index was found, but it cannot be concluded that this will generally be the case. In fact, such a correlation might be an indication that there is a structural deficiency in the model formulation. The positive correlation between apparent transmissivity and topographic index, for example, might be an indication that the topographic analysis was overestimating the effective upslope contributing areas to each point.

The issues are not specific to this TOPMODEL study but are generic to any application of any distributed model, including most process-based models, for which some internal data are available for evaluation of the model predictions. One approach to model evaluation might then be a type of split sample test in which only part of the internal data are used to test whether local parameter calibration might be necessary and the remainder are held back to test the resulting, improved, predictions. Those held back, however, will also have their local characteristics, implying a limit to how far such distributed predictions can be validated. Some uncertainty in such predictions is, then, inevitable and should, if possible, be quantified. Uncertainty in the predictions for the Saeternbekken study have been made by Lamb *et al.* (1998b).

6.5 TOPKAPI

An interesting variant on the TOPMODEL approach to distribution function models is TOPographic Kinematic Approximation and Integration – TOPKAPI (Todini, 1995; Ciarapica and Todini, 2002; Liu and Todini, 2002). TOPKAPI attempts to account for two additional features of hillslope flow processes, relative to the TOPMODEL approach. The first is that downslope flows in the unsaturated zone might also contribute to the storage at any point in the catchment. By including such fluxes, the possibility that there is no downslope saturated zone flow and very small downslope unsaturated zone effectively allows a more dynamic formulation of the upslope contributing area to a point. Secondly, the assumption of instantaneous redistribution of soil water storage on the hillslope, assumed for the saturated zone in TOPMODEL, is relaxed. In TOPKAPI, an approximate relationship between downslope flux and the integrated profile of soil moisture is assumed as:

$$q = \tan \beta \; K_s L \tilde{\theta}^{\alpha} = C\eta^{\alpha} \tag{6.8}$$

where K_s is the saturated hydraulic conductivity of the soil (assumed constant with depth), L is the depth of the soil layer, $\tilde{\theta}$ is the profile integrated average relative soil moisture content ($= \frac{1}{L}\int_0^L (\theta - \theta r)/(\theta_s - \theta_r)$), α is a coefficient in the Brooks–Corey relationship between relative hydraulic conductivity and moisture

content ($\alpha = 3 + 2\lambda$ in Equation (B5.4.2), see Box 5.4) and the total profile moisture content is $\eta = (\theta_s - \theta_r)L\tilde{\theta}$ and the coefficient

$$C = \frac{LK_s \tan\beta}{(\theta_s - \theta_r)^{\alpha} L^{\alpha}} \tag{6.9}$$

At a point, the continuity equation may then be written as

$$\frac{\partial \eta}{\partial t} = -\frac{\partial q}{\partial x} + p = -C\frac{\partial \eta}{\partial x} + p \tag{6.10}$$

where x is plan distance and p is the rainfall rate. This is a kinematic wave equation for the change in total profile soil moisture with time. The solution to Equation (6.10) is greatly simplified if it assumed that the rate of change of η is everywhere constant in space. This allows an integration to the basin scale to derive an expression for the rate of change of the total storage volume with time as

$$\frac{dV}{dt} = -\bar{C}\left[\frac{\alpha + 1}{\alpha N x}\right]^{\alpha} V^{\alpha} + Nxp \tag{6.11}$$

where

$$\frac{1}{\bar{C}} = \left[\sum_{i=1}^{N} \frac{\left(\frac{k}{N}\right)^{\frac{\alpha+1}{\alpha}} - \left(\frac{k-1}{N}\right)^{\frac{\alpha+1}{\alpha}}}{C_i^{1/\alpha}}\right]^{\alpha} \tag{6.12}$$

where the summation is taken over all the N pixels in the catchment, k represents the total number of pixels contributing to point i and V is the total storage in the catchment. In the earliest versions of TOPKAPI (Ciarapica, 1998; Ciarapica and Todini, 2002), the equations were solved numerically but later Liu and Todini (2002) developed analytical solutions for the storage at successive times. The theory, then, allows the calculation of local storage given the total storage on the catchment and local values of the $1/\bar{C}$ index in a similar way to the original TOPMODEL formulation. Ciarapica (1998) has applied the TOPKAPI model to the Montano del Reno basin in Italy and the Can Vila basin in Spain with comparisons to the ARNO and SHE models. Liu and Todini (2002) report on an application to the large Arno basin at over 8000 km^2, using grid elements of 1 km^2. They suggest that the TOPKAPI theory ensures that parameters retain their physical significance, even when used at larger scales. Liu *et al.* (2005) have used TOPKAPI for flood forecasting at even larger scales, on the 10 000 km^2 Upper Xixian catchment in China. These newer versions use integrations of the equations over the grid square and the addition of layers in the vertical to allow for interception and deeper groundwater fluxes. Sinclair and Pegram (2010) have used TOPKAPI to estimate soil moisture for 1 km scale cells across South Africa, for comparison with satellite-derived estimates.

6.6 Semi-Distributed Hydrological Response Unit (HRU) Models

As noted in the introduction to this chapter, one method of relating hydrological responses to characteristics of the landscape has arisen naturally out of the use of GIS in hydrological modelling. A GIS is commonly used to store data derived from soil maps, geological maps, a digital elevation map and a vegetation classification. They can also be used to organise remote sensing and other types of image. These different maps do not generally provide information of direct use in hydrological modelling but they certainly provide information that is relevant to hydrological modelling. By overlaying the different types of information, a classification of different elements of the landscape into hydrological response

units (HRUs) can be obtained (e.g. Figure 2.7). This is a relatively easy task with a modern GIS – at least, relatively easy once all the different sources of information have been stored and properly spatially registered in the GIS database (which can be very time consuming). The HRUs defined in this way may be irregular in shape where overlays of vector data are used, or based on regular elements where a raster (grid or pixel) database is used. A variation on a grid-based calculation scheme is the hexagonal discretisation used in the DHMS system of Vinogradov *et al.* (2011). Calculations might be made on every grid element in space, e.g. LISFLOOD (De Roo *et al.*, 2000; van der Knijff *et al.*, 2008), or similar HRUs within the catchment discretisation will often be grouped together into a single unit for calculation purposes, as in the Grouped Response Units of the SLURP model (Kite, 1995) and the Runoff Formation Complexes of the DHMS model (Vinogradov *et al.*, 2010). It is these groupings that are then used to predict the distribution of responses within the catchment.

The difficult part of this type of modelling is working out how to represent the hydrological response of each HRU and its contribution of discharge to the stream network. This varies significantly between different models. Some models use a conceptual storage model to represent each HRU element (e.g the SLURP, LISFLOOD and DHMS models; the USGS PRMS system (Leavesley and Stannard, 1995; Flügel, 1995); the HYPE model (Lindstrom *et al.*, 2010); the ARC/EGMO model (Becker and Braun, 1999); the ECOMAG model (Motovilov *et al.*, 1999); the HYDROTEL model (Fortin *et al.*, 2001); the ARC Hydro model (Maidment, 2002); the PREVAH model (Viviroli *et al.*, 2009); and the WATFLOOD model (Cranmer *et al.*, 2001; Bingeman *et al.*, 2006), which is included in the Green Kenue modelling system of Environment Canada).

Other models use a loss function to calculate a rainfall excess which is then routed to the catchment outlet, in some cases assuming a distribution of storage capacities within each HRU (e.g. Schumann and Funke, 1996). As the HRU element scale becomes finer, and the hydrological description becomes more continuum physics-based, this type of model approaches the fully distributed models of Chapter 5; the distinction we draw here, by including HRU models in this chapter, is that they do not explicitly aim to solve the continuum flow equations for surface and subsurface flow but allow the grouping of elements to reduce the number of calculations required. Finally, there have been attempts to define HRUs directly in terms of the dominant hydrological processes in each unit of the landscape (e.g. Uhlenbrook and Leibundgut, 2002; Scherrer and Naef, 2003).

One widely used example is the USDA Soil and Water Assessment Tool (SWAT) model of Arnold *et al.* (1998) which is based on the USDA Soil Conservation Service (SCS) curve number method for fast runoff generation (although other options are also provided). A particular advantage of SWAT is that, because it was developed by the United States Department of Agriculture, it can be freely downloaded (this is also the case with the Green Kenue system from Environment Canada, see Appendix A). It also has many additional components than just predicting runoff, including the transport of sediments, nutrients, pathogens and pesticides, and the prediction of crop growth based on the soil moisture, temperature and nutrient predictions. The model is also provided with databases of default parameter values for different types of soils, crops and natural vegetation. Many other organisations have taken advantage of this to organise workshops and training courses as a way into using the model. The SWAT model has been so widely used in the last decade that a more complete description is provided in Box 6.2.

In fact, there are a number of examples of this type of model that use the SCS curve number method for predicting runoff generation, for example AFFDEF (Brath *et al.*, 2004, 2006). The SCS method has an interesting history and it will continue to be used because of the way in which databases of the SCS curve number can be related to distributed soil and vegetation information stored within a GIS. The SCS method has its origins in empirical analyses of rainfall–runoff data on small catchments and hillslope plots. It is commonly regarded as a purely empirical method for predicting runoff generation with no basis in hydrological theory. It is also commonly presented in hydrology texts as an infiltration equation

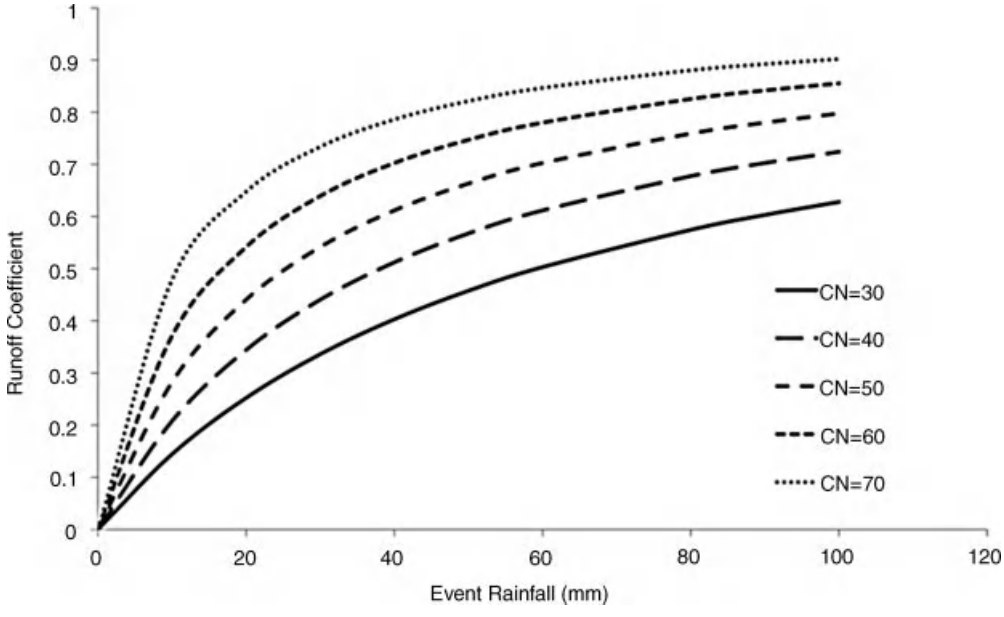


Figure 6.10 *Relationship between storm rainfall and runoff coefficient as percentage runoff predicted by the USDA-SCS method for different curve numbers.*

or a way of predicting Hortonian infiltration excess runoff (e.g. Bras, 1990), and a study by Yu (1998) has attempted to give it a basis in physical theory by showing that partial area infiltration excess runoff generation on a statistical distribution of soil infiltration characteristics gives similar runoff generation characteristics to the SCS method (see Box 6.3).

This is of some interest in itself, but the method becomes even more interesting if we return to its origins as a summary of small catchment rainfall–runoff measurements by Mockus (1949). Mockus related storm runoff to rainfalls and showed that the ratio of cumulative discharge to cumulative storm rainfall shows a characteristic form (see Figure 6.10). In the past, the storm runoff may have been widely interpreted as infiltration excess runoff, but this is not a necessary interpretation. At the small catchment scale, the measured runoff in some of the original experiments will almost certainly have included some subsurface-derived water due to displacement, preferential flows or subsurface contributions from close to the channel. Certainly the method has since been applied to catchments and hydrological response units that are not dominated by infiltration excess runoff generation. Steenhuis *et al.* (1995) have already interpreted the SCS method in terms of a variable saturated contributing area, excluding, in their analysis, some data from high intensity events that might have produced an infiltration excess runoff (see also Box 6.3). A sufficient view of the method is that it incorporates some empirical knowledge of fast runoff generation, by whatever method, at the small catchment scale into a simple functional form. It may be necessary to check whether that form is appropriate in any particular application but it may be an appropriate method to use at the HRU scale since it encapsulates knowledge gained at similar scales. In considering the scale dependency of an HRU model, therefore, it might be more appropriate than, say, any of the point scale infiltration equations described in Box 5.2, even though they would normally be considered more "physically based".

Scale is an issue in HRU modelling. The type of HRU representation used to predict runoff generation might be expected to vary both with the hydrological environment and with the spatial scale at which the HRU elements are defined (see also Chapter 9). Each HRU is generally considered as homogeneous in its parameter values and response so that, for example, if surface runoff is calculated to occur, it will do so over the whole HRU. The HRUs are also often treated independently with no explicit routing of downslope surface or subsurface flows between HRU elements, only routing of runoff to the nearest channel. This assumption of independence of position in the catchment is, in fact, necessary if HRUs with similar characteristics are to be grouped together.

One advantage of the HRU approach is that the calculated responses can be mapped back into space using visualisation routines in a GIS so that this can, in principle at least, provide information for a spatial evaluation of the predictions. The major disadvantage is the way in which each HRU is considered to be spatially homogeneous, an assumption that must become poorer as the HRU gets larger. At very large scales, new methods may be needed, as in the macroscale hydrological models that are discussed in Chapter 8. At the current time, we have no theory for predicting the appropriate model structure or parameter values at one scale, given information at another scale. Indeed, it has been argued that we will never have such a theory and that we will need to resort to models that are essentially scale dependent (Beven, 1995, 1996, 2006b). There is an alternate view, expressed by Vinogradov *et al.* (2010) that we should only use models that work well with common parameter values at all scales and that, if it is shown that they do not, then there is something wrong with the model structure. There is work in progress to show that their DHMS model can be successfully applied in this way across a range of catchments and climatic regions (though, as we see in Chapter 7, the declaration of "success" in this respect depends on how the simulations are evaluated).

In applying HRU models, I still feel that the scale dependence of the parameter values used in representing each unit will generally need to be considered (see also Chapter 9 for more discussion of the scale issue). Because there are many HRUs and several (or many) parameters are needed for each HRU, it is not easy to calibrate parameters by an optimisation process (though see Section B6.2.5). In this, such models face similar problems to fully distributed physically based models. The GIS may store soil type and vegetation type but the information about model parameters for each classification may be highly uncertain and may not be independent (e.g. rooting depth of a certain vegetation type may depend on soil type or the hydraulic characteristics or erosivity of a soil type might depend on the type of land use). The real hydrological response of an HRU may depend on the heterogeneity within the element, which might not be well represented by homogeneous "effective" parameter values. This has to be an important limitation on this type of model structure but, as discussed in Chapter 5, it is essentially a limitation of all types of model given the limitations of our knowledge of how to represent the detailed variability of hydrological systems. Again, it suggests that the predictions of such models should be associated with some estimate of uncertainty.

6.7 Some Comments on the HRU Approach

With some exceptions, the semi-distributed HRU models that use conceptual storage representations as a basis for their runoff predictions are the direct continuation of the lumped conceptual or explicit soil moisture accounting models of the past. There are many such models that have taken conceptualisations developed at the catchment scale to smaller spatial units, taking advantage of all the modern possibilities of GIS overlays. In some cases, for example the SWAT model of Box 6.2, the pedigree of the conceptualisation from earlier models is quite explicit. The exceptions are those few studies that have started anew by allowing that the process representations required in different areas of the landscape might be quite different. The idea then is that if an assessment of the dominant processes can be made on the basis of

a hydrological survey, then appropriate process representations can be used for different types of source areas (e.g. Uhlenbrook and Leibendgut, 2002; Scherrer and Naef, 2003).

The disadvantage of this approach is the fundamental problem of calibration of the parameters. As with any distributed model there may be hundreds or thousands of parameter values that must be defined before a simulation can be run. Databases of parameter values are available for such models to provide a working first estimate, but there is no guarantee that they will be the effective values required for particular places and at a particular discretisation scale to give good predictions (Beven, 2000). Indeed, there is no guarantee that the conceptual nature of the process representations will be able to give a good simulation of the surface and subsurface runoff generation in a particular application. Calibration of parameter values will help, but clearly such models pose great difficulties for calibration given the very large number of parameters of different types that could be changed. It is possible to reduce the dimensionality of the calibration problem, for example, by fixing the relative magnitudes of the spatial pattern of the initial estimates and adjusting the whole field of values by a calibration multiplier. There is no firm assurance, however, that this will provide the right runoff predictions in the right places and for the right reasons, even if the predictions of the hydrograph are improved by calibration.

Others have claimed generality for their process representations, suggesting even that they might be more generally representative than those of the continuum distributed models of Chapter 5. Vinogradov *et al.* (2010) make this claim on the basis that the model has produced good results over a wide range of scales using parameter values estimated only on the basis of soil and vegetation characteristics. They are correct when they assert that any model that claims generality should be widely tested in this way, but it is difficult to accept a model as general that ignores the effects of topographic convergence and divergence on surface and subsurface flows and uses point scale parameter values to predict runoff production over each HRU as if it was homogeneous (though in this respect their model is similar to many others). But this takes us back to the perceptual model of the individual hydrologist as to what is important: on large catchments in the steppes of Russia, topographic convergence and divergence might seem much less important than in small humid temperate catchments in the uplands of the UK.

Does all this matter if good predictions of discharge can be achieved over a wide range of catchments, especially if good predictions can be achieved while avoiding the problems intrinsic to catchment-specific model calibration with such models? This class of models is, for the most part, empirical in nature (even though the term "physically based" is now seen quite often as a descriptor for SWAT, including in the recent review by Gassman *et al.*, 2010). There is then an *instrumentalist* argument that empirical success in prediction is an adequate justification. But, what then should be considered as a hydrological success in this respect, given that we know that there are limitations of input data, model structures and the observations with which a model might be compared (e.g. Harmel *et al.* 2006). This is a question that we return to several times in the chapters that follow in considering model calibration and evaluation, predicting future change in catchment runoff, considering what might be expected of the next generation of rainfall–runoff models, predicting the response of ungauged catchments, and models of everywhere.

6.8 Key Points from Chapter 6

- In considering the variability of hydrological responses within a catchment area, it may be difficult, and may not be necessary, to simulate those responses in a fully distributed way if a simpler way can be found of representing the distribution of responses in the catchment. This implies finding a way to define whether different points in a catchment act in hydrologically similar ways.
- The Probability Distributed Moisture (PDM) model does this in terms of a purely statistical representation of conceptual storage elements. It is an attempt to represent the variability in the catchment but does not allow any mapping of the predictions back into the catchment space for comparison with

perceived hydrological processes in the catchment, so can be considered as an extension of a lumped storage model. It has an advantage over lumped storage models that the parameters may be easier to estimate in an optimisation exercise. An extension that includes explicit routing between grid elements (the Grid to Grid, G2G, model) was also described.

- The rainfall–runoff model, TOPMODEL, makes use of an index of hydrological similarity based on topography and soils information that allows the model predictions to be mapped back into space (see Box 6.1). Calculations are made based on the distribution of the index, which greatly reduces the computer resources required. The TOPMODEL concepts are not, however, applicable everywhere, particularly in catchments subject to strong seasonal drying when the basic assumptions underlying the index break down.
- The simplicity of the TOPMODEL calculations have allowed the interaction between grid resolution of the topographic analysis and calibrated parameter values to be studied in a number of applications. A similar interaction between scale of discretisation and effective parameter values should hold for more complex models, including physically based fully distributed models, but may not be so readily apparent.
- A number of recent attempts have been made to improve the theory of TOPMODEL, while retaining its advantages of simplicity.
- Models based on the definition of hydrological response units (HRUs) derived by overlaying different soil, geology, topography and vegetation characteristics in a geographical information system are becoming increasingly popular (the USDA SWAT model (Box 6.2) is particularly widely used). These models differ widely in the representation for each of the HRUs defined in this way and in the routing of flows to the catchment outlet. Such models do allow the mapping of the distribution of responses back into the catchment space using the GIS. With fine discretisations of the catchment and routing of flow between HRUs, such models may be considered as simplified fully distributed models. Where similar HRU elements are grouped together for calculation purposes, they may be considered as distribution function models.
- HRU models might involve specifying hundreds or thousands of parameter values. Default values might be available in associated databases for different models but we should expect that effective values required to get good predictions might vary with discretisation scale and model structure. The question of how to define hydrologically good performance given uncertainties in input data, model structure and discharge observations recurs in the remaining chapters.

Box 6.1 The Theory Underlying TOPMODEL

B6.1.1 Fundamental Assumptions of TOPMODEL

The development of the TOPMODEL theory presented here is based on the three assumptions outlined in the main text:

A1 There is a saturated zone that takes up a configuration as if it was in equilibrium with a steady recharge rate over an upslope contributing area a equivalent to the local subsurface discharge at that point.
A2 The water table is near to parallel to the surface such that the effective hydraulic gradient is equal to the local surface slope, $\tan\beta$.
A3 The transmissivity profile may be described by an exponential function of storage deficit, with a value of T_o when the soil is just saturated to the surface (zero deficit).

B6.1.2 Steady Flow in the Saturated Zone and the Topographic Index

Under these assumptions, at any point i on a hillslope the downslope saturated subsurface flow rate, q_i, per unit contour length (m^2/h) may be described by the equation:

$$q_i = T_o \tan\beta \exp(-D_i/m) \tag{B6.1.1}$$

where D_i is the local storage deficit, m is a parameter controlling the rate of decline of transmissivity with increasing storage deficit, and T_o and $\tan\beta$ are local values at point i. Note that $\tan\beta$ is used to represent the hydraulic gradient because the slope is calculated on the basis of elevation change per unit distance in plan (rather than along the hillslope).

Then, under the assumption that, at any time step, quasi-steady-state flow exists throughout the soil, so that local subsurface discharge can be assumed equivalent to a spatially homogeneous recharge rate r (m/h) entering the water table, the subsurface downslope flow per unit contour length q_i is given by:

$$q_i = ra \tag{B6.1.2}$$

where a is the area of the hillslope per unit contour length (m^2) that drains through point i.

By combining (B6.1.1) and (B6.1.2) it is possible to derive a formula for any point relating local water table depth to the topographic index $\ln(a/\tan\beta)$ at that point, the parameter m, the local saturated transmissivity, T_o, and the effective recharge rate, r:

$$D_i = -m \ln\left(\frac{ra}{T_o \tan\beta}\right) \tag{B6.1.3}$$

Note that when the soil is saturated, the local deficit is zero; as the soil dries and the water table falls, numerical values of storage deficit get larger. An expression for the catchment lumped, or mean, storage deficit ($\overline{D}$) may be obtained by integrating (B6.1.3) over the entire area of the catchment (A) that contributes to the water table. In what follows, we express this areal averaging in terms of a summation over all points (or pixels) within the catchment:

$$\overline{D} = \frac{1}{A}\sum_i A_i \left[-m \ln \frac{ra}{T_o \tan\beta}\right] \tag{B6.1.4}$$

where A_i is the area associated with the i point (or group of points with the same characterisitcs). In spatially integrating the whole catchment, it is also implicitly required that (B6.1.4) holds even at locations where water is ponded on the surface ($D_i < 0$). This assumption can be

justified on the basis that the relationship expressed by (B6.1.1) is exponential and that, for many catchments, surface flow is likely to be relatively slow due to vegetation cover. However, it is also possible to correct the average deficit calculation to take account of the varying contributing area, as shown below.

By using (B6.1.3) in (B6.1.4), if it is assumed that r is spatially constant, $\ln r$ may be eliminated and a relationship found between mean water table depth, local water table depth, the topographic variables and saturated transmissivity. This has the form:

$$D_i = \overline{D} + m\left[\gamma - \ln\frac{a}{T_o \tan\beta}\right] \tag{B6.1.5}$$

where $\ln(a/T_o \tan\beta)$ is the soil-topographic index of Beven (1986a) and

$$\gamma = \tfrac{1}{A}\sum_i A_i \ln\tfrac{a}{T_o \tan\beta}.$$

A separate areal average value of transmissivity may be defined thus:

$$\ln T_e = \tfrac{1}{A}\sum_i A_i \ln T_o.$$

Equation (B6.1.5) may now be rearranged to give:

$$\frac{\left(\overline{D} - D_i\right)}{m} = -\left[\lambda - \ln\frac{a}{\tan\beta}\right] + \left[\ln T_o - \ln T_e\right] \tag{B6.1.6}$$

where $\lambda = \tfrac{1}{A}\sum_i A_i \ln\tfrac{a}{\tan\beta}$ is a topographic constant for the catchment.

Equation (B6.1.6) expresses the deviation between the catchment average water table depth (or deficit) and the local water table depth (or deficit) at any point in terms of the deviation of the local topographic index from its areal mean, and the deviation of the logarithm of local transmissivity from its areal integral value. The relationship is scaled by the parameter m. In particular, we are interested in where the local deficit is zero: this is the area that will act as a saturated contributing area. As the mean deficit changes as the catchment wets and dries, so will the contributing area.

Similar relationships can be derived for other transmissivity profile assumptions, relaxing assumption A3 above. Ambroise *et al*. (1996a) for example, analyse the case of linear and parabolic transmissivity functions; Iorgulescu and Musy (1997) and Duan and Miller (1997) have generalised the analysis to all power law transmissivity functions. Different assumptions about the transmissivity result in different topographic index functions. For the power law function,

$$q_i = T_o \tan\beta(1 - D_i/M)^n \tag{B6.1.7}$$

where M is a maximum gravity drainage storage in the soil profile. The equivalent soil-topographic index is

$$(a/T_o \tan\beta)^{1/n} \tag{B6.1.8}$$

and the equation relating mean storage deficits to local deficits is

$$\frac{(1 - D_i/M)}{\left(1 - \overline{D}/M\right)} = \left[\frac{a}{T_o \tan\beta}\right]^{1/n} \Big/ \left[\frac{1}{A}\sum_i A_i \left(\frac{a}{T_o \tan\beta}\right)^{1/n}\right] \tag{B6.1.9}$$

B6.1.3 The Topographic Index as an Index of Hydrological Similarity

The implication of (B6.1.6) is that each point in the catchment with the same value of the topographic index, $\ln(a/T_o \tan\beta)$, is predicted as responding in a hydrologically similar way.

Thus it is not necessary to make calculations for every point in space but only for different values of the topographic index. In most applications of TOPMODEL, the distribution function of the topographic index (see, for example, Figure 6.5) is discretised into a number of increments representing appropriate proportions of the catchment. At every time step, those increments with high values of the topographic index are predicted as being saturated or having low storage deficits. This is shown schematically in Figure B6.1.1. The calculations for each increment are completed by an unsaturated zone component.

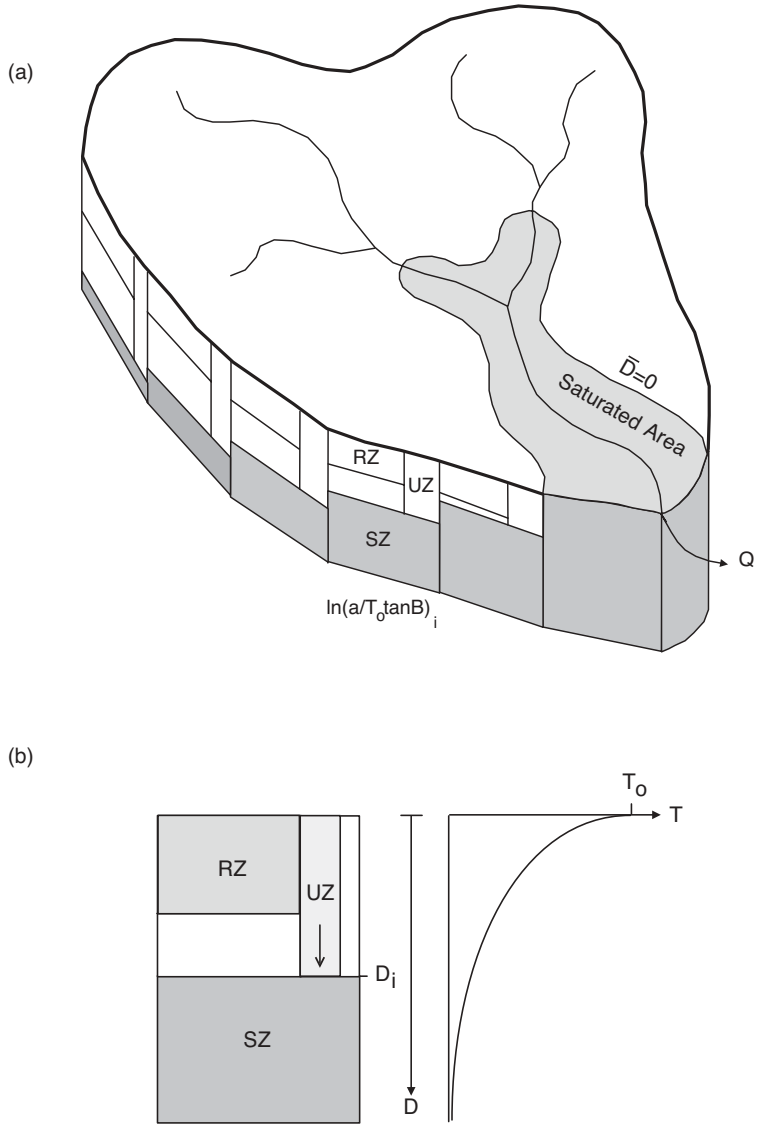


Figure B6.1.1 *Schematic diagram of prediction of saturated area using increments of the topographic index distribution in TOPMODEL.*

B6.1.4 Moisture Accounting: Unsaturated Zone Fluxes

The basic soil structure illustrated in Figure B6.1.1 may be used to accommodate a variety of unsaturated zone process descriptions as defined by the modeller. One formulation that has been adopted in past TOPMODEL applications assumes that the root zone store for each topographic index value is depleted only by evapotranspiration and that water is added to the unsaturated zone drainage store only once the root zone reaches field capacity. The drainage is assumed to be essentially vertical and a drainage flux q_v [LT^{-1}] is calculated for each topographic index class.

Expressed in terms of storage deficit, Beven and Wood (1983) suggested that a suitable functional form for the vertical flux q_v at any point i is:

$$q_v = \frac{S_{uz}}{D_i t_d} \tag{B6.1.10}$$

where S_{uz} [L] is storage in the unsaturated (gravity drainage) zone; D_i is the local saturated zone deficit due to gravity drainage and dependent on the depth of the local water table [L]; t_d is a time constant, expressed as a mean residence time for vertical flow per unit of deficit [TL^{-1}]. Equation (B6.1.10) is the equation of a linear store but with a time constant $\{D_i t_d\}$ that increases with increasing depth to the water table. Note that there is no physical justification for this functional form, but it has the advantages that it allows for longer residence times and slower drainage rates for lower values of the index where the water table is predicted as being deeper below the surface and yet it only introduces one parameter value. It has generally been found that modelling results are not very sensitive to this parameter.

Accounting for evapotranspiration with a minimal number of parameters poses a problem of similar complexity to that of the unsaturated zone drainage. TOPMODEL follows the widely adopted practice of calculating actual evapotranspiration, E_a, as a function of potential evaporation, E_p, and root zone moisture storage for cases where E_a cannot be specified directly. In the TOPMODEL description of Beven (1991a), evaporation is allowed at the full potential rate for water draining freely in the unsaturated zone and for predicted areas of surface saturation. When the gravity drainage zone is exhausted, evapotranspiration may continue to deplete the root zone store at the rate E_a, given by:

$$E_a = E_p \frac{S_{rz}}{S_{r\max}} \tag{B6.1.11}$$

where the variables S_{rz} and $S_{r\max}$ are, respectively, root zone storage [L] and maximum available root zone storage [L]. If some effective root zone depth z_{rz} [L] can be assumed, $S_{r\max}$ can be estimated approximately from:

$$S_{r\max} = z_{rz}\left(\theta_{fc} - \theta_{wp}\right) \tag{B6.1.12}$$

where θ_{fc} [-] is moisture content at *field capacity* and θ_{wp} [-] is moisture content at wilting point. For calibration it is only necessary to specify a value for the single parameter $S_{r\max}$. An effective value for $S_{r\max}$ might be greater than that suggested by Equation (B6.1.12) due to capillary rise of water into the root zone under dry conditions.

The flux of water entering the water table locally at any time is q_v. This drainage is also a component of the overall recharge of the lumped saturated zone. To account for the catchment average water balance, all the local recharges must be summed. If Q_v is the total recharge to the water table in any time step, then:

$$Q_v = \sum_i q_{v,i} A_i \tag{B6.1.13}$$

where A_i is the fractional area associated with topographic index class i as a fraction of total catchment area.

B6.1.5 Moisture Accounting: Saturated Zone Fluxes

Output from the saturated zone is given by the baseflow term, Q_b. This may be calculated in a distributed sense by the summation of sub-surface flows along each of M stream channel reaches of length one. Recalling (B6.1.1), we may write:

$$Q_b = \sum_{j=1}^{M} l_j (T_o \tan \beta) e^{-D_j/m} \tag{B6.1.14}$$

Substituting for D_j using (B6.1.5) and rearranging, it can be shown that:

$$Q_b = \sum_{j} l_j a_j e^{-\gamma - \overline{D}/m} \tag{B6.1.15}$$

Since a_j represents contributing area per unit contour length, then:

$$\sum_{j=1}^{m} l_j a_j = A \tag{B6.1.16}$$

Therefore:

$$Q_b = A e^{-\gamma} e^{-\overline{D}/m} \tag{B6.1.17}$$

where A is the total catchment area (m^2). It is therefore possible to calculate baseflow in terms of the average catchment storage deficit ($\overline{D}$):

$$Q_b = Q_o e^{-\overline{D}/m} \tag{B6.1.18}$$

where $Q_o = A e^{-\gamma}$ is the discharge when $\overline{D}$ equals zero. This is the same form as that originally assumed by Beven and Kirkby (1979). Solution of (B6.1.18) for a pure recession in which recharge is assumed to be negligible shows that discharge has an inverse or first-order hyperbolic relationship to time as:

$$\frac{1}{Q_b} = \frac{1}{Q_o} + \frac{t}{m} \tag{B6.1.19}$$

Thus, if (B6.1.18) is an appropriate relationship to represent the subsurface drainage of a given catchment, a plot of $1/Q_b$ against time should plot as a straight line (e.g. Figure B6.1.2) with slope $1/m$. Then, given at least some recession curves that are not greatly influenced by evapotranspiration or snowmelt processes, it should be possible to set the value of m which will need minimal calibration.

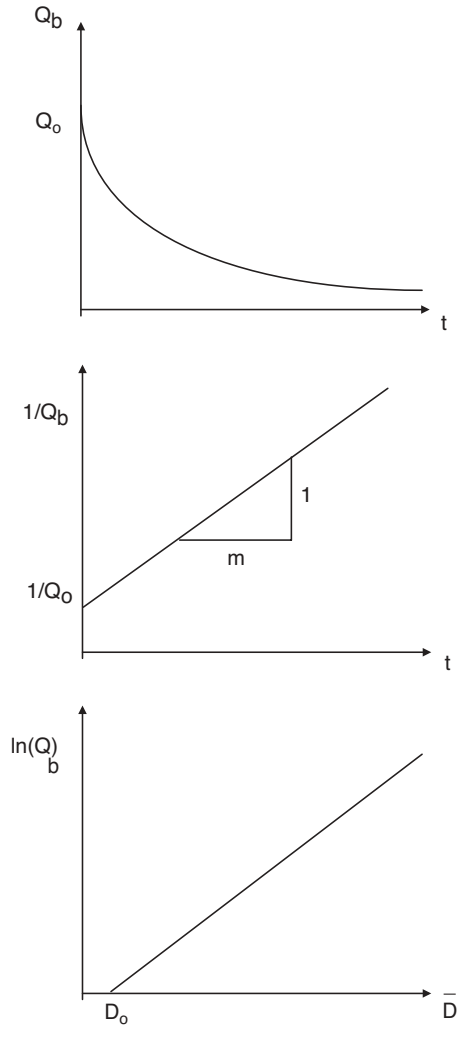


Figure B6.1.2 *Derivation of an estimate for the TOPMODEL m parameter using recession curve analysis under the assumption of an exponential transmissivity profile and negligible recharge.*

The catchment average storage deficit before each time step is updated by subtracting the unsaturated zone recharge and adding the baseflow calculated for the previous time step, thus:

$$\overline{D}_t = \overline{D}_{t-1} + \left[Q_{b_{t-1}} - Q_{v_{t-1}}\right] \tag{B6.1.20}$$

Equation (B6.1.18) can be used to initialise the saturated zone of the model at the start of a run. If an initial discharge, $Q_{t=0}$, is known and assumed to be only the result of drainage from the saturated zone, Equation (B6.1.18) can be inverted to give a value for $\overline{D}$ at time $t = 0$ as:

$$\overline{D} = -m \ln\left(\frac{Q_{t=0}}{Q_o}\right) \tag{B6.1.21}$$

Once $\overline{D}$ is known, local values of initial storage deficit can be calculated from (B6.1.5). Other forms of transmissivity function can also be used to derive different forms of index and recession curves. Lamb *et al.* (1997) show how an arbitrary recession curve can be used in a generalised TOPMODEL.

B6.1.6 Runoff Routing and Subcatchment Structure

For many catchments, especially large ones, it may be inappropriate to assume that all runoff reaches the catchment outlet within a single time step. In the original TOPMODEL formulation by Beven and Kirkby (1979), an overland flow delay function was calculated as a distance related delay. The time taken to reach the basin outlet from any point was assumed to be given by:

$$t_j = \sum_{i=1}^{N} \frac{x_i}{v^* \tan \beta_i} \tag{B6.1.22}$$

where t_j is the time delay for point j, x_i is the plan flowpath length and $\tan \beta$ the slope of the ith segment of a flow path comprising N segments between point j and the catchment outlet. If the velocity parameter v^* (m/h) is assumed constant then this equation allows a unique time delay histogram to be derived on the basis of basin topography for any runoff contributing area extent. This is, in effect, a variation of the classic time-area routing method (see Figure 2.2), but developed so as to relate the runoff time delay histogram dynamically to the size of the source area.

Channel routing effects were considered by Beven and Kirkby (1979) using an approach based on an average flood wave velocity for the channel network, this being related non-linearly to total outflow. The approach was an explicit approximation to a kinematic wave channel routing algorithm and is not recommended, since it is not always stable. Most applications have been based on a simple constant wave speed routing algorithm, equivalent to the algorithms based on channel network width function used by Surkan (1969), Kirkby (1976), Beven (1979b) and Beven and Wood (1993), which has the advantage that it introduces only a single wave speed parameter. In a single event version of TOPMODEL, Saulnier *et al.* (1997b) adopted a routing method based on a unit hydrograph derived by the DPFT-ERUHDIT method of Duband *et al.* (1993). Since the unit hydrograph can be expected to reflect the time response of subsurface as well as surface flow processes, they show that this choice of routing algorithm has an effect on the other parameters of the model. Some other linear routing algorithms are discussed in Section 4.5.

B6.1.7 Recent Variations on TOPMODEL

The simplicity of the TOPMODEL formulation as a way of reflecting the topographic controls on runoff generation has proven attractive now that digital terrain models for catchments are more widely available. The simplicity of the ideas has also encouraged both an assessment of the assumptions in relation to perceptions of the processes controlling the hydrological responses in different catchments and attempts to reformulate and improve the theory.

The first category of variations on TOPMODEL includes the idea of the reference level for deeper water tables proposed by Quinn *et al.* (1991), in which the hydraulic gradient is based on a characteristic water table surface rather than the soil surface. This idea was used with an exponential transmissivity profile, which may not be appropriate for a depth water table, but could be extended to more realistic transmissivity profiles. Lamb *et al.* (1997, 1998a) showed how a generalised transmissivity function could be developed on the basis of a recession curve analysis (maintaining the simplifying assumption that the pattern of saturated zone hydraulic

gradient in the catchment stays constant). Lamb *et al.* (1998b) have shown how data on the spatial distribution of water table depth in a catchment can be used to modify the index distribution to reflect heterogeneity of effective transmissivity values.

Recent attempts to reformulate the index theory include the modified index of Saulnier and Datin (2004), which avoids the assumption that a steady recharge rate and transmissivity function also apply to water ponded on the surface by averaging the local soil moisture deficits only over the non-saturated part of the catchment. Thus Equation (B6.1.5) becomes:

$$\overline{D}' - D_i = -m\left[\gamma_t' - \ln\frac{a}{T_o\tan\beta}\right] \tag{B6.1.23}$$

where $\overline{D}'$ is now the average deficit over the non-saturated area at time t, $\gamma_t' = \frac{1}{A-A_t^c}\sum_i A_i \ln\frac{a}{T_o\tan\beta}$, A_t^c is the contributing area at time t, A_i is the area of a pixel (or group of similar pixels) as before, and the summation is now taken only over unsaturated pixels.

Since, by definition, the local deficit on the saturated area is everywhere zero then:

$$\overline{D}_t = \frac{A - A_t^c}{A}\overline{D}_t' \tag{B6.1.24}$$

and, combining Equations (B6.1.23) and (B6.1.24):

$$\left(\frac{A}{A-A_t^c}\right)\overline{D} - D_i = -m\left[\gamma_t' - \ln\frac{a}{T_o\tan\beta}\right] \tag{B6.1.25}$$

with $D_i = 0$ at the critical index value:

$$\left(\ln\frac{a}{T_o\tan\beta}\right)_c = \gamma_t' + \left(\frac{A}{A-A_t^c}\right)\frac{\overline{D}}{m} \tag{B6.1.26}$$

This can be rearranged to define a function $G(A_t^c)$ that can be calculated from the cumulative distribution function of the soil-topographic index such that:

$$G(A_t^c) = \frac{\overline{D}}{m} \tag{B6.1.27}$$

where

$$G(A_t^c) = \left[\left(\ln\frac{a}{T_o\tan\beta}\right)_c - \gamma_t'\right]\left[\frac{A}{A-A_t^c}\right] \tag{B6.1.28}$$

An example of the function $G(A^c)$ is shown as $G(\kappa)$ in Figure B6.1.3. $H(\kappa)$ is the original topographic index function. This function is specific to a catchment but may be calculated during the topographic analysis prior to a simulation. At any time step, knowledge of $\overline{D}$ can be used to calculate the function $G(A^c)$ and thence the contributing area and the average deficit over the non-saturated area $\overline{D}'$. Figure B6.1.3 also shows the relative correction to the original TOPMODEL function. Thus, this turns out to be quite a simple modification and results in a more satisfactory representation of the time variation in the average (non-saturated) deficit

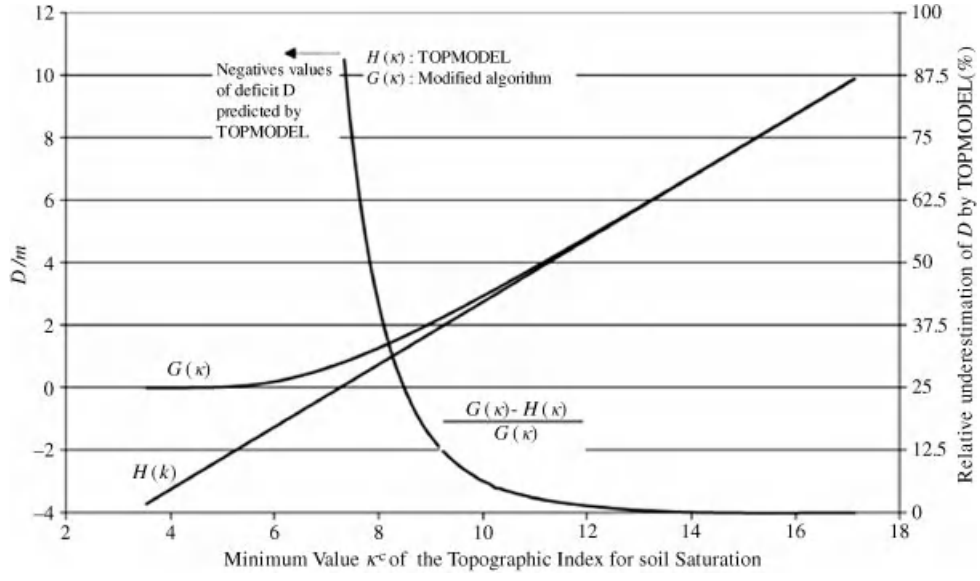


Figure B6.1.3 *Use of the function $G(A^c)$ to determine the critical value of the topographic index at the edge of the contributing area given $\frac{D}{m}$, assuming a homogeneous transmissivity (after Saulnier and Datin, 2004, with kind permission of John Wiley and Sons).*

and pattern of deficits in the catchment. This has been incorporated into the TOPDYN version used, for example, by Vincendon *et al.* (2010).

Another alternative index formulation is provided by the TOPKAPI model of Section 6.5. This attempts to relax the assumption of instantaneous redistribution of the saturated zone at each time step in formulating the new index (see Liu and Todini, 2002; Liu *et al.*, 2005). The MACAQUE model of Watson *et al.* (1999) addresses the same problem in a different way by introducing a "lateral redistribution factor" that limits the redistribution towards the steady state water table configuration allowed at each time step.

One of the limitations of the topographic index approach is the assumption that there is always downslope flow from an upslope contributing area that is constant for any point in the catchment. Improved predictions might be possible if this area was allowed to vary dynamically. Barling *et al.* (1994) showed that an index based on travel times could improve prediction of saturated areas for a single time step prediction, but did not suggest how this might be extended to a continuous time model. A dynamic TOPMODEL can also be derived by an explicit redistribution of downslope fluxes from one group of hydrologically similar points to another, where "hydrologically similar" can be based on more flexible criteria than the original topographic index. In the extreme case of every pixel in a catchment being considered separately, this approach would be similar to the distributed kinematic wave model of Wigmosta *et al.* (1994). Grouping of similar pixels results in computational efficiency that might be advantageous in application to large catchments or where large numbers of model runs are required to assess predictive uncertainty. This is the basis for a new, more dynamic version of TOPMODEL (Beven and Freer, 2001). A simpler way of allowing for changes in upslope contributing area is suggested by Vincendon *et al.* (2010) in the TOPDYN implemententation, essentially by continuously recalculating the number of upslope cells that contribute to the topographic index.

B6.1.8 Linking TOPMODEL to the Power Law Rainfall Nonlinearity

In Section 4.3, it was suggested that there is a link between the bilinear power law that has been shown to be useful in defining the effective rainfall nonlinearity in the data-based mechanistic (DBM) modelling methodology and the prediction of saturated areas in TOPMODEL. The starting point for demonstrating this link is the equation for the local deficit calculation in TOPMODEL which, assuming that the transmissivity is homogeneous in the catchment in Equation (B6.1.5) may be simplified to:

$$\frac{\left(\overline{D} - D_i\right)}{m} = -\left[\lambda - \ln \frac{a}{\tan \beta}\right] \tag{B6.1.29}$$

The edge of the saturated contributed area A_c for any mean deficit $\overline{D}$ is defined by the locus of points where $D_i = 0$ so that for those points that are just saturated

$$\frac{\overline{D}}{m} = -\left[\lambda - \ln \frac{a}{\tan \beta}\right] \tag{B6.1.30}$$

and, using B6.1.17 with $\gamma = \lambda$, the subsurface discharge is given by:

$$Q_b = Ae^{-\lambda} e^{-\left[\ln \frac{a}{\tan \beta} - \lambda\right]} \tag{B6.1.31}$$

We then assume that we can approximate the upper tail of the topographic index curve as an exponential function of the form:

$$A_c = A_o exp\left[n(\lambda - \ln \frac{a}{\tan \beta}\right] \tag{B6.1.32}$$

From Equation (B6.1.31)

$$\frac{Q_b}{Ae^{-\lambda}} = \left[\lambda - \ln \frac{a}{\tan \beta}\right] \tag{B6.1.33}$$

and combining with Equation (B6.1.32)

$$A_c = \frac{A_o}{(Ae^{-\lambda})^n} Q_b^n \tag{B6.1.34}$$

or

$$A_c \propto Q_b^n \tag{B6.1.35}$$

This is equivalent to the power law transformation of the rainfall inputs suggested in a number of DBM SDP analyses as derived from the observations (though not all the SDP transformations result in this power law form).

Box 6.2 The Soil and Water Assessment Tool (SWAT) Model

B6.2.1 The Pedigree of SWAT

The Soil and Water Assessment Tool or SWAT model is a product of the Agricultural Research Service (ARS) of the US Department of Agriculture. It is the latest generation of a long line of USDA ARS models with rainfall–runoff components that includes the Chemical, Runoff and

Erosion from Agricultural Management Systems (CREAMS), the Groundwater Loading Effects of Agricultural Management Systems (GLEAMS) and the Erosion Productivity Impact Calculator (EPIC). EPIC later evolved into the Environmental Impact Policy Climate model and is a direct descendant of the Simulator for Water Resources in Rural Basins (SWRRB) model which aimed to predict the responses of runoff and sediments to management impacts of ungauged rural basins across the USA. The latest version of the model is SWAT2009.

As their names suggest, however, these models were aimed at improving agricultural management, of which runoff prediction was only one component – albeit an important one, in that runoff acts as a driver for many of the transport processes considered. In the USA, the model has been used to simulate all of the catchment areas in the conterminous states for more than a decade in support of national policy development. It has also been widely used in different parts of the world, because it provides a flexible tool for rainfall–runoff modelling and agricultural management that, very importantly, has been made freely available (see Appendix A). At the time of writing, some 600 articles have been published that describe using SWAT. A publication list can be found at the SWAT web site. Gassman *et al.* (2007) give a tabular review of many different applications of SWAT.

B6.2.2 Definition of Hydrological Response Units in SWAT

Each catchment in a SWAT application can be subdivided into subcatchments linked by the channel network. In the USA, the subcatchments are based on the eight-digit hydrologic cataloguing units (HCU) that subdivide the conterminous USA into 2149 catchment units.

Each subcatchment can then be subdivided into a number of hydrological response units (HRUs) of homogeneous land use, management and soil characteristics. The HRUs are treated as fractions of the local subcatchment and need not be differentiated in space. Runoff predicted from each HRU is routed directly to the channel network. The hydrological components of each HRU include interception; partitioning of precipitation, snowmelt and irrigation water between surface runoff and infiltration; redistribution of water in the soil profile; evapotranspiration from the root zone; lateral subsurface flow in the soil profile; and return flow from shallow groundwater.

Fast runoff is predicted by using the Soil Conservation Service Curve Number method (SCS-CN) at daily or sub-hourly time intervals or the Green–Ampt infiltration equation using sub-hourly rainfall rates. Redistribution in the soil profile uses a storage-flux method between specified soil layers, with the possibility of preferential flow in soils with secondary structure (Arnold *et al.*, 2005). Three methods of estimating evapotranspiration are also available, depending on available weather data: the Penman–Monteith equation, the Priestly–Taylor equation and the Hargreaves equation.

Recharge below the soil profile is partitioned between shallow and deep groundwater systems. The shallow groundwater can produce return flow back to the surface and can also lose water to the atmosphere through deep rooting plants. Water that reaches the deep groundwater is assumed not to contribute to streamflow but is treated as a loss to the catchment water balance. A routine is included for accounting for the effects of field drainage on soil water and runoff (Green *et al.*, 2006). Runoff from the different mechanisms included in the hydrology component of the model are summed and routed through the channel network using either a variable-rate storage method or the Muskingum method.

The availability of the SWAT code has led to a number of variants on the model with modified components and interfaces with different GIS systems and databases (including the ESRI Arc packages) and Windows GUIs (see Gassman *et al.*, 2007) in defining HRUs. Santra *et al.* (2011) have suggested a fuzzy set approach to defining HRUs that they test using the SWAT model. Recent developments include a simplified HRU representation in which the partitioning of fluxes is based on the water balance for use in data scarce regions (Easton *et al.*, 2010; White *et al.*, 2011).

B6.2.3 Enumerating the Assumptions of SWAT

As recommended elsewhere in this book, the user of a model should always aim to list and evaluate the assumptions made by a modelling package but in the case of SWAT there is a certain difficulty in listing all the assumptions (especially if going beyond the rainfall–runoff components to all the other dependent functions). The list depends on the specific options chosen (and possibly calibrated) for a specific application.

Two critical assumptions are, however, worth making explicit:

A1 Each HRU within a subcatchment is assumed to respond homogeneously. Thus if an HRU produces surface runoff, it does so over 100% of its area, regardless of scale. Any effects of subcatchment heterogeneity (and processes such as runon infiltration) are therefore assumed to be treated implicitly by the use of effective parameter values.

A2 SWAT is intended only to predict runoff above some "baseflow" (which implies that some baseflow separation is required for model evaluation, e.g. Arnold and Allen, 1999). This then means that even the water balance assumption can be relaxed for particular applications, by allowing losses to deeper groundwater which then make no contribution to the predicted flow.

As discussed in Box 6.3, the original derivation of the SCS-CN method for predicting fast runoff was made at the small watershed scale and would therefore have allowed for heterogeneity implicitly. It is not clear, however, how far the default estimates of curve numbers and Green–Ampt infiltration parameters for different soils do so. Lyon *et al.* (2004) and Easton *et al.* (2008) have used the reconceptualisation of the curve number approach suggested by Steenhuis *et al.* (1995) as a way of representing variable source areas within SWAT. They use the TOPMODEL topographic index as a way of defining the HRUs for the SWAT model. However, most applications still make the homogeneous response assumption (A1).

B6.2.4 Additional Components Available in SWAT

The intended use of SWAT for agricultural management means that many other component processes have been added to the hydrological components: crop and forest growth routines; fertiliser, manure and pesticide additions; tillage operations and grazing; erosion and transport of sediments (using a Modified Universal Soil Loss equation), nutrients, pesticides, and pathogens; in-stream processes affecting water quality; and management strategies including irrigation and conservation practices. Many of these components require the specification of time schedules for the operations and all require the specification of additional parameter sets.

B6.2.5 Calibration and Uncertainty in the SWAT Model

Any application of SWAT requires the specification of a large number of parameter values. Whittaker *et al.* (2010), for example, report on an application to the Blue River basin in Oklahoma with 55 sub-basins and 193 HRUs each of which requires many parameters depending on the number of components included. One of the attractive features of SWAT is that it can be downloaded with a database of default parameter values on the basis that the parameters are **physically based** and therefore can be defined with some realistic range. The default parameters can be used to provide predictions of catchments for which no observational data are available for calibration (e.g. Srinivasan *et al.*, 2010). This, of course, assumes that parameter values do not need to compensate for model structural errors, heterogeneities within HRUs and scale effects. So it is perhaps not surprising that the great majority of applications include some form of manual or automatic calibration of selected model parameters and, more recently, attempts to account for uncertainty in the estimated parameter values.

It is always best to concentrate calibration effort on those parameter values to which the model outputs are most sensitive and there have been a number of sensitivity analyses of

the SWAT model in different applications. The results suggest that the sensitivity of different parameters might be different for different catchment areas (e.g. van Griensven *et al.*, 2006).

A variety of automatic calibration routines have been used with SWAT. However, the sheer number of parameters also raises issues of over-parameterisation and over-fitting (see Chapter 7). Whittaker *et al.* (2010) suggest that this is not an important problem with SWAT because of the natural regularisation of the range of feasible values that the parameters can take. They used a multiple objective genetic search algorithm to identify over 4000 parameter values in the Blue River application and suggested that a comparison of calibration and evaluation period performance indicated that over-fitting was not a problem. However, the calibrated values still surely depend on the performance measures chosen and how they are combined, and physically there is clear potential for an increase in runoff generation in one HRU to be off-set by a decrease in another HRU. Given a large number of HRUs, the potential for interaction amongst parameters is great.

SWAT has also been used with a number of uncertainty estimation approaches including PEST (Lin and Radcliffe, 2006) and GLUE (Muleta and Nicklow, 2005; Arabi *et al.*, 2007). More details of these methods can be found in Chapter 7.

B6.2.6 An Application of SWAT

Kannan *et al.* (2007) give an interesting example of the application of SWAT (in the ArcView SWAT2000 version) to the small Colworth catchment (141.5 ha) in England. It is concerned with getting the hydrological partitioning correct so as to model nutrient and contaminant transport more correctly. Most of the catchment is covered by an arable rotation, with one section of forest. Soil information was provided from the National Soils Resources Institute database. Many of the fields had been drained in the 1960s with clay tile drains, with some secondary mole drainage and sub-soiling. Raster GIS overlays of soil and land use were used to define

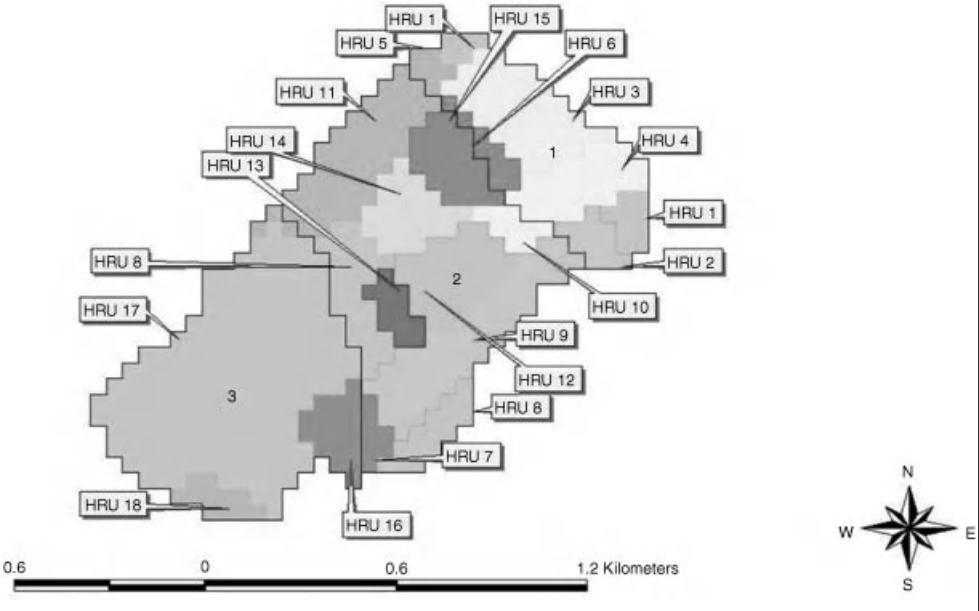


Figure B6.2.1 *Definition of hydrological response units for the application of the SWAT model to the Colworth catchment in England as grouped grid entities of similar properties (after Kannan* et al.*, 2007, with kind permission of Elsevier).*

the SWAT HRUs (Figure B6.2.1), as far as possible keeping each field as a separate HRU. This resulted in three sub-basins and 18 HRUs. Although this is a very small catchment, the model was used with a daily time step. In an initial test of the hydrological performance, baseflow was removed using the Arnold and Allen (1999) filter, potential evapotranspiration was estimated using the Hargreaves equation, and a manual calibration of seven parameters was performed, including the curve number, available water capacity, soil evaporation compensation factor and hydraulic conductivity of the top soil.

Reasonable results were obtained but, by comparison with measurements at other sites in the region, predicted surface runoff seemed to be overestimated and loss to evapotranspiration underestimated. The latter appeared to be related to problems with the crop growth model, which is based on accumulated heat in the growing season. The model was modified so that plant heat units were calculated separately for each HRU outside of the model, while crop parameters, such as maximum leaf area index, canopy height and rooting depth, were specified based on published information. The model was then recalibrated by running best estimate and upper and lower perturbed values of the most sensitive parameters. The revised model was subject to a split sample evaluation. The results from the validation period are shown in Figure B6.2.2. The proportion of surface runoff predicted was felt to be more reasonable, but the model still tends to underpredict peak flow values. This has been reported in many SWAT model applications, but here might partly be induced by the discretisation problems of using a daily time step in simulating a small catchment. The authors suggest that if the curve numbers were updated based on the water content of the upper layer rather than the whole soil profile (in a way that might also reflect surface effects such as crusting), it might lead to further improvements in the predictions.

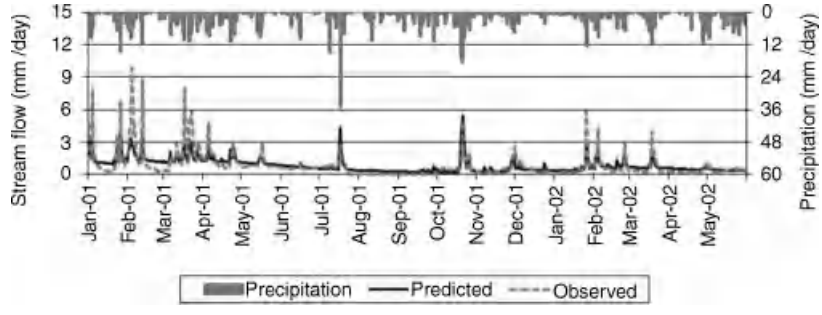


Figure B6.2.2 *Predictions of streamflow for the Colworth catchment using the revised ArcView SWAT2000 model: validation period (after Kannan* et al.*, 2007, with kind permission of Elsevier).*

B6.2.7 Some Comments on SWAT

Until the introduction of the sub-daily time step options in SWAT, it had always been presented as a long term yield model rather than a hydrograph prediction model (Arnold *et al.*, 1998; Peterson and Hamlett, 1998). This is, however, somewhat at odds with the requirement of getting the partitioning of the runoff processes correct when these processes are drivers for the other components predicted by the full model. The results of Figure B6.2.2 are not untypical of those obtained with the SWAT model; the predictions are good in places and not so good in others. It is difficult to know how much belief to put into the partitioning of the runoff processes since there is rarely any independent information with which to evaluate the predictions.

It is, however, understandable why the SWAT model has been so widely used given the range of facilities it provides, the ability to freely download the model, the links to GIS databases, the way it can reflect catchment characteristics and land management strategies, and the

way in which can be applied to ungauged catchments using default parameters. These are also reasons why it has been so often used for studies of the impacts of changes in land management and future climate change. It is relatively simple to make a run to represent current conditions and then modify the parameter values or inputs to represent potential future change.

It is relatively simple to do but the question then is how meaningful the resulting predictions are, given that the SWAT model, like all rainfall–runoff models, is an approximation to the complexity of the perceptual model of runoff processes in a particular catchment and when there may be many thousands of parameter values involved. This question of meaningfulness of such simulations is dependent on how far we can believe that the complexity can be represented by the effective values of parameters and their stationarity over time. There is still much to be learned about model calibration and evaluation (see Chapter 7). For the moment it is suggested that, where possible, the model predictions and particularly predictions of the impacts of change should be associated with an uncertainty analysis so that the significance of the difference between simulations can be assessed.

Box 6.3 The SCS Curve Number Model Revisited

B6.3.1 The Origins of the Curve Number Method

The USDA Soil Conservation Service Curve Number (SCS-CN) (now, strictly, the Natural Resources Conservation Service, NRCS-CN) method has its origins in the unit hydrograph approach to rainfall–runoff modelling (see Chapter 4). The unit hydrograph approach always requires a method for predicting how much of the rainfall contributes to the "storm runoff". The SCS-CN method arose out of the empirical analysis of runoff from small catchments and hillslope plots monitored by the USDA. Mockus (1949) proposed that such data could be represented by an equation of the form:

$$\frac{Q}{P - I_a} = \left[1 - (10)^{-b(P - I_a)}\right] \tag{B6.3.1}$$

or

$$\frac{Q}{P - I_a} = \left[1 - \exp -B(P - I_a)\right] \tag{B6.3.2}$$

where Q is the volume of storm runoff; P is the volume of precipitation, I_a is an initial retention of rainfall in the soil, and b and B are coefficients. Mockus suggested that the coefficient b was related to antecedent rainfall, soil and cover management index, a seasonal index, and storm duration.

Mishra and Singh (1999) show how this equation can be derived from the water balance equation under an assumption that the rate of change of retention with effective precipitation is a linear function of retention and with the constraint that $B(P - I_a) < 1$. Approximating the right hand side of Equation (B6.3.2) as a series expansion results in an equation equivalent to the standard SCS-CN formulation

$$\frac{Q}{P - I_a} = \frac{P - I_a}{S_{max} + P - I_a} \tag{B6.3.3}$$

where $S_{max} = 1/B$ is some maximum volume of retention. Mishra and Singh (1999) propose a further generalisation resulting from a more accurate series representation of Equation (B6.3.2) (and giving better fits to data from five catchments) as

$$\frac{Q}{P - I_a} = \frac{P - I_a}{S_{max} + a(P - I_a)} \tag{B6.3.4}$$

This is equivalent to assuming that the cumulative volume of retention $F(t)$ can be predicted as

$$\frac{F(t)}{S_{max}} = \frac{Q}{P - I_a} \tag{B6.3.5}$$

$F(t)$ is often interpreted as a cumulative volume of infiltration, but it is not necessary to assume that the predicted stormflow is all overland flow, since it may not have been in the original small catchment data on which the method is based (application of the method to one of the permeable, forested, Coweeta catchments (Hjelmfelt *et al.*, 1982) is a good example of this).

A further assumption is usually made in the SCS-CN method that $I_a = \lambda S_{max}$ with λ commonly assumed to be ≈ 0.2. Thus, with this assumption, the volume of storm runoff may be predicted from a general form of the SCS-CN equation:

$$Q = \frac{(P - \lambda S_{max})^2}{P + (1 - \lambda) S_{max}} \tag{B6.3.6}$$

B6.3.2 The Curve Number

The SCS-CN method can be applied by specifying a single parameter called the curve number (CN) and the popularity of the method arises from the tabulation of CN values by the USDA for a wide variety of soil types and conditions (see Table B6.3.1). In the standard SCS-CN method, the value of S_{max} (in mm) for a given soil is related to the curve number as

$$\frac{S_{max}}{2.54} = \left(\frac{1000}{CN} - 10\right) \tag{B6.3.7}$$

where the constant 2.54 mm converts from the original units of inches to millimetres. Ponce (1989) and others suggest that this is better written as

$$\frac{S_{max}}{C} = \left(\frac{100}{CN} - 1\right) \tag{B6.3.8}$$

where S_{max} now varies between 0 and C. The original equation implies a value of C of 10 inches (254 mm). It is easily seen from Equation (B6.3.8) that the range 0 to C is covered by curve numbers in the range from 50 to 100. Although values less than 50 are sometimes quoted in the literature, this implies that they are not physically meaningful for the original C of 254 mm.

Curve number range tables are provided for the estimation of the curve number for different circumstances of soil, vegetation and antecedent conditions (e.g. Table B6.3.1). For areas of complex land use, it was normally suggested in the past that a linear combination of the component curve numbers, weighted by the area to which they apply, should be used to determine an effective curve number for the area (e.g. USDA-SCS, 1985). However, because of the nonlinearity of the runoff prediction equation, this will not give the same results as applying the equation to each individual area and weighting the resulting runoff by the components' fractional areas. Grove *et al.* (1998) show how using a distributed curve number calculation increases the volume of predicted runoff, by as much as 100% in their examples, relative to using the composite method. Now that computer limitations are much less restrictive, the latter procedure should be adopted and, in fact, arises naturally when the SCS-CN method is applied to an HRU classification of a catchment based on GIS overlays.

Table B6.3.1 *Curve Numbers (after USDA-SCS, 1985): soils of Group A have high infiltration rates even if thoroughly wetted; group B soils have moderate infiltration rates and are moderately well drained; group C soils have low infiltration rates with an impeding layer for downward drainage; group D soils have low infiltration rates when wetted, including swelling clays and soils with a permanently high water table*

Land Use	Hydrologic Condition	Hydrologic Soil Group A B C D
Fallow		77 86 91 94
Row Crops: Contoured	Poor	70 79 84 88
	Good	65 75 82 86
Small grain: Contoured	Poor	63 74 82 85
	Good	61 73 81 84
Rotation Meadow: Contoured	Poor	64 75 83 85
	Good	55 69 78 83
Pasture or Range	Poor	68 79 86 89
	Fair	49 69 79 84
	Good	39 61 74 80
Meadow	Good	30 58 71 78
Woods	Good	25 55 70 77
Farmsteads	–	59 74 82 86
Roads: Dirt	–	72 82 87 89
Roads: Surfaced	–	74 84 90 92

B6.3.3 Curve Numbers, Antecedent Moisture Condition, and Contributing Areas

The curve number approach, in its simplest form, takes no explicit account of the effect of the antecedent moisture condition of the catchment in calculating the storm runoff. Yet it is known that there can be a highly nonlinear dependence of runoff on antecedent state. The SCS-CN model allows for this by suggesting that the curve number be modified according to whether the catchment is wet or dry (Table B6.3.2), noting that the curve number is inversely related to

Table B6.3.2 *Adjusting curve numbers for antecedent moisture condition (after USDA-SCS, 1985): Condition 2 represents the normal curve number condition prior to the annual maximum flood; Condition 1 represents dry conditions; Condition 3 represents soil that is nearly saturated*

Curve Number (Condition 2)	Condition 1 (Dry)	Condition 3 (Wet)
100	100	100
90	78	96
80	63	91
70	51	85
60	40	78
50	31	70
40	22	60
30	15	50

the effective S_{max} value. Michel *et al.* (2005) note that this is only a way of trying to compensate for a structural deficiency in the method and suggest a modification so that it can be used as a continuous soil moisture accounting method.

The background to the SCS-CN method is purely empirical. That is primarily its strength, but also a limitation in that it does not allow any direct interpretation in terms of process, except in so far as Mockus originally suggested that the value of S_{max} would be the volume of infiltration or available storage, whichever was the smaller. More recent studies have attempted to extend this interpretation. Steenhuis *et al.* (1995) have interpreted the method as equivalent to assuming a variable contributing area of runoff generation over which the storm rainfall volume is sufficient to exceed the available storage prior to an event (see also the work of Lyon *et al.* (2004) and Easton *et al.* (2008) for applications of this interpretation). The implied proportion of the catchment contributing to runoff for a given effective rainfall value is then directly related to the value of S_{max} (Figure B6.3.1). They show reasonable fits to observations for several Australian and American catchment areas (ranging from 16.5 to 7000 ha) with permeable soils. Adjusting the value of S_{max} for each catchment resulted in a range of values from 80 to 400 mm (see Figure B6.3.2). It is also worth noting that Ponce (1996) recorded, in an interview, that Vic Mockus later stated that "saturation overland flow was the most likely runoff mechanism to be simulated by the method".

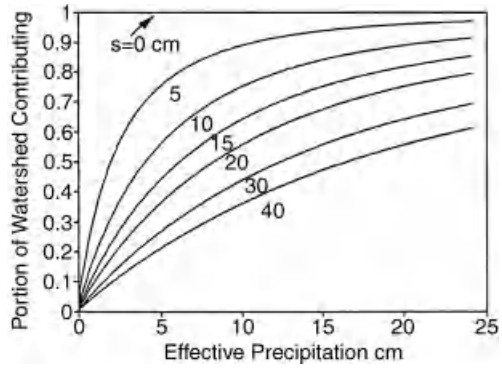


Figure B6.3.1 *Variation in effective contributing area with effective rainfall for different values of* $S_{\max}$ *(after Steenhuis* et al.*, 1995, with kind permission of the American Society of Civil Engineers); effective rainfall is here defined as the volume of rainfall after the start of runoff,* $P - I_a$.

Yu (1998) suggests that a model of the form of the SCS-CN method can be derived on the basis of assumptions about the spatial variability of (time constant) infiltration capacities and the temporal variability in rainfall intensities. Under these assumptions, runoff will be produced anywhere on a catchment where the time varying rainfall rate exceeds the spatially variable but time constant infiltration capacity (making no allowances for any runon process). For an exponential distribution of infiltration capacity in space and rainfall intensity in time, Yu shows that the runoff generated Q is given by the SCS-CN equation.

B6.3.4 Applications of the Curve Number Method

A wide variety of models have been based on the curve number method. It has been used widely in procedures recommended by the USDA, notably in the TR20 and TR55 methods for estimating peak runoff and hydrographs (USDA-SCS, 1986). A detailed summary of the method is given by McCuen (1982). It has also provided a runoff component for a succession of water

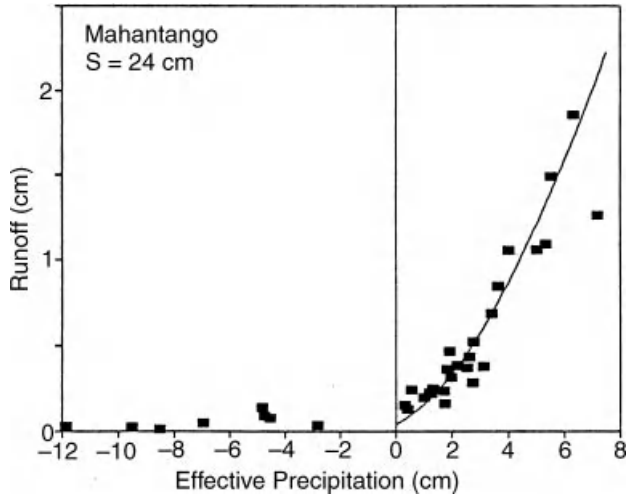


Figure B6.3.2 *Application of the SCS method to data from the Mahatango Creek catchment (55 ha), Pennsylvania (after Steenhuis* et al.*, 1995, with kind permission of the American Society of Civil Engineers); effective rainfall is here defined as the volume of rainfall after the start of runoff.*

quality and erosion models including the Areal Nonpoint Source Watershed Environment Response System (ANSWERS; Beasley *et al.*, 1980); the Chemicals, Runoff and Erosion from Agricultural Management Systems model (CREAMS) and its companion model for estimating pesticide loadings to groundwater, GLEAMS (Knisel and Williams, 1995); the Simulator for Water Resources in Rural Basins (SWRRB; Arnold and Williams, 1995); Erosion Productivity Impact Calculator (EPIC; Williams, 1995); The Pesticide Root Zone model (PRZM; Carsel *et al.*, 1985); the Agricultural Nonpoint Source model (AGNPS; Young *et al.*, 1995); the Water Erosion Prediction Project (WEPP; Nearing *et al.*, 1989) and, most recently, the distributed Soil and Water Assessment Tool (SWAT; Arnold *et al.*, 1998; Gassman *et al.*, 2007; see Box 6.2).

B6.3.5 Limitations of the Curve Number Method

It is worth remembering what is being predicted by the SCS-CN approach. It is the volume of "storm runoff" in a given storm, after some initial retention before runoff begins, that is then to be routed by the unit hydrograph or some other routing method to predict a storm runoff hydrograph. It is therefore subject to all the problems and limitations associated with separating storm runoff from the total discharge hydrograph, both in the analyses that underlay the original model formulation and in the calculation of curve numbers for particular situations.

The curve number approach to predicting runoff generation has been the subject of a number of critical reviews (e.g. Hjelmfelt *et al.*, 1982; Bales and Betson, 1982; Michel *et al.*, 2005). Further work is required to clarify under what conditions the method gives good predictions. Mishra and Singh (1999) show that their generalised version of the method, Equation (B6.3.6) gives better results than the original formulation (with $\lambda = 0.2$ and $a = 0$), as it should, since it has two additional fitting parameters. Hjelmfelt *et al.* (1982) suggest that the curve number, rather than being considered as a characteristic for a given soil-land cover association, might better be considered as a stochastic variable. Their analysis, of the annual maximum storms for two small catchments in Iowa, suggested that the storage capacity parameter, S_{max}, derived for individual storms was approximately log normally distributed with a coefficient of variation of the order of 20%. The 10% and 90% quantiles of the distributions corresponded well to the

modified curve numbers for dry and wet antecedent conditions, following the standard SCS procedures based on the preceding five-day rainfalls. However, they found no strong correlation between curve number and antecedent condition for the individual storms, suggesting that interactions with individual storm characteristics, tillage, plan growth and temperature were sufficient to mask the effect of antecedent rainfall alone.

Despite its limitations, the SCS-CN method has been widely used because the tabulated curve number values provide a relatively easy way of moving from a GIS dataset on soils and vegetation to a rainfall–runoff model (see for example Berod *et al.*, 1999; Brath *et al.*, 2004, 2006). It does have the important advantage in this respect that it is not formulated on the basis of point scale measurements, but directly from small catchment measurements.

7

Parameter Estimation and Predictive Uncertainty

Far better an approximate answer to the right question, which is often vague, than an exact answer to the wrong question, which can always be made precise.

John W. Tukey, 1962

The weight of evidence for an extraordinary claim must be proportional to its strangeness.

Pierre Simon Laplace (1749–1827)

On the basis of the quantity and quality of information that is available today, equifinality (from practical or mathematical perspectives) cannot be totally eliminated from hydrologic-response simulation for any model, simple or complex. One promising protocol for reducing parameter uncertainty for physics-based simulation is an improved dialogue between experimentalists and modelers . . . , so that future field experiments and long-term observations will better capture the non-intuitive nuances associated with the distributed response for real systems.

Brian Ebel and Keith Loague, 2006

7.1 Model Calibration or Conditioning

It should be clear from the preceding chapters that limitations of both model structures and the data available on parameter values, initial conditions and boundary conditions, will generally make it difficult to apply a hydrological model (of whatever type) without some form of calibration. In a very few cases reported in the literature, models have been applied and tested using only parameter values measured or estimated *a priori* (e.g. Beven *et al.*, 1984; Parkin *et al.*, 1996; Bathurst *et al.*, 2004); in the vast majority of cases, the parameter values are adjusted to get a better fit to some observed data. This is the model calibration problem discussed in Section 1.8. The question of how to assess whether one model or set of

Rainfall–Runoff Modelling: The Primer, Second Edition. Keith Beven.

parameter values is better than another is open to a variety of approaches, from a visual inspection of plots of observed and predicted variables to a number of different quantitative measures of goodness of fit, known variously as objective functions, performance measures, fitness (or misfit) measures, likelihood measures or possibility measures. Some examples of such measures that have been used in rainfall–runoff modelling are discussed in Section 7.3.

All model calibrations and subsequent predictions are subject to uncertainty. This uncertainty arises in that no rainfall–runoff model is a true reflection of the processes involved, it is impossible to specify the initial and boundary conditions required by the model with complete accuracy and the observational data available for model calibration are not error free. Discussions of the impact of these sources of uncertainty may be found in the work of Beven (2006a, 2010); see also the exchanges summarised by Beven (2008). There is a rapidly growing literature on model calibration and the estimation of predictive uncertainty for hydrological models and this is an area that has greatly developed in the last decade (at least in the research domain; there has been, as yet, somewhat less impact on practice). This chapter can give only a summary of the topics involved and a more extensive analysis of the methodologies available can be found (Beven, 2009). For the purposes of the discussion here, we differentiate several major themes as follows:

- Methods of model calibration that assume an optimum parameter set and ignore the estimation of predictive uncertainty. These methods range from simple trial and error, with parameter values adjusted by the user, to the variety of *automatic optimisation* methods discussed in Section 7.4.
- Methods of uncertainty estimation that are based only on prior assumptions about different sources of uncertainty. These methods are grouped under the name "forward uncertainty analysis" and discussed in Section 7.5.
- Methods of model calibration and uncertainty estimation that use Bayesian statistical methods to condition posterior parameter distributions given some observations about the catchment. These methods are grouped under the name "Bayesian conditioning" and discussed in Section 7.7.
- Methods of model conditioning that reject the idea that there is an optimum parameter set in favour of the idea of *equifinality* of models, as discussed in Section 1.8. Equifinality is the basis of the GLUE methodology discussed in Section 7.10. In this context, it is perhaps more appropriate to use model "conditioning" rather than model "calibration" since this approach attempts to take account of the many model parameter sets that give acceptable simulations. As a result, the predictions are necessarily associated with some uncertainty.

In approaching the problem of model calibration or conditioning, there are a number of very basic points to keep in mind. These may be summarised as follows:

- There is most unlikely to be one right answer. Many different models and parameter sets may give good fits to the data and it may be very difficult to decide whether one is better than another. In particular, having chosen a model structure, the optimum parameter set for one period of observations may not be the optimum set for another period. This is because so many of the sources of error in rainfall–runoff modelling are not simply statistical in nature (*aleatory uncertainties*). They are much more the result of a lack of knowledge (*epistemic uncertainties*, what Knight (1921) called the **real** uncertainties). In principle, epistemic uncertainties can be reduced by more observations and experiment. In practice, this tends to increase our appreciation of complexity without greatly improving predictions.
- Calibrated parameter values may be valid only inside the particular model structure used. It may not be appropriate to use those values on different models (even though the parameters may have the same name) or in different catchments.
- Model results will be much more sensitive to changes in the values of some parameters than to others. A basic sensitivity analysis should be carried out early on in a study (see Section 7.2).

- Different performance measures will usually give different results in terms of both the "optimum" values of parameters and the relative sensitivity of different parameters.
- Sensitivity may depend on the period of data used and, especially, on whether a particular component of the model is "exercised" in a particular period. If it is not, for example if an infiltration excess runoff production component is only used under extreme rainfalls, then the parameters associated with that component will appear insensitive.
- Model calibration has many of the features of a simple regression analysis, in that an optimum parameter set is one that, in some sense, minimises the overall error or residuals. There are still residuals, however, and this implies uncertainty in the predictions of a calibrated model. As in a statistical regression, these uncertainties will normally get larger as the model predicts the responses for more and more extreme conditions relative to the data used in calibration.
- Both model calibration (in the sense of finding an optimum parameter set) and model conditioning (finding a joint posterior parameter distributions) will depend on the shape of a response surface in the parameter space for the chosen performance or likelihood measure. The complexity of that surface depends on the interaction of the model with errors in the input data and, in some cases, on the implementation of the numerical solution to the model equations.
- Because of epistemic uncertainties, there can be no completely objective analysis of uncertainty in rainfall–runoff modelling. The analysis, therefore, depends on a set of assumptions that a modeller is prepared to accept and justify with a particular purpose in mind. Thus, the results of such an analysis should always be associated with a clear exposition of the assumptions on which it is based.

7.2 Parameter Response Surfaces and Sensitivity Analysis

Consider, for simplicity, a model with only two parameters. Some initial values are chosen for the parameters and the model is run with a calibration data set. The resulting predictions are compared with some observed variables and a measure of goodness of fit is calculated and scaled so that, if the model was a perfect fit, the goodness of fit would have a value of 1.0 and if the fit was very poor, it would have a value of zero (specific performance measures are discussed in Section 7.3). Assume that the first run resulted in a goodness of fit of 0.72, i.e. we would hope that the model could do better (get closer to a value of 1). It is a relatively simple matter to set up the model, change the values of the parameters, make another run and recalculate the goodness of fit. This is one of the options provided in the TOPMODEL software (see Appendix A). However, how do we decide which parameter values to change in order to improve the fit?

One way is by simple trial and error: plotting the results on screen, thinking about the role of each parameter in the model and changing the values to make the hydrograph peaks higher, the recessions longer or whatever is needed. This can be very instructive but, as the number of parameters gets larger, it becomes more and more difficult to sort out all the different interactions of different parameters in the model and decide what to change next (try it with the demonstration TOPMODEL software in which up to five parameters may be changed). In the very early days of rainfall–runoff modelling, there was a story that the only person who could really calibrate the Stanford Watershed Model, with all its parameters, was Norman Crawford who wrote the original version of the model as part of his PhD thesis.

7.2.1 Defining a Response Surface

Another way is to make enough model runs to evaluate the model performance in the whole of the *parameter space*. In the simple two-parameter example, we could decide on a range of values for each parameter, use 10 discrete increments on each parameter range and run the model for every combination

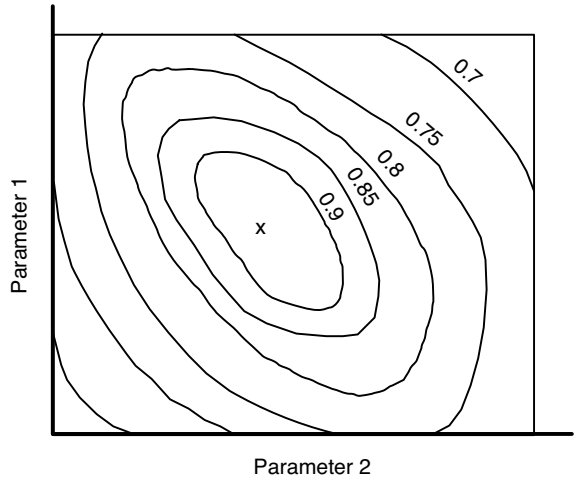


Figure 7.1 *Response surface for two parameter dimensions with goodness of fit represented as contours.*

of parameter values. The ranges of the parameters define the parameter space. Plotting the resulting values of goodness of fit defines a parameter *response surface*, such as that shown as contours in Figure 7.1 (see also the three-dimensional representation in Figure 1.7). In this example, 10 discrete increments would require $10^2 = 100$ runs of the model. For simple models this should not take too long, although complex fully distributed models might take much longer. The same strategy for three parameters is a bit more demanding: 10^3 runs would be required; for six parameters, 10^6 or a million runs would be required and 10 increments per parameter is not a very fine discretisation of the parameter space. Not all those runs, of course, would result in models giving good fits to the data. A lot of computer time could therefore be saved by avoiding model runs that give poor fits. This is a major reason why there has been so much research into automatic optimisation and density dependent sampling techniques, which aim to minimise the number of runs necessary to find an optimum parameter set or joint posterior distribution of the parameter values.

The form of the response surface may become more and more complex as the number of parameters increases and it is also more and more difficult to visualise the response surface in three or more parameter dimensions. Some of the problems encountered, however, can be illustrated with our simple two-parameter example. The form of the response surface is not always the type of simple hill shown in Figure 1.7. If it was, then finding an optimum parameter set would not be difficult; any of the "hill-climbing" automatic optimisation techniques of Section 7.4.1 should do a good job finding the way from any arbitrary starting point to the optimum. Density dependent sampling algorithms (such as Monte Carlo Markov Chain methods) would also find the shape of the response surface rather efficiently

One of the problems commonly encountered is parameter insensitivity. This occurs when a parameter has very little effect on the model result in part of the range. This may result from the component of the model associated with that parameter not being activated during a run (perhaps the parameter is the maximum capacity of a store in the model and the store is never filled). In this case, part of the parameter response space will be "flat" (Figure 7.2a) and changes in that parameter in that area have very little effect on the results. Hill-climbing techniques may find it difficult to find a way off the plateau and towards higher goodness of fit functions if they get onto such a plateau in the response surface. Different starting points may, then, lead to different final sets of parameter values.

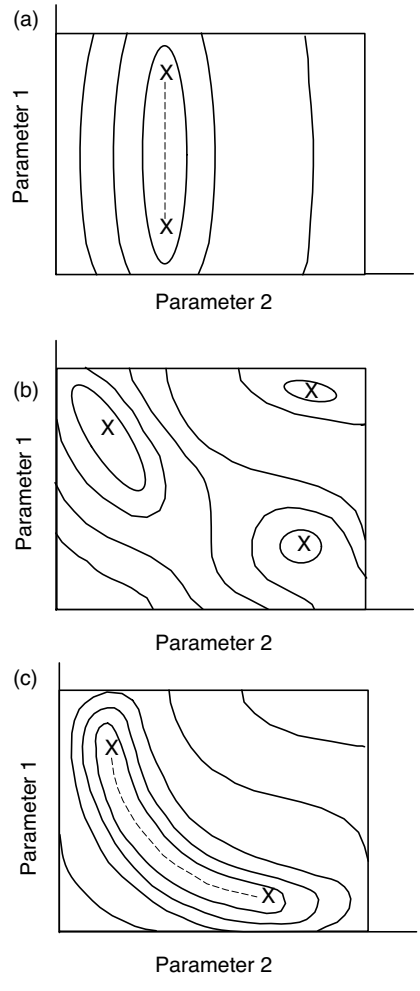


Figure 7.2 *More complex response surfaces in two parameter dimensions: (a) flat areas of the response surface reveal insensitivity of fit to variations in parameter values; (b) ridges in the response surface reveal parameter interactions; (c) multiple peaks in the response surface indicate multiple local optima.*

Another problem is parameter interactions. In simple cases this can lead to "ridges" in the response surface (Figure 7.2b), with different pairs of parameter values giving very similar goodness of fit. In such cases a hill-climbing technique may find the ridge very easily but may find it difficult to converge on a single set of values giving the best fit. Again, different starting values may give different final sets of parameter values.

There may also be a number of different peaks and ridges in the response surface (Figure 7.2c) giving rise to a number of *local optima*. One of these will be the *global optimum* but there may be a number of local optima that give similar goodness of fit. The response surface may also be very irregular or jagged (see the work of Blackie and Eeles (1985) for a good two-parameter example and also the

discussion by Sorooshian and Gupta, 1995). The irregularity of the surface might be dependent both on the choice of performance measure and other numerical issues. Again, different starting points for a hill-climbing algorithm might lead to very different final values. Most such algorithms will find the nearest local optimum, which may not be the global optimum. This is not just an example of mathematical complexity: there may be good physical reasons why this might be so. If a model has components for infiltration excess runoff production, saturation excess runoff production or subsurface stormflow (we might expect more than two parameters in that case), then there are likely to be sets of parameters that give a good fit to the hydrograph using the infiltration excess mechanism; sets that give a good fit using a saturation excess mechanism; sets that give a good fit by a subsurface stormflow mechanism; and even more sets that give a good fit by a mixture of all three processes (see Beven and Kirkby (1979) for an example using the original TOPMODEL). The different local optima may then be in very different parts of the parameter space.

The types of behaviour shown in Figure 7.2 can make finding the global optimum difficult, to say the least. Most parameter optimisation problems involve more than two parameters. To get an impression of the difficulties faced, try to imagine what a number of local optima would look like on a three-parameter response surface; then on a four-parameter response surface .... Some advances have been made in computer visualisation of higher dimensional response surfaces but trying to picture such a surface soon becomes rather taxing for bears of very little brain (or even expert hydrological modellers). The modern hill-climbing algorithms described in Section 7.4 are designed to be robust with respect to such complexities of the response surface.

There is, however, also another way of approaching the problem by designing hydrological models to avoid such calibration problems. A model could be structured, for example, to avoid the type of threshold maximum storage capacity parameter that is activated only for a small number of time steps. Early work on this type of approach in rainfall–runoff modelling was carried out by Richard Ibbitt using conceptual ESMA-type models (Ibbitt and O'Donnell, 1971, 1974) and, as noted in Section 6.2, the PDM model was originally formulated by Moore and Clarke (1981) with this in mind (see also Gupta and Sorooshian, 1983). Normally, of course, hydrological models are not designed in this way. The hydrological concepts are given priority rather than the problems of parameter calibration, particularly in physics-based models. However, for any model that is subject to calibration in this way, these considerations will be relevant. Care should also be taken with the numerical methods used in solving the model equations. It has been suggested that some complexities of the response surface might be induced only because of poor approximations in solving the model equations (Kavetski *et al.*, 2006; Kavetski and Kuczera, 2007; Kuczera *et al.*, 2010a; Kavetski and Clark, 2010; Schoups *et al.*, 2010).

There are particular problems in assessing the response surface and sensitivity of parameters in distributed and semi-distributed models, not least because of the very large number of parameter values involved and the possibilities for parameter interaction in specifying distributed fields of parameters. This will remain a difficulty for the foreseeable future and the only sensible strategy in calibrating distributed models would appear to be to insist that most, if not all, of the parameters are either fixed (perhaps within some feasible range, as in Parkin *et al.*, 1996) or calibrated with respect to some distributed observations and not catchment discharge alone (as in the work of Franks *et al.*, 1998; and Lamb *et al.*, 1998b). Claims have been made that optimisation of many thousands of parameter values in a semi-distributed model is possible (e.g. Arabi *et al.* 2007; Whittaker *et al.*, 2010) but, if this is done only on the information contained in observations on the outputs from a catchment area, it cannot be a well-posed problem. Rather obviously, an increase in runoff generation in one part of the domain can be offset by a decrease in another part (or vice versa) and it is impossible to tell which is the correct spatial representation. Thus it might be possible to obtain an **acceptable** fit to some set of observations, but this solution might be neither optimal nor unique. This has been one reason why Beven (1993, 2006a) has argued for hydrological modellers to consider the *equifinality* thesis: that many different model representations (both parameter sets and model structures) might be found that demonstrate acceptable fits to any set of

observations. Even then, however, because of the epistemic nature of the many sources of error in the modelling process, there may still be surprises in making predictions outside of the calibration data set.

7.2.2 Assessing Parameter Sensitivity

The efficiency of parameter calibration would clearly be enhanced if it was possible to concentrate the effort on those parameters to which the model simulation results are most sensitive. This requires an approach to assessing parameter sensitivity within a complex model structure. Sensitivity can be assessed with respect to both predicted variables (such as peak discharges, discharge volume, water table levels, snowmelt rates, etc.) or with respect to some performance measure (see Section 7.3). Both can be thought of in terms of their respective response surfaces in the parameter space. One definition of the sensitivity of the model simulation results to a particular parameter is the local gradient of the response surface in the direction of the chosen parameter axis. This can be used to define a normalised sensitivity index of the form:

$$S_i = \frac{dQ/dx_i}{Q/x_i} \tag{7.1}$$

where S_i is the sensitivity index with respect to parameter i with value x_i and Q is the value of the variable or performance measure at that point in the parameter space (see, for example, McCuen, 1973). The gradient will be evaluated locally, given values of the other parameters, either analytically for simple models or numerically by a finite difference, i.e. by evaluating the change in Q as x_i is changed by a small amount (say 1%). Thus, since the simulation results depend on all the parameters, the sensitivity S_i for any particular parameter i will tend to vary through the parameter space (as illustrated by the changing gradients for the simple cases in Figure 7.2). Because of this, sensitivities are normally evaluated in the immediate region of a best estimate parameter set or an identified optimum parameter set after a model calibration exercise.

This is, however, a very local estimate of sensitivity in the parameter space. A more global estimate might give a more generally useful indication of the importance of a parameter within the model structure. There are a number of global sensitivity analysis techniques available. The method proposed by van Griensven *et al.* (2006) extends the local sensitivity approach by averaging over a Latin Hypercube sample of points in the parameter space, consistent with prior estimates of the joint distribution of the parameter values, but a simple average might be misleading. A technique that makes minimal assumptions about the shapes of the response surface is variously known as Generalised Sensitivity Analysis (GSA), regionalised sensitivity anlysis (RSA) or the Hornberger–Spear–Young (HSY) method (Hornberger and Spear, 1981; Young, 1983; Beck, 1987) which was a precursor of the GLUE methodology described in Section 7.10. The HSY method is based on *Monte Carlo simulation*, which makes many different runs of a model with each run using a randomly chosen parameter set. In the HSY method, the parameter values are chosen from uniform distributions spanning specified ranges for each parameter. The ranges should reflect the feasible parameter values in a particular application. The idea is to obtain a sample of model simulations from throughout the feasible parameter space. The simulations are classified in some way into those that are considered *behavioural* and those that are considered *nonbehavioural* in respect of the system being studied. Behavioural simulations might be those with a high value of a certain variable or performance measure, nonbehavioural those with a low value.

HSY sensitivity analysis then looks for differences between the behavioural and nonbehavioural sets for each parameter. It does so by comparing the cumulative distribution of that parameter in each set. Where there is a strong difference between the two distributions for a parameter, it may be concluded that the simulations are sensitive to that parameter (Figure 7.3b). Where the two distributions are very similar, it may be concluded that the simulations are not very sensitive to that parameters (Figure 7.3c). A quantitative measure of the difference between the distributions can be calculated using the nonparametric

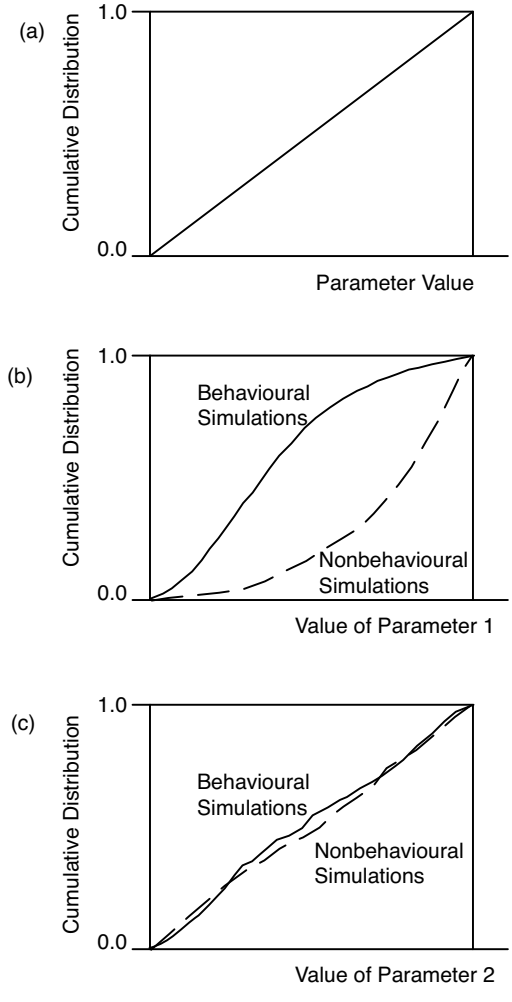


Figure 7.3 *Generalised (Hornberger–Spear–Young) sensitivity analysis – cumulative distributions of parameter values for: (a) uniform sampling of prior parameter values across a specified range; (b) behavioural and nonbehavioural simulations for a sensitive parameter; (c) behavioural and nonbehavioural simulations for an insensitive parameter.*

Kolmagorov–Smirnov d statistic, although for large numbers of simulations this test is not robust and will suggest that small differences are statistically significant. The d statistic can, however, be used as an index of relative difference. This approach may be extended, given enough Monte Carlo simulation samples, to more than two sets of parameters (the GLUE software, for example, uses 10 different classes in assessing sensitivity). Other examples of the use of the HSY approach in rainfall–runoff modelling include those by Hornberger *et al.* (1985), using TOPMODEL, and Harlin and Kung (1992), using the HBV model.

The HSY approach is essentially a *nonparametric* method of sensitivity analysis in that it makes no prior assumptions about the variation or covariation of different parameter values, but only evaluates sets of parameter values in terms of their performance. Other nonparametric methods include approaches based on fuzzy set theory, in which fuzzy measures are used to reflect the imprecision of knowledge about individual parameter values and boundary conditions (e.g. Dou *et al.*, 1995; Schulz and Huwe, 1997, 1999). There is also a variance based Sobol' Global Sensitivity Analysis, which aims to decompose the variance of a predicted variable from a sample of model runs into primary effects (associated with individual parameters), secondary effects (associated with the joint variation of two parameters) and so on (see Saltelli *et al.*, 2004, 2006). It would be useful to have such a decomposition, and in some applications it might be very instructive, particularly where within the structure of the model there is a strong hidden interaction of two or more parameters. However, it is a linear decomposition and, in nonlinear models, parameter interactions also exhibit strong nonlinearities such that the analysis is sometimes misleading (for example, where two parameters show a positive interaction in one part of the parameter space and a negative interaction in another part).

There have been some comparative studies of sensitivity analysis methods as applied to hydrological models (e.g. Pappenberger *et al.*, 2006a, 2008). These have tended to show that sensitivities of different parameters depend on the method of sensitivity analysis used, the predicted variable to which sensitivity is examined and the period of data used in the comparison (although with some consistency in identifying a small number of the most sensitive parameters).

7.3 Performance Measures and Likelihood Measures

The definition of a parameter response surface as outlined above and shown in Figures 7.1 and 7.2 requires a quantitative measure of performance, goodness of fit or likelihood. It is not too difficult to define the requirements of a rainfall–runoff model in words: we want a model to predict the hydrograph peaks correctly, to predict the timing of the hydrograph peaks correctly, and to give a good representation of the form of the recession curve to set up the initial conditions prior to the next event. We may also require that, over a long simulation period, the relative magnitudes of the different elements of the water balance should be predicted accurately. The requirements might be somewhat different for different projects, so there may not be any universal measure of performance that will serve all purposes.

Most measures of goodness of fit used in hydrograph simulation in the past have been based on the sum of squared errors or error variance. Taking the squares of the residuals results in a positive contribution of both overpredictions and underpredictions to the final sum over all the time steps. The error variance (assuming a zero mean error), σ_ε^2, is defined as

$$\sigma_\varepsilon^2 = \frac{1}{T-1}\sum_{t=1}^{T}\left(\widehat{y}_t - y_t\right)^2 \tag{7.2}$$

where $\widehat{y}_t$ is the predicted value of variable y at time step $t = 1, 2, \ldots, T$. Usually the predicted variable is discharge, Q (as shown in Figure 7.4), but it may be possible to evaluate model performance with respect to other predicted variables so we use the general variable y in the following discussion. A widely used goodness-of-fit measure based on the error variance is the modelling "efficiency" of Nash and Sutcliffe (1970), defined as

$$E = \left[1 - \frac{\sigma_\varepsilon^2}{\sigma_o^2}\right] \tag{7.3}$$

where σ_o^2 is the variance of the observations. The efficiency is like a statistical coefficient of determination. It has the value of 1 for a perfect fit when $\sigma_\varepsilon^2 =$ zero; it has the value of 0 when $\sigma_\varepsilon^2 = \sigma_o^2$, which is equivalent

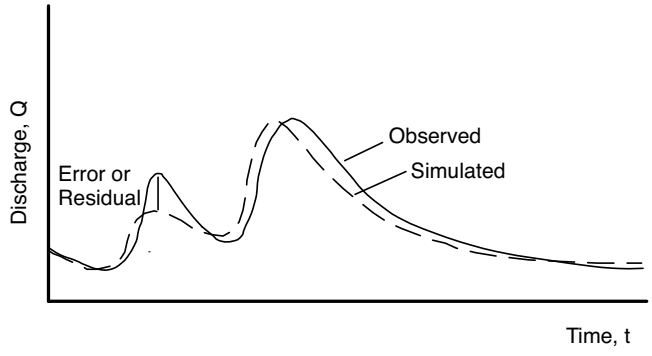


Figure 7.4 *Comparing observed and simulated hydrographs.*

to saying that the hydrological model is no better than a one parameter "no-knowledge" model that gives a prediction of the mean of the observations for all time steps! Negative values of efficiency indicate that the model is performing worse than this "no-knowledge" model (which has one parameter of the mean flow).

The sum of squared errors and modelling efficiency are not ideal measures of goodness of fit for rainfall–runoff modelling for three main reasons. The first is that the largest residuals tend to be found near the hydrograph peaks. Since the errors are squared, this can result in the predictions of peak discharge being given greater weight than the prediction of low flows (although this may clearly be a desirable characteristic for some flood forecasting purposes). Secondly, even if the peak magnitudes were to be predicted perfectly, this measure may be sensitive to timing errors in the predictions. This is illustrated for the second hydrograph in Figure 7.4 which is well predicted in shape and peak magnitude but the slight difference in time results in significant residuals on both rising and falling limbs.

Figure 7.4 also illustrates the third effect, that the residuals at successive time steps may not be independent but may be *autocorrelated* in time. The use of the simple sum of squared errors as a goodness-of-fit measure has a strong theoretical basis in statistical inference, for cases where the samples (here, the residuals at each time step) can be considered as independent and of constant variance. In many hydrograph simulations there is also a suggestion that the variance of the residuals may change in a consistent way over time, with a tendency to be higher for higher flows. This has led to significant criticism of the Nash–Sutcliffe efficiency and sum of squared errors as performance measures (see Beran, 1999; Houghton-Carr, 1999; McCuen *et al.*, 2006; Schaefli and Gupta, 2007; Smith *et al.* 2008a; Reusser *et al.*, 2009). In general, measures based on the sum of squared errors lead to biased parameter estimates when there is structure in the residual series, such as autocorrelated values in time. This has not stopped a flow of papers in rainfall–runoff modelling that still use the Nash–Sutcliffe efficiency measure as the basic measure of performance (almost certainly because of its rather natural range of values aiding interpretation).

So, it should be possible to do better and the obvious thing to do is to borrow from the theory of likelihood in statistics, which attempts to take account of structure in the residuals, such as correlation and changing variance of the errors (*heteroscedastic errors*, e.g. Sorooshian *et al.*, 1983; Hornberger *et al.*, 1985). Maximum likelihood, in frequentist statistics, aims to maximise the probabilty of predicting an observation, given the model. These probabilities are specified on the basis of a likelihood function, which is a goodness-of-fit measure that has the advantage that it can be interpreted directly in terms of such prediction probabilities. Bayesian statistics combines this likelihood function with some prior distributions of the parameters to produce posterior distributions for the parameters that can be used

in estimating the probability of predicting a new observation. However, the likelihood function that is appropriate depends on defining an appropriate structure for the modelling errors.

Underlying the development of the likelihood functions used in statistical likelihood approaches is the idea that there is a "correct" model, focussing attention on the nature of the errors associated with that model. Ideally, we would hope to find a model with zero bias and purely random errors with minimum variance and no autocorrelation. The likelihood for this and more complex cases is developed in Box 7.1. Statisticians have also suggested incorporating a *model inadequacy function* into the analysis for the case where there is a recognisable and consistent structure in the residuals (Kennedy and O'Hagan, 2001). This is usually a form of additive function that serves as a statistical compensation for epistemic error in the model or input data. The predictions might be improved in this way but it means adding a non-physical model component. The rainfall–runoff modeller might better use an inadequacy function as a guide to how a model structure might be improved. More discussion of statistical approaches to model calibration, within the Bayesian framework, is given in Section 7.7.

Remember that all these performance or likelihood measures are aimed at providing a relative measure of model performance. Statistical likelihoods also attempt to estimate the probability of a new observation in an objective way. That objectivity is undermined by the epistemic nature of some of the sources of uncertainty, so that we cannot be sure that the error characteristics in either calibration or prediction have a stationary structure. In general, therefore, we should aim to use measures that reflect the aims of a particular application (and any identified structure in the residuals) in an appropriate way. Even qualitative or "soft" information of different types that cannot easily be reduced to a statistical likelihood might be useful in model calibration (Seibert and McDonnell, 2002; Vaché and McDonnell, 2006; Winsemius *et al.*, 2009). There is no universal performance measure and, whatever choice is made, there will be an effect on the relative goodness-of-fit estimates for different models and parameter sets, particularly if an optimum parameter set is sought. Section 7.4 examines the techniques for finding optimal parameter sets, after which a more flexible approach to model calibration is discussed.

7.4 Automatic Optimisation Techniques

A full description of all the available techniques for automatic optimisation is well beyond the scope of this book, particularly since we have already noted that the concept of the optimum parameter set may not be a particularly useful one in rainfall–runoff modelling. In this section, we give just a brief outline of the algorithms available. For more specifics, descriptions of different algorithms are available from Press *et al.* (1992) and Sen and Stoffa (1995); a discussion of techniques in respect of hydrological models is given by Sorooshian and Gupta (1995).

7.4.1 Hill-Climbing Techniques

Hill-climbing techniques for parameter calibration have been an important area of research since the start of computer modelling in the 1960s. Hill climbing from any point on the response surface requires knowledge of the gradient of that surface so that the algorithm knows in which direction to climb. The available techniques may be classified into two basic types.

Gradient algorithms require the gradient of the response surface to be defined analytically for every point in the parameter space. Mathematically, this requires that an analytical expression be available for the differential of the model output with respect to each parameter value. These methods are not generally used with hydrological models since it is often impossible to define such differentials analytically for complex model structures. Much more commonly used are "direct search" algorithms that search along trial directions from the current point with the aim of finding improved objective function values. Different

algorithms vary in the search strategies used. Algorithms that have been widely used in rainfall–runoff modelling include the Rosenbrock (1960) method and the Simplex method (Nelder and Mead, 1965). The latter is explained in a hydrological context by Sorooshian and Gupta (1995).

Hill climbing is, of course, much easier on smooth response surfaces than on flat or jagged surfaces. Many hydrological models do not give smooth response surfaces but, as noted above, with three or more parameter values it may be difficult to evaluate or visualise the full shape of the surface. If a hill-climbing technique is used for parameter calibration, a **minimal** check on the performance of the algorithm in finding a global optimum is to start the algorithm from a number of very different (or randomly chosen) starting points in the parameter space and check the consistency of the final sets of parameter values found. If the final sets are close, then it may be implied that there is a single optimum. If not, then consider one of the algorithms in the sections that follow; they have all been developed to be robust with respect to complexities in the response surface.

7.4.2 Simulated Annealing

Another way of using random starting points to find a global optimum is simulated annealing. The name arises from an analogy between the model parameters included in the optimisation and particles in a cooling liquid. If the particles are initially all in a liquid state, they are randomly distributed through the space occupied by the fluid. As the liquid is cooled to a lower temperature, annealing takes place in a way that minimises the energy of the system. If the cooling is too fast, the energy minimisation occurs locally; if it is very slow then eventually a global minimum energy state results. The idea of simulated annealing is to mimic this cooling process, starting from randomly distributed sets of parameters in the parameter space to find a global optimum state with respect to the performance measure of the optimisation problem.

There are a number of variants on simulated annealing, including very fast simulated reannealing and mean field annealing (see Tarantola, 1987; Ingber, 1993; Sen and Stoffa, 1995). The essence of all of the methods is a rule for the acceptance of new parameter sets. Given a starting parameter set, a perturbation of one or more parameter values is generated and the new performance measure is calculated. If it is better than the previous one, the new model is accepted. If it is not better, it may still be accepted with a probability based on an exponential function of the difference in the performance measure value scaled by a factor that is equivalent to the temperature in the annealing analogy. As the "temperature" is gradually reduced over a number of iterations, this probability is reduced. This way of allowing parameter sets with worse performance to be accepted ensures that the algorithm does not get trapped by a local optimum, at least if the rate of "cooling" is slow enough. The choice of the cooling schedule is therefore important and varies from problem to problem. The various simulated annealing methods differ in the ways that they attempt to increase the number of accepted models relative to those rejected and therefore increase the efficiency of the search. In hydrology, an application of simulated annealing may be found in the work of Thyer *et al.* (1999).

There are similarities between simulated annealing and some of the Monte Carlo Markov Chain (MC^2) methods for parameter estimation that have seen a rapid recent development in statistics. Sen and Stoffa (1995) note that the Metropolis MC^2 algorithm is directly analogous to a simulated annealing method. This was introduced into rainfall–runoff model parameter estimation by Kuczera and Parent (1998) and has been widely used since (see Section 7.7 and Box 7.3).

7.4.3 Genetic Algorithms

Genetic algorithm (GA) methods are another way of trying to ensure that a global optimum is always found, but are based on a very different analogy, that of biological evolution. A random population

of "individuals" (different parameter sets) is chosen as a starting point and then allowed to "evolve" over successive generations or iterations in a way that improves the "fitness" (performance measure) at each iteration until a global optimum fitness is reached. The algorithms differ in the operations used to evolve the population at each iteration, which include selection, cross-over and mutation. A popular description has been given by Forrest (1993); more detailed descriptions are given by Davis (1991). Sen and Stoffa (1995) show how some elements of simulated annealing can be included in a genetic algorithm approach. GA optimisation was used by Wang (1991) in calibrating the Xinanjiang model, by Kuczera (1997) with a five-parameter conceptual rainfall–runoff model and by Franchini and Galeati (1997) with an 11-parameter rainfall-runoff model.

One form of algorithm that has been developed for use in rainfall–runoff modelling, which combines hill-climbing techniques with GA ideas, is the shuffled complex evolution (SCE) algorithm developed at the University of Arizona (UA) by Duan *et al.* (1992, 1993). In this algorithm, different Simplex searches are carried out in parallel from each random starting point. After each iteration of the multiple searches, the current parameter values are shuffled to form new Simplexes which then form new starting points for a further search iteration. This shuffling allows global information about the response surface to be shared and means that the algorithm is generally robust to the presence of multiple local optima. Kuczera (1997) concluded that the SCE algorithm was more successful in finding the global optimum in a five-parameter space than a classical crossover GA algorithm. The SCE-UA algorithm has become one of the most widely used optimisation algorithms in rainfall–runoff modelling because of its robustness in finding the global optimum on complex surfaces. The methodology has also been incorporated into more general response surface search algorithms (see Box 7.3).

7.5 Recognising Uncertainty in Models and Data: Forward Uncertainty Estimation

The techniques of Section 7.4 are designed to find an optimum parameter set as efficiently as possible. A run of the model using that optimum parameter set will give the best fit to the observations used for the calibration, **as defined by the performance measure used**. It has long been recognised that different performance measures, and different calibration datasets, generally result in different optimum parameter sets. Thus, as far as is possible, the performance measure should reflect the purpose of the modelling. The optimum parameter set alone, however, will reveal little about the possible uncertainty associated with the model predictions.

There are many causes of uncertainty in a modelling study of which the most important are as follows (we consider the additional uncertainties associated with calibration data later):

- uncertainties in initial and boundary conditions, including model inputs;
- uncertainties in the model structure;
- uncertainties in the model parameter estimates;
- uncertainties that have been overlooked (including known omissions and unknown unknowns).

All tend to induce uncertainty in the model predictions that should, as far as possible, be assessed. As noted earlier, not all of the uncertainties will be statistical or aleatory in nature. Very often, for example, we suspect that patterns of rainfall have been such that an event has been under-recorded or over-recorded by the available raingauges. Such uncertainties are neither random or systematic, but rather changing over time as is typical of uncertainties resulting from lack of knowledge. In extreme cases, data subject to such errors might not be informative in model and parameter identification (Beven and Westerberg, 2011).

In this section, we consider only cases where no observations are available for constraining uncertainty through a calibration or conditioning process. In that case, uncertainty estimation has to depend on prior assumptions about the different sources of uncertainty. This is then a "forward uncertainty estimation" exercise. In most practical applications, it involves an analysis conditional on the choice of a certain model structure. We do not try to evaluate the effects of uncertainties that have been knowingly or unknowingly overlooked. We are therefore left with the uncertainties in the initial and boundary conditions and uncertainties in the parameter estimates. To carry out a forward uncertainty estimation we need to decide on the nature of those uncertainties and how they might be represented (for example, as statistical distributions or fuzzy variables).

Once those decisions are made, the uncertainties must be propagated through the model. In the unusual case of a model that is linear in its responses and parameters, this can be done analytically (see, for example, Beven, 2009). Rainfall–runoff models (with the exception of simple linear transfer functions for runoff routing) are nonlinear, so it is more usual to propagate the uncertainty using a Monte Carlo sampling technique. This involves taking random samples from the specified distributions, taking account of any interactions if they have been specified as joint distributions, and running the model with the chosen values. If the distributions are sampled according to the probability density of the specified distributions (for example, using the Latin Hypercube technique of Iman and Conover (1982) then this can be a relatively efficient process even for a large number of uncertain variables. More approximate methods (for example, those mentioned in the first edition of this book and reviewed by Melching, 1995) were valuable when computer power was much more limited but have now almost completely disappeared from use.

The output from such a Monte Carlo study is a set of model "realisations", each of which is associated with a probability (equal if the samples have been drawn to reflect the probabilities defined by the prior distributions) and a set of output variables. Thus the distributions of the output variables can be formed and different types of uncertainty intervals extracted.

7.6 Types of Uncertainty Interval

The aim of uncertainty estimation is to assess the probability of a certain quantity, such as the peak discharge of an event, being within a certain interval but it is worth noting that different types of interval might be required. Hahn and Meeker (1991), for example, distinguish three different types of interval:

- A confidence interval contains the estimate of an unknown characteristic of the quantity of interest, for example the mean peak discharge of the event. Since we cannot estimate the peak discharge precisely from the sample of model runs available, even the estimate of the mean is uncertain. The confidence interval can be used to define the mean estimate with specified probability. Most often, 5% and 95% limits are used to define a confidence interval (i.e. a 90% probability that the value lies within the interval). Confidence limits can also be calculated for other summary quantities for the distribution of peak discharge, such as the variance or even a quantile value.
- A tolerance interval is defined so as to contain a certain proportion of the uncertain model estimates of an observation used in model calibration. For the peak discharge example, tolerance intervals could be defined for the model predictions of a particular observed peak used in model calibration.
- A prediction interval can be defined as the interval containing a certain proportion of the uncertain model estimates of peak discharge (or any other predicted variable) for a future event. In rainfall–runoff modelling, we are mostly interested in prediction intervals after calibration or conditioning of a model.

Uncertainty limits are related to the changes in the predicted variable in the parameter space or, more precisely, if a predicted variable (rather than the performance measure) is represented as a surface in the parameter space, to the *gradient* or slope of the surface with respect to changes in the different parameter values. If the slope is steep then the uncertainty in the predictions is large. If the slope is quite small, however, then the uncertainty is predicted as small since the predicted variable changes little if the parameter is considered to be uncertain. Recalling Equation (7.1), the slopes are an indication of the local sensitivity of the predictions to errors in the estimation of the parameter values.

7.7 Model Calibration Using Bayesian Statistical Methods

Forward uncertainty estimation can be a useful technique but it is essentially a form of sensitivity analysis since the outputs depend totally on the assumptions about the prior distributions that represent the sources of uncertainty. We might try to make those assumptions as realistic as possible, but very often there is little information on which to justify either the form of a distribution or the values of its characteristic parameters (the assumptions are themselves subject to epistemic uncertainties).

The whole process becomes much more interesting when there are some observations available to try and condition the uncertainty estimation. In principle, the more observations that can be made available about the system response, the more we should be able to constrain the uncertainty in the model predictions. As we will see in the following discussion, this is also generally the case in practice, but not all the observations might be as informative as we might have hoped or expected. The observations are themselves subject to error, are not always hydrologically consistent with other data and might be *incommensurate*. This is a technical term that describes the case when observed and predicted variables might have the same names and appear to be comparable, but in fact the observations are at a different time and space scale to the model predicted variable. Thus, a soil moisture measurement, for example, is usually a point sample. A model might predict soil moisture, but as a bulk measure over some grid element or hydrological response unit. The two have the same name but are not the same quantity. Similar commensurability issues arise with model parameters. We can go into the field and measure hydraulic conductivity at a point, stomatal resistance for a collection of leaves or the roughness coefficient for a channel cross-section, but the model needs values at the scale of a grid element, vegetation stand or channel reach. Again these are not commensurate quantities.

Thus, we cannot assume that every observation is equally informative in any analysis. But the observations do provide information that can be used both in model identification or calibration and in constraining uncertainty in the predictions. This has led many hydrologists to adopt formal statistical methods, where both parameter values and observations are treated as random variables and uncertainty estimation is an intrinsic part of the analysis. In most cases, the methods adopted have been within a Bayesian statistical framework. The Bayes equation (see Box 7.2) provides a formalism to combine prior distributions that allow input of prior knowledge about the problem with a likelihood based on the model predictions of the observations to form posterior distributions of parameters and model errors that can be used to estimate the probability of predicting the next observation conditional on the model. The equation derives from an article written by the Reverend Thomas Bayes (1702–1761) and found amongst his papers after his death by his friend Richard Price. The paper was read at and published by the Royal Society of London (Bayes, 1763) and formed the basis of statistical methods for some 200 years. The development of frequentist statistics was an attempt to provide a basis for statistics that did not involve the subjectivity inherent in defining the prior distributions. More recently, however, Bayesian methods have come to dominate statistical analysis again, primarily as a result of the more sophisticated and powerful computer methods that can be brought to bear in estimating the posterior distribution and resulting prediction uncertainties for nonlinear models.

A Bayesian analysis requires three basic elements:

- a definition of the prior distribution for the uncertain quantities to be considered (in some cases, "noninformative" priors might be defined where there is no prior knowledge about a quantity);
- a definition of the likelihood function that reflects how well a model can predict the available observations;
- a method for integrating the product of prior probabilities and likelihoods to calculate the posterior distribution.

The definition of the prior is much the same requirement as for a forward uncertainty estimation (see Section 7.5). The definition of the likelihood depends on making formal assumptions about the structure of the residuals (see Box 7.1). The third requirement is generally satisfied now by the use of different forms of Monte Carlo Markov Chain (MC^2) methods (see Box 7.3). It is equvalent to trying to search for the high likelihood areas of the posterior parameter space without knowing where they might be beforehand (though the definition of the prior is intended to be helpful in this respect). To do so efficiently means concentrating the effort on those areas where past samples have indicated high likelihoods (but with some additional random sampling to try to avoid missing areas of high likelihood that have not yet been sampled, especially when a large number of parameters might be involved). MC^2 methods provide tools for making this type of sampling efficient. For complex problems, this still requires tens or hundreds of thousands of runs of a model. As with the hill-climbing optimisation techniques, they work best when the dimensionality of the space is low and the surface is simple in shape.

The critical issue with the application of Bayesian methods to rainfall–runoff models is really the definition of the likelihood function. This has been the subject of significant debate in the hydrological literature (Beven and Young, 2003; Mantovan and Todini, 2006; Beven, 2006a, 2008; Hall *et al.*, 2007; Montanari, 2007; Todini and Mantovan, 2007; Beven *et al.*, 2008; Stedinger *et al.*, 2008; Beven and Westerberg, 2011). It involves questions of belief about the nature of errors (and therefore differences in philosophy). A formal definition of a likelihood function requires the specification of a statistical model of different sources of an error (though often applied only to the residuals between observed and predicted values). Different assumptions about the errors lead to different likelihood functions (see Box 7.1). Generally this also requires assuming that the statistical model has the same structure for all the rainfall–runoff models tried, and that the structure does not change through time. The objectivity of the approach is then predicated on two conditions. The first is that if these assumptions are valid, then the probability of predicting an observation conditional on the model is mathematically justified. The second is that the validity of the assumptions can be checked against the characteristics of the actual model residuals (and the error model changed, if necessary).

This check for the validity of the assumptions to ensure that an appropriate error model is being used is, therefore, an important part of the process. In practice, it should really be an iterative process since, under the assumption that there is a best possible model (even if not a "true" model because all models are approximations), it is the structure of the errors of that maximum likelihood model that must be checked, but finding the maximum likelihood model depends on defining a likelihood function for an error structure. There are many examples in the hydrological literature of modelling studies using the Bayesian framework where it is only too obvious that the likelihood function that has been used is not valid (e.g. Thiemann *et al.*, 2001; Feyen *et al.*, 2007). This is bad practice, in that the resulting posterior parameter distributions will be biased. In fact, even small departures from the validity of the assumptions can lead to biased parameter inference (Beven *et al.*, 2008a).

Experience suggests that hydrological models do not, in general, conform well to the requirements of the classical techniques of statistical inference. This is, at least in part, because so many of the uncertainties encountered are the result of lack of knowledge rather than the result of random realisation effects (i.e. epistemic rather than aleatory errors, as noted earlier in this chapter). If epistemic errors are treated as if

they can be represented as a stochastic model, then the result is generally to overestimate the information content in a set of observation data. This is because epistemic errors are expected to be non-stationary in nature (varying over time) so that the assumption of an error model of constant form and parameters does not hold. Ignoring the nonstationarity of error characteristics also leads to biased parameter inference and certainly over-conditioning in the sense of an underestimation of parameter variances.

This suggests that a more flexible and application-oriented approach to model calibration is required. There are certainly many other performance measures that could be used that are not related to specific statistical assumptions about the errors. Some examples, for the prediction of single variables, such as discharges in hydrograph simulation, are given in Box 7.1. It may also be necessary to combine goodness-of-fit measures for more than one variable, for example both discharge and one or more predictions of observed water table level. Again a number of different ways of combining information are available; some examples are given in Box 7.2. Some of the more interesting recent developments are based on a set theoretic approach to model calibration (see Section 7.9). However, something is also lost in using these other models: the theoretical advantage that it is possible to estimate the probability of a new observation conditional on a model. However, this advantage **only** holds when all the assumptions required can be justified.

7.8 Dealing with Input Uncertainty in a Bayesian Framework

There has been an important development in the last decade in trying to allow for input uncertainties within the Bayesian statistical framework. It uses a hierarchical Bayes approach in which the model inputs, parameters and observations are all treated as random (aleatory) variables. This recognises that, in general, our information about the true inputs to a catchment area during a time step or a rainfall or snowmelt event are generally poor because of the limitations of both raingauge and radar rainfall techniques. The nature of the errors might, however, vary from event to event, and an obvious way of dealing with this type of variability is to treat the error in the inputs as a multiplier, drawn from a random distribution (equivalent to an additive error on the logs of the rainfalls). The prior expectation is that the multiplier is positive, centred on a value of one and with a moderate variance.

Estimation of the multipliers for each event then becomes part of the Bayesian calibration process. Each event requires the estimation of an appropriate multiplier which now serves as a parameter. Where there are many events in the calibration data, there are often many more multiplier parameters than model parameters to be identified, but this can be done sequentially event by event since the multiplier for an event cannot have an effect on the predictions from previous events (though the value of a multiplier can have an effect on subsequent events, albeit gradually dying away). The results of this type of calibration methodology can lead to very good results in model calibration, with accurate predictions of discharge, small predictive uncertainties and very well-defined model parameter values (Thyer *et al.*, 2009; Renard *et al.*, 2010; Vrugt *et al.*, 2009). The users claim that this provides an objective way of separating sources of error in the modelling process.

I do not believe that this claim is justified. The good results are certainly, in part, a result of compensating for the deficiencies in the estimation of catchment inputs but we should be very careful not to believe that the resulting estimates of corrected inputs represent the true inputs. This is because the rainfall multipliers are correcting not only for errors in estimating the true inputs but also for deficiencies in the model structure. For example, it is very often the case that hydrological models are poor at estimating observed discharges during the first few events of the wetting-up period after a long dry period. Let us say that a model underestimates the runoff generation in the first event (this is not uncommon in this situation); this can be corrected in calibration by increasing the rainfall multiplier. That, however, also has the effect of putting more water into storage prior to the next event. The model might then overestimate the runoff generation for the second event, which is corrected by decreasing the event multiplier for the second event, and so on.

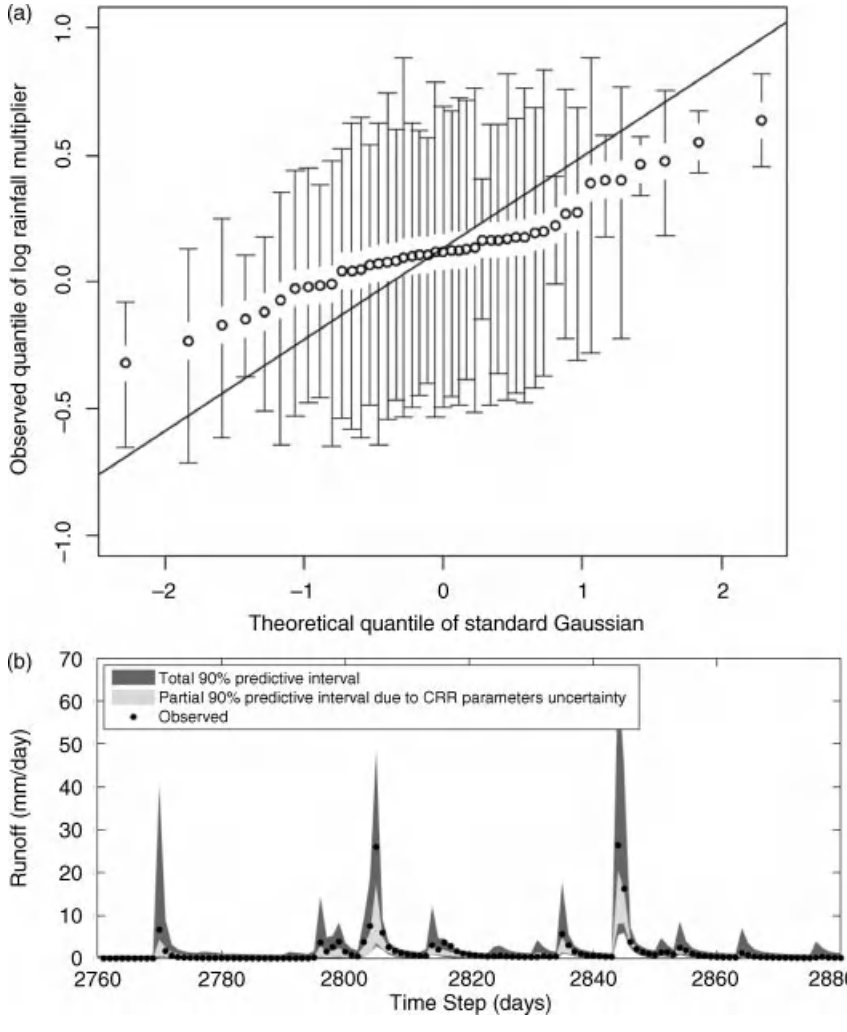


Figure 7.5 *(a) Empirical distribution of rainfall multipliers determined using BATEA in an application of the GR4 conceptual rainfall–runoff model to the Horton catchment in New South Wales, Australia; the solid line is the theoretical distribution determined from the identification process; note the log transformation of the multipliers: the range −1 to 1 represents values of 0.37 to 2.72 applied to individual rainstorms in the calibration period; the difference from the theoretical distribution is attributed to a lack of sensitivity in identifying the multipliers in the mid-range, but may also indicate that the log normal distribution might not be a good assumption in this case; (b) validation period hydrograph showing model and total uncertainty estimates (reproduced from Thyer* et al.*, 2009, with kind permission of the American Geophysical Union).*

So event multipliers identified in this way are not independent of model structural errors. This gives very much better results in calibration but gives rise to some issues in predicting the next event (and the next and . . .). Given the state of the model, and the recorded rainfall, what multiplier should be used? This is important because the identified range of multipliers in calibration has been quite large (see Figure 7.5). Using such a range over all events will lead to some rather wide prediction uncertainties,

even if the full range is truncated, as has been the case in some studies. This is currently the state of the art in Bayesian inference as applied to hydrological models. It has been applied as part of the Bayesian Total Error Analysis (BATEA) approach by Kuczera *et al.* (2006, 2010b; Thyer *et al.*, 2009; Renard *et al.*, 2010) and the Differential Evolution Adaptive Metropolis (DREAM) approach by Vrugt *et al.* (2008b, 2009). Identification of the posterior distributions of multipliers and model parameters in these cases involves the use of efficient Monte Carlo techniques (see Box 7.3). It is worth noting that the implementation of these types of method needs to be done with care, even in the case that the assumptions about the various sources of uncertainty in the hierarchical structure can be considered valid (e.g. Renard *et al.*, 2009).

7.9 Model Calibration Using Set Theoretic Methods

There is another approach to model calibration that relies much less on the specification of a statistical likelihood function and the idea of a maximum likelihood or optimal model. It was noted in Section 1.8 that detailed examination of response surfaces reveals many different combinations of parameter values that give good fits to the data, even for relatively simple models. The concept of the maxiumum likelihood parameter set may then be ill-founded in hydrological modelling, carried over from concepts of statistical inference. A basic foundation of the theory of statistical inference is that there is a correct model; the problem is to estimate the parameters of that model given some uncertainty in the data available. In hydrology, it is much more difficult to make such an assumption. There is no correct model, and the data available to evaluate different models may have large uncertainty associated with it, especially for the extreme events that are often of greatest interest.

An alternative approach to model calibration is to try to determine a set of acceptable models. Set theoretic methods of calibration are generally based on Monte Carlo simulation. A large number of runs of the model are made with different randomly chosen parameter sets. Those that meet some performance criterion or criteria are retained, those that do not are rejected. The result is a set of acceptable models, rather than a single optimum model. Using all the acceptable models for prediction results in a range of predictions for each variable of interest, allowing an estimation of prediction intervals. This type of method has not been used widely in rainfall–runoff modelling (with the exception of the GLUE variant described in Section 7.10) but there were a number of early studies in water quality modelling (see, for example Klepper *et al.*, 1991; Rose *et al.*, 1991; van Straten and Keesman, 1991).

An interesting development in set theoretic approaches has been the multi-criteria calibration strategy of Yapo *et al.* (1998) and Gupta *et al.* (1998). Their approach is based on the concept of the Pareto optimal set, a set of models with different parameter sets that all have values of the various performance criteria that are not inferior to any models outside the optimal set on any of the multiple criteria. In the terminology of the method, the models in the optimal set are "dominant" over those outside the set. Yapo *et al.* (1998) have produced an interesting method to define the Pareto optimal set, related to SCE optimisation (Section 7.4). Rather than a pure Monte Carlo experiment, they start with N randomly chosen points in the parameter space and then use a search technique to modify the parameter values and find N sets within the Pareto optimal set (Figure 7.6). They suggest that this will be a much more efficient means of defining the Pareto optimal set.

They demonstrate the use of the model and the resulting prediction limits with the Sacramento ESMA-type rainfall–runoff model, used in the US National Weather Service River Forecasting System, in an application to the Leaf River catchment, Mississippi. The model has 13 parameters to be calibrated. Two performance measures were used in the calibration, a sum of squared errors and a heteroscedastic maximum likelihood criterion. 500 parameter sets were evolved to find the Pareto optimal set, requiring 68 890 runs of the model. The results are shown in Figure 7.7, in terms of the grouping of the 500 final parameter sets on the plane of the two performance measures (from Yapo *et al.*, 1998) and the associated

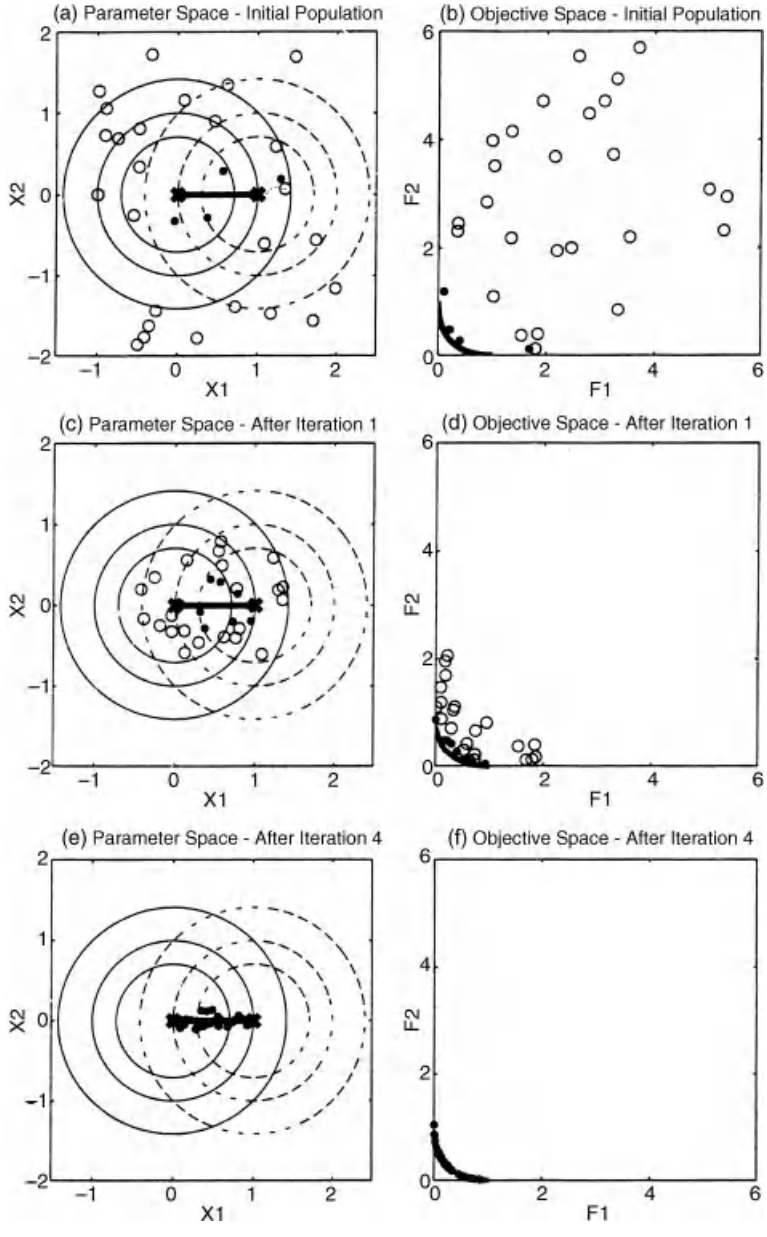


Figure 7.6 *Iterative definition of the Pareto optimal set using a population of parameter sets initially chosen randomly: (a) in a two-dimensional parameter space (parameters X1, X2); (b) in a two-dimensional performance measure space (functions F1, F2); (c) and (d) grouping of parameter sets after one iteration; (e) and (f) grouping of parameter sets after four iterations; after the final iteration, no model with parameter values outside the Pareto optimal set has higher values of the performance measures than the models in the Pareto set (after Yapo* et al.*, 1998, with kind permission of Elsevier).*

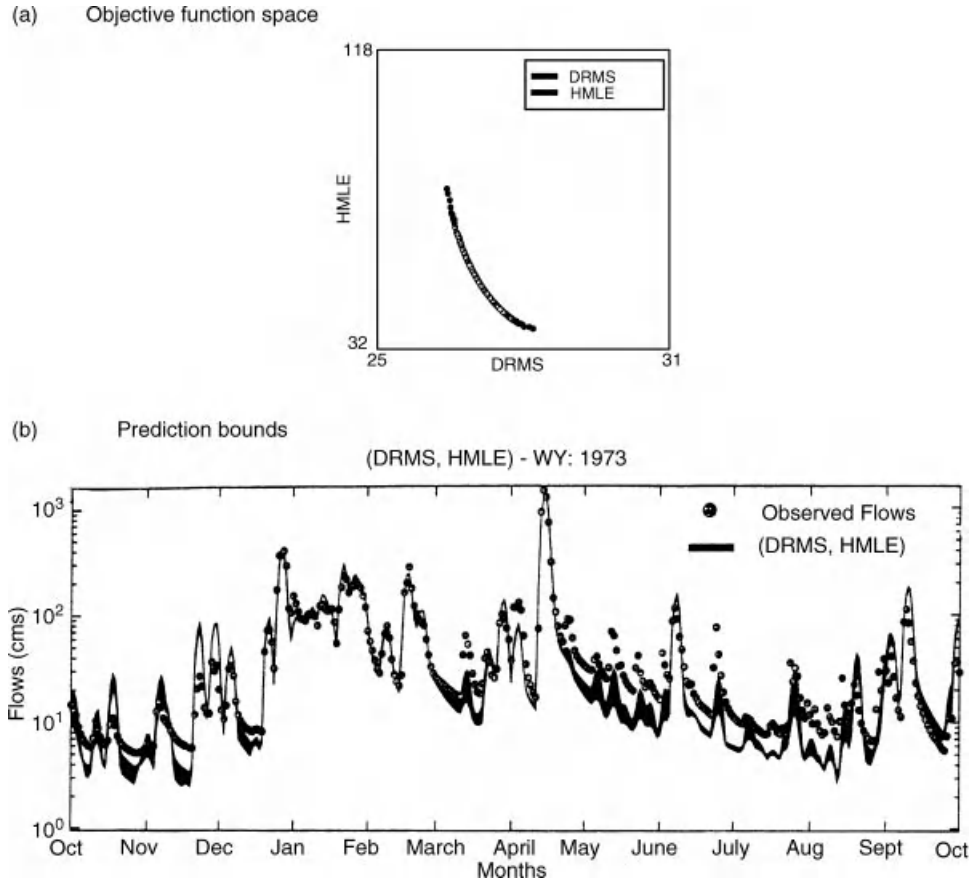


Figure 7.7 *Pareto optimal set calibration of the Sacramento ESMA rainfall–runoff model to the Leaf River catchment, Mississippi (after Yapo* et al.*, 1998, with kind permission of Elsevier): (a) grouping of Pareto optimal set of 500 model parameter sets in the plane of two of the model parameters; (b) prediction limits for the 500 Pareto optimal parameter sets.*

ranges of discharges predicted by the original randomly chosen parameter sets and the final Pareto optimal set (from Gupta *et al.*, 1998). A major advantage of the Pareto optimal set methodology is that it does not require different performance measures to be combined into one overall measure. Gupta *et al.* (1999) suggest that this method is now competitive with interactive methods carried out by a modelling expert in achieving a calibration that satisfies the competing requirements on the model in fitting the data.

As shown in Figure 7.7a the set of models that is found to be Pareto optimal reflects the sometimes conflicting requirements of satisfying more than one performance measure. Figure 7.7b, however, shows that this does not guarantee that the predictions from the sample of Pareto optimal models will bracket the observations since it cannot compensate completely for model structural error or discharge observations that are not error free. The original randomly chosen sets do bracket the observations, but with limits that are considerably wider (note the log discharge scale in Figure 7.7b). It must be remembered that the method is not intended to estimate prediction limits in any statistical sense, but one feature of this approach is that it does seem to result in an over-constrained set of predictions in comparison with the

observations. The user would have difficulty in relating the range of predictions with any degree of confidence that they would match any particular observation.

7.10 Recognising Equifinality: The GLUE Method

If we accept that there is no single correct or optimal model, then another approach to estimating prediction limits is to estimate the degree of belief we can associate with different models and parameter sets: this is the basic idea that follows from recognising the equifinality of models and parameter sets (Beven, 1993, 2006a, 2009). Certainly we will be able to give different degrees of belief to different models or parameter sets, and many we may be able to reject because they clearly do not give the right sort of response for an application. The "optimum", given some data for calibration, will have the highest degree of belief associated with it but, as we discuss in this section, there may be many other models that are almost as good. This can be seen in the dotty plots of Figure 7.8a which represent an application of TOPMODEL to the Maimai catchment in New Zealand.

The dotty plots are scatter diagrams of parameter value against some single performance measure value. Each dot represents one run of the model from a Monte Carlo experiment using many simulations with different randomly chosen parameter values. They essentially represent a projection of a sample of points from the goodness-of-fit response surface onto individual parameter dimensions. In Figure 7.8a, the good models are those that plot near the top. For each parameter, there are good simulations across a wide range of values. We commonly find with this type of Monte Carlo experiment that good simulations go all the way up to the edge of the range of parameters sampled. There are generally also poor simulations across the whole range of each parameter sampled. In another early study of this type, Duan *et al.* (1992) show similar behaviour for a completely different model. Whether a model gives good or poor results is not a function therefore of individual parameters but of the whole set of parameter values and the interactions between parameters. As a projection of the response surface, the dotty plots cannot show the full structure of the complex parameter interactions that shape the surface. In one sense, however, that does not matter too much since we are really primarily interested in where the good parameter sets are **as a set**.

All of these good parameter sets give different predictions, but if we associate a measure of belief with each set of predictions (highest for the optimum, zero for those models that have been rejected) then we can estimate the resulting uncertainty in the predictions in a conceptually very simple way by weighting the predictions of all the acceptable models by their associated degree of belief. Such an approach allows the nonlinearity of the response of acceptable models using different parameter sets to be taken into account in prediction and uncertainty estimation.

This appears to lead quite naturally to a form of Bayesian analysis, but in the form proposed in the original paper of Bayes (1763) more than in more recent usage. Bayes himself talked about the way in which evidence (of some sort) might modify the odds on a particular hypothesis, given some prior estimates of the odds (see, for example, Howson and Urbach, 1993). Thus, if any model and its associated parameter set are considered to be a hypothesis about how the catchment system functions, then any measure of belief in representing the evidence of how well the model fits the available observations fits rather naturally into this form of Bayes (although the way in which different measures of belief are combined need not be restricted to Bayes multiplication). This is the essence of the generalised likelihood uncertainty estimation (GLUE) methodology proposed by Beven and Binley (1992), which has now been used in a variety of hydrological modelling contexts with a variety of likelihood measures. Updating of the model likelihood distributions as new calibration data become available is handled easily within this Bayesian framework.

In the GLUE methodology, a prior distribution of parameter values is used to generate random parameter sets for use in each model using Monte Carlo simulation. An input sequence is used to drive

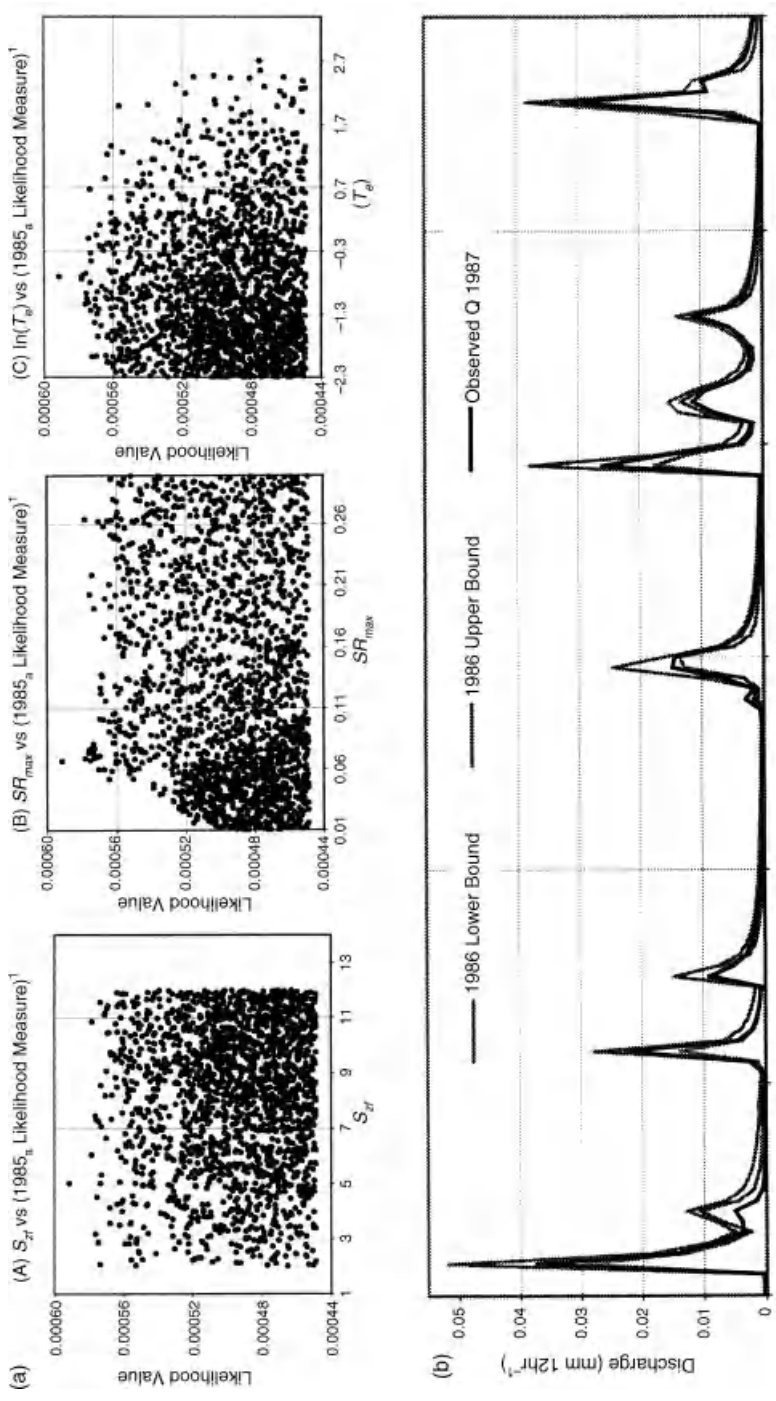


Figure 7.8 *An application of TOPMODEL to the Maimai M8 catchment (3.8 ha), New Zealand, using the GLUE methodology: (a) dotty plots of the Nash–Sutcliffe model efficiency measure (each dot represents one run of the model with parameter values chosen randomly by uniform sampling across the ranges of each parameter); (b) discharge prediction limits for a period in 1987, after conditioning using observations from 1985 and 1986.*

each model and the results are compared with the available calibration data. A quantitative measure of performance is used to assess the acceptability of each model based on the modelling residuals. Any of the likelihood measures in Box 7.1 or combinations of the measures in Box 7.2 could serve this purpose. The only requirements are that the measure should increase monotonically with increasing goodness of fit and that "nonbehavioural" models should have a likelihood of zero. Different likelihood measures or combinations of likelihood measures will, however, lead to different estimates of the predictive uncertainty.

In using the model for predictions, all simulations with a likelihood measure greater than zero are allowed to contribute to the distribution of predictions. The predictions of each simulation are weighted by the likelihood measure associated with that simulation. The cumulative likelihood weighted distribution of predictions can then be used to estimate quantiles for the predictions at any time step.

Implementation of the GLUE methodology requires a number of decisions to be made as follows:

- which model or models to include in the analysis;
- a feasible range for each parameter value;
- a sampling strategy for the parameter sets;
- an appropriate likelihood measure or measures, including conditions for rejecting models that would not be considered useful in prediction on the basis of their past performance, so leaving those that are considered *behavioural*.

These decisions are all, to some extent, subjective but an important point is that they must be made explicit in any application. Then the analysis can be reproduced, if necessary, and the decisions can be discussed and evaluated by others. Some sources of GLUE software are listed in Appendix A.

7.10.1 Deciding on Which Models to Include

Given a large enough sample of Monte Carlo simulations, the range of likelihood weighted predictions may be evaluated to obtain prediction quantiles at any time step. This is most easily done if the likelihood values are renormalised such that $\Sigma\, L\,[M\,(\Theta_i)] = 1$, where $M(\Theta_i)$ now indicates the ith Monte Carlo sample, so that at any time step t:

$$P\left(\hat{Q}_t < q\right) = \sum_{i=1}^{N} L\left[M\,(\Theta_i) \,|\, \left(\hat{Q}_{i,t} < q\right)\right] \tag{7.4}$$

where $\hat{Q}_{i,t}$ is the variable of interest predicted by the ith Monte Carlo sample and N is the number of samples. The prediction quantiles, $P\left(\hat{Q}_t < q\right)$, obtained in this way (as shown, for example, in Figure 7.8b) are conditioned on the inputs to the model, the model responses for the particular sample of parameter sets used, the subjective choice of likelihood measure and the observations used in the calculation of the likelihood measure. They are, therefore, empirical but note that, in such a procedure, the simulations contributing to a particular quantile interval may vary from time step to time step, reflecting the nonlinearities and varying time delays in model responses. It also allows for the fact that the distributional characteristics of the likelihood weighted model predictions may vary from time step to time step (see Freer *et al.* (1996) and the case study in Section 7.11).

Any parameter interactions in the model, and any effects of errors in the input data and observational data, are implicitly reflected in the likelihood measure associated with each simulation and do not therefore have to be considered separately. This makes an assumption that the effects will be similar during a prediction, but avoids the problem that they are very difficult indeed to consider separately.

7.10.2 Deciding on Feasible Parameter Ranges

Although deciding on feasible parameter ranges is analogous to defining the prior distributions for the type of forward uncertainty estimation of Section 7.5 or the Bayes priors of Section 7.7, it is not necessarily easy, even given some experience of previous applications of a model. The aim is to have a parameter space wide enough that good fits of the model are not excluded, but not so wide that the parameter values have no sense or meaning. It is often found, however, that even if the ranges are drawn quite wide, good fits are found right up to the boundary of some parameters (as in Figure 7.8a). This may be because the model predictions are not very sensitive to those parameters, or it may be that the range has not been drawn wide enough since it implies that there will be good fits beyond the edge of the range. The best suggestion is to start with quite wide ranges and see if they can be narrowed down after an initial sampling of the parameter space.

There may, of course, be some prior information about parameters. This information may take a number of forms. The first would be some sense of expected distribution and covariance of the parameter values. Some parameter sets, within the specified ranges, may be known *a priori* as not being feasible on the basis of past performance or mechanistic arguments. Then each parameter set could still be formed by uniformly sampling the parameter space but could be given a prior likelihood (perhaps of zero). If the prior likelihood is zero, it will not be necessary to run the model – such a model is considered as infeasible.

An interesting question arises when there are measured values available for one, some or all parameter values in the model. In some (rare) cases, it may even be possible to specify distributions and covariances for the parameter values on the basis of measurements. These could then be used to specify prior likelihood weights in the (uniformly) sampled parameter space. Although it is often the case that such measurements are the best information that we have about parameter values, there is, however, no guarantee that the values measured at one scale will reflect the effective values required in the model to achieve satisfactory functional prediction of observed variables. As with observed and predicted variables, the measured and effective values of a parameter may be incommensurate. It might then be possible to feed disinformation into the prior parameter distributions but, if the parameter space is sampled widely enough to include suitable effective parameter values, the repeated application of Bayes equation or some other way of combining likelihood measures (see Box 7.2) should result in the performance of the model increasingly dominating the shape of the response surface relative to the initial prior estimates of parameter distributions.

7.10.3 Deciding on a Sampling Strategy

The choice of a sampling strategy may be very important: if a large number of parameters are included in the analysis, a very large number of model runs is required to define the form of the response surface adequately in a high-dimension parameter space. The idea of using randomly chosen parameter sets is to at least get a large sample from this space. In most of the applications of GLUE to date, a uniform independent sampling of parameters in the parameter space has been used. This ensures the prior independence of the parameter sets before their evaluation using the chosen likelihood measure and is very easy to implement but can be a relatively inefficient strategy if large areas of the parameter space result in nonbehavioural simulations.

The computational expense of making many thousands of simulations so that an adequate definition of the response surface is obtained is the major reason why Monte Carlo methods have not been more widely used in hydrological modelling. The greater the number of parameters and the greater the complexity of the response surface, then the greater the number of simulations that are required. This constraint is becoming less limiting, at least for relatively simple models, as computer power continues to increase and prices continue to fall. The recent development of low-cost, Ethernet-linked, parallel PC systems using off-the-shelf boxes will mean that Monte Carlo simulations will become increasingly feasible in both

research and application projects. Such parallel systems are ideally suited to this type of calculation. If a single run of the model fits in the memory of a single processor, then there is very little loss in efficiency in the parallel system: essentially each processor runs at full calculation capacity during a run, except for some short periods in which the results from a run are passed to the master processor or written to disk and a new run is initiated. A recent alternative is the use of off-the-shelf graphics processing units (GPUs) for scientific calculations using, for example, the CUDA extension of the C programming language. GPUs consist of hundreds of parallel processors on a single board and can be used to greatly speed up run times for certain types of modelling problem (e.g. Lamb *et al.*, 2009).

However, with large models, there are still some advantages in trying to make the Monte Carlo sampling more efficient. In the GLUE methodology, for example, there is little advantage in sampling regions of the parameter space with low likelihood measure values once those regions have been established. It would be better to concentrate the sampling in the regions of high likelihoods. This is a subject that has been studied in a variety of different fields, resulting in extensive literature on what is often called *importance sampling*. A number of methods have been developed to try to exploit the knowledge gained of the response surface in refining an adaptive sampling strategy. These include the MC^2 and DREAM techniques described in Box 7.3, which attempt to sample the response surface according to likelihood density, so that regions of high likelihood are sampled more frequently. The hope is that considerable savings in computer time will be made in defining the likelihood surface. Such methods may work well when there is a well-defined surface but for surfaces with lots of local maxima or plateaux, the advantages may not be so great. The efficiency of uniform sampling techniques can also be improved by running the model only in areas of the parameter space where behavioural models are expected on the basis of previous sampling. The nearest neighbour method of Beven and Binley (1992), the tree-structured search of Spear *et al.* (1994), and the Guided Monte Carlo method of Shorter and Rabitz (1997) can all be used in this way.

7.10.4 Deciding on a Likelihood Measure

There are many measures that can be used to evaluate the results of a model simulation. They, in part, depend on what observational data are available to evaluate each model, but even if only one type of data is available (such as observed discharges in evaluating a rainfall–runoff model) there are different ways of calculating a model error and using those model errors to calculate a likelihood measure. What is certain is that if we wish to rank a sample of models by performance, different likelihood measures will give different rankings and the same measure calculated for different periods of observations will also give different rankings.

The choice of a likelihood measure should clearly be determined by the nature of the prediction problem. If the interest is in low flows, then a likelihood measure that gives more weight to the accurate prediction of low flows should be used. If the interest is in water yields for reservoir design, then a likelihood measure based on errors in the prediction of discharge volumes would perhaps be more appropriate. If we are interested in predicting flood peaks, then a likelihood measure that emphasises accurate prediction of measured peak flows should be chosen. In flood forecasting, a likelihood measure that takes account of accuracy in predicting the timing of flood peaks might be chosen. If an evaluation of a distributed model is being made then a likelihood measure that combines both performance on discharge prediction and performance on prediction of an internal state variable, such as water table level, might be appropriate. A summary of various likelihood measures is given in Box 7.1.

Most recently, applications of GLUE have been using the "limits of acceptability" approach suggested by Beven (2006a). This is a way of trying to allow for epistemic errors in the modelling process by estimating how accurately we might expect a model to make predictions given the multiple sources of error in the modelling process. The limits should, ideally, be set prior to making any runs of the model

so that they are considered independently of any model residuals (we say "ideally" because it is difficult to make an assessment of the potential impact of input errors independently of the model). It is therefore a rejectionist approach to model evaluation such that, for any observation in space and time, $O(x, t)$, a model with parameter set θ, is only retained as behavioural if:

$$O_{min}(x, t) \leq M(\theta, x, t) \leq O_{max}(x, t) \tag{7.5}$$

where $M(\theta, x, t)$ is the model prediction. Within that range, the performance of the model can also be scaled so that observations of quite different characteristics can be assessed in a similar manner. To retain information about whether the model is over- or under-predicting, a normalised score in the range -1 to $+1$ can be used, with 0 at the value of the observation. The limits do not need to be symmetric on either side of the observation since, given the limits defined for that observation, the prediction can always be scaled back from the score.

The scores can also be used to construct a performance or likelihood measure within the acceptable range (see Box 7.1). The simplest index of performance would be the triangular function from zero at the upper and lower limits to one at the observed value. As with analogous fuzzy measures, however, other functions might be suitable, such as a trapezoidal measure allowing for an equally likely range around the observation (Beven, 2006a). Different types of observation and soft information about performance is also easily incorporated into this approach (see, for example, the work of Blazkova and Beven, 2009). An example application is developed in Section 7.12. Other earlier GLUE studies have effectively used limits of acceptability for individual observations of this type in model evaluation (e.g. Iorgulescu *et al.*, 2005, 2007; Pappenberger *et al.*, 2006b, 2007a; Page *et al.*, 2007).

A rejectionist approach of this type will reveal periods where models do not provide adequate predictions. This might be because of limitations of the model structure or driving the model with inadequate input data. We do not want to make the error of rejecting a good model because of poor data. Thus the limits of acceptability approach also focuses attention on the quality of calibration data, some of which might not be informative in deciding on which models are good hypotheses about catchment response.

7.10.5 Updating Likelihood Measures

If more than one period of data is available for evaluating the model, or if new data become available, then the likelihood measures from each period can be combined in a number of different ways, as shown in Box 7.2. This can be viewed as an updating procedure. At each stage, including after the first period, there is a prior likelihood associated with each parameter set that is combined with the value of the likelihood measure for the period being used for evaluation to calculate a posterior value. Bayes equation is one way of doing such calculations that is well known in statistical theory but it is not the only one (Box 7.2). The posterior from one period then becomes the prior for the next application. The likelihood measures for a given parameter set for the periods may be correlated; indeed, one should hope it is the case that if a model performs well in one calibration period, it will continue to perform well in other periods. If this is not the case then its combined likelihood measure will be reduced.

It is possible that, in combining two measures from different observed variables during the same calibration period, there will be a correlation in model performance against different variables, i.e. a model that produces good simulations of an output variable might equally produce good simulations of an internal state variable (although it has to be said that this does not necessarily follow in many environmental models). If a model produces good simulations on both variables, its relative likelihood is raised; if it does not, its relative likelihood is lowered.

The choice of method of combining likelihood measures may have implications for the choice of the measure itself, in particular if it is required that multiple combinations (for example, of measures from different periods of data) have the same result as treating the data as a single continuous period (where this

is possible). Repeated application of Bayes equation would not lead to this end if the likelihood measure was a linear function of the inverse error variance. The successive multiplications would result in the most recent period of data having the greatest weight in the determination of the posterior likelihoods (which may, of course, give the desired effect if the system is thought to be changing over time). However, the use of a likelihood measure that is a linear function of the inverse exponential of the error variance would result in an equivalence of final posterior likelihood. This type of choice has been the subject of some discussion in the literature (Mantovan and Todini, 2006; Beven *et al.*, 2008; Smith *et al.*, 2008a). It is intrinsically related to beliefs about the information content of a series of model residuals (see Box 7.2).

7.11 Case Study: An Application of the GLUE Methodology in Modelling the Saeternbekken MINIFELT Catchment, Norway

The approach of the GLUE methodology is probably best understood by means of an example. In this example, we consider only one model, TOPMODEL, in an extension of the application to the small Saeternbekken MINIFELT catchment in Norway (used in the case study of Section 6.4). Although this application is now a little old, it is still very interesting since it involves the use of spatial information about catchment responses as well as discharge observations. There are still relatively few studies of this type. The version of TOPMODEL used was based on the original exponential transmissivity profile assumptions. It is worth pointing out that the choice of just a single model is equivalent to assigning a positive prior likelihood of one to that model (my model!) and zero to all other models. This is, of course, quite common practice, but there is no reason why more than one model structure should not be included within the GLUE framework (apart from the computational expense of making even more Monte Carlo simulations).

The use of both discharge and borehole measurements in conditioning the uncertainty in the predictions of TOPMODEL for the Saeternbekken catchment has been considered by Lamb *et al.* (1998b). They first studied the use of global (catchment scale) parameter values. Five parameters were varied in the Monte Carlo simulations. The ranges chosen for each parameter reflect past modelling experience using TOPMODEL and an initial analysis of recession curves in setting the range for the m parameter. An example of the responses to the individual parameters has already been shown in the dotty plots of Figure 7.7, which use the Nash–Sutcliffe efficiency as a likelihood measure.

Lamb *et al.* (1998b) chose to use a different likelihood measure, though also one based on the variance of the residuals, as:

$$L = \exp\left(-W\sigma_\varepsilon^2\right) \tag{7.6}$$

where W is a weighting coefficient. This gives values close to zero for larger error variance but, as with the Nash–Sutcliffe efficiency, has a limit of one if the error variance is very small. In the case of the Saeternbekken catchment, a number of different measures of performance could be calculated using discharge and different borehole observations. This form of performance measure can be combined easily using Bayes equation, since multiplying the likelihood measures is equivalent to taking a weighted average of the different residual variances within the exponential (see Equation (B7.2.3) in Box 7.2). In this particular application, this allowed different weights to be used in the 1987 calibration period, when one of the recording boreholes did not appear to produce hydrologically meaningful responses.

An impression of the sensitivity of the individual TOPMODEL parameters can be gained by plotting the cumulative likelihood weighted distributions for each parameter after evaluating each parameter set on data from the 1987 simulation period (see Figure 7.9). For each parameter, a rather wide initial range of values was used. Because of its operation within the TOPMODEL structure, the transmissivity at saturation was sampled on a log scale as $ln(T_o)$. Lacking any prior information about the covariation of

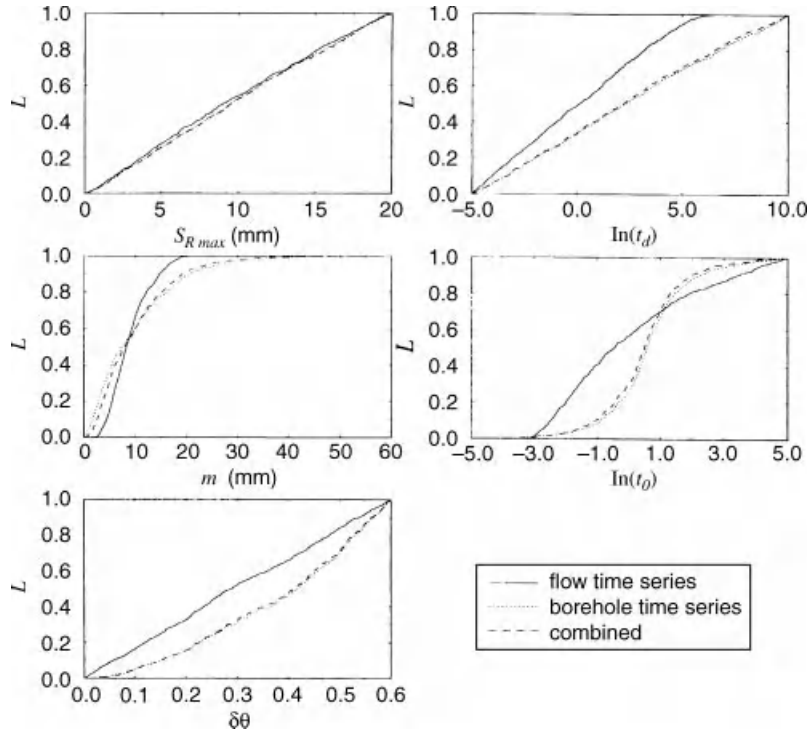


Figure 7.9 *Rescaled likelihood weighted distributions for TOPMODEL parameters, conditioned on discharge and borehole observations from the Saeternbekken MINIFELT catchment for the autumn 1987 period (after Lamb et al., 1998b, with kind permission of Elsevier).*

the individual parameters, each was sampled independently from uniform distributions across the range shown. The uniform prior distributions would be represented by a straight diagonal line on each plot. The strongest departures from the prior distributions are shown by the two parameters of the transmissivity profile, m and $ln(T_o)$. The other three parameters are much less well-conditioned by the observations. This is to be expected since the shallow wet soil in the catchment would not be expected to strongly limit evapotranspiration (controlled by the S_{Rmax} parameter), nor the time delay for vertical recharge (controlled by the t_d parameter). The relatively poor constraint on an effective storage capacity of the soil ($\delta\theta$) was not surprising in the conditioning on discharges alone but was more surprising when the borehole data were added as this parameter controls the calculation of a water table level from the storage deficits calculated by the model. It may be a reflection of the interaction in the model between this and other parameters in predicting the borehole levels. It must be remembered that, in the GLUE methodology, it is the individual parameter **sets** that are assessed in terms of the chosen likelihood measure. Figure 7.9 represents only a summary or marginal distribution for each parameter over all the parameter sets being considered. Prediction bounds for the 1987 period are not shown here but they showed that, in predicting the catchment discharges, very little additional constraint was provided by adding the information from the recording borehole information.

The resulting prediction bounds are shown in Figure 7.10 for discharges and Figure 7.11 for the recording boreholes. In both cases, bounds are shown for the 1989 simulation period using parameter

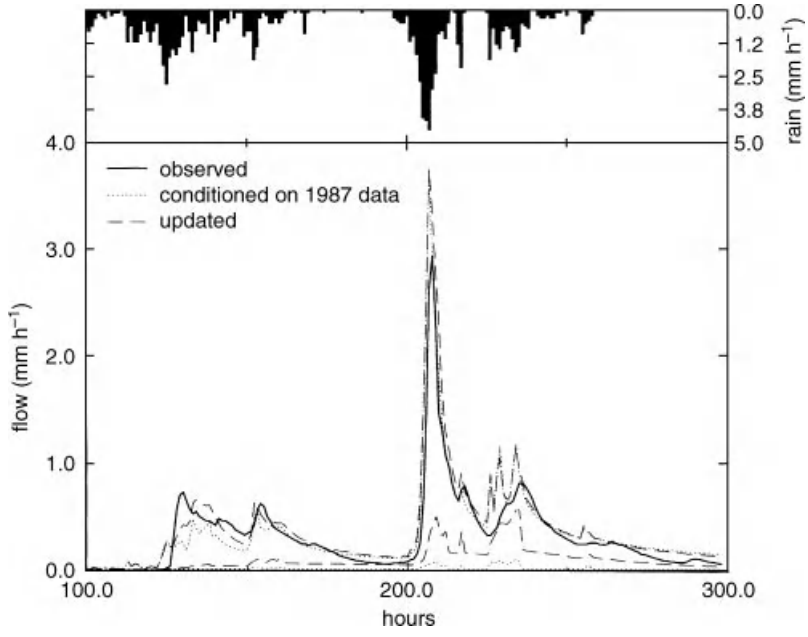


Figure 7.10 *Prediction bounds for stream discharge from the Saeternbekken catchment for the 1989 simulations, showing prior bounds after conditioning on the 1987 simulation period and posterior bounds after additional updating with the 1989 period data (after Lamb et al., 1998b, with kind permission of Elsevier).*

sets conditioned on the earlier 1987 period, and also after combining with likelihoods calculated for the 1989 period itself. These images, therefore, represent prior and posterior likelihoods for this period. In both cases, it can be seen that adding the information from the 1989 period leads to a narrowing of the prediction bounds. For most of both sets of simulations, the observations are bracketed by the prediction bounds but there are periods when the observations fall outside of the prediction bounds, indicating some deficiencies in the model structure or input data used. The predicted borehole responses do not show the same dynamic variation in the prediction bounds as the observations (although, as already seen in Figure 6.8, individual simulations can reproduce the dynamics much more closely).

Lamb *et al.* (1998b) also examined the prediction of the spatially distributed piezometer data which was available for five different discharges. All of the piezometers not indicating saturated conditions were used in the conditioning process. A weight was given to the residual variance for each set of measurements proportional to the number of points included. The results are shown in Figure 7.12. No account was taken of the possibility of local variations in transmissivity (see discussion in Section 6.4). Interestingly, the prediction bounds based on discharges and recording borehole measurements are much narrower than those based on the piezometer data themselves. Even the latter, however, are not wide enough to encompass a significant number of the data points. This indicates that either the model dynamics cannot adequately reproduce the pattern of water tables in the catchment or local soil heterogeneity cannot adequately be represented by catchment scale parameters. Certainly Lamb *et al.* (1997) have shown that improved predictions can be achieved by allowing a local effective transmissivity value for each piezometer site.

This study is of interest because of the way it used both discharge and internal water table measurements to evaluate the model predictions and constrain the predictive uncertainty of TOPMODEL (see also

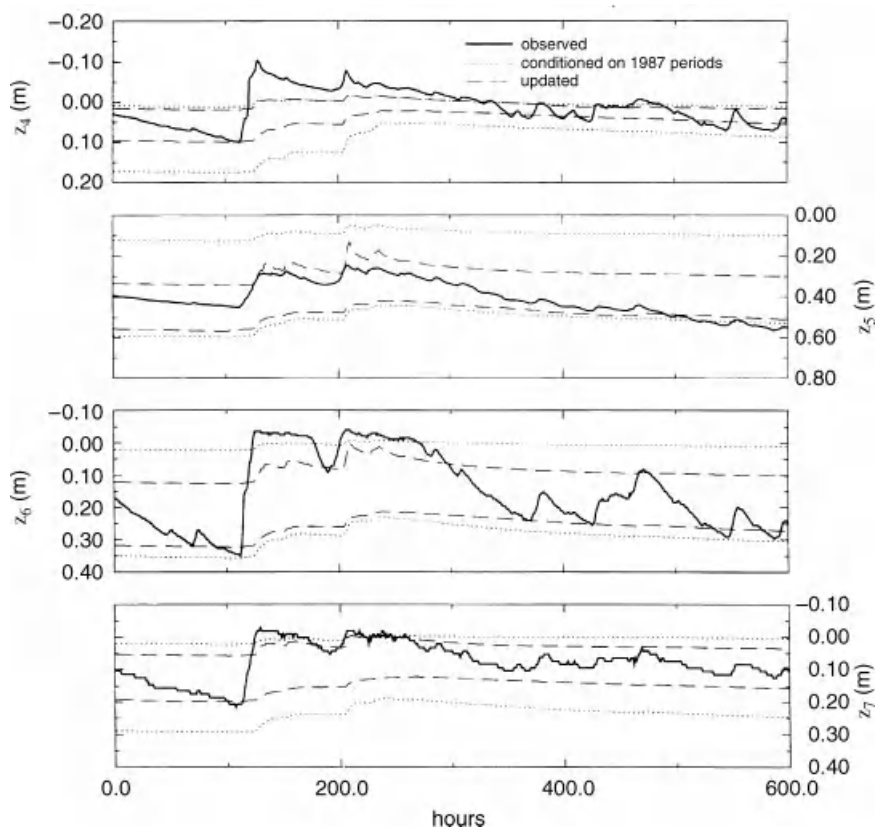


Figure 7.11 *Prediction bounds for four recording boreholes in the Saeternbekken catchment for the 1989 simulation period, showing prior bounds after conditioning on the 1987 simulation period and posterior bounds after additional updating with the 1989 period data (after Lamb* et al.*, 1998b, with kind permission of Elsevier).*

Blazkova *et al.*, 2002a, 2002b; Gallart *et al.*, 2007). It is, perhaps, representative in showing that we might not expect a rainfall–runoff model to reproduce all of the observations all of the time, even when the predictions are associated with some uncertainty bounds. It is also perhaps representative in revealing that there may be difficulties associated with using information from internal state measurements in model calibration or conditioning. Firstly, the use of such data may require adding additional local parameter values; secondly, such local data may not have great value in conditioning the prediction bounds for catchment discharges.

7.12 Case Study: Application of GLUE Limits of Acceptability Approach to Evaluation in Modelling the Brue Catchment, Somerset, England

This recent application of the GLUE methodology makes use of the limits of acceptability approach to model evaluation in an application of Dynamic TOPMODEL to the Brue catchment (135 km^2) in

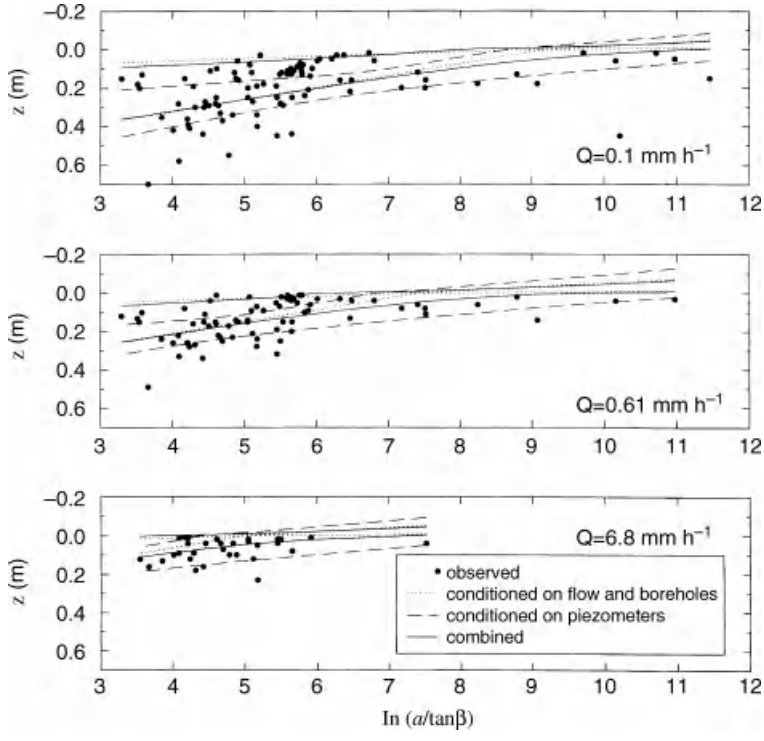


Figure 7.12 *Prediction bounds for spatially distributed piezometers in the Saeternbekken catchment for three different discharges, showing prior bounds after conditioning on discharge and recording borehole observations, bounds based on conditioning on the piezometer data alone and posterior bounds based on a combination of both individual measures (after Lamb et al., 1998b, with kind permission of Elsevier).*

Somerset, England. This is an interesting catchment in a number of ways, but was chosen because it was used in the HYREX experiment (Moore *et al.*, 2000; Wood *et al.*, 2000) which involved a network of 49 raingauges in the catchment over an extended period. Errors in the input data should therefore be expected to be relatively small, except perhaps in some localised summer convective events. Another point of interest was that when we investigated the rating curve for the gauging station, it proved to be rather uncertain, so limits of acceptability were defined prior to running the model to span the range of discharge observations (Figure 7.13; see also Pappenberger *et al.* (2006b), Blazkova and Beven (2009), Liu *et al.* (2009) and Westerberg *et al.* (2010a) for other applications of discharge uncertainties within the limits of acceptability GLUE framework). No internal state observations were available for model evaluation in this case. It is worth noting at this point that other studies have looked at representing discharge observational error within a statistical framework (e.g. Thyer *et al.*, 2009; Di Baldassarre and Montanari, 2009; McMillan *et al.*, 2010).

Each model run was evaluated with respect to the limits of acceptability. Figure 7.14 shows the results of two model runs with different parameter sets in the same dynamic TOPMODEL structure, expressed in terms of the standardised scores (with zero at the value of the observation). Two things are quite striking. The first is the nonstationarity in the residual characteristics when expressed in this way; the second is the quite different characteristics of the two models. In fact, neither of the models is behavioural in making

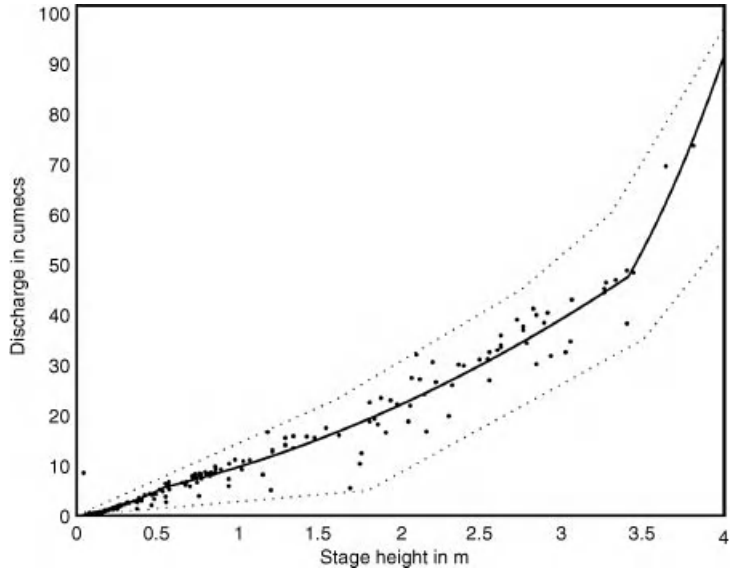


Figure 7.13 *Uncertainty in the rating curve for the River Brue catchment with assumed limits of acceptability bounds for model simulations (calculations by Philip Younger).*

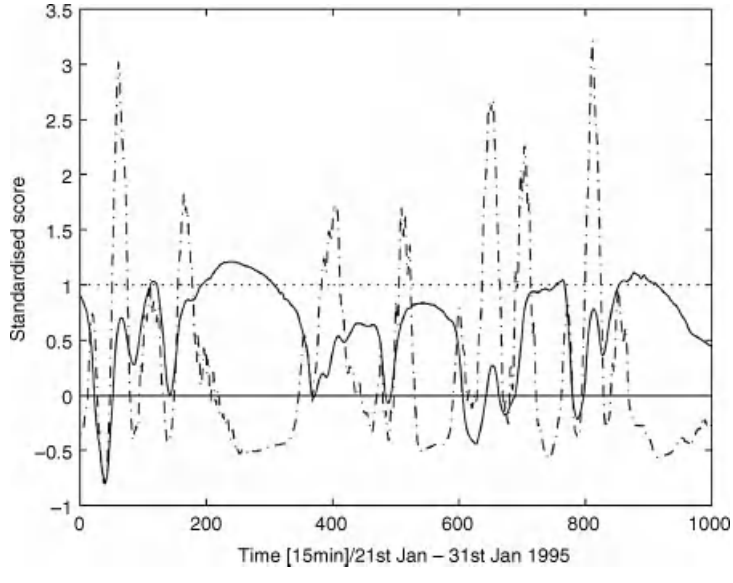


Figure 7.14 *Model residuals as time series of scaled scores for two runs of dynamic TOPMODEL as applied to the River Brue catchment, Somerset, England. The horizontal lines at −1, 0, and +1 represent the lower limit of acceptability, observed value and upper limit of acceptability respectively (calculations by Philip Younger).*

predictions within the limits specified for all time steps in the calibration period. Indeed, of the 50 000 model runs made, not one model was always within the limits. Since no explicit account had been taken of input error in setting the limits, the rejection criterion was relaxed to allow models that were within the limits 95% of the time to be considered behavioural (by analogy with the confidence limits of a statistical evaluation). There is some danger in doing this because, given that the deviations are not random, the remaining 5% could conceivably be the periods of greatest interest (e.g. around the peaks). This is also sometimes seen in the results of statistical error models, giving further support to the idea that random error assumptions may not hold in practice. In the GLUE limits of acceptability approach, no random error assumptions are made, but there is still a potential impact of errors that have not been accounted for in setting the limits.

Under this relaxed rejection criterion, some 413 models were retained as behavioural. An attempt was also made in this study to see whether there was any structure in the residuals that could lead to improved predictions by correcting for bias in the simulations of individual behavioural models. Since this was expected to be nonstationary, the model outputs were classified for different parts of the hydrograph. Five classes were used (rising limbs, peaks, falling limbs, baseflows and troughs). The standardised scores for each of these periods for each behavioural model were examined for consistent deviations with a view to reconstructing the observations in a way that could then be used in prediction. Only in the baseflow periods were consistent deviations found. For the other periods, applying the distributions of scores made little difference to the final uncertainty bounds; in the baseflow periods some improvement in predicting the observations could be found. The resulting uncertainty in the simulated hydrographs for a prediction period (not used in the choice of behavioural models and reconstruction) is shown in Figure 7.15.

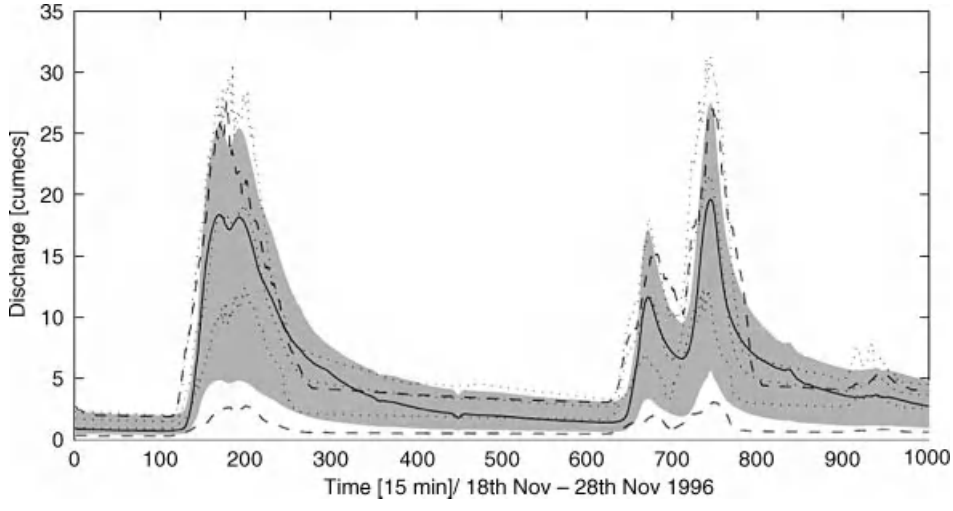


Figure 7.15 *A comparison of reconstructed model flow and observed flow for the validation period, 18–28 November 1996: the reconstructed flow is shown by the middle dotted line and was created by taking the median value of the reconstructed flows from all of the behavioural models at each time step; the outer dotted lines show the 5th and 95th percentiles of the reconstructed model flow; the dashed lines show the 5th and 95th percentiles of the unreconstructed behavioural models; the continuous line is the observed flow; the shaded grey area shows the limits of acceptability applied to the observed discharges for this period (calculations by Philip Younger).*

7.13 Other Applications of GLUE in Rainfall–Runoff Modelling

To summarise, the GLUE methodology provides one way (and not necessarily the only way) of recognising the possible equifinality of models and parameter sets. The approach is conceptually simple, easily implemented, and leads quite naturally to an assessment of the uncertainty of model predictions. There are undoubtedly limitations of the method, of which the most important seems to be computing constraints that restrict the number of runs that can be made in sampling the parameter space (although see the two billion runs, of which 213 were accepted as behavioural, in the study by Iorgulescu *et al.*, 2005, 2007). However, the type of calculations required is ideally suited to parallel computers. The GLUE methodology also involves a number of subjective decisions, so that any uncertainty bounds or prediction limits derived in this way are intrinsically qualitative. However, as noted earlier, all the subjective decisions must be made explicit in the analysis so that they can be discussed or disputed and the analysis repeated, if necessary, with alternative assumptions. Thus, despite the qualitative nature of the procedure, there is some scientific rigour associated with it (and there are good reasons to suggest that the claims that a Bayesian statistical approach might be more rigorous or objective are misleading when the sources of error are epistemic rather than statistical in nature (Beven, 2006a, 2010).

Some recent applications of the GLUE methodology to rainfall–runoff modelling include those by Beven and Freer (2001); Younger *et al.* (2009); Choi and Beven (2007); and Liu *et al.* (2009). Blazkova *et al.* (2002a, 2002b) and Gallart *et al.* (2007) have used internal measurements as well as discharges in model evaluation. Thorndahl *et al.* (2008) have used it in the context of urban runoff modelling; Dean *et al.* (2009) have used it in runoff and water quality modelling. Iorgulescu *et al.* (2005, 2007) and Page *et al.* (2007) have combined runoff modelling with the use of tracers to identify the sources of runoff, while Zhang *et al.* (2006b) examined uncertainty in the parameters of soil pesticide transport models. Blazkova and Beven (2009) used a limits of acceptability approach to the prediction of flood frequencies using continuous simulation in a catchment with multiple stream gauges. In an exercise that involved several months of computation, Vázquez *et al.* (2009) used GLUE in an evaluation of the SHE model. Buytaert and Beven (2009) have shown how the method can be applied in a regionalisation context to predict the uncertain responses of ungauged catchments.

In other areas of hydrological and hydraulic modelling, Franks and Beven (1997a) and Schulz and Beven (2003) have applied GLUE to the problem of estimating land surface to atmosphere fluxes; Brazier *et al.* (2000) looked at erosion modelling; and Schulz *et al.* (1999) have addressed the problem of estimating local nitrogen balances for regulation purposes. Romanowicz and Beven (2003) and Pappenberger *et al.* (2007a, 2007b) have used GLUE in an evaluation of the spatial predictions of flood inundation models. Such a wide variety of applications has necessarily involved a great variety of likelihood measures, depending on the type of data available. This is another reason why Beven (2006a) suggested the limits of acceptability approach to model evaluation as a common method that can be applied to a wide range of different types of calibration data.

One interesting aspect of a number of these applications of GLUE is that the best model available might not attain reasonable levels of acceptability (e.g. Choi and Beven, 2007, Page *et al.*, 2007; Brazier *et al.*, 2000, amongst others; see also Mroczkowski *et al.*, 1997). This suggests that all the models tried could be rejected given the data with which they are being driven and the observations with which they are being evaluated. This would not happen within a statistical framework (except as a result of judgement by the modeller): the final error variance can always expand to satisfy the condition of bracketing the observations to the required level, at least in calibration. A good example is provided by Freer *et al.* (1996) where an error in predicting the onset of snowmelt in one of the simulated years of data results in the observations being outside the prediction limits of the simulations for several weeks (see Figure 7.16). Thus, there would appear to be limitations on how far models can be validated as representations of the real runoff processes. This is not to imply, however, that the model predictions might not be sufficiently accurate to be useful for water resources management, flood forecasting or other purposes.

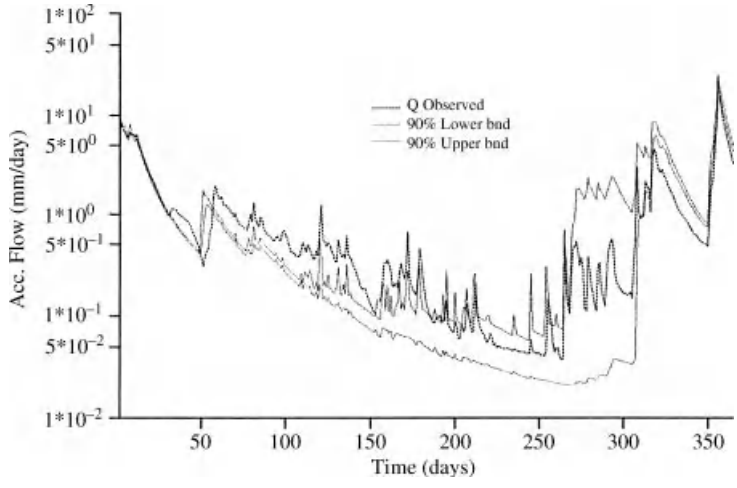


Figure 7.16 *Observed and predicted daily discharges simulated by a version of TOPMODEL for the small Ringelbach catchment (34 ha), Vosges, France: the model was run using an 18-minute time step; note the logarithmic discharge scale; prediction limits are estimated using the GLUE methodology (after Freer* et al. *1996, with kind permission of the American Geophysical Union).*

Three points should be made in respect of such model rejections. The first is that the rejection of a model might be a result of using it with poor data. Thus, to avoid rejecting a good model, just because it has been used with poor data, the data should be examined carefully (see Section 7.17). Secondly, in a complex model space of high dimensions it might be difficult to find the areas yielding the best models, so the modeller should ensure that the space has been searched adequately. Finally, if it is concluded that even the best models are not fit for purpose, this is a good thing! It suggests that something has been missed in defining either the perceptual or conceptual models, or in the numerical implementation of the model equations. Improvements should therefore be possible. We suggest in Chapter 12 that this is one important way in which rainfall–runoff modelling will progress in the future.

7.14 Comparison of GLUE and Bayesian Approaches to Uncertainty Estimation

By this point, the reader may be feeling a little confused about so many approaches to uncertainty estimation (and there are yet more, see Beven, 2009). Despite the considerable advances that have been made in the last decade in understanding the issues involved, there is indeed still much uncertainty about uncertainty estimation in hydrology. The methods we have concentrated on here, GLUE and Bayesian statistical approaches, in fact, represent quite different philosophies in addressing the problem. GLUE is an attempt at formulating a nonstatistical approach to uncertainty, based on the concept of equifinality of different model structures and parameter sets in providing acceptable fits to calibration data. It allows that sources of uncertainty are epistemic rather than statistical in nature. Bayesian statistical methods, on the other hand, try to represent all sources of uncertainty within a coherent statistical framework assuming epistemic uncertainties can be represented as if they were statistical in nature. The latter has an important advantage that if (and only if) the assumptions required hold in prediction then the probability of predicting an observation conditional on the model can be estimated. This is only the case in GLUE if the same

statistical assumptions are made (though as noted earlier, very similar results should then be obtained subject only to sampling of the parameter space). More usually, the uncertainties estimated by GLUE are the empirical probabilities of the predictions of model output by the *ensemble* of behavioural models. It has been suggested, however, that the GLUE approach might be better at revealing nonstationary deficiencies in model structure (and input data) because such deficiencies are not obscured by a statistical error variance. Thus the choice between approaches is not so easy.

There is, however, one feature that is common to both approaches. It is important to distinguish between calibration or conditioning and the prediction of new events. For prediction, both methods need to assume that the characterisation of the uncertainties in calibration or conditioning will hold when carried over to prediction. In the statistical methods, this means carrying over the parameters of the structural model of the different uncertainties. In GLUE this is more implicit. The errors associated with a behavioural model in calibration are assumed to be similar in prediction. Thus, since the epistemic errors in prediction might be different to those in calibration, particularly if the model is used to predict outside the range of the calibration data, both approaches might not necessarily provide good estimates of the uncertainty associated with new events.

This should not, however, suggest that it is not worth trying to make an assessment of uncertainty, whichever method is used. It is, I would suggest, far more dangerous **not** to associate predictions with some estimate of uncertainty and rely only on single deterministic simulations in assessing catchment responses. Unfortunately, there are very many published contributions that still do so, particularly in estimating the impacts of future catchment changes (see Chapter 8). However, taking account of the uncertainty in such predictions might make a difference to the decisions that such predictions are intended to inform.

7.15 Predictive Uncertainty, Risk and Decisions

Let us assume that it has been possible to make a realistic assessment of the uncertainty associated with the predictions of a rainfall–runoff model, by whatever method (e.g. Figures 7.5, 7.7, 7.10 or 7.15). It is important to remember that rainfall–runoff modelling, in practical applications, is done for a purpose: to help some decision. How best then to interpret and make use of those uncertain predictions? One interpretation is that the uncertainty represents the model error in representing the data but a better interpretation is in terms of the risk of a certain outcome in certain circumstances, given the model as a means of extrapolating knowledge and understanding to those circumstances. Both statistical and nonstatistical approaches to uncertainty estimation can be interpreted in this way. In essence, evaluating the risk of an outcome based on the model predictions is also an evaluation of the risk of the model predictions being wrong (as they may well prove to be).

Risk, however, is something that is readily incorporated into modern decision-making processes, when it may also be necessary to take account of the costs of mitigating that risk, for example, in enlarging a reservoir spillway or raising flood embankments. The important point is that the risk associated with the model predictions should be included in the decision analysis. In risk-based decision making, a common technical definition of risk is as:

$$risk(outcome) = probability(outcome) * consequences(outcome) \tag{7.7}$$

In the general case, both the probability and the consequences should be considered as uncertain quantities themselves. The consequences are usually some expected loss (often expressed in monetary terms). Such uncertainties might affect the decision that is made about reducing the risk, so it is important that they be incorporated into the analysis. One rather simple way of doing so is to integrate over all the uncertainties to produce a cumulative distribution of risk. This can then (as in the simple case of estimating a flood frequency) be expressed in terms of the exceedance probability (EP) of a given level

of loss. Thus, the $EP(x > X)$ can be defined in terms of the cumulative density function of the risk of an outcome $F(X)$, obtained by integrating over all probabilities of a consequence less than X, given the probabilities of possible outcomes (both of which might involve some uncertainty):

$$EP(x > X) = 1 - F(X) \tag{7.8}$$

Different loss functions, resulting from different decision options, give different levels of risk as expressed in terms of exceedance probability in this way. Given the design lifetime of a project, the risk can also be expressed in terms of expected annual damages (Pingel and Watkins, 2010). Other approaches to decision making under uncertainty are also available (see Beven (2009), Chapter 6).

7.16 Dynamic Parameters and Model Structural Error

It was noted earlier that a completely objective identification of different sources of uncertainty in the modelling process is very difficult, if not impossible, because of the strong interaction of different sources in the modelling process and the limited information in a series of model residuals. In particular, the potential interaction between input errors and model structural errors was noted (see also Beven, 2005). Model structural error in statistical analysis of uncertainties is generally ignored (unless some identifiable functional form of model inadequacy or discrepancy function can be defined); model identification is generally carried out as if the model was correct. But of course, in hydrology, we know very well that the model is only a very approximate representation of the complexity of the processes occuring in a catchment, and in some cases might not be at all good. Thus, assuming the model is correct might not be a good assumption.

So the question arises as to whether information can be gained about model structural error given the uncertainties in the modelling process. In some studies, simple intercomparisons between different model structures have been made (e.g. Refsgaard and Knudsen, 1996; Butts *et al.*, 2004). In one study, Perrin *et al.* (2001) compared 19 different models on 429 catchments. These were daily lumped catchment models but some of their conclusions should be expected to hold more generally. They showed that more complex models generally do better in simulating the observed discharges in calibration but do not necessarily provide better predictions out of calibration, suggesting the more complex models are subject to overfitting. Refsgaard *et al.* (2006) also give a nice example of how the implementation of a single model structure by different groups of modellers can be very different. Commonly, this type of study has suggested that there may be little to choose between the performance of quite different models, and that combinations of models (for example using Bayesian Model Averaging), can perform better than any single model (e.g. Hsu *et al.*, 2009). As noted earlier, multiple model structures are easily incorporated into the equifinality principle and the GLUE framework, as long as each model structure is subject to the same type of evaluation to infer a likelihood weighting.

More recent approaches to this problem have been based on the idea that if model structural error is important it should lead to consistent patterns in the model residuals. One way of trying to identify such consistent patterns is to allow the distributions of parameters to be time variable. This can be done in a simple way, for example for investigating whether there is consistency in parameter distributions between different types of hydrological conditions. Choi and Beven (2007) did this within the GLUE framework for a small humid catchment in Korea, with 15 classes of hydrological period from wet to dry. They found that there was no overlap between the posterior distributions of parameters in TOPMODEL between the wet and dry periods but suggested that, in making predictions, different sets of parameters might be used in different periods to compensate for the apparent model structural error.

Another approach has been to use filtered estimates of parameter distributions that are allowed to evolve over time. In the dynamic identifiability (DYNIA) approach of Wagener *et al.* (2003), it was reasoned that the distributions of different parameters would be conditioned during different periods and different time

scales. Those that were involved in the generation of fast runoff would be sensitive to wet periods over short time scales; those that were controlling baseflows would be sensitive during dry periods and over a longer time scale. This was used to allow parameter distributions to evolve over time within the GLUE framework. Within a more statistical uncertainty framework, particle-filtering techniques have also been used to allow posterior parameter distributions to evolve over time (e.g. Smith *et al.*, 2008b; Bulygina and Gupta, 2009, 2010; Salamon and Feyen, 2009). In Chapter 4, it was also shown how the recursive estimation of parameters could be used to define the structure of a model within the DBM methodology (see Box 4.3).

7.17 Quality Control and Disinformation in Rainfall–Runoff Modelling

The GLUE approach to uncertainty estimation has been criticised in a number of articles. The main criticism has been that it is not objective because it does not make use of the full power of statistical theory (Mantovan and Todini, 2006; Stedinger *et al.*, 2008). Mantovan and Todini (2006) go further and suggest that the GLUE methodology is not "coherent" in the technical sense that additional observations should lead to further refinement of the posterior parameter distributions. The demonstrations of the advantages of a formal statistical approach to uncertainty estimation, however, have generally relied on hypothetical test cases, in which the assumptions of a formal statistical analysis are valid, because the experiment is designed to ensure that they are valid. Thus the model structure is known to be correct and the model residuals are known to have a simple statistical structure.

In response, Beven *et al.* (2008) argue that this is a special case in GLUE. If it is **known** that the residuals have a simple statistical structure then that knowledge can be used in the analysis and a formal statistical likelihood function can be used within GLUE. The outputs from the uncertainty analysis should then be identical, subject only to sampling differences. This is not, however, the type of situation that GLUE was intended to deal with and they also show that even small changes to the assumptions of a hypothetical example can lead to bias in the posterior parameter distributions, even for cases when the model structure is still known to be correct. This is never the case in real applications, of course, and Beven (2006a) differentiates between ideal cases (where the strong assumptions required by a formal statistical analysis can be accepted) and non-ideal cases (where such strong assumptions are not justified, which will be case for most real applications).

But this also raises the issue as to just how far each model residual should be considered as informative in real applications. If, as suggested for example by Figure 7.5, some event inputs are poorly estimated, then it might be difficult for a good model to predict the observed discharges. The residuals might not then be so informative in identifying a good model or parameter set. In extreme cases, for example where the volume of discharges is greater than the observed volume of inputs for a simple rainstorm, the residuals could actually be disinformative in the calibration process. However, model calibration and uncertainty estimation methods have not generally tried to identify disinformative periods of data.

There is, of course, a difficulty in doing so. In the most extreme cases, perhaps, it might be obvious that there is a problem, but the effect is often more subtle. In Figure 7.17, for example, at the end of the period the model predicts a hydrograph when no change in discharge is observed. This can happen if the model underestimates the available storage prior to that hydrograph, but in this case it follows a relatively wet period. It would therefore appear as if either the observed inputs have overestimated the real inputs to the catchment or there is a problem with the discharge observations. Earlier in the simulation, in the largest observed hydrograph peak, the model residuals are also significant. This might also be a problem of the observed inputs and discharges but here it could also be informative about whether the model structure itself is adequate.

Disinformation in data and how it affects hydrological inference is a topic that has not been discussed much in the hydrological literature. There are a number of measures for information content that are conditional on an assumed model structure in being, directly or indirectly, derived from series of model

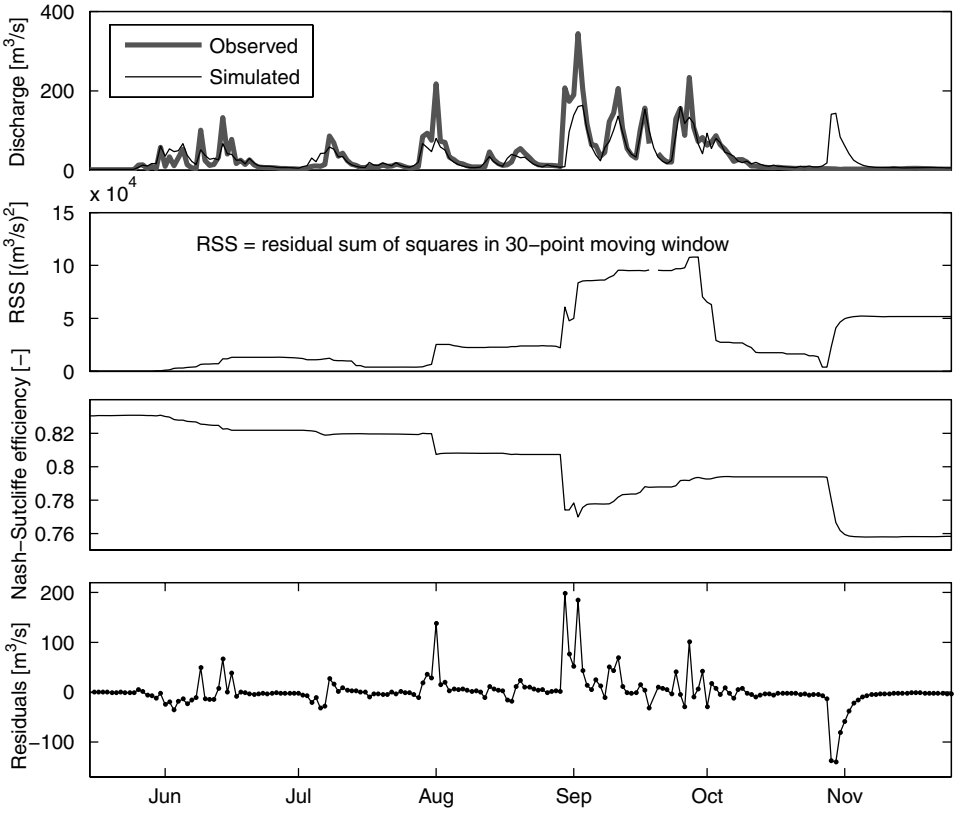


Figure 7.17 *Observed and WASMOD simulated discharge (top) for 1984 in the Paso La Ceiba catchment, Honduras; at the end of October, a period of disinformative data can be seen; the second, third and fourth plots from the top show the effect of the disinformative data on the residual sum of squares, the Nash–Sutcliffe efficiency and the magnitude of the residuals (from Beven and Westerberg, 2011, with kind permission of John Wiley and Sons).*

residuals. Classical information measures (Akaike, Bayes, Deviance, Young, etc.) depend critically on the posterior variance estimates of the model parameter values and the residual variance.

These classical measures effectively assume, however, that the sources of error are essentially stochastic or aleatory in nature, whereas disinformation is a form of knowledge or epistemic error. The classical measures also assume that every residual is informative in conditioning the model parameters and uncertainty. The sample of residuals should asymptotically approach the true distribution for the prediction uncertainty, where generally it is assumed that this true distribution is of simple form and stationary. The residuals may be correlated, in which case the information content will be smaller and the approach to the true distribution will be slower. Some sources of uncertainty may induce long-term correlation (an error in a rainfall input that will also affect the response to later events), but statistical estimation still requires that the structure of the errors be asymptotically stationary.

The nature of the problem of disinformation resulting from substantial nonstationary epistemic error is, however, different. Disinformation tends to increase both the posterior parameter and residual variance (and thereby reduce the apparent information content) but it cannot be represented by some simple form

as if it was aleatory. Disinformation induces nonstationarity in the characteristics of the residuals, not just simple biases, inadequacy functions, heteroscedasticity, or extended autocorrelation that might not be easily compensated for by some error structure model or model inadequacy function (Kennedy and O'Hagan, 2001). It may be one reason why some catchments seem resistant to success in modelling (Andréassian *et al.*, 2007; Kuczera *et al.*, 2010a). The effect may persist over some period of time (particularly where a disinformative input error will affect model outputs over a sequence of more than one event during some relaxation period). It may also be superimposed on some underlying stationary error structure, but the effects on the residual series of these types of epistemic error would not be expected to be simple or stationary.

There is some literature on distinguishing aleatory from epistemic errors and different ways of representing epistemic error (see, for example, the special issue edited by Helton and Oberkampf (2004) and the discussion in the context of groundwater modelling by Ross *et al.*, 2009). The issue of identifying disinformation does not, however, appear to have received much attention. Ideally, we would wish to have an assessment of both information content and periods of disinformation that was independent of the choice of model so that the effective information content for the conditioning of any chosen model could be increased by rejecting some periods of data as disinformative prior to running the model. But how can informative and disinformative periods of data then be distinguished? We could define informative periods as those that have stationary error characteristics that should not, given a large enough sample, lead to bias in inference. We would therefore expect disinformative periods of data to lead to incorrect model conditioning and calibration.

Thus, not all data – and not all model residuals – might be informative in model calibration and hypothesis testing. It is possible that nearly all data sets used in rainfall–runoff modelling will have some periods of disinformative data. This is rather important if we want to get the right results for the right reasons. Inference is always a question of balancing Type I (false positive; i.e accepting a poor model representation because of uncertainties in the conditioning data) and Type II (false negative; i.e. rejecting a good model representation because of uncertainties in the conditioning data) errors. If there are inconsistent or disinformative data being used in the modelling process, then correct inference could only be achieved by a fortuitous balance of disinformation in interaction with a chosen likelihood measure. It would therefore be better to try to identify disinformative periods **prior** to running the model so as to avoid the *reductio ad absurdam* of excluding all periods that the model does not fit as disinformative.

There would appear to be only one way to try to make an assessment of disinformative data independent of a rainfall–runoff model by calculating event water balances and evaluating departures that are outside some acceptable limits of uncertainty. The water balance equation itself, of course, is also a model but one that does not depend on the preferences of an individual hydrologist. The problems with this strategy are the difficulty of separating out the effects of different events on total discharge; the difficulty of evaluating the water balance itself with its many unmeasured components and uncertainties; and the definition of what might constitute "normal" behaviour.

Consider a humid temperate catchment (here the South Tyne at Featherstone, 322 km^2). Rainfall is estimated by interpolating measurements at five sites in the catchment. Storm events were defined as periods separated by at least 12 hours with less than 0.2 mm of rain. To separate the effect of each event on discharge, the technique of extrapolating the recession curve to estimate the volume of discharge that might have occurred if there had been no later rainfall is used (e.g. Reed *et al.*, 1975). To do this, a master recession curve has been constructed from segments of observed winter recession curves (Lamb and Beven, 1997). Discharge measurements at this site are made using a compound Crump weir structure that is not normally overtopped and so should be relatively accurate over the range of observed flows. Figure 7.18 shows the master recession curve and the discharge separation for a single event. A runoff coefficient can then be calculated for each event. These are plotted for multiple events against total event input in Figure 7.19. The range of runoff coefficients is surprisingly wide, even allowing for the approximations involved in their estimation. Many storms show significantly more discharge output than

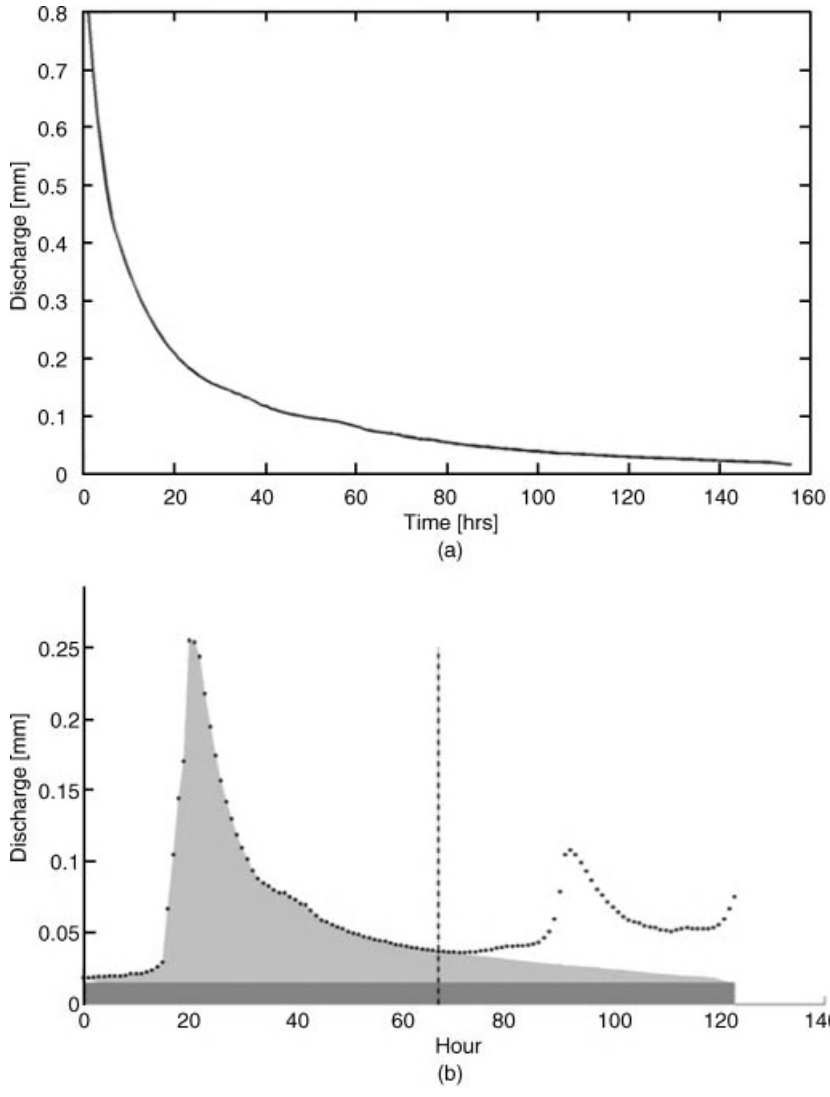


Figure 7.18 *The South Tyne at Featherstone: (a) Master recession curve and (b) example of event separation (after Beven et al., 2011).*

input (runoff coefficients greater than one). For these periods, even a perfect rainfall–runoff model would necessarily underestimate the observed discharges (unless the type of rainfall multiplier of Figure 7.5a is invoked). There are also some very unexpectedly low runoff coefficients for a catchment in which the long-term average discharge is 73% of the observed rainfalls.

Similar issues might be expected in very many catchments. It seems that disinformative periods of data might be a generic problem in rainfall–runoff modelling. Even if it will remain difficult to be secure about identifying disinformative periods because they do not result from simple statistical variability but

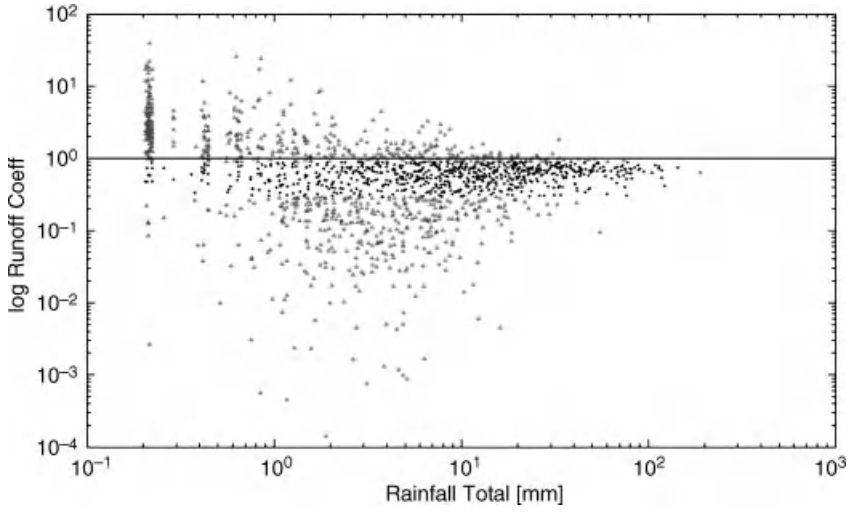


Figure 7.19 *The South Tyne at Featherstone: estimated runoff coefficients for events; black circles have values between 0.3 and 0.9 (after Beven et al., 2011).*

from error characterstics that are changing over time, it should be possible to identify periods that are physically inconsistent with the water balance. We can be sure that they will have an effect on all model calibration strategies and increase the possibility of rejecting a good model because of epistemic errors in the calibration data. This possibility makes a very strong case for quality assurance of the calibration data prior to running any model. It is clear that if the issue of disinformation in model calibration is ignored, and all residuals are treated as being aleatory and informative, we should expect to be wrong in our inference. We should try to avoid that in trying to do good hydrological science. We should, therefore, try to find ways of at least mitigating the worst effects of disinformative data on inference about models and parameters (see Beven et al., 2011).

There is one further issue about the calibration data in rainfall–runoff modelling and that is the additional information that might be added by longer periods of calibration data. Statistical likelihood functions will generally result in further stretching of the response surface and shaping of posterior parameter distributions as more calibration data and longer residual series are added (this is again the coherence argument of Mantovan and Todini, 2006). The new residuals, if of similar characteristics to those that have been seen before, are reinforcing the inference based on earlier periods. This is, however, a choice and an argument could also be made that if the new data has similar characteristics to previous calibration data then not much real additional information is being added (and certainly much less than a statistical likelihood function would suggest, see Box 7.1). If the new data is of quite different characteristics, however, then there is an opportunity to test a model against out-of-previous-sample predictions. The information then provided is much greater and should be given greater weight in conditioning (see also Chapter 12).

What are the implications of these issues about how the likelihood of a particular model and parameter set should be assessed? The answer is that we do not really know. We would wish to use as much information as possible from as wide a range of observations as possible in deciding on a relative likelihood, but avoid periods of disinformation and avoid over-conditioning on data that is subject to epistemic errors. There is still very much to be learned about this part of the modelling process. It can

also be important during flood forecasting, though in a forecasting situation it can be at least partially compensated for by the use of data assimilation (see Chapter 8).

7.18 The Value of Data in Model Conditioning

As in the quotation from Brian Ebel and Keith Loague at the head of this chapter, there have been many exhortations for improved interaction between the field hydrologist and the modeller in providing data of different types for improved model conditioning (Seibert and McDonnell, 2002; Weiler and McDonnell, 2004; Sidle, 2006; Fenicia *et al.*, 2007b, 2008a). However, a search of the literature will reveal very little in the way of advice or guidance as to the value of different types of data in conditioning rainfall–runoff models. In fact, there is also very little advice or guidance on the quality of the rainfall, evapotranspiration and discharge data that are used to drive rainfall–runoff models. Even data sets that have been used in many modelling studies might still have deficiencies (e.g. the Leaf River catchment example discussed in Beven, 2009b). This is, perhaps, not surprising, given the epistemic nature of the errors associated with such data (they are differentiated as epistemic exactly because we do not know too much about their nature).

There were some early studies about how much data was required to optimise a hydrological model (e.g. Gupta and Sorooshian, 1985; Jakeman and Hornberger, 1993) but there has been little in the context of constraining prediction uncertainty (but see Seibert and Beven (2009) for a case of a daily conceptual model within a GLUE framework). This is, however, actually a rather important problem because it is relevant to the issue of planning a measurement campaign for understanding the response of an ungauged catchment (where providing information for a decision might justify the expense).

There is also the possibility of using different types of data within a multiple criterion evaluation. A number of authors have suggested that environmental tracer data might help constrain the representation of runoff processes (e.g. Seibert and McDonnell, 2002; Iorgulescu *et al.*, 2005, 2007; Vaché and McDonnell, 2006; Fenicia *et al.*, 2008b). In the short term such data can distinguish (with some uncertainty) between water in the hydrograph that is event water and that which is pre-event water displaced from storage. In the longer term, tracer data can give some indication of the residence times of water in the catchment system (eg. McGuire and McDonnell, 2006; Botter *et al.*, 2009, 2010). Back in Chapter 1, however, the point was made that distinguishing between event and pre-event water was not the same as distinguishing surface and surbsurface processes; that while an ideal tracer follows the water velocities, the hydrograph is controlled by the wave celerities (see also the kinematic wave analysis in Section 5.5.3). In terms of constraining models, therefore, the information provided by tracer observations might be relevant to different types of parameters (or effective storage volumes) than that provided by discharge observations and hydrographs (see also Chapter 11). This has not always been recognised in studies that have used tracer information in model conditioning. There are other types of data that might also need additional model components to be useful in conditioning. Stream temperatures, for example, have been used to infer discharge increments in river reaches but require an intepretative model, with its own uncertain parameters to be useful. Most remote sensing imaging is also of this type.

This is an area that still demands much more research. This is probably best posed in the context of funding for data collection. If some expense can be justified to try to reduce the uncertainty associated with rainfall–runoff model predictions, how should that money be spent?

7.19 Key Points from Chapter 7

- Limitations of both model structures and the data available on parameter values, initial conditions and boundary conditions generally make it difficult to apply a rainfall–runoff model (of whatever type) without some form of calibration or conditioning of model parameter sets.

- In evaluating the goodness of fit of models to observations we often find that the response surface in the parameter space is very complex. There may be flat areas due to a lack of excitation of certain parameters, long ridges due to interactions between parameters, and many local optima as well as the point of global best fit. In general, carefully designed, simple models with a small number of parameters and good numerical implementation avoid the problem of overparameterisation and have smoother parameter response surfaces but this may be difficult to achieve in rainfall–runoff modelling.
- A review of automatic optimisation techniques reveals that many have difficulties in finding the global optimum on a complex response surface. Techniques such as simulated annealing or genetic algorithms, such as the Shuffled Complex Evolution algorithm, have been designed to be more robust in finding the global optimum.
- Set theoretic techniques based on Monte Carlo simulation suggest, however, that the idea of an optimum parameter set might be illusory and would be better replaced by a concept of equifinality in simulating a catchment indicating that there may be many different model structures or parameter sets that could be considered as acceptable in simulation.
- The concept of the Pareto optimal set allows that multi-criteria parameter optimisation might result in a set of models, each of which achieves a different balance between the different performance measures, but all of which are better than models outside the optimal set. This results in a range of different predictions from the different models in the set, but the range of predictions may not bracket the observations.
- Bayesian statistical methods make specific assumptions about the structure of model residuals from which the definition of a likelihood function follows. This generally leads to over-conditioning of posterior parameter distributions when the errors are due to epistemic errors rather than only aleatory errors.
- Generalised likelihood uncertainty estimation (GLUE) is one technique for conditioning of model parameter sets based on the equifinality concept. In the GLUE methodology, many different model runs are made using randomly chosen parameter sets. Each run is evaluated against observed data by means of a likelihood measure. If a model is rejected, it is given a likelihood value of zero. The likelihood measures are then used to weight the predictions of the retained models to calculate uncertainty estimates or prediction limits for the simulation. Likelihood values from different types of data may be combined in different ways or updated as more data are collected.
- This approach to modelling focuses attention on the value of different types of data in rejecting or falsifying models. Hypothesis tests may be formulated to refine the set of acceptable models in a truly scientific way. Some compromise in such tests is generally necessary, however, since if the criteria for acceptability are made too strict, all models are rejected.
- The assessment of the uncertainty associated with a set of model predictions is also an assessment of the risk of a certain outcome that can be used in a risk-based decision analysis for the problem under study. Taking account of uncertainty might make a difference to the decision that is made and it might therefore not be good practice to rely on deterministic simulations in informing the decision process.
- There is still much to be learned about the information content in periods of calibration data of different types for model evaluation and the constraint of prediction uncertainty. Because of the limitations of hydrological data, some periods might actually be disinformative if used in this way. There will also be more information in periods of hydrologically consistent data that show quite different characteristics than there is in additional data of similar characteristics to the data already available.
- There is little guidance available on the value of different types of data in constraining the uncertainty associated with model predictions. This is best posed as a problem of how to spend a budget on observational data.

Box 7.1 Likelihood Measures for use in Evaluating Models

B7.1.1 Likelihood Measures and Likelihood Functions

The definition of a likelihood function for use in statistical inference is based explicitly on the expected statistical structure of the errors conditional on the idea that there is a correct (even if not true) model and that the error is *aleatory* in nature. This is readily appreciated if the simple example of linear regression is considered. A set of data is observed for two (or more) variables. A scatter plot suggests that there is a linear relationship between two of the variables. A linear regression is used to calculate the values of the intercept (a) and gradient (b) parameters of the model:

$$Y = a + bX \tag{B7.1.1}$$

The line does not go through all the data points but the line is calculated so as to minimise the sum of the squared errors between observed and predicted values of Y. The data points are then in error, relative to the model. The form of the likelihood function used then depends on assumptions about the nature of the errors (whether they are normally distributed, whether they are independent or correlated, etc.). A common set of assumptions for the case of linear regression is that the residuals are independent and identically distributed (iid), with a Gauss distribution (after some transformation of the Y and X variables if necessary). The theory of linear regression then provides equations for the identification of a, b and their covariance **under those assumptions** and the uncertainty of any predicted Y given a value of the independent variable X.

The errors associated with a rainfall–runoff model are not quite so simple (actually this is even true for statistical regressions, see the interesting example provided by Draper, 1995). They arise from multiple sources, including the model inputs and model structure itself. The data with which the model is being compared might also be associated with significant observation error. There is also a tendency for the errors at successive time steps to be correlated, in some cases with correlation that remains high for many time steps (see, for example, Feyen *et al.*, 2007). Thus, defining a statistical structure for the model residuals or for the individual sources of uncertainty is difficult (and Chapter 7 makes the case that many of the errors in the modelling processes might not be statistical in nature at all).

There have been two approaches within the Bayesian statistical framework of model conditioning. The first has been to base a likelihood function only on the nature of the model residuals. Thus at any point in space or time, the residual $\epsilon(x, t)$ can be defined as:

$$\epsilon(x, t) = O(x, t) - M(\theta, I, x, t) \tag{B7.1.2}$$

where $O(x, t)$ is an observed value and $M(\theta, I, x, t)$ is the prediction of a model with parameter set θ and inputs I. In this form, the error is assumed to be additive. A multiplicative error can also be easily implemented in this framework by taking logs of the values $O(x, t)$ and $M(\theta, I, x, t)$.

It is then common to follow Gauss and base the contribution of a single residual to the likelihood as being proportional to the square of its value so that:

$$L(\epsilon|M(\theta, I)) \propto exp\left[-\frac{\epsilon^2}{\sigma_\epsilon^2}\right] \tag{B7.1.3}$$

Over a time series of N such residuals, assuming independence, the individual contributions combine multiplicatively so that

$$L(\epsilon|M(\theta, I)) \propto \prod^{N} exp\left[-\frac{\epsilon^2}{\sigma_\epsilon^2}\right] \tag{B7.1.4}$$

leading to a likelihood function for iid Gaussian errors of the form:

$$L(\epsilon|M(\theta, I)) = \left(2\pi\sigma_\epsilon^2\right)^{-N/2} exp\left[-\frac{1}{2\sigma_\epsilon^2}\left(\sum_{t=1}^{N}\left[\epsilon_t^2\right]\right)\right] \tag{B7.1.5}$$

There is a large body of likelihood function theory developed for different sets of assumptions. Assuming, for example, that the errors have a mean bias and are correlated in time yields a likelihood function of the form:

$$L(\epsilon|M(\theta, I)) = \left(2\pi\sigma_\epsilon^2\right)^{-N/2}(1-\rho^2)^{1/2} exp\left[-\frac{1}{2\sigma_\epsilon^2}\left((1-\rho^2)(\epsilon_1-\mu)^2 + \sum_{t=2}^{N}\left[\epsilon_t - \mu - \rho(\epsilon_{t-1}-\mu)\right]^2\right)\right] \tag{B7.1.6}$$

Note that in Equations (B7.1.5) and (B7.1.6) there is a term in $\left(\sigma_\epsilon^2\right)^{-N/2}$. When the calibration data are based on time series of observations, the value of N can be very large. This means that a formal likelihood of this type will take models of similar error variance and stretch the difference between them. The effect is reduced by the other bias and correlation terms in Equation (B7.1.6) but can still involve small differences in error variance being stretched to orders of magnitude differences in likelihood. In fact, the calculations are normally made in log likelihood transformation to avoid small likelihoods being lost in computer rounding error. The result is to produce a highly peaked likelihood surface and focus attention on the models with the highest likelihood (even if that model might have only a very slightly smaller error variance than many other models). Such a stretching of the surface is a consequence of assuming that every residual contributes to the overall likelihood in a multiplicative (but not zero) way (as in the product of Equation (B7.1.5)).

The question is whether such a stretching is justified in real cases? Why should models with only slight differences in residual variance have orders of magnitude difference in likelihood? That does not seem reasonable (even if consistent with the theory) when we know little of the various sources of error in the modelling process. The sensitivity of the stretching does mean that different realisations of error in the calibration data (such as provided by different calibration periods) might result in peaks in the surface being in quite different places. Indeed, even in hypothetical cases where we can ensure that the model is correct, that we know the true parameter distributions, and the assumptions about the structure of the error are correct, different error realisations can result in biased parameter estimates using a statistical likelihood of this type (see the work of Kuczera, 1983, and Beven *et al.*, 2008a). In the hypothetical case, the addition of more observations should result in convergence on the true parameter distributions. In real applications, we cannot be sure of the same asymptotic behaviour.

Another common feature of actual residual series is that they are *heteroscedastic*. This means that the variance of the errors changes with the magnitude of the prediction (in rainfall–runoff modelling, it is common for the residuals at higher discharges to be greater for a variety of reasons). It is possible to make some assumption about the nature of the dependence of residual variance in deriving a likelihood function (e.g. Schoups and Vrugt, 2010), but a more common strategy is to transform the residuals such that they have a more constant variance and are more nearly Gaussian in distribution. The most common transformation used is the Box–Cox transformation (Box and Cox, 1964). This is defined as:

$$\begin{aligned} \epsilon_t^* &= \left(\epsilon_t^\lambda - 1\right)\lambda && \text{for } \lambda \neq 0 \\ &= \ln \epsilon_t && \text{for } \lambda = 0 \end{aligned} \tag{B7.1.7}$$

where ϵ_t^* is the transformed values of the variable ϵ_t and λ is a constant chosen so as to minimise the skewness of the transformed values. Note that if the original residual series exhibits autocorrelation, the transformed residuals are autocorrelated.

For extreme cases of "unusual" distributions of residuals, a technique called "normal quantile transform" (NQT) can be used to ensure that the transformed residuals have a Gaussian distribution (see, for example, the work of Kelly and Krzysztofowicz (1997); Montanari and Brath (2004); Götzinger and Bárdossy, 2008). This is achieved by taking a time series of residuals and ranking them from high to low. The quantiles of the resulting cumulative distribution are then taken as the quantiles of a standard Gaussian distribution (see Figure B7.1.1). Gaussian likelihood theory can then be applied to the transformed values (though again the series of transformed residuals might still be correlated). Personally, I find this shoehorning of complex residuals that might have non-stationary statistics into Gaussian form somewhat disturbing. It is true that it allows the power of the statistical theory to be applied but in a way that might have the effect of overstimating the information content of the residuals (including the extreme stretching of the likelihood surface noted above) and therefore lead to overconditioning. Other forms of transformation are also possible. Montanari and Toth (2007), for example, use a spectral domain decomposition of the residuals leading to a likelihood function derived by Whittle; while Schaefli and Zehe (2009) use a wavelet decomposition of the residuals.

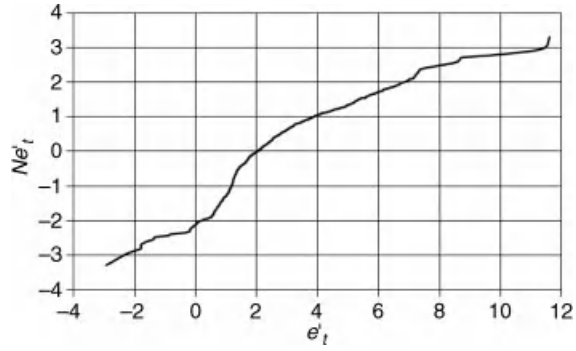


Figure B7.1.1 *A normal quantile transform plot of the distribution of the actual model residuals (horizontal axis) against the standardised scores for the normal or Gaussian distribution (vertical axis) (after Montanari and Brath, 2004, with kind permission of the American Geophysical Union).*

There are two important points to remember about likelihood functions of this type. The first is that they are based on treating the residual errors as only varying statistically or in an aleatory way. The second is that the assumptions about the structure should always be checked against the actual series of residuals for a model run. This is only usually done for the maximum likelihood model found (which does not necessarily imply that the same error structure applies throughout the parameter space). Engeland *et al.* (2005) show an example of good practice in this respect; Feyen *et al.* (2007) do not. The latter used an assumption of independent and uncorrelated errors in the identification of the maximum likelihood model but then showed that the residuals were highly correlated in time. They did not go back and repeat the analysis using a likelihood function that included autocorrelation. This means that the resulting posterior parameter distributions are biased.

There is actually a further point that should be recalled here. The proportionality of Equation (B7.1.3) derives from the work of Gauss in the early 19th century. It is the basis for all the widely used statistical theory based on squared errors. It is, however, a choice. At much the same time, Laplace proposed an alternative error norm based on a proportionality to the absolute error (see Tarantola, 2006). Before the days of digital computers, this was much less analytically tractable than the Gauss assumption, so did not attract so much attention. There are arguments

for using the squared error (see, for example, Halpern, 2003), but once it is realised that it is a choice there is no real reason why it should be an exclusive choice. This wider view does open up the possibility of using *likelihood measures* based on different types of model evaluation, including qualitative and fuzzy measures that might reflect the complex sources of error in the modelling process in different ways.

The definition of a likelihood measure for use within the GLUE methodology is more relaxed. It requires only a measure that is zero for nonbehavioural models and increases monotonically as performance or goodness of fit to the available observational data increases. There are many different measures that could be used (including performance measures that have been used in optimisation of model parameters in the past). The choice of measure should reflect the purposes of the study and the nature of the errors being considered but, ultimately, the choice of a likelihood measure is subjective. In this, it has much in common with recent developments in approximate Bayesian computation (ABC) that allow for model conditioning without likelihoods using rejection criteria (see, for example, Beaumont *et al.*, 2002, 2009; Toni *et al.*, 2009).

B7.1.2 Hierarchical Statistical Likelihoods

Recent developments in model conditioning using statistical theory have been concerned with trying to take more explicit account of different sources of uncertainty, rather than dealing only with the lumped model residuals of Equation (B7.1.2) (e.g. Kuczera *et al.*, 2006; Liu and Gupta, 2007; Götzinger and Bárdossy, 2008; Huard and Mailhot, 2008; Vrugt *et al.*, 2008; Thyer *et al.*, 2009; Renard *et al.*, 2010). This requires a hierarchical Bayesian approach to identification with the specification of separate error models for input uncertainties, observation uncertainties, parameter uncertainites and residual uncertainties. These specifications will have parameters (often called *hyperparameters* of the analysis) that are identified as part of the conditioning process. A typical assumption for input error, as shown in Section 7.8, is that the uncertainty associated with the rainfall estimation for an event can be treated as a multiplier on the observed rainfalls, where the multipliers are chosen from a log normal distribution defined by mean and variance hyperparameters (e.g. Thyer *et al.*, 2009; Renard *et al.*, 2009; Vrugt *et al.*, 2008). Errors in discharge observations might be represented by a heteroscedastic model in which an additive error is assumed to have a variance that is proprtional to the value of discharge. The simplest heteroscedastic model has a single hyperparameter, a constant relative standard error (Huard and Mailhot, 2008; Thyer *et al.*, 2009; McMillan *et al.*, 2010) but more complex models might also be defined (Moyeed and Clarke, 2005). It might also be necessary to allow that the rating curve is non-stationary in mobile bed rivers that are subject to extreme discharges (e.g. Westerberg *et al.*, 2010a). A residual variance is usually assumed to follow the simple additive form of Equation (B7.1.2).

Note that a number of additional degrees of freedom have been added here in explaining the uncertainty in the total model residuals, while constraining the representation of different sources of uncertainty to specific assumed forms. Such a disaggregation will, ideally, assign the effects of different sources of uncertainty correctly but, as with any form of statistical inference of this type, the validity of the inference will depend on the validity of the assumptions. In this case, it might be quite difficult to test assumptions about the sources of uncertainty and their stationarity independently of the analysis and, as noted in Section 7.8, it is possible that the sources of uncertainty interact; rainfall multipliers may interact with model structural error, for example. Where the assumptions are not strictly valid, this will almost certainly lead to over-conditioning of the posterior parameter distributions.

B7.1.3 Informal Likelihood Measures

It follows from this discussion that treating all sources of error as aleatory tends to result in overconditioning of posterior model likelihoods when some epistemic sources of error are

not well behaved in this statistical sense. This is most obvious in the way that models that have very similar error variance can have orders of magnitude difference in likelihood when Equations (B7.1.5) or (B7.1.6) are applied to large numbers of residual errors. The problem then is what type of likelihood measure might more properly reflect the real information content of the conditioning data for these non-ideal cases (Beven, 2006a). The GLUE methodology, for example, has been used with a variety of informal likelihood measures, such as the Nash–Sutcliffe efficiency measure or the inverse exponential variance measure used in the case study in Section 7.11.

Smith *et al.* (2008a) have made an analysis of a number of informal likelihood measures, in particular the extent to which an informal likelihood can be consistent with the axioms of probability as an expression of belief in a particular model realisation. Such belief should reflect how far the simulated variables are unbiased with respect to the observations, rise and fall with a similar pattern to the observations, and show similar variability as the observations. Note that expressing belief in this way is not inconsistent with the original Bayes (1763) concept of taking account of the evidence in assessing the odds for belief in a particular hypothesis.

The essential requirements of such a measure is that it should increase monotonically as the fit to the observations improves (though the goodness of fit is usually defined with respect to the particular measure being used). Within the GLUE framework, the likelihood can also be set to zero for models that are rejected as *nonbehavioural*, which might involve defining some additional threshold conditions as local (for each single observation) or global (over all observations) limits of acceptability.

Concentrating for the moment on the Nash–Sutcliffe efficiency measure (Equation (7.3)), Smith *et al.* show how this can be thought of as the combination of three positive components in the form:

$$E = \left(\rho^2 - \left[\rho - \frac{\sigma_s}{\sigma_o} \right]^2 - \left[\frac{\bar{o} - \bar{s}}{\sigma_o} \right] \right) \text{ for } E > 0 \qquad \text{(B7.1.8)}$$

where ρ is the linear correlation coefficient between the observed o and simulated s variables with means $\bar{o}$ and $\bar{s}$ and variances σ_o and σ_s respectively.

The first term summarises the ability to reproduce the pattern of the data through the linear correlation. The second and third terms penalise differences in the variance and mean of the two data series. The difference in the mean is penalised relative to the standard deviation of the observed series. The second term penalises deviations of the ratio of standard deviations away from ρ, rather than 1 which would be desirable. The second term combined with the first term indicates that a positive efficiency value is only possible when ρ is positive. This type of decomposition provides the basis for evaluating different measures with respect to the axioms of probability. It is shown that the efficiency measure (when scaled such that the sum over all the ensemble of models considered is unity) does satisfy the axioms, whereas some other objective functions, such as Willmott's Index of Agreement or, more importantly, a total volume error measure, do not (see the work of Smith *et al.* (2008a) for details and Guinot *et al.* (2011) for an argument in favour of using the balance error as a model evaluation criterion).

They also show, in a hypothetical example for which the true results are known, that such informal measures do not produce the correct conditioning of posterior parameter distributions for residuals with a consistent bias. In this, however, they are no worse than a formal likelihood that does not explicitly account for the bias (e.g. Equation (B7.1.5)), reinforcing the fact that the assumptions of a formal likelihood measure should always be checked against the characteristics of the actual model residuals.

A further feature of informal likelihood measures considered was the way in which they continue to condition the posterior parameter distributions as more data are added. This was the essence of the criticism of the use of the Nash–Sutcliffe efficiency in GLUE by Mantovan and Todini (2006). When used as a global measure, the efficiency will asymptotically reduce the rate of conditioning as new data are added. However, as pointed out by Beven *et al.* (2008) this is a choice. An alternative choice would be to use the efficiency as a common sense

informal measure for periods of "similar" data (when the asymptotic limit on learning from new data might be justified) but to combine evaluations for periods of different characteristics, or from different types of observations, using one of the methods in Box 7.2.

B7.1.4 Fuzzy Measures for Model Evaluation

There have been some attempts to take a different approach to model evaluation based on non-statistical measures. One approach is that of using fuzzy measures (Franks *et al.*, 1998; Aronica *et al.*, 1998; Pappenberger *et al.*, 2007b; Jacquin and Shamseldin, 2007). It has been used particularly in situations where observational data are scarce and statistical measures might be difficult to evaluate. The basic fuzzy measure can be defined as a simple function of the error between observed and predicted variables, such as discharge. If the error is zero (or within a certain range of zero) then the fuzzy measure is assumed to be at a maximum, say unity. The measure then declines to zero as the error gets larger in some defined way. A linear decline towards zero at some maximum allowable error is often assumed. The maximum allowable error might be different for overprediction compared with underprediction. It might also be different for different time steps or variables. The individual fuzzy measures for different time steps or variables may then be combined in some way (linear addition, weighted addition, multiplication, fuzzy union or fuzzy intersection; see Box 7.2) to create an overall measure for model evaluation. Such measures can be treated as likelihood measures within the GLUE methodology and arise naturally in the "limits of acceptability" approach to model evaluation suggested by Beven (2006a). They have no foundation in statistical theory but they can be used to express likelihood as a measure of belief in the predictions of a model.

Within the limits of acceptability approach, one way of making evaluations against different types and magnitudes of observations comparable is to scale the performance of a model in terms of a standardised score (Figure B7.1.2). The score has the value of zero at the observed value, −1 at the lower limit of acceptability and +1 at the upper limit. Overprediction therefore leads to positive scores and underprediction to negative scores. Series of these standardised scores are shown for an application of dynamic TOPMODEL in Figure 7.14. These scores can

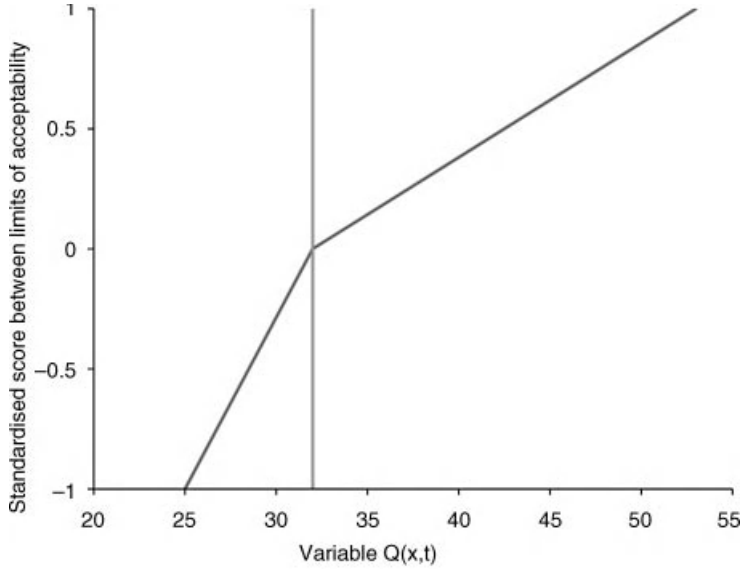


Figure B7.1.2 *Residual scores.*

be transformed simply into likelihood measures (Figure B7.1.3). The simplest transformation would use a triangular function, analogous to a triangular fuzzy measure, but other forms of transformation are also easily implemented according to the characteristics of particular observations. Note that the resulting likelihood measures apply to single observations. They need to be combined in some way to form an overall likelihood weight for that model simulation.

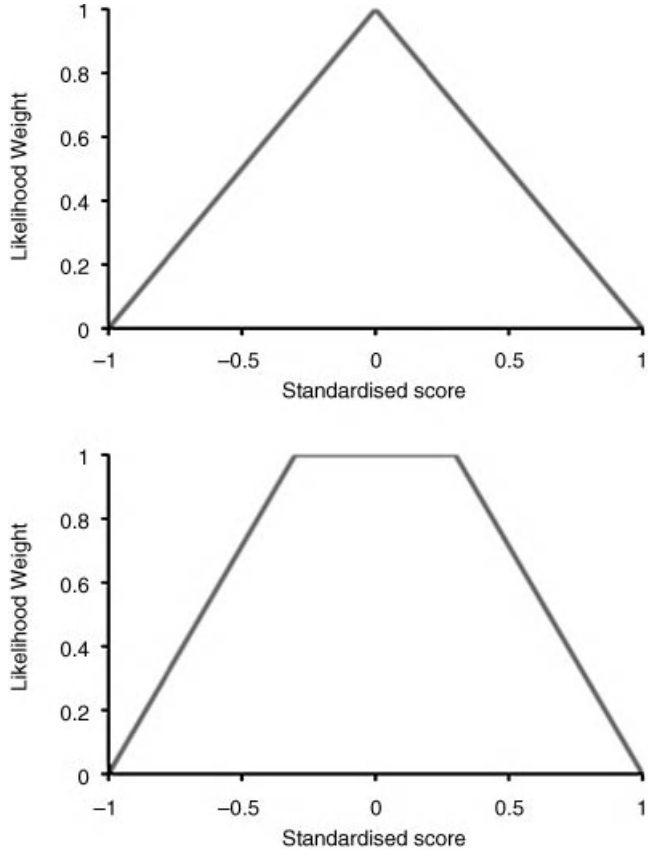


Figure B7.1.3 *Conversion of scores to likelihoods.*

B7.1.5 Qualitative Measures for Model Evaluation

Model evaluation need not be based solely on quantitative measures of model fit to observations. Some qualitative or "soft" measures might also be useful (e.g. Seibert and McDonnell, 2002; Vaché and McDonnell, 2006). An example would be in evaluating the predictions of a distributed model. The distributed model might, with different parameter sets, be able to obtain good fits to observed discharges using different runoff generation mechanisms, e.g. infiltration excess runoff alone, saturation excess runoff alone, subsurface stormflow alone or a mixture of them. If, however, observations suggest that there is negligible overland flow generated on the hillslopes on a catchment, then models that give good simulations based on such mechanisms should be rejected, even though those models might actually give the best simulations. Other soft measures might be used based on acceptability in reproducing various hydrological

signatures (e.g. Sivapalan *et al.*, 2003; Sivapalan, 2005; Winsemius *et al.*, 2009). Most qualitative measures of performance can be set up as either a binary likelihood measure (one if behavioural according to the qualitative measure, zero if not) or as a fuzzy measure. They may be inherently subjective, in that different modellers might choose to emphasise different aspects of the behaviour in making such a qualitative assessment (see, for example, Houghton-Carr, 1999).

Box 7.2 Combining Likelihood Measures

There may be a need to combine different likelihood measures for different types of model evaluation (such as, one measure calculated for the prediction of discharges and one calculated for the prediction of water table or soil moisture levels) or to update an existing likelihood estimate with a new measure calculated for the prediction of a new observation or set of observations.

Most cases can be expressed as successive combinations of likelihoods, where a prior likelihood estimate is updated using a new likelihood measure to form a posterior likelihood. This is demonstrated by one form of combination using *Bayes equation*, which may be expressed in the form:

$$L_p(M(\underline{\Theta})|\underline{Y}) = \frac{L_o(M(\underline{\Theta}))L(M(\underline{\Theta})|\underline{Y})}{C} \tag{B7.2.1}$$

where $L_o(M(\underline{\Theta}))$ is the prior likelihood of a model with parameters $\underline{\Theta}$, $L(M(\underline{\Theta})|\underline{Y})$ is the likelihood calculated for the current evaluation given the set of observations $\underline{Y}$, $L_p(M(\underline{\Theta})|\underline{Y})$ is the posterior likelihood and C is a scaling constant to ensure that the cumulative posterior likelihood is unity. In the GLUE procedure, this type of combination is used for the likelihoods associated with individual parameter sets so that for the ith parameter set:

$$L_p(M(\underline{\Theta}_i)|\underline{Y}) = \frac{L_o(M(\underline{\Theta}_i))L(M(\underline{\Theta}_i)|\underline{Y})}{C} \tag{B7.2.2}$$

and C is taken over all parameter sets. Strictly, in the theory of Bayesian statistics, this form should be used only where a sampled likelihood (for a particular parameter set) can be assumed to be independent of other samples. In GLUE, the parameter sets are chosen randomly so as to be independent samples of the parameter space.

The use of Bayes equation to combine likelihoods has a number of characteristics that may, or may not, be attractive in model evaluation. Since it is a multiplicative operation, if any evaluation results in a zero likelihood, the posterior likelihood will be zero regardless of how well the model has performed previously. This may be considered as an important way of rejecting nonbehavioural models; it may cause a re-evaluation of the data for that period or variable; it may lead to the rejection of all models.

Statistical likelihoods, however, never go to zero but only get very, very small (in some cases with time series of residuals leading to hundreds of orders of magnitude difference in likelihood between good and poor models which is why statisticians tend to work with log likelihood rather than likelihood so as to make the calculations feasible within the rounding error of digital computers).

The successive application of Bayes equation will tend to gradually reduce the impact of earlier data and likelihoods relative to later evaluations. This may be an advantage if it is thought that the system is changing over time. It has the disadvantage that if a period of calibration is broken down into smaller periods, and the likelihoods are evaluated and updated for each period in turn, for many likelihood measures the final likelihoods will be different from using a single evaluation for the whole period. An exception is a likelihood measure based on

an exponential transformation of an error measure E since for $j = 1, 2, \ldots, m$ periods with $L_j \propto \exp(E_j)$:

$$L_p(M(\underline{\Theta}_i)|\underline{Y}) = \exp(E_1)\exp(E_2)\ldots\exp(E_m) = \exp(E_1 + E_2 + \ldots E_m) \quad \text{(B7.2.3)}$$

However, Bayes equation is not the only way of combining likelihoods. A simple weighted addition might be considered more appropriate in some cases such that for m different measures:

$$L_p(M(\underline{\Theta}_i)|\underline{Y}) = \frac{W_o L_o(M(\underline{\Theta}_i)) + W_1 L(M(\underline{\Theta}_i)|\underline{Y}_1) + \ldots W_m L(M(\underline{\Theta}_i)|\underline{Y}_m)}{C} \quad \text{(B7.2.4)}$$

where W_o to W_m are weights and $\underline{Y}_1$ to $\underline{Y}_m$ are different evaluation data sets. A weighted addition of this form has the effect of averaging over any zero likelihood values for any individual periods. Again C is calculated to ensure that the cumulative posterior likelihoods sum to one.

Further forms of combination come from fuzzy set theory. Fuzzy operations might be considered appropriate for the type of fuzzy likelihood measures introduced in Box 7.1. A fuzzy union of several measures is effectively the maximum value of any of the measures, so that

$$\begin{aligned} L_p(M(\underline{\Theta}_i)|\underline{Y}) &= \frac{L_o(M(\underline{\Theta}_i)) \cap L_1(M(\underline{\Theta}_i)|\underline{Y}) \cap \ldots L_1(M(\underline{\Theta}_i)|\underline{Y})}{C} \\ &= \frac{\max\left(L_o(M(\underline{\Theta}_i)), L_1(M(\underline{\Theta}_i)|\underline{Y}), \ldots L_1(M(\underline{\Theta}_i)|\underline{Y})\right)}{C} \end{aligned}$$

The fuzzy intersection of a set of measures is the minimum value of any of the measures:

$$\begin{aligned} L_p(M(\underline{\Theta}_i)|\underline{Y}) &= \frac{L_o(M(\underline{\Theta}_i)) \cup L_1(M(\underline{\Theta}_i)|\underline{Y}) \cup \ldots L_1(M(\underline{\Theta}_i)|\underline{Y})}{C} \\ &= \frac{\min\left(L_o(M(\underline{\Theta}_i)), L_1(M(\underline{\Theta}_i)|\underline{Y}), \ldots L_1(M(\underline{\Theta}_i)|\underline{Y})\right)}{C} \end{aligned}$$

Thus, taking a fuzzy union emphasises the best performance of each model or parameter set over all the measures considered; taking a fuzzy intersection emphasises the worst performance. In particular if any of the measures is zero, taking a fuzzy intersection leads to the rejection of that model as nonbehavioural. All of these possibilities for combining likelihood measures are included in the GLUE software (see Appendix A).

Box 7.3 Defining the Shape of a Response or Likelihood Surface

In estimating the uncertainty associated with a set of model predictions, it is generally necessary to evaluate the outputs associated with different sets of parameter values and their associated likelihood weight. In a forward uncertainty analysis, the likelihood weights are known from the definition of the prior distributions of the parameters and it is relatively easy to define an efficient sampling scheme. When the likelihood weights need to be determined from a conditioning process of comparing predicted and observed variables, however, it is not known exactly where in the parameter space the areas of high likelihood might be, even if in a Bayesian analysis some prior likelihoods have been specified. This gives rise to a sampling problem, since there is no point in repeatedly sampling areas of low likelihood in the space. A simple random Monte Carlo sampling technique might then be very inefficient, especially in high-dimensional spaces where areas of high likelihood are very localised. Ideally, therefore

an adaptive search strategy will seek out those areas of higher likelihood and concentrate the sampling in those areas with a local sampling density proportional to the local magnitude of the likelihood.

This is effectively a process of trying to define the shape of a likelihood surface in the parameter space, concentrating on the peaks in that surface. Those peaks are easiest to find if they are not too localised and if they are close to the prior regions of high likelihood. This is not always the case, however, and one of the important features of an adaptive sampling scheme is that it should not concentrate exclusively on those areas of high likelihood that have already been found but should continue to sample other parts of the space in case there are other areas of high likelihood that have not yet been found. Different search algorithms have different ways of doing this. In high-dimensional spaces of complex structure there is still a possibility of missing areas of high likelihood just because they never get sampled. It has also recently been appreciated that the complexity of the surface, and consequently the difficulty of finding areas of high likelihood, might also be the result of the numerics of a particular model implementation (see, for example, Schoups *et al.*, 2008, 2010; Kavetski and Clark, 2010). Much more on sampling and search techniques is discussed by Beven (2009). Note that the techniques presented in the following discussion can be extended to multiple model structures.

An important issue that arises in the implementation of such methods is the generation of random numbers. Since it might be necessary to generate very large numbers of random numbers, the particular characteristics of different random number schemes might have an effect on the sampling. In fact, there are no random number schemes on digital computers, there are only algorithms for generating pseudo-random numbers (see Beven, 2009). As such there are some good algorithms (with long repeat sequences and low correlation) and some poor algorithms (including some of the standard calls made available in programming languages). Even the choice of parameters for a particular algorithm can have an important impact on the apparent randomness of the resulting sequences. A modern algorithm that, with suitable parameters, can give an enormous period before repeating is the Mersene Twister (Matsumoto and Nishimura, 1998).

B.7.3.1 Structured Sampling for Forward Uncertainty Analysis

In a forward uncertainty analysis, the areas of high likelihood in the space to be sampled will be known explicitly from the definition of the prior distributions for the parameter values. This will include any correlations between different parameters if necessary (either specified as a covariance matrix if the distributions are assumed to be multivariate Gaussian or, more generally, in terms of copula functions Beven, 2009). An obvious sampling strategy is then to sample each parameter randomly in a way that reflects the probability density of that parameter (as well as any effects of covariation). However, even if only 10 samples are taken per parameter this very rapidly builds up to a large number of samples when there are more than a small number of parameters.

More efficient, if more approximate, sampling techniques have been proposed of which the most commonly used is Latin Hypercube sampling (Iman and Conover, 1982). In this approach the number of samples is decided beforehand. Each parameter axis is then split into the same number of values but using a cumulative probability scale so that for each parameter each sample will be of equal probability. For the simplest case of independent parameters (and it is often very difficult to specify parameter correlations *a priori* except in some special circumstances), then a value of each parameter is chosen randomly to create a parameter set. The next sample chooses a new set of values, but without replacement, so that once all the samples have been chosen all the parameter values have been used. Iman and Connover (1982) show how rank correlation between parameters can be introduced into the Latin Hypercube sampling process.

B7.3.2 Simple Monte Carlo Sampling

The sampling problem is much more difficult when we do not know the shape of the likelihood surface before starting sampling. There may also be little prior information to guide the search, since it can be difficult to estimate beforehand what **effective** values of the parameters (and their potential interactions) might be needed to get good results from a model in matching the available observations. It is quite common, for example, to only estimate some prior upper and lower limits for particular parameters, without any real idea of what distribution to assume between those limits.

Without such prior information there is initially little to guide the sampling strategy. So a rather obvious choice of strategy is to use simple Monte Carlo sampling, in which random values of each parameter are chosen independently across the specified ranges. Where prior information is available, this is easily modified to samples of equal probability by sampling across the cumulative probability range; this will result in a sampling density proportional to the prior probability. The only problem with this very simple strategy is that of taking enough samples. Similar issues arise as for the forward uncertainty estimation problem, except that now we do not know where to concentrate the search. If only a small number of samples are taken, areas of higher likelihood on the surface might be missed. The method is therefore only useful where model runs that might be retained as behavioural (higher likelihood) are spread through the parameter space. Otherwise, very large numbers of samples might be required to define the shape of local areas of high likelihood. However, it is the commonly the case where non-statistical likelihood measures are used within the GLUE methodology and many GLUE applications have used this type of simple Monte Carlo sampling. Refinement of the simple sampling strategy is also possible by discretising the space into areas where high likelihoods and low likelihoods have been found by some initial search and then concentrating sampling in the sub-spaces of high likelihood. A variety of methods are described by Beven and Binley (1992), Spear *et al*. (1994), Shorter and Rabitz (1997), Bárdossy and Singh (2008) and Tonkin and Doherty (2009).

B.7.3.3 Importance Sampling: Monte Carlo Markov Chain

As noted elsewhere, statistical likelihood functions tend to stretch the response surface greatly, resulting in one or more areas of high likelihood that are highly localised. This means that simple Monte Carlo search algorithms would be highly inefficient for such cases and a more directed strategy is needed to define the shape of the surface with any degree of detail. Most strategies of this type are adaptive in the sense of using past samples to guide the choice of new samples and have the aim of finishing with a set of samples that are distributed in the parameter space with a density that is directly proportional to the local likelihood. This is a form of importance sampling. The most widely used techniques for importance sampling in hydrological modelling are those of the Monte Carlo Markov Chain (MC^2) family. They have been used in rainfall–runoff modelling at least since Kuczera and Parent (1998).

The concept that underlies MC^2 sampling is quite simple to understand (see also Beven, 2009). The scheme starts with a set of random samples chosen according to some proposal scheme. Each chosen point represents a parameter set. The model is run with that parameter set and a posterior likelihood for that point is calculated. A new set of points around that point is then chosen, consistent with the proposal distribution. Whether a model run is made at the new point depends on the likelihood of the original point and a random number, so that there is a probability of making a run even if the likelihood of the original point was low. This is to guard against not sampling regions of the space where a new high likelihood area might be found. Once the sampling is complete, the chain is checked to see if the sample of points is converging on a consistent posterior distribution. If not, another iteration is carried out, which might involve adapting the proposal distribution to refine the sampling. The process is effectively a chain of random walks across the likelihood surface, where the probability of

choosing a new sample is dependent on knowledge of the local values of likelihood, but where it is possible to start off a new chain in areas that have not been sampled before.

The most commonly used form of MC^2 sampling is the Metropolis–Hastings algorithm (see the work of Kuczera and Parent (1998), Engeland *et al.* (2005), Gallagher and Doherty (2007) and Smith and Marshall (2008) for examples in rainfall–runoff modelling). MC^2 sampling is an essential part of the BATEA uncertainty estimation methodology of Kuczera *et al.* (2006; see also Thyer *et al.*, 2009; Renard *et al.*, 2010) and the DREAM approach of Vrugt *et al.* (2008a, 2009; Schoups *et al.*, 2010). A variant is the Gibbs sampler, where the **form** of the posterior distributions is assumed known (e.g. multivariate Gaussian) but the details of the posterior marginal distributions for each parameter and any covariance structure are not known beforehand.

The potential user needs to be aware that the implementation of such schemes needs some expertise. It is possible to implement them in such a way that the sampling of the posterior is biased (or "unbalanced") in important ways (see Renard *et al.*, 2009; Vrugt *et al.*, 2008a). Improvements to MC^2 sampling strategies for hydrological applications are continuing to be made (Kuczera *et al.*, 2010b; Vrugt *et al.*, 2008a, 2009; Schoups and Vrugt, 2010). MC^2 sampling has been most often used in the context of defining the shape of posterior likelihood surfaces in a Bayesian analysis of uncertainty. They can also be used to improve the efficiency of finding behavioural models within a GLUE analysis (Blasone *et al.*, 2008; McMillan and Clark, 2009). Recent developments in search algorithms for approximate Bayesian computation (ABC) might also be usefully applied to some GLUE applications (Sisson *et al.*, 2007; Beaumont *et al.*, 2009; Toni *et al.*, 2009).

8

Beyond the Primer: Models for Changing Risk

However the endeavour to produce such a theory [of flood hydrology] would be well worthwhile. It would improve our understanding of hydrologic phenomena, improve our decision making in relation to water resources, and improve our standing among geophysicists. To accomplish it, we require a broad background knowledge of our own subject and of cognate disciplines and a real capacity to think both imaginatively and to work hard.

Jim Dooge, 1986

In view of the magnitude and ubiquity of the hydroclimatic change apparently now under way, however, we assert that stationarity is dead and should no longer serve as a central, default assumption in water-resource risk assessment and planning. Finding a suitable successor is crucial for human adaptation to changing climate.

Chris Milly *et al.* 2008

8.1 The Role of Rainfall–Runoff Models in Managing Future Risk

Catchment systems are not stationary systems; both their characteristics and the inputs which drive the hydrology are changing over time (Milly *et al.*, 2008). There are two types of future risk that need to be managed in catchment hydrology where rainfall–runoff models can play a useful role. The first is the short-term forecasting problem of whether an important flood discharge with the potential to pose a threat to life or property will occur. The second is the longer term seasonal or decadal prediction problem of whether changes in catchment characteristics or climate might pose a threat to water resources or flood and drought frequencies. Both of these problems are dependent on the inputs from weather and climate prediction models, which are associated with significant uncertainties in their predictions, including epistemic uncertainties. Managing these risks therefore requires consideration of a cascade of uncertainties through multiple model components and seeking ways to constrain the resulting uncertainty using forms of data assimilation wherever possible (e.g. Pappenberger *et al.*, 2005b).

Rainfall–Runoff Modelling: The Primer, Second Edition. Keith Beven.

Predicting the longer term hydrological effects of change of both climate and land use is certainly a currently fashionable indoor sport. Concerns about the impacts of global warming, deforestation and other changes have been enhanced by public and political awareness of the greenhouse effect, the ozone hole over Antarctica, the El Niño effect, satellite pictures of burned and burning forests in Amazonia and the increasing demands of a growing global population on the world's freshwater resources. Hydrological modelling contributes only a small part of the global atmosphere–ocean–land modelling systems that are being used to predict changes into the next century. It is not, however, an insignificant part. Global circulation models (GCMs) are known to be sensitive to how the hydrology of the land surface is represented. There is also no doubt that both global and local changes will have feedback effects on hydrological processes in the next century, but the problem is how best to predict whether those impacts will be significant when hydrological systems are subject to so much natural variability.

It follows that, as with models of present-day hydrological responses, we should be qualifying our predictions by estimating uncertainties of the impacts of change or, put in another way, the risk of seeing a certain degree of impact on flood peaks, minimum flows or the usable water resource in the future. Translating uncertainty into a future risk can provide a valuable contribution to the decision-making process. Risk is the more acceptable face of uncertainty. The decision maker is then faced, rather than a prediction surrounded by a fuzzy cloud of uncertainty, with the more manageable problem of assessing an acceptable risk. The information provided by the simulations is essentially the same (see Section 7.15).

8.2 Short-Term Future Risk: Flood Forecasting

Managing risk in the short term is one of the most important applications of rainfall–runoff modelling, particularly for flood forecasting which requires decisions to be made as to whether flood warnings should be issued on the basis of the data coming in from raingauges, radar rainfall images, stream gauges and the model predictions as the event happens in "real time". This is a risk management problem because of the uncertainties inherent in the modelling process. It is a case where taking account of the uncertainties might make a difference to the decision that is made (Krzysztofowicz, 2001a; Todini, 2004).

There are a number of simple principles for real-time operational forecasting that can be expressed as follows:

- The event of greatest interest is the next event when a warning might (or might not) need to be issued.
- The next event is likely to be different from all previous events in some of the details of rainfall pattern, volumes and observation uncertainty; rainfall forecast uncertainty; antecedent condition and runoff generation uncertainty; and rating curve uncertainty where discharges for overbank flows are of interest.
- Allowing for uncertainty means being right more often in terms of bracketing when warning thresholds are crossed. It also gives a more realistic representation of forecasting capability in communicating with professional partners.
- Estimating the uncertainty associated with any forecast should not be the end point of the analysis. There are still issues about communicating the assumptions of the analysis to those who have to act on predictions, of managing uncertainty by post-event analysis, and of trying to reduce the uncertainty by more effective measurements (and longer term improvements in the underlying science).

The requirement is for the forecasts and warnings for an area at risk of flooding to be made as accurately as possible and as far ahead as possible or with the greatest possible *lead time*. We can use a model for these predictions that has been calibrated on historical data sets but, because every event has its own peculiarities, we should expect that during a flood event the model predictions of river stage or discharge will start to deviate from those values being received online at a flood forecasting office from the telemetry

outstations. Thus, in any flood forecasting situation it is useful to have data assimilation methods that both allow for *real-time updating* of the forecasts as the event proceeds and constrain the uncertainty in the forecasts. This is particularly the case where the sources of uncertainty are not simply random but can be the result of lack of knowledge. Krzysztofowicz (2002), for example, suggests that

> operational uncertainty is caused by erroneous or missing data, human processing errors, unpredictable interventions (e.g. changes in reservoir releases not communicated by a dam operator to the forecaster), unpredictable obstacles within a river channel (e.g. ice jams), and the like.

To these we might add the potential for unexpected breaches of flood defences during an event. In forecasting (rather than simulation), these sources of epistemic uncertainty can be at least partially compensated by the use of adaptive modelling methods.

Forecasting methods based on transfer function modelling are ideally suited to these requirements and have been quite widely implemented (see, for example, Sempere Torres *et al.*, 1992; Cluckie, 1993; Moore *et al.*, 1990; Lees *et al.*, 1994; Young, 2002; Romanowicz *et al.*, 2006, 2008). An example application using the DBM methodology is provided by Leedal *et al.* (2008), who describe the implementation of a flood forecasting system for the town of Carlisle on the River Eden in Cumbria (see Section 8.5).

8.3 Data Requirements for Flood Forecasting

One of the most important ways of mitigating the costs of flood damage is the provision of adequate warnings, allowing people to act to protect their property and themselves. The responsibility for flood warnings varies between countries. In most cases, the system is based on local flood warning offices that become operational as soon as a potential flood-forming rainfall is forecast. The offices use rainfall–runoff modelling in real time to predict the likely flood discharges and stage in different areas as a basis for decisions about whether to issue flood warnings.

This process is very much easier in large catchments where the build up of a flood and the transmission of the flood wave downstream can take days or even weeks. Recent extreme cases include the Tennessee floods in 2010; the Pakistan floods in 2010; and the Toowoomba and Brisbane, Australia floods of 2011. In small catchments, with short times to peak, real-time flood forecasting, is much more difficult. The most extreme discharges in such catchments tend to occur as a result of localised convective rainfalls or high intensity cells within larger synoptic weather systems. Even if there are telemetering raingauges or a rainfall radar monitoring the area, the response times may be too short to issue warnings in real time. Events where the required lead time for issuing a warning is less than the catchment response time are called "flash floods". Recent examples include the 2009 floods in Cumbria, England, and in Arkansas, Singapore, Ladakh and the Var, France, in 2010. The best option for such flash floods is generally to issue warnings on the basis of forecasts of extreme rainfalls (e.g. Hapuarachchi *et al.*, 2011), but such warnings tend to be very general since the precipitation forecasts of weather models are not sufficiently accurate in either time or space (Pappenberger *et al.*, 2008b). The potential for a flood-producing rainfall might be recognised but it may be very difficult to specify exactly where. A good example of this occurs in southern France and northern Italy where, particularly in September, weather systems moving inland from the Mediterranean and being subjected to orographic rise can give rise to very intense rainfalls and floods. There is at least a minor flood somewhere in this broad region virtually every year. Well-known examples are the 1988 flood in Nice and the 1992 flood on the 560 km^2 Ouvèze catchment at Vaison-la-Romaine in France.

In the case of Vaison-la-Romaine, the French Bureau de Météorologie (now Météo-France) issued a warning of the potential for flash flooding for the region. It is difficult in such a case for individual

communities to react, since the rain may not fall on the catchment that would affect them. In Vaison, 179 mm were recorded in 24 hours, with higher amounts elsewhere on the catchment and intensities of up to 200 mm/hr during six-minute periods. The time between the start of the rainfall and the peak of the flood was just 3.5 hours. The estimated peak discharge was 600–1100 m^3/s, with a peak stage in the town of Vaison of 21 m. The ancient Roman bridge in the centre of Vaison, was overtopped but survived the flood, although there is impressive video footage of floating caravans being crushed beneath it. The majority of the fatalities in this event were tourists at a campsite situated in the valley bottom upstream of Vaison. There was much discussion after this event about whether development on the flood plain upstream of the town had made the depth of flooding and the impact of the flood much worse (Arnaud-Fassetta *et al.*, 1993). What is clear is that many lives could have been saved if an accurate flood warning had been possible.

To make adequate warnings requires knowledge of the rainfalls as they occur or, even better, accurate forecasts of potential rainfall intensities ahead of time, which would allow an increase in the forecast lead time that might be important for small catchments and flash floods, such as that at Vaison-la-Romaine. The advantage of weather radar in these situations is that the radar will often pick up the most intense cells of rainfall in the weather system (e.g. Smith *et al.*, 1996). Such cells may be smaller than the spacing between telemetering raingauges and might therefore be missed by ground-based systems. The problem with radar, as an input to predicting the resulting flood discharges, is that the relationship between the radar signal and the rainfall intensity may not always give an accurate estimate of the absolute intensity (see Section 3.1), particularly when there are attenuation effects due to heavy rainfall close to the radar masking the signal from further afield. Thus, for both radar and raingauge systems, it may be possible to recognise that intense rainfall is occurring over a catchment area but not exactly how intense.

Forecasting rainfalls generally now involves a combination of projecting radar rainfall images into the future and ensemble weather predictions of future precipitation (for example the UK Met Office NIMROD radar and MOGREPS forecast products). Projection of radar images is easier with newer Doppler radar systems that also provide information about local wind speed within the image but the degradation of the radar projections is usually quite rapid because of the way in which individual rain cells grow and decay in complex ways. This has also led to attempts to increase lead times using radar by relating the development of the current pattern of rainfalls to historical analogues (e.g. Obled *et al.*, 2002). The MeteoSwiss Nowcasting Orographic Rainfall in the Alps (NORA) system is based on this approach (Panziera and Germann, 2010).

A number of weather forecasting agencies are also now producing ensemble numerical weather predictions of precipitation that are being used for flood forecasting (see the recent review by Cloke and Pappenberger, 2009). The European Centre for Medium Range Weather Forecasts (ECMWF), for example, provides a control run and an ensemble of 50 "stochastic physics" runs up to 10 days ahead at a grid scale of 40 km. The ensemble is re-initialised using a 4D variational data assimilation methodology and re-run every day. The ECMWF have also run reanalyses of past periods that can be used in "hindcasting" studies. A past reanalysis, ERA40, covers the period from 1957 but the most recent reanalysis, ERA-Interim, which uses an updated forecasting model, started in 1989.

The ECMWF forecasts are used in the European Flood Alert System (EFAS) that is run at the European Community Joint Research Centre in Ispra, Italy (see http://floods.jrc.ec.europa.eu/efas-flood-forecasts). This system is run operationally to provide flood alerts for larger basins in Europe with lead times from three to 10 days. The ECMWF flood rainfalls are processed through a version of the LISFLOOD rainfall–runoff model to provide forecasts of river flow on a 5 km grid (see Thielen *et al.*, 2009; Bartholmes *et al.*, 2009). Although some calibration of the LISFLOOD model has been carried out, the EFAS system recognises that there may be uncertainty in predicting runoff and uses a system of alerts based on the current situation at any grid square relative to the ERA40-reanalysis-driven LISFLOOD simulations. This gives an estimate of the number of ensemble members that suggest extreme flows in a river relative to the reanalysis period. The US National Weather Service is making more direct operational use of ensemble

precipitation forecasts in driving hydrological forecasting models (e.g. Wu *et al.* 2011). It is an on-going research question as to how to best make use of such forecasts when they are known to have deficiencies, especially in mountain areas and in the representation of local convective rainfalls.

At least for cases where the catchment response time is greater than the lead time for decision making, this makes it important that any rainfall–runoff model used be capable of real-time adaptation to take account of any errors in the forecasts resulting either from errors in the inputs, whether from radar or rainfall, or from error in the model structure. This requires, however, that the flood warning centre also receive information about river stages in real time, at one or more gauging stations in a catchment, so that model forecasts can be compared with observed stages or discharges in real time and the model adapted to produce more accurate forecasts (at least until the gauge is washed away or the telemetering system fails). Some ways of doing this are discussed in Section 8.4.

There is another reason why observed stage information might be useful in flood warning, particularly in larger catchments where the time delays in the channel system are sufficiently long compared with the required lead time for a forecast. In general, two to six hours would be the minimum feasible lead time to allow a warning to be transmitted to the public but longer lead times might be needed for decisions about deploying demountable flood defences, for example. A measurement of the discharge or stage upstream can be used as part of the system for forecasting the stage and discharge and timing of the flood peak further downstream. Such observations can also help constrain the uncertainty in propagating forecasts from the headwaters of a catchment in predicting the risk to flood-prone areas downstream.

In general, flood warnings are issued in relation to the forecast stage of the river at a critical gauging point without modelling the detailed pattern of inundation upstream of that point. In many situations this may be adequate, since if flooding is predicted to occur somewhere in the flood plain, then a general warning can be issued. In large rivers, however, such as the Mississippi, the progress of the flood wave downstream may be very much controlled by the pattern of inundation during the flood, including the effects of dyke failures which are inherently difficult to predict ahead of time. Thus it may be necessary to use a hydraulic routing model in forecasting the expected depths downstream, continually revising the calculations as conditions change (although transfer function methods can also be used for this purpose for specific sites; see, for example, the work of Beven *et al.* (2008b), which includes the use of data assimilation where measured levels are available, and Section 8.5). The use of a hydraulic model adds the requirement of knowing the channel and flood plain topography, together with parameters such as effective resistance coefficients. Topography is normally provided as a series of cross-sectional profiles surveyed across the flood plain and channel at different sites, but the increasing use of two-dimensional, depth-integrated models will lead to the use of topographic data in the form of detailed digital elevation maps of the flood plain. Channel form, can of course change during a flood event due to erosion and deposition. Models of sediment transport in rivers have not advanced to the stage where they can be used operationally and most current hydraulic flood routing models use a "fixed bed" assumption.

8.4 Rainfall–Runoff Modelling for Flood Forecasting

Any rainfall–runoff model that has been calibrated for a particular catchment can be used in the prediction of flood discharges. The US National Weather Service River Forecast System (Burnash, 1995), for example, is a development of the Sacramento model, a form of lumped explicit soil moisture accounting model with many parameters to be calibrated (see, for example, Sorooshian *et al.*, 1992; and Gupta *et al.*, 1999). Using methods such as those discussed in Chapter 7, the predictions may also be associated with an estimate of uncertainty in the predictions. Qualifying the estimates in this way may be important. Experience suggests that uncertainty in both measurements and predictions of flood peaks increases with peak magnitudes. In addition, even if a model has been calibrated for a certain range of discharges, uncertainty is bound to increase as predictions are made outside this calibration range for extreme events.

Thus, it may not be possible to predict ahead of time whether a flood stage will definitely be exceeded in a event; it may, however, be possible to assess the risk that the flood stage will exceeded by consideration of the distribution of (uncertain) predictions.

As noted earlier, the propensity for error in the predictions during extreme events also suggests that it would be advantageous to use an adaptive modelling strategy, so that if a comparison of observed and predicted discharges reveals that the model predictions are in error, then a strategy for adjusting the model predictions can be implemented. This is clearly only possible where discharge or river stage measurements can be made available in real time. Adaptation is also more easily implemented for simpler models.

An early comparison of real-time forecasting methods, including adaptive schemes, was carried out by the World Meteorological Organisation (WMO, 1975) and a review of approaches has been given by Moore (1999). Ensemble forecasting systems have been reviewed by Cloke and Pappenberger (2009) including experience from the Hydrological Ensemble Prediction Experiment (HEPEX) (Schaake *et al.*, 2007).

It is also possible to take an approach that makes no attempt at all to model runoff generation during flood forecasting. As noted in Section 4.6, neural net models and support vector machine models have recently been popular as a means of estimating N-step ahead flood discharges, using inputs that include rainfalls and previous values of discharge or water levels and a training set of historical events (see Figure 4.13). In discussing neural network models, however, it has already been noted that such models are grossly over-parameterised and there has to be some concern then that methods based only on data analysis might not be accurate in predicting events more extreme than those included in the training set.

We consider some simple strategies for real-time forecasting here. The first is an adaptive deterministic method due to Lambert (1972) that is very difficult to beat in forecasting the response of small catchments. This Input–Storage–Output (ISO) model has been used in a number of UK flood forecasting schemes, particularly in the River Dee catchment in north Wales. The second is an adaptive form of the transfer function models considered in Section 4.3.2. Adaptive transfer functions can be used for both rainfall–runoff and upstream discharge (or stage) to downstream discharge (or stage), depending on what data are available. The example application presented in the case study of Section 8.5 is for an operational forecasting model for the town of Carlisle in Cumbria, that uses both rainfall–flow and discharge–discharge transfer functions. Adaptation of such models can be implemented in a number of different ways. In the Carlisle model, a simple adaptive gain parameter is used, i.e. the transfer function is scaled up or down in real time without changing its form. This simple approach has proven very effective in this and other applications. We also briefly consider the Bayesian forecasting system and quantile regressions as other ways of adding uncertainty to the forecasts of deterministic rainfall–runoff models.

These types of model can be made part of a larger flood forecasting system that includes flood routing components. For the River Dee catchment, for example, ISO models were developed for all the gauged subcatchments and linked to a flood routing model. The Forecasting and Early Warning System (FEWS) developed by Deltares in the Netherlands (e.g. Werner *et al.*, 2004) provides a unified framework for networking different model components in this way that has been implemented in a number of other countries including the UK Environment Agency National Flood Forecasting System. The FEWS software is freely available under licence (see Appendix A).

8.4.1 The Lambert ISO Model

The idea behind Alan Lambert's ISO model is wonderfully simple. It is based on the development of a master recession curve for a catchment or subcatchment where a period of discharge measurements are available, by piecing together partial recession curves from individual events. In general, the shape of such a recession curve may not be easily represented by a simple mathematical function but Lambert

suggested using simple linear and logarithmic functions to represent different parts of the discharge range (Lambert, 1969, 1972). Then, at each time step during an event, average rainfall (less an estimate of evapotranspiration if necessary) is added to the relative catchment storage (the absolute storage does not need to be estimated) and, using the representation of the storage–discharge function, an incremental change in discharge is easily predicted. The discharge is subtracted from the storage, and the model is ready for the next time step. This is about as simple a rainfall–runoff model as is possible. Calibration of the model is simply a matter of deriving the master storage–discharge curve for as wide a range of discharges as possible. There are effectively no other parameters (see the similar more recent use of this concept in Section 4.2).

In continuous simulation, such a model would not be very accurate since we know that there is not a simple relationship between storage and discharge. This is why we use more complex models that attempt to reflect the complexities of the rainfall–runoff processes more realistically. In real-time forecasting, however, the ISO model can be used in a way that reduces the impact of the errors inherent in using such a simple model. Its first advantage is in initialising the model at the start of an event. In general, the best index of the antecedent wetness of a catchment area is the discharge at the start of an event. The ISO model is easily initialised if the discharge is known at the first time step, since this can be used to infer the initial relative storage. The second advantage is that, as soon as errors are detected between the observed and predicted discharges, the model can be reinitialised using the current measured discharge. This procedure can be implemented at every time step as soon as the measured discharges are received and used to update the predictions into the future.

Thus this is also the very simplest possible adaptive modelling scheme. No complex mathematics is required for the adaptation, it is easily understood, easily calibrated and easily implemented. It can be a very effective real-time forecasting scheme but clearly has some limitations. In particular, extrapolation outside the range of measured recession curves is uncertain. In the model, the relationships used for the upper and lower sections of the master recession curves are simply assumed to continue for more extreme conditions. In addition, adaptive use of the model to change the current relative storage invalidates any overall water balance for the model, but this is not important in real-time forecasting if it leads to improved predictions.

In small catchments, this is effectively the model to beat for real-time forecasting!! More complex models that are less easy to make adaptive may not necessarily produce better real-time flood forecasts.

8.4.2 Adaptive Transfer Function Models for Real-Time Forecasting

It is worth noting that the simplest linear ISO model element is effectively a first-order transfer function model, equivalent to Equation (B4.3.1). The difference is that the inputs used with the ISO model are the rainfalls, whereas the transfer function models of Section 4.3.2 use the rainfalls filtered in some way to produce an effective rainfall. For a filter that is directly and only dependent on the current discharge, such as that used by Young and Beven (1994), simple ISO model type updating can still be used directly, but this is not possible where effective rainfall components introduce additional storage elements. There is, however, a simple way of making such models adaptive, as noted above, by using an adaptive gain or multiplier. The best initial estimate of the gain parameter is normally 1.0 but it is then allowed to vary as the event progresses to correct for any differences detected between the forecasts and the observations supplied to the flood warning system. If an underprediction is detected, the gain can be increased for the next time step; if overprediction, the gain can be reduced. The changes in the gain are filtered so that the changes from time step to time step stay relatively smooth. The adaptive gain approach is a simple way of compensating for any errors in the data or transfer function model structure that might affect the accuracy of the forecasts. It will generally lead to greatly improved forecasts. One example of a simple but successful adaptive algorithm that can, in fact, be applied to any deterministic model output is given in Box 8.1.

There is one important limitation of transfer function models in flood forecasting applications. Transfer functions are designed to make predictions of output variables for which historical data are available for calibration. They are an empirical modelling strategy in this sense. Thus although they can provide predictions, easily updated in real time, of river stage at specific sites in the catchment they cannot predict the extent of flooding in the catchment, except in so far as this is possible from knowledge of the flood stage at the measurement sites. Transfer models can also be used to emulate the outputs from hydraulic flood routing models at sites that are not gauged. They can do so very accurately, including taking account of the hysteresis in the flood wave celerity (Beven *et al.*, 2008b). However, that does not mean that the predictions of the original hydraulic model will necessarily provide accurate forecasts of inundation everywhere in an area at risk of flooding (Pappenberger *et al.*, 2007a) and clearly it is not possible to update such local forecasts if no local level information is available.

8.4.3 The Bayesian Forecasting System (BFS) and Quantile Regression

A significant research programme to develop a flood forecasting system that took account of the uncertainties in the forecasting processes has been led by Roman Krzysztofowicz (1999, 2001a, 2001b, 2002; Krzysztofowicz and Kelly, 2000; Herr and Krzysztofowicz, 2010). The Bayesian Forecasting System (BFS) is based on combining a representation of the rainfall input uncertainties (Kelly and Krzysztofowicz, 2000) with a representation of the hydrological modelling uncertainties (Krzysztofowicz and Herr, 2001) within a Bayesian framework. Thus a good estimate of the uncertainty in the rainfalls can be propagated through the hydrological model to provide a posterior distribution for the forecasts (Krzysztofowicz, 2001b). The uncertainty in the hydrological model is derived from simulations of past data for which the rainfalls are asssumed known. There is an expectation, however, that the errors will be complex in nature, and the BFS uses mixed distributions for under and overpredictions that are trasformed into a multivariate gaussian space using a normal quantile transform, otherwise known as a metagaussian transform (Kelly and Krzysztofowicz, 1997). As new data become available, the uncertainty associated with the predictions can be updated before calculating new forecast uncertainties for river levels. Another application of the BFS, to forecasting on the River Rhine, is reported by Reggiani and Weerts (2008).

In propagating the forecast uncertainties to longer and longer lead times, the BFS makes use of prior distributions of forecast errors based on historical performance. This technique can, in fact, be used to provide a simple off-line estimate of uncertainty in the forecasts for any deterministic forecasting model. In effect, past performance is summarised in terms of the residual quantiles for different forecast lead times over a number of past events. Regressions are then calculated, for each required quantile, that smooth the relationship between quantile and lead time (see, for example, Weerts *et al.*, 2011). The expectation is that the occurrences of residuals for new events in the future will (more or less) behave in a similar way to the quantiles for past events. The method is simple and cheap to update after each new event. It relies on new events being statistically similar to past events in their residual behaviour (see also the error correction method of Montanari and Grossi, 2008). This is not necessarily the case but at least gives a guide to the magnitude of potential errors. Where it is possible to use data assimilation during an event, however, there is some advantage in dealing with the particularities of that event.

There are other dynamic error correction and Bayesian data assimilation schemes (such as the Ensemble Kalman Filter and forms of particle filter) that can be used to make more complex models adaptive (see Beven (2009) for more details). However, in complex models, with many different interacting variables or components that could be adjusted, it might be difficult to decide what to adjust, especially when the information content of a new observation residual at a gauging site is certainly going to be limited (and potentially reduced by correlation with the same observation residual at the previous time step).

8.5 Case Study: Flood Forecasting in the River Eden Catchment, Cumbria, England

A flood forecasting system for the River Eden catchment upstream of the town of Carlisle has been described by Leedal *et al.* (2008). In January 2005, Carlisle suffered a significant flood (with an annual exceedance probability of less than 0.01), which caused three deaths and damage of the order of £250 million. The inundation of the town was caused by a combination of both pluvial and fluvial flooding in this event but the flood stage was higher than anything previously observed (Mayes *et al.*, 2006). It was significantly underpredicted by the Environment Agency forecasting models at the time, primarily because the inputs to the upland parts of catchment appeared to be significantly underestimated by the available raingauge data.

The models used by Leedal *et al.* (2008) were similar to the type of data-based mechanistic (DBM) transfer function models described in Chapter 4, but with some additional features specific to the flood forecasting situation. Transfer functions were fitted both for rainfall–river flow modelling and for flow routing in the main channel, the type of cascade of DBM models first used by Lees *et al.* (1994); see also the case study on flood forecasting for the town of Dumfries in the first edition of this book and Box 8.1). In the case of the Eden, a simplifed representation of the catchment was implemented that used two headwater basin models and one flow routing model. Thus only six raingauges and three river gauging sites were used of those available (Figure 8.1).

Recent applications of the DBM approach in the flood forecasting context have made use of rainfall to water level models in headwater basins and upstream water level to downstream water level models for flood wave routing. This clearly means that the models being used are not constrained by mass balance, as would be the case if discharges were used. However, given that in many major flood events, the input rainfall volumes are not precisely known and the rating curve to convert water levels to discharges at

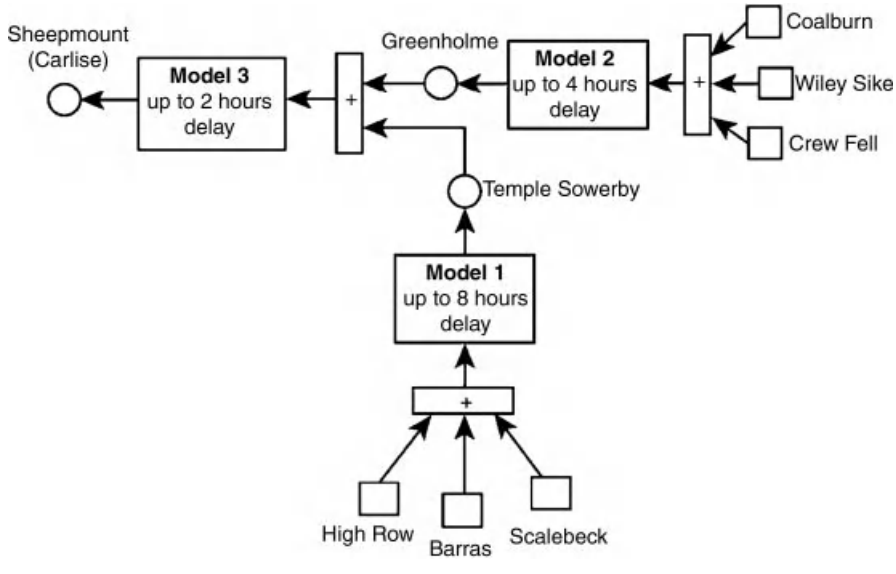


Figure 8.1 *Cascade of flood forecasting system components in the River Eden, Cumbria, UK: The model assimilates data at each gauging station site (circles) and generates forecasts for Sheepmount with a six-hour lead time (after Leedal et al., 2008, with kind permission of the CRC Press/Balkema).*

flood stages are highly uncertain, imposing mass balance on a forecasting model might actually be a disadvantage. This is reinforced by many other studies that suggest that the major source of uncertainty in flood forecasting is knowledge of the inputs (e.g. Krzysztofowicz, 2002).

There are other advantages to using water level information directly in forecasting. One is that level is actually measured directly so that measurement errors are, in general, small. No additional uncertainty is introduced by the application of the rating curve. Rating curve uncertainties can be significant for flood stages (the estimate of the peak discharge for the Carlisle flood was revised from 900 m^3s^{-1}, obtained by extrapolating the rating curve prior to the event, to 1500 m^3s^{-1} after a post-event re-evaluation). In addition, for flood warning purposes, the predicted variable that is required is very often the water level. Decisions about warnings are usually made on the basis of forecast levels (although velocities derived from discharges might be of secondary interest in assessing risk to life). It is also worth remembering that the forecasting problem is different from the simulation problem. The requirement for forecasting is to have an accurate estimate of water level with minimal uncertainty at the lead time of interest. It does not really matter how that is achieved and because of all the potential for operational epistemic errors during an event it might be more advantageous to use a simple model that is easily made adaptive than a complex model that aims to simulate the runoff processes more correctly.

Both the headwater rainfall to level models and the level to level flow routing models should be nonlinear. The nonlinearities were identified using the SDP approach described in Box 4.3 and represented using a Takagi–Sugeno first-order fuzzy inference system (T-S FIS). The resulting nonlinearities for the three components of the forecasting cascade are shown in Figure 8.2. Note how that for the flow routing (Figure 8.2c) is quite different from the headwater rainfall to level models. The models are completed by

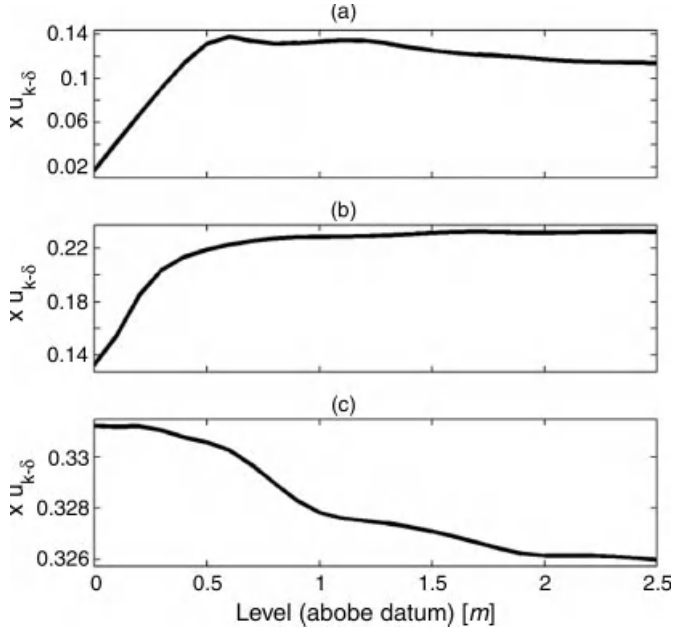


Figure 8.2 *Identified input nonlinearities for each of the forecasting system components in the River Eden, Cumbria, UK: (a) input nonlinearities for Model 1 in Figure 8.1; (b) input nonlinearities for Model 2 in Figure 8.1; (c) input nonlinearities for Model 3 in Figure 8.1 (after Leedal et al., 2008, with kind permission of the CRC Press/Balkema).*

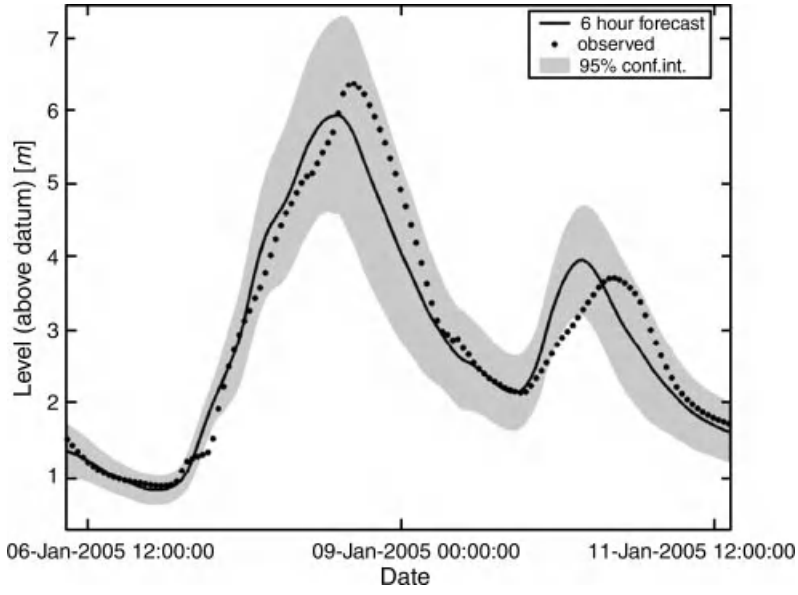


Figure 8.3 *Adaptive six-hour ahead forecasts for Sheepmount gauging station, Carlisle, with 5% and 95% uncertainty estimates for the January 2005 flood event (after Leedal et al., 2008, with kind permission of the CRC Press/Balkema).*

linear transfer functions conditional on the identified nonlinearities (see Box 4.1). The models are made adaptive using the simple form of modified Kalman filter explained in Box 8.1.

Calibration of the model was carried out for the period March to December 2004, when the water level at the Sheepmount gauging station in Carlisle did not exceed 4 m above base level. The model parameters were then fixed. Figure 8.3 shows the predictions at Sheepmount for the January 2005 flood when the water level reached 7 m above base level. Application of the model cascaded without data assimilation underpredicts the peak (this type of forecast model is not immune to underestimation of the inputs noted above when used without adaptation). With data assimilation both the rising limb and the peak are predicted reasonably well (a Nash–Sutcliffe efficiency of 0.93 in the period 4th to 10th January) and the uncertainty bounds reflect the accuracy of the prediction.

8.6 Rainfall–Runoff Modelling for Flood Frequency Estimation

Part of the problem of managing changing risk in catchments is to have an adequate estimation of the expected frequency of extremes (either floods or droughts). Where data are available at long-term gauging stations, it is normal to fit a statistical distribution to either the annual maximum peaks or the peaks over a chosen threshold as a way to define the frequencies of extremes. This statistical approach has a number of limitations. One is that we do not know what the correct distribution should be. Log Normal, Wakeby, Generalised Extreme Value, Log Pearson Type III and a number of other distributions have been used in the past (Cunnane, 1989; Hosking and Wallis, 1997; IH, 1999). Several different distributions might give acceptable fits to the data available, all producing different frequency estimates in extrapolation. Even in developed countries, such data are available at only a limited number of sites, and for limited periods, so that the calibration may not be robust and the extrapolation to more extreme,

lower exceedance probability events rather uncertain. Fitting any statistical distribution also does not take any account of how the runoff generation processes might change under more extreme conditions (Wood *et al.*, 1990; Struthers and Sivapalan, 2007). Rainfall records, however, tend to be longer and more widespread. Thus, the possibility of using rainfall–runoff models to estimate frequencies of flow extremes is attractive.

In a seminal paper on the dynamics of flood frequency, Pete Eagleson (1972) outlined a method of "derived distributions" for the calculation of flood frequency. Eagleson's idea was to look at the flood frequency problem as a transformation of the probability distribution of rainstorms into a probability distribution of flood peaks using a rainfall–runoff model. The advantage of this is that rainfall data is available for many more sites and generally much longer periods than data for stream discharges. The disadvantages are that rainstorms vary in intensity, duration and storm profile so that, if it is necessary to model a sequence of rainstorms, a more complicated stochastic model is required; the results will be very dependent on the type of rainfall–runoff model and the set of parameter values used. As we have already seen, this is an important source of uncertainty in model predictions. Observed rainfalls had been used previously with conceptual rainfall–runoff models to simulate sequences of flood peaks (Fleming and Franz, 1971) and were shown to be competitive with statistical methods. Eagleson's initial proposal, however, was not so much concerned with practical prediction for individual sites as to provide a structure for thinking about the problem.

In the method of derived distributions, the probability density function of flood magnitudes is derived from the density functions for climate and catchment variables through a model of catchment response to a rainstorm. Eagleson managed to come up with a set of simplifying assumptions that allowed (in a paper of only 137 equations) the analytical integration of the responses over all possible rainstorms to derive the distribution of flood peak magnitudes. The method has been extended by later workers to include different stochastic characterisations of the rainstorms and different runoff-producing mechanisms at the expense of losing the mathematical tractability that allows analytical solutions. This is not perhaps so important given current computer resources when a simple rainfall–runoff model can easily be driven by stochastic inputs for periods of hundreds or thousands of years to derive a distribution of predicted flood magnitudes numerically. However, Eagleson's (1972) attempt is of interest in that it provided the framework that later models have followed of starting with a stochastic rainstorm model which is used to drive, in either a probabilistic or numerical way, a runoff generation model, which might then be linked to a routing model to predict the flood peak at a desired location.

8.6.1 Generating Stochastic Rainstorms

One possibility for the inputs to a rainfall–runoff model for estimating flood frequency is simply to use a long record of observed rainfalls (Calver and Lamb, 1996; Lamb, 1999). However, many studies, including the Eagleson (1972) paper, have chosen to use a stochastic rainfall generator. Eagleson's rainstorm model was based on independent distributions for storm duration and mean storm intensity at a point. Both probability distributions were assumed to be of exponential form. In his study, Eagleson assumes that the storms are independent so that in calculating flood frequencies he only needs to know the average number of storms per year. Other later studies have taken account of the variation in antecedent conditions prior to different events and this requires the introduction of a third distribution of inter-arrival times between storms. To allow the analytical calculations, Eagleson's stochastic rainstorms are of constant intensity over a certain duration but, in using rainfall statistics from individual stations, he did take account of the difference between point rainfall estimates and the reduction in mean intensity to be expected as catchment area increases using an empirical function from the US Weather Bureau.

Later studies that have used direct numerical simulation, either on a storm by storm basis or by continuous simulation, have not been so constrained in their rainstorm model component. Beven (1987)

added a rainstorm profile component to the Eagleson model based on the statistics of observed cumulative rainstorm profiles normalised to a unit duration and unit rainstorm volume. This was later adapted by Cameron *et al.* (1999) to use a database of observed normalised profiles from 10 000 real storms, classified by different duration classes. When a new storm profile is required, a normalised profile is chosen randomly from the database. An interesting variation on this technique of using normalised profiles that will produce rainfall intensities with fractal characteristics has been suggested by Puente (1997). Another way of using actual storm data in flood frequency analysis is to use the random transposition of real storms recorded in a region over a catchment of interest (Franchini *et al.*, 1996). Finally a number of rainstorm models have been proposed based on pulse or cell models, where the time series of generated rainfalls is based on the superposition of pulses of randomly chosen intensity and duration by analogy with the growth and decay of cells in real rainfall systems. Thus rainstorms are not generated directly but arise as periods of superimposed cells separated by dry periods. The Bartlett–Lewis (Onof *et al.*, 1996) and Neyman–Scott (Cowpertwait *et al.*, 1996a, 1996b) models are of this type while Blazkova and Beven (1997, 2004, 2009) have used mixed distributions of high intensity and low intensity events. Such models require a significant number of parameters, especially when the model is fitted to individual months of the year. This type of model has been included in the UKCP09 weather generator for future climate scenarios (see Section 8.9).

One interesting feature of driving flood frequency models by stochastic rainstorm generators is that the period of simulation is not restricted to the lengths of historical records. In fact, when checking against actual flood frequency data for limited observation periods, the length of the record should be taken into account since there will be an important realisation effect for the extremes in these limited periods (Beven, 1987; Blazkova and Beven, 1997, 2004). The historical data is, in fact, only one realisation of all the possible sequences of events and resulting flood extremes that could have occurred for a period of the length of the observations. That realisation might be subject to significant variability in the occurrences of floods (see the example of Blazkova and Beven (2009) in Section 8.5), particularly where climate is subject to el Niño oscillations (Kiem *et al.*, 2003). Thus, finding a model that is consistent with the observed frequencies for such a period will depend not only on the model parameters but also on the particular realisation of inputs.

Actually this is also an issue for models driven by observed rainfalls. Although potentially a model might be able to represent any changes in the hydrological response of a catchment under more extreme conditions, there is no guarantee that a model calibrated by a prior parameter estimation or by calibration against a period of observed discharges will, in fact, produce accurate simulations for extreme flood peaks (Fontaine, 1995). Lamb (1999), for example, has shown how the PDM model, calibrated over a two-year continuous simulation period, can reproduce the magnitudes of the largest observed flood peaks when driven by a longer period of observed rainfall data, but not necessarily from the same events that produce the recorded peaks.

Running much longer periods of stochastic inputs to a model can also have another interesting effect. Since the distributions that are used in a stochastic rainfall model are often assumed to be standard statistical distributions with infinite tails some very, very large events might be generated. They will occur with low probability, but under the infinite tail assumption there is always a small but finite probability of very large events occurring. Such events might be inconsistent with the potential meteorological conditions at the site (sometimes expressed in terms of a "probable maximum precipitation" (PMP), although this is a value that is always difficult to define for a site). However, such events might also be much higher than anything that has been observed in a region. The infinite tail assumption is not, however, a necessary assumption. Cameron *et al.* (2001) showed how it could be changed within the Barlett–Lewis rainfall generator to reflect the upper envelope of observed rainfalls for given durations in the UK. Verhoest *et al.* (2010) suggest that stochastic rainfall models should be evaluated not only in terms of reproducing point rainfall statistics, but also in terms of reproducing flood statistics allowing for the dynamics of antecedent conditions in the periods between rainstorms.

8.6.2 The Runoff Generation Model Component

Given a series of rainstorms, the next stage is to model how much of that rainfall becomes streamflow. Eagleson did this by estimating effective rainfalls using a simple Φ-index model (see Section 2.2), with a constant value of Φ, to allow an analytical solution to the derived distribution of flood peaks. His model did also recognise that runoff might be generated over only part of the catchment, on a contributing area that would vary both between catchments, because of different catchment characteristics of soils, vegetation and topography, and within a catchment, because of variation in antecedent conditions. He modelled these variations by assuming a probability density function for contributing area of triangular form, which "provides the bias towards small fractions of the catchment areas observed by Betson (1964)" (Eagleson, 1972, p. 885). Random selection of a contributing area for a storm is a way of effectively introducing the antecedent condition control on runoff generation into the predictions. Eagleson then routes the effective rainfall produced using kinematic wave routing for both overland and channel flows (see Section 5.5). The runoff generation component produces, for each rainstorm with $i_o > \Phi$, a constant effective rainfall of duration t_r for that storm. This very simple time distribution allows analytical solutions of the kinematic wave equation to be performed.

More recent studies using direct numerical simulation have not been so constrained and a variety of runoff generation models and routing methods have been used, including the PDM model of Section 6.2 (Lamb, 1999; Lamb and Kay, 2004; Calver *et al.*, 2009) and TOPMODEL of Section 6.3 (Beven, 1987; Blazkova and Beven, 2000, 2004, 2009). Those that are based on event by event simulation, as in the original Eagleson study, require some way of reflecting the effect of varying antecedent conditions on runoff production. Those that use continuous simulation over long periods of time, including the drying of the catchment during interstorm periods, account for antecedent conditions directly. Berthet *et al.* (2009) recently compared event by event and continuous simulation approaches (in the flood forecasting, rather than frequency estimation, context) and suggested that continuous simulation would generally perform better in predicting peak discharges.

Sivapalan *et al.* (1990) produced a scaled flood frequency model based on the TOPMODEL concepts and showed that catchment runoff production could be compared on the basis of eight similarity variables. Their flood frequency curves were derived from storm by storm simulations and showed a transition between saturation excess overland flow dominated flood peaks to infiltration excess overland flow dominated flood peaks for more extreme events. However, they presented only a sensitivity analysis and this similarity theory still has to be tested against real data sets, although a simplified version was shown to have value in differentiating the hydrology of seven small Australian catchments by Robinson and Sivapalan (1995). More on estimating the flood frequency characteristics of ungauged catchments using rainfall–runoff models is to be found in Chapter 10.

8.7 Case Study: Modelling the Flood Frequency Characteristics on the Skalka Catchment, Czech Republic

This study, reported by Blazkova and Beven (2009) is the latest of a number of applications of using continuous simulation to estimate flood frequency for Czech catchments, including for catchments treated as ungauged (Blazkova and Beven, 1997, 2004, 2009). The Skalka catchment (Figure 8.4) which spans the German–Czech border was much larger at 672 km^2 than previous applications, with an elevation range of 460 to 1041 m asl, a number of different gauging stations for subcatchment areas and multiple raingauge sites with a common measurement period of 28 years. The gauging stations had records of between 28 and 67 years and there were four snow measurement sites with records of between 10 and 20 years. This catchment has the potential for floods caused by winter rainfalls, spring snowmelt, and summer thunderstorms.

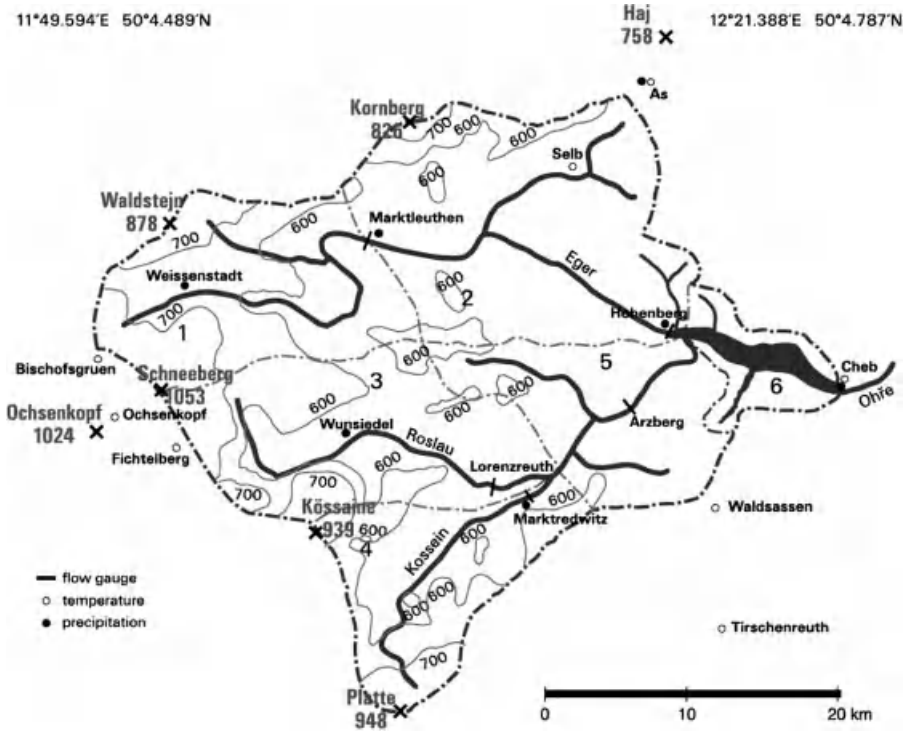


Figure 8.4 *Subcatchments and observation sites in the Skalka catchment in the Czech Republic (672 km^2) (after Blazkova and Beven, 2009, with kind permission of the American Geophysical Union).*

A modified version of TOPMODEL was implemented for the different subcatchments, driven by a weather generator that was fitted to temperature, rainfall and snow accumulation data for sites in the catchment, taking account of the elevation range in the catchment. The stochastic rainfall component of the weather generator included joint distributions for both high intensity and low intensity events at the subcatchment scale, and took account of the observed spatial correlation of rainfall amounts in the subcatchments. Snowmelt calculations were made for 13 different snow accumulation areas spanning the elevation range in the catchment using a degree-day method.

The purpose of the frequency estimation was to evaluate the safety of the dam at Cheb at the outlet of the catchment. A particular feature of the study was that there were discharge measurements at Cheb for 60 years (1887–1959) before the dam was built. This includes 50 years prior to the earliest measurements at any of the other sites, the oldest of which, at Marktleuthen on the Eger subcatchment, started in 1937. The Cheb data were not used in calibrating the model but were retained as a validation check on the frequency predictions.

The modifications to TOPMODEL were made because early runs of the model showed that it could not reproduce the low flow characteristics of some of the subcatchments. A master recession curve was derived from the observed discharges for each subcatchment, using the method of Lamb and Beven (1997), and used to modify the subsurface store in the model. This would be even more important if it was the distribution of extreme low flows that was of interest. The model was then applied within the GLUE framework of Chapter 7, using the limits of acceptability approach in model evaluation described

by Beven (2006a). Each model realisation was compared against 114 different limits of acceptability. The evaluation of discharge predictions included both peak frequency quantiles (up to an annual exceedance probability of 0.1 to be robust in the estimation from relatively short periods of observations) and quantiles of the flow duration curve for each site, taking account of the observational errors in discharges derived from rating curve interpolation and extrapolation using fuzzy regression. Other limits were set for the distribution of maximum winter snow accumulation in each accumulation area, estimated by fuzzy regression from the four observation sites. For the rainfall model, limits of acceptability were defined by putting bounds on the subcatchment hourly and daily rainfall frequency information estimated from the point raingauge sites by bootstrapping from the distributions of estimates of the average subcatchment rainfalls from the point measurements. Excluding one measurement point at a time and reestimating the subcatchment rainfalls allows a range of possible subcatchment rainfalls to be defined directly from the measurements. A final limit of acceptability was imposed for the proportion of floods occurring in the winter season. This was not a strong constraint, but was set because in early runs too many floods were found to be generated in summer.

Within this version of GLUE, all models that survive the multicriteria tests of acceptability are then used in prediction of the frequency characteristics for the catchment, weighted by a likelihood measure that depends on performance of the model within the limits of acceptability, represented as a standardised score (see Figure B7.1.2). For each criterion, the likelihood measure was defined as a trapezoidal fuzzy measure around each observation, reducing to zero at the limits of acceptability (see Figure B7.1.3). The full model had 46 parameters that could be varied and, as noted above, whether a particular model run matches the limits of acceptability derived from the observations might depend on the particular realisation of the inputs. In fact out of 610 000 model runs of 67 years duration, only 39 were found to satisfy all 114 limits. These were run to generate flood frequency curves for the Cheb validation site (see Figure 8.5).

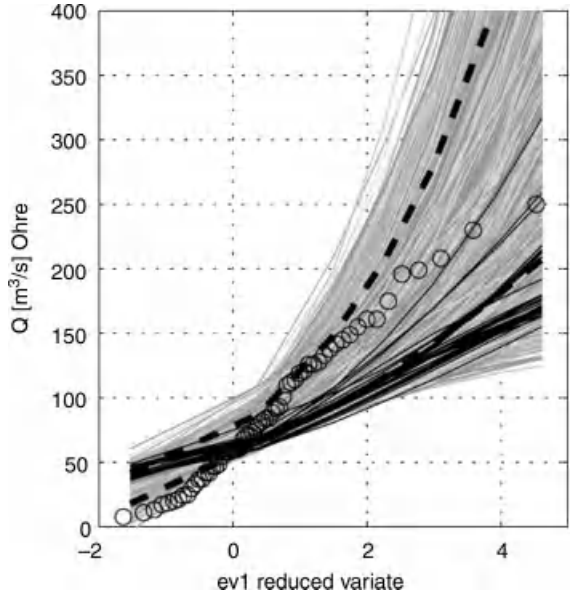


Figure 8.5 *Flood frequency curve at the Cheb gauging station: circles are the observed annual flood peaks (1887–1959); grey lines are the 4192 simulations, each of 10 000 years of hourly data, with scores on all criteria <1.48; dashed lines are the 5% and 95% possibility bounds from the trapezoidal weighting; thin black lines are the behavioral simulations with scores on all criteria <1 for the initial 67-year simulations (after Blazkova and Beven, 2009, with kind permission of the American Geophysical Union).*

Realisations of 10 000 years of hourly data were then used to get good estimates of the peak discharge with an annual exceedance probability of 0.001 (1:1000 year event) for the dam safety evaluation. One interesting feature of this methodology is that it is possible to investigate which of the limits of acceptability causes most model rejections. In this case, it was the underprediction of low-magnitude floods at the Arzberg site.

The best model in satisfying the limits was also run with 10 000 realisations of input data generated for the same period of 67 years. Less than 1% of those runs survived the limits of acceptability. This suggested that the initial limits might be relaxed to avoid rejecting models that might be useful in prediction over longer runs simply because of this *realisation effect*. The grey frequency curves in Figure 8.5 represent the estimates from a larger set of 4192 models selected in this way, all of which have standardised scores of less than $+/-1.48$. Each of these models was run with inputs of 10 000 years of hourly data. The dashed lines in Figure 8.5 represent 5% and 95% possibility limits using an average of the trapezoidal weightings over all evaluation measures.

Figure 8.5 also demonstrates another realisation effect, this time associated with the observations themselves. The fit of the model to flood peak frequency data for the subcatchment sites was better than that demonstrated for the Cheb site in Figure 8.5, which is outside the 5% and 95% possibility limits for small floods. As noted earlier, that site was not used in calibration but also largely represents a different period to the data used in calibration. It would appear as if there has been a change in flood characteristics between the late 19th to early 20th century period and the later period.

8.8 Changing Risk: Catchment Change

All catchments have a history. Some of this history has a long timescale, such as soils developed on till or fluvioglacial deposits from the last ice age, or deep lateritic soils resulting from weathering over long periods. Some changes are the much more recent results of human activities, such as urbanisation, deforestation, reforestation, reservoir or detention pond construction, the effects of wild fires, or the installation of field drainage systems. Both natural and anthropogenic changes in land use have affected catchments in the past and are continuing today. Some changes are documented, others may not be (I know of one experimental catchment in the UK where it was discovered, after the installation of the instrumentation, that the headwater hollows were drained by old field drains built of stone slabs and estimated to be more than 100 years old).

In most gauged catchments, certainly larger gauged catchments, such changes have been on-going during the period of historical records. However, hydrological analyses, for example flood frequency analyses, do not often attempt to take any account of the possible effects of such changes. Why not? Primarily because it may be very difficult to detect such changes in the historical record given the difficulties of closing the water balance and the natural year to year and storm to storm variability in hydrological responses, especially where the changes are only gradual. Even where a distinct change in behaviour is detected it can be difficult to differentiate between different potential causes (Fenicia *et al.* (2009) describe a large-scale example in the Meuse catchment). In a study carried out for the UK government, we examined the hydrological records for a number of catchments which were expected to show the greatest effects of land use change and management on discharges (Beven *et al.*, 2008c). This analysis showed that discharge was generally increasing over the 30-year period studied, but that these changes were matched by increases in rainfalls. This is consistent with other UK work that has indicated an increase in the number of storm peaks above different thresholds over time (e.g. Archer, 2007; Climent-Soler *et al.*, 2009). By classifying individual storm events into classes that depended on storm rainfall volumes and antecedent conditions, some effects on peak timing and volumes were detected in the parameters of DBM models for different periods – but only for the class with small input volumes and dry antecedant conditions, not those that produce floods. The other classes showed a lot of variability but no apparent trends across the period (the classification allows similar hydrological conditions to be

compared regardless of changes in rainfall over time). McIntyre and Marshall (2010) have used a similar approach for the Pontbren catchments in Wales to compare subcatchment characteristics for a shorter data set.

Thus there are difficulties in detecting change in the analysis of historical data. The major problem in predicting the effects of change in the future is that at least some of the model parameters should be expected to change. This is evidently the case if there is a change of land use, but changes may also be induced by a change in inputs as a result of climate change. It has been argued, for example, that an increase in CO_2 concentrations in the atmosphere will cause a decrease in the density of stomata on leaf surfaces because a plant can achieve the necessary exchanges of CO_2 for photosynthesis. The result would be an increase in the effective canopy resistance, leading to smaller evapotranspiration losses as a result of climate change. The argument is plausible and is supported by stomatal density measurements for some plants in the historical collections at Kew Gardens, but it is not clear if this will be a general response or a significant factor in the response of plants to climate change.

It has long been suggested that one of the most important reasons to develop a process-based distributed modelling capability is precisely that it might be easier to estimate such changes (Abbott *et al.*, 1986a; Ewen and Parkin, 1996; Dunn and Ferrier, 1999). It is also rare for the whole of a catchment to change suddenly so that the changes in characteristics may be gradual and often quite local. Thus a spatially distributed model can implement any changes in parameter values in their correct spatial context, if we can be secure in estimating changes in the effective values of the parameters required by such models. Some models include components that attempt to represent the growth of vegetation communities and their interactions with hydrological processes (e.g. RHESSys Band *et al.*, 1993; TOPOG-IRM Dawes *et al.*, 1997; and MACAQUE Watson *et al.*, 1999) but with the consequent need to estimate even more parameters.

The advantages of process-based models in this context are still widely supported and for good reasons since the only alternative would seem to be to try to reason about potential changes in bulk catchment scale parameters, such as the mean residence time of a transfer function, that necessarily integrate the impacts of change on a variety of interacting processes. This would appear to be much more difficult than reasoning about the impacts on the parameters of individual processes. However, this belief might be a little naive since it has not yet been consistently demonstrated that the parameters of process-based models can be estimated *a priori* and produce successful simulations under current conditions. It is actually not clear that such models are always able to reproduce the behaviour of vegetation communities at individual sites (see, for example, Mitchell *et al.*, 2009, 2011).

The one area where significant advances have been made in predicting catchment changes is in modelling the impacts of urbanisation, driven by regulations that require that peak runoff rates from new developments should not be any greater than the natural state (the concept of "sustainable urban drainage systems" (SUDS). Although modern equivalents of the rational method for predicting peak flows are still often used in such calculations, there are also sophisticated commercial packages that represent every impermeable area, drain, channel and detention basin in making the predictions needed (e.g. HydroWorks from Wallingford Software, UK, and MOUSE from DHI, Denmark). Such packages can be used for design purposes as well as simulating existing urban drainage systems. The representation of runoff generation in urban areas is still difficult, however, and Burges *et al.* (1998) argue that continued monitoring of small catchments subject to change remains necessary to resolve some of the uncertainties associated with the prediction of the hydrological impacts of development.

Catchment change is clearly an important issue. There is no doubt that land management can have an important impact on runoff generation processes. There was much more on predicting different types of catchment change in the first edition of this book but, even after another decade, understanding the nature and impacts of catchment change is still very much in its early stages, limited by the short length of hydrological records which can make it difficult to detect change given the natural variability of hydrological systems. Experimental work at the plot scale has shown that land management practices

can have an important effect on runoff generation but it has proven very difficult to upscale such information to larger scales (though see Jackson *et al.* (2008) for a recent attempt to do so). There have been many studies that report predictions of the hydrological impacts of catchment change but these have been almost entirely based on fitting new parameters with the hindsight that change is known to have occurred or speculative scenarios of future change without later verification (and mostly without allowing for uncertainty in the modelling process). Uncertainties in basic hydrological data can make it difficult to identify change (Brath *et al.*, 2006). In fact, the experience in groundwater modelling, where there have been a number of such *post-audit* verification studies, is not too promising in this respect. Most predictions of groundwater responses have been shown to be wrong (Konikow and Bredehoeft, 1992). The result has been a continuing discussion about the possibility of validation of groundwater models, much of which is also relevant for rainfall–runoff modelling (see, for example, the work of Oreskes *et al.*, 1994).

8.9 Changing Risk: Climate Change

There are literally hundreds of papers that have been published in peer reviewed journals that have applied rainfall–runoff models to the problems of assessing the impacts of climate change on stream discharges. Some of this work is summarised in the latest hydrological impacts review of the Intergovernmental Panel on Climate Change (IPCC) (Parry *et al.*, 2007). Such work has attracted significant amounts of funding, driven by the requirements of policy makers who need to respond to the potential impacts of future changes. I will not review that body of work here, as I believe that it is not generally (as yet) fit for purpose (see also Hall, 2007; Wilby *et al.*, 2008; Beven, 2011).

For it to be fit for purpose, it should reflect the uncertainties in the future scenarios for catchment and climate change and the uncertainties in representing rainfall–runoff processes. Very few of the published studies do this in an adequate way, in part because they are constrained by the limitations of the various components in the uncertainty cascade. As noted in Section 8.8, there are very important limitations in the representation of the impacts of land management and other catchment changes on runoff processes. These are probably dominated, however, by the limitations of predicting future climate inputs. The science of climate prediction has progressed dramatically in the last decade. There are now finer resolution coupled ocean–atmosphere models at the global scale; there are high-resolution regional climate models nested within these global models; and there are ensembles of model predictions from both multiple implementations of models in different research centres and "stochastic physics" ensembles for single models.

As an example, the UK Climate Impacts Programme now provides, in its UKCP09 product, a probabilistic interpretation of potential future changes under different IPCC emission scenarios, with maps of the 10%, 50% and 90% projections for various variables. These are based on 11 fine-resolution dynamic regional climate model runs and 300 coarser resolution decadal regional model runs that have been merged and bias-corrected relative to a 1961 to 1990 baseline period using a Bayesian Gaussian emulator method to provide spatial posterior distributions for predicted variables on a 25 km grid for the whole of the UK (see http://ukcp09.defra.gov.uk/). These projections have been used directly as inputs to hydrological models (e.g. Bell *et al.* (2007a) used the G2G model of Section 6.2). These probabilistic projections are given a qualitative interpretation. It is expected that the change is **very unlikely** to be less than the mapped 10% projection and **very unlikely** to be greater than the mapped 90% projection; while the 50% projection can be taken as a current best estimate of the projected changes. That is the best current assessment available, but making direct use of these probabilities is to treat them as if the model was correct and the range of potential outcomes was complete. This is not the case (see also Hall, 2007).

UKCP09 also provides a weather generator that can be used to generate weather sequences for any 5 km grid in the country. The weather generator has been fitted to interpolated weather data for the

baseline period. Weather realisations for future decades can be generated by implementing change factors based on the probabilistic projects from the climate models. Such a weather generator is clearly suited for driving a rainfall–runoff model to represent future conditions, including the possibility of running multiple realisations to reflect the uncertainty in the future conditions. This has been done, for example, by Manning *et al.* (2009).

So why would I suggest that such simulations are not fit for purpose? We have seen earlier in this book, particularly in Chapter 7, that there are real concerns about how well we can constrain the representation of runoff processes in rainfall–runoff models when so many of the uncertainties in the modelling process are epistemic, due to lack of knowledge. This is even more the case for climate models which are known to be inadequate predictors of the baseline period in many parts of the world. In particular, for a hydrologist, they do not adequately represent rainfalls to the extent that many hydrological studies of the impacts of climate change need to use significant bias corrections with respect to current conditions (see, for example, Leith and Chandler, 2010). Underpredictions of 50% of observed rainfalls are not uncommon in upland areas, with even greater problems in parts of the Alps and Andes. These bias corrections are then simply assumed to hold in the future. Climate models also do not adequately predict monsoon rainfalls, el Niño and la Niña effects, or the North Atlantic Oscillation in the baseline period. For hydrologists interested in floods, it is a particular concern that they do not do well in predicting extreme rainfalls. Like our hydrological models, they are wrong and are known to be wrong (whether that be the result of scale effects, sub-grid rain, snow and cloud parameterisations, the simplicity of land surface parameterisations, inadequate representations of heat exchange with the oceans, anthropogenic forcings other than greenhouse gases (Pielke *et al.*, 2009) or other causes).

We expect, of course, that with more research money devoted to climate modelling, more computer power devoted to climate modelling at finer grid scales, better land surface parameterisations (that includes the hydrology and hydrologists should still wince about how it is being represented!) and improved understanding of other process representations in the models, the projections of the next generation of climate models might well be better. But fit for purpose in assessing the impacts of climate change still seems a long way off.

That is not to say I do not believe in climate change. I believe in climate change. I am also worried about the possibility that the climate system, as a nonlinear dynamic system, might be subject to mode of behaviour shifts instigated by variability that is not being predicted by the current generation of global climate models (GCMs) (see, for example, Smith, 2000). We know that there have been rapid modal shifts in the past, before any significant anthropogenic greenhouse gas inputs to the atmosphere (Alley *et al.*, 2003). That suggests that we should plan to adapt to the possibility of change, despite the fact that we should be sceptical of climate model projections (see also Prudhomme *et al.*, 2010).

How, therefore? This depends on how risk averse or risk accepting we are prepared to be and that will often be a matter of how much we are prepared to spend on an adaptation strategy. Being risk averse generally requires more expensive measures than being risk accepting. But we can consider how expensive the required adaptation might be for different scenarios of future change, more or less extreme, quite independently of any climate model projections. Making such decisions does not require specifying the probabilities of potential outcomes; there are other decision frameworks for adapting to change (e.g. Beven, 2009, Chapter 6). In that way, it is possible to plan a response to different magnitudes of change in terms of costs (and benefits) that might be robust with respect to future change. This does not necessarily mean the over-design of hard infrastructure: other societal policies that permit the ability to adapt over time might well be more cost-effective.

A general principle commonly followed in policy formulation is that the response should be proportionate to the risk so that ideally we would wish to evaluate the probability associated with each magnitude of change. UKCP09 is presented in this way, but remember that these projections are not a representation of the odds of climate actually turning out that way – they are, rather, the probabilities of the model

projections within the ensemble sample itself. This difference is important. Indeed, such are the known scientific limitations of representing precipitation in climate models that these probabilities might have little or no relevance to the policy response. To ignore those probabilistic estimates and deal with magnitudes of change directly (without the need for climate simulations) precludes a complete risk-based strategy but does place the focus directly on what is considered to be affordable in being precautionary.

A particular case in point is protection against flooding. If a changing climate is intensifying the hydrological cycle, we expect the frequency of floods of a given magnitude to change (even if, given the nature of extremes, this might be difficult to demonstrate from the available observations, e.g. Wolock and Hornberger, 1991; Robson, 2002; Kundzewicz *et al.*, 2005; Wilby, 2006; Wilby *et al.*, 2008; Fowler and Wilby, 2010). A number of studies have invoked the change factors produced by climate models to examine how flow frequencies might change. This is straightforward to do if it can be assumed that the parameters calibrated to represent catchment response might not change with changing inputs. It is much more difficult to do if it is thought that the change in inputs or land use and management might require that parameter sets be changed to represent new sets of conditions.

A number of strategies are possible so as not to exacerbate the flooding problem: avoiding new developments on flood plains; improving flood defences; flood proofing of existing buildings; breaching of existing defences to make more storage; building flood detention basins. In most cases, these solutions are robust in the sense of not precluding future adaptive management strategies but they all have a greater or lesser cost. So what is the cost–benefit of protecting against different levels of change. How precautionary are we prepared to pay to be? This is, essentially, a political decision. The science comes in estimating the costs and benefits of different policy options, which might be considered to be a far more realistic goal than the accurate prediction of future change (Beven, 2011).

In addition, there are other factors that might affect future hydrological responses (societal change, urbanisation, agricultural intensification, energy security, deforestation–afforestation, river training and re-naturalisation, ...) that might be far more important on the decadal time scales over which we might achieve some adaptive management strategy. There are certainly model-based scenario projections of the effects of potential changes in different factors, mostly deterministic in nature (e.g Bronstert *et al.*, 2007; Viney *et al.*, 2009), even though we know that process representations of such factors are subject to considerable uncertainty (and should really be embarrassed about some of the studies that have been published that purport to assess the impacts of change using only deterministic model predictions). Such changes could also be evaluated in the form of the precautionary cost–benefit strategy suggested above.

It is clear that more science and understanding is required to reduce the uncertainties in assessing the impacts of change as an input to an informed and open policy framing debate. Decision makers require "evidence" of how great the impacts of change might be; but it would be much more intellectually honest to change the nature of the game into something more overtly political before the "evidence" comes to be seen as based on insubstantial foundations. Future food and energy security might, for example, provide far more politically compelling arguments for climate mitigation policies than uncertain climate projections. We need better ways of deciding how precautionary to be in planning for the future.

8.10 Key Points from Chapter 8

- There are two uses of rainfall–runoff modelling in flood prediction: one in forecasting discharges in real time during flood periods, the other in predicting the frequencies of different flood peak magnitudes.
- Real-time forecasts are very dependent on the accuracy of input data, particularly of spatial patterns of rainfall intensities. The availability of radar rainfall data has greatly improved the potential for forecasting flood peaks. Improvements in the lead time of forecasts are dependent on improvements

in near future forecasts of rainfalls from either projections of radar images into the future or numerical weather prediction, which at present are not very reliable.

- Flood forecasting is best carried out using adaptive models that can take account of prediction errors if discharge or stage data is available in real time by telemetry. This is most easily done with relatively simple models, such as the Lambert ISO model for small catchments, or transfer function type models.
- Rainfall–runoff modelling is also being increasingly used for prediction of flood frequencies in combination with either long rainfall records or a stochastic rainstorm model as an alternative to a purely statistical estimation of flood frequency. This approach has the potential to take account of the changing nature of the hydrological response of a catchment with changing antecedent conditions and rainstorm volume or intensities, but at the present time is associated with significant uncertainties.
- There are strong realisation effects in flood frequency estimation, both between different observation periods and in realisations of stochastic rainfall models. The observed period is but one realisation of many possible sequences of floods in a length of record and this needs to be taken account of in model calibration.
- An important use of rainfall–runoff models is to assess the impact of future land use and climate change on catchment hydrology, particularly extreme events and water yields for water resource evaluation. Such scenario projections should be associated with an estimation of uncertainty to provide a reality check on the prediction of future changes.
- With some exceptions of dramatic change in small catchments, the natural variability of hydrological systems in space and time, and generally short periods of available observations, make it very difficult to detect, study and understand the impacts of past change as evidenced in hydrological observations.
- The spatial complexities of land use change, for example in urban areas and in agricultural field drainage systems make it particularly difficult to predict the impacts of change.
- The impacts of climate change have, to date, been assessed using scenario simulations that are conditional both on the rainfall–runoff model being used and on the particular climatic scenario being used as forcing data. Such projections should be associated with an estimate of uncertainty or risk, but there have been very few studies in which any attempt has been made to assess uncertainty of the impacts of change. The uncertainties are such that it becomes a political issue as to how far to protect against the impacts of future change.

Box 8.1 Adaptive Gain Parameter Estimation for Real-Time Forecasting

Consider the simple first-order transfer function model:

$$\widehat{y}_t = G_t \frac{b_o + b_1 z^{-1}}{1 - a_1 z^{-1}} u_{t-\delta} \tag{B8.1.1}$$

where $\widehat{y}_t$ is the predicted output variable, $u_{t-\delta}$ is the input variable (effective rainfall or upstream river level or flow), z is the backward difference operator (see Box 4.1), a_1, b_o and b_1 are parameters fixed by previous calibration and G_t is a time variable *gain*.

The time variable gain can be estimated in real time by the methods of time variable parameter estimation developed in Young (1984; see also Young (2011a) and Box 4.3). The algorithm is a form of recursive filtering of the gain parameter G in that the gain at time t, G_t, depends on G_{t-1} at the previous time step, together with a function of the current prediction error $\left(y_t - \widehat{y}_t\right)$ where y_t is the observed flow at time t.

The algorithm has the form of predictor–corrector equations as follows:

Predictor step:

$$\widehat{G}_{t|t-1} = G_{t-1} \tag{B8.1.2}$$

$$P_{t|t-1} = P_{t-1} + Q \tag{B8.1.3}$$

Corrector step:

$$G_t = \widehat{G}_{t|t-1} + \frac{P_{t|t-1}\widehat{y}_t \left[y_t - \widehat{G}_{t|t-1}\widehat{y}_t\right]}{1 + P_{t|t-1}\widehat{y}_t^2} \tag{B8.1.4}$$

$$P_t = P_{t|t-1} - \frac{\left[P_{t|t-1}\widehat{y}_t\right]^2}{1 + P_{t|t-1}\widehat{y}_t^2} \tag{B8.1.5}$$

The degree to which the time variable gain G_t is allowed to change between time steps is controlled by the value of Q which is a noise variance ratio (NVR). The higher the value of Q, the faster the memory effect of previous values of G at times, $t-1$, $t-2$, ... is allowed to decay away. Thus a large value of NVR means that there is a short memory and less filtering of the changes in G. In the Dumfries study by Lees *et al.* (1994) the value of Q was allowed to vary with the magnitude of the prediction error so that the larger the prediction error at a time step, the more rapidly G was allowed to change. Romanowicz *et al.* (2008) have shown how a representation of a heteroscedastic noise variance ratio can be included in this approach.

One advantage of this type of methodology is that an estimate of the uncertainty in the predictions can be calculated and updated recursively. If we define the n step ahead prediction error as

$$\epsilon_{t+n} = \left(y_{t+n} - \widehat{y}_{t+n}\right) \tag{B8.1.6}$$

then the variance of ϵ_n is given by

$$var(\epsilon_n) = \widehat{\sigma}_t^2 \left[1 + P_{t+n}\widehat{y}_{t+n}^2\right] \tag{B8.1.7}$$

where $\widehat{\sigma}_t^2$ is the current forecast error variance which is estimated from the further recursion

$$\widehat{\sigma}_t^2 = \widehat{\sigma}_{t-1}^2 + P_t \left[e_t^2 - \widehat{\sigma}_{t-1}^2\right] \tag{B8.1.8}$$

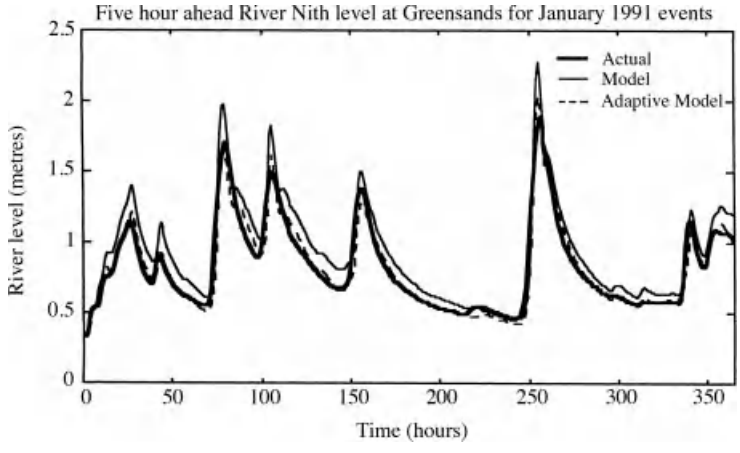


Figure B8.1.1 *Comparison of adaptive and non-adaptive five-hour-ahead flood forecasts on the River Nith (after Lees et al., 1994, with kind permission of John Wiley and Sons).*

where e_t is the scaled prediction error

$$e_t = \frac{y_t - \widehat{y}_t}{\sqrt{1 + P_t \widehat{y}_t^2}} \tag{B8.1.9}$$

The variance of ϵ_n can be used to estimate prediction limits for the forecasts based on the uncertainty in the time variable estimates of G_{t+n} as reflected in the magnitude of P_{t+n}. Further details of this algorithm, and its extension to recursive updating of multiple parameters, is provided by Young (2011a).

Note that in the predictor–corrector equations, only the error $\left(y_t - \widehat{y}_t\right)$ at the current time t can be known, but that this can be fixed in extrapolating the estimates of G and P forwards using the predictor–corrector equations to time $t + n$. In general, the magnitude of the prediction limits increase with increasing discharge and with extrapolation further into the future. In the Dumfries forecasting system described by Lees *et al.* (1994) it was found that reasonable forecasts could be made for five hours ahead (see Figure B8.1.1), even though the natural time delay in the river network was of the order of three hours.

The time variable gain in this type of model is a way of compensating for both data errors and any nonlinearities that are not properly represented in the model structure. Adaptive methods can also be used with nonlinear models (Beven, 2009). Time variable parameter estimation can also be used to investigate the nature of the nonlinearities of an input–output system and is the basis of the state-dependent parameter approach to model identification (see Box 4.3 and Young, 2011a).

9

Beyond the Primer: Next Generation Hydrological Models

We must admit that our paper is more of an artist's conception than a blueprint

Freeze and Harlan, 1969

Predictions of hydrologic responses now need to allow for adaptive temporal evolution of vegetation, soils, and river networks (among other things) under human-induced environmental changes, although the changes might occur at different and varying rates. This requires hydrologists to develop a new understanding of how all the associated components (climate, soils, vegetation, and topography) have coevolved in the past and how they might do so in the future

Thorsten Wagener *et al.* 2010

9.1 Why are New Modelling Techniques Needed?

In considering the next generation of hydrological models, it is important to distinguish between different reasons for modelling. What the next generation model looks like might well be different for different purposes. If the interest is only to show that we "understand our science and its complex interelated phenomena" (see the quotation at the head of Chapter 1) then a model structure might look very different to that needed to make flood forecasts at a particular site of interest. Using a complex process model for flood forecasting might not be sensible either in improving the accuracy of predictions or in providing forecasts with sufficient lead time for people to react to a flood warning. In other cases, a model that predicts the fluxes in different flow pathways might be needed. Such models might be used to explore the effects of land management on water fluxes and water quality, for example. There are also increasing demands to understand the links between hydrology and ecology in effecting improvements to the water environment and maintaining biodiversity, demands given impetus in Europe by the requirements of the Water Framework Directive.

Rainfall–Runoff Modelling: The Primer, Second Edition. Keith Beven.

It seems unlikely that all these different reasons for using models can be met by a single modelling approach in the future. The next generation of hydrological models is likely therefore to continue to be diverse. What does seem clear is that in the future there will continue to be ever more powerful computers available to modellers, and ever more spatial information available through remote sensing and geographical information systems. This will be the context for a new generation of models. The aim of this chapter is to explore the limitations of what has gone before and what the new generation might look like.

One alternative is to use "physically based" models. It is now more than 40 years since Freeze and Harlan published their seminal "blueprint for a physically-based digitally-simulated hydrologic response model" (Freeze and Harlan, 1969). As explained in Chapter 5, this was the first attempt to define representations of flow processes in a catchment based on physical theory (the Darcy–Richards equation for partially saturated subsurface flow and the depth averaged St Venant equations for surface flows). At the time, implementation of those process descriptions was severely limited by the computer power available. The expansion of computer power in the last 40 years has, however, greatly relaxed this constraint and it is possible now to apply such models with a fine discretisation to both small and large catchments. A number of implementations of the Freeze and Harlan blueprint are discussed in Chapter 5.

I have suggested in the past, however, that this blueprint needs reconsidering (Beven, 2001, 2002a). In fact, it really needs to be abandoned in favour of a new generation of modelling concepts. The reasons for this lie in the very nature of the physical concepts on which it is based, particularly in the representation of subsurface flows. The Freeze and Harlan blueprint leads to a system of coupled nonlinear partial differential equations. Solution of those equations is, in itself, an interesting technical problem to ensure that the model predictions are a consistent approximation to the orginal equations. But the equations also have parameters (saturated hydraulic conductivity, porosity, soil moisture characteristic functions, roughness parameters, etc.) that must be specified for all the calculation elements in the solution domain.

There is an interesting comparison to be made here with the somewhat similar problems encountered in atmosphere and ocean models. These are also based on physical principles in representating the dynamics of fluid motion. The difference arises because of the importance of turbulence in cascading momentum from large scale to small scale motion. In flow domains where the boundary conditions are not so locally dominant, the approach has been to take a "top-down" approach to solution of the equations. Once a grid scale for the solution is defined (and computer constraints are still a constraint on how fine a grid scale can be used), all processes at smaller scales are parameterised in some approximate way.

Hydrological modelling of this type has, in contrast, taken a "bottom-up" approach, using small-scale theory from laboratory experiments and trying to scale up to larger scales. Indeed, in most hydrological models of this type the only scaling attempted is to take a locally predicted flux and multiply by the area of a calculation element. In the unsaturated zone in particular, the physics tells us that is wrong. Even if the Darcy–Richards equation did apply locally, the nonlinearity inherent in its formulation means that, in any calculation element with heterogeneous soil properties, there can be no simple averaging that leads to the integral of the heterogeneous fluxes over the element being well represented by the same equation applied as if the element was homogeneous. We should be using a **different** equation that reflects the nature of the heterogeneity of the soil (if only we could know what that might be for heterogeneous properties that are essentially unknowable). The approximation will be worse if the same Darcy–Richards parameters are applied to all elements classified as having the same soil type (as is commonly done in this modelling approach to reduce the number of parameters to be specified).

The approximation will also be worse if the flow is not well represented by the Darcy–Richards equation. This might be the case if, for example, there are macropores or other structures inducing preferential flow or by-pass flow in the soil or by-passing due to film flow in finer pores (see Section 1.4 and Figure 1.1; Germann and Hensel, 2006; Hincapié and Germann, 2009; Gerke *et al.*, 2010). There is an interesting historical point to be made here. L. F. Richards, in his classic 1931 paper extending Darcy's law to unsaturated flow (see Section 5.1.1), carried out laboratory experiments in which he used air

pressure applied to an enclosed sample of a porous medium in order to create unsaturated conditions that immediately empty all the connected larger pores. They stay empty even if there is a hydraulic gradient inducing flow. In the field, of course, there is no such imposed pressure and by-passing would be more likely to take place during wetting if (locally) gravity is not dominated by capillary gradient effects. The classical treatment of unsaturated soil water flows is, therefore, effectively predicated on an unrealistic experiment! Air pressure effects could still be important perhaps when there is a continuous saturated wetting front infiltrating into uniform soil, though even in this case air might escape through larger pore pathways. This was recognised by Robert Horton, who also argued that the decline in infiltration rates seen in experiments was more likely to be due to changes at the surface rather than a control due to capillary gradient effects in the profile (see Beven (2004a) and Box 5.2).

Actually, the limitations of the theory of the Freeze and Harlan blueprint is not the only problem in applying it in actual model applications. The greater problem is in the characterisation of the uniqueness of individual catchments in terms of the parameters of the model. The argument, commonly made, that the parameter values required by such models are more easily estimated because of the physical basis of the concepts, does not really hold if the concepts are (at best) an approximation to the physics of flow in heterogeneous domains (Beven, 2000, 2002a). As noted earlier, what is then needed is *effective* values of the parameters that help compensate for the limitations of the physics incorporated into the model. The effective values might be quite different from what can be measured or estimated from *pedotransfer functions* derived from small scale measurements (see Box 5.5). Indeed, numerical experiments suggest that where there is a strong interaction between surface and subsurface processes, it might not be possible to find effective values consistent for all conditions (Binley *et al.*, 1989).

So, what should we conclude? That the Darcy–Richards equation might not be the best description of flow in unsaturated soil in the field and that applying the Darcy–Richards equation as if the soil had uniform soil properties is to ignore the implication of the physics that some other equation should be being used. This, together with the difficulty of knowing the true boundary conditions and characteristics of the subsurface in the field, may certainly be why it has proven rather difficult to show that models based on the Freeze and Harlan blueprint can successfully reproduce the behaviour of real catchments (see the case studies of Sections 5.4, 5.6 and 5.7). The real question then is can we do better? I think we can.

9.2 Representative Elementary Watershed Concepts

In fact, the basis for a new generation of hydrological models already exists in the "representative elementary watershed" (REW) concepts developed by Paolo Reggiani and others (Reggiani *et al.* 1998, 1999, 2000, 2001; Reggiani and Shellekens, 2003; Reggiani and Rientjes, 2005). The REW framework is based on physics, in that it sets out the fundamental balance equations for the processes underlying the hydrological response of a control volume element of the landscape (Figure 9.1). It is normal for hydrological models to consider the local or catchment water balance equation, of course, and some models also make use of a surface energy balance equation in estimating evapotranspiration or snowmelt fluxes. In this sequence of papers, Reggiani *et al.* take this further to consider the water balance, energy balance and momentum balance equations for the processes in control volumes underlying any landscape unit. This framework is discrete in space, resulting in ordinary differential equations in time, rather than the continuum partial differential equation approach of the Freeze and Harlan blueprint. It is also scale independent: soil profiles, small plots, hillslopes, or a zero-order or higher catchments can be treated as REWs. Larger catchments could be made up of multiple REW elements, though these need not be of the form of elements with simple geometry such as the grid squares of SHE or SWAT, or the triangular or hexagonal elements of other distributed models. Within the REW framework they can be hillslope or

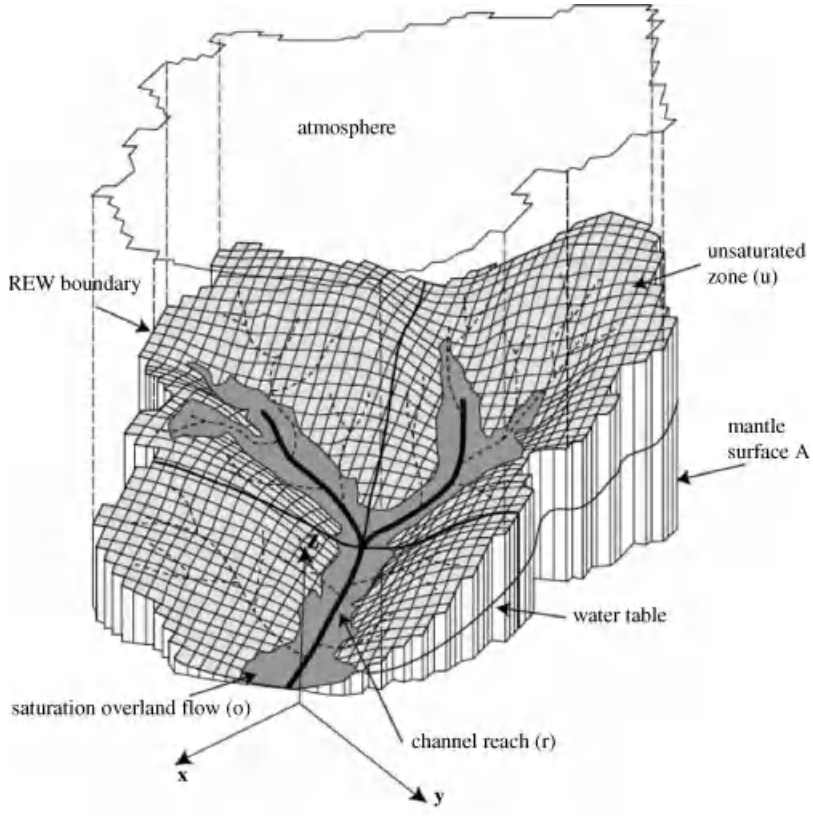


Figure 9.1 *Three-dimensional view of an ensemble of three REWs, including a portion of atmosphere (after Reggiani and Rientjes, 2005, with kind permission of the American Geophysical Union).*

catchment shaped, linked by channel reach control volumes and reflecting the topographic structure of the catchment (Reggiani and Rientjes, 2005).

This scale independence is a major advantage of this framework but we also expect that the nature of hydrological fluxes should be scale dependent and exhibit some nonlinearities (Beven, 2006b). The REW balance equations include these scale dependencies and nonlinearities but only implicitly. They do so through the flux terms in the balance equations at the boundaries of the chosen landscape units. In applying a model based on the REW concepts, these flux terms (of mass, energy and momentum) must be specified in some way. We expect hydrological fluxes to be nonlinearly related to storage (perhaps with some hysteresis between wetting and drying) and to be dependent on the scale of the element. This is what is called the *closure problem*. Addressing the closure problem is exactly what is required to develop the new generation of hydrological models; indeed, Beven (2006b) has called the solution to the closure problem the "Holy Grail" of scientific hydrology. It is worth saying that none of the current implementations of the REW concepts have provided entirely satisfactory solutions to the closure problem (see Section 9.4); hence further research is required. The framework provides, however, a general structure for thinking about how the complex nature of hydrological responses within a REW landscape unit might best be represented by the functional relationships specified for the fluxes at different scales.

There is an analogy here with the top-down sub-grid parameterisations of atmosphere and ocean models mentioned above; but rather than the representation of the momentum dissipation due to turbulence, the hydrologist is confronted with energy and momentum dissipation due to friction and viscosity at all the internal boundaries for surface and subsurface flows. It is difficult to theorise about the latter because of the sheer complexity of all the internal boundaries across a wide range of pore sizes in the soil and roughness conditions on the surface. Thus specifying the "physics" at this scale is inherently difficult and inherently uncertain. Indeed, the hydrological closure problem might therefore actually be considered harder than the representation of atmospheric turbulence (even if the problem of representing the effects of inhomogeneous turbulence is accepted as one of the most difficult problems in geophysics). That does not mean that this is not a valuable strategy to take. While inital solutions to the closure problem might be lacking in both sophistication and accuracy, hydrological science can work to gradually develop improvements.

9.2.1 The REW Balance Equations

Examples of the mass, energy and momentum balance equations have already appeared earlier in this book (see for example Equations (3.2), (5.4) and (5.5)) but, before the formulation of the REW framework, no-one had developed these equations in a consistent way for multiple processes. In fact, all the balance equations can be reduced to the simple form (Reggiani and Rientjes, 2005):

$$\frac{\partial \psi}{\partial t} = \sum_i Q_\psi + R + G \tag{9.1}$$

where ψ is the mass, energy or momentum, t is time, Q_ψ is a flux of ψ for the ith internal or external boundary, R is a source or sink term, and G is an internal production term. Examples of such balance equations are given by Reggiani and Rientjes (2005) for multiple processes within a REW. The expansion of the terms of these equations can be done in a variety of ways but the essence of the closure problem immediately becomes apparent. Even this simplest representation of the catchment requires the flux terms $\sum_i Q_\psi$ to be defined for all i to close the system of equations. In some cases, it will be possible to make a plausible argument that some of these terms are negligible but, in most cases, representations of these boundary fluxes are required, hence the closure problem is rather important.

9.2.2 The REW Closure Problem

It is evident that closure of all the mass and energy balance equations will not be easy in most applications. (What, for example, would be the effective kinetic energy added to a control volume by the input of rainfall? What is the effective dissipation of energy in macropore flows over a control volume?) This does not affect the physical principles involved, only their practical implementation and application (Beven, 2006b). We should, in principle, be able to close the mass, energy and momentum balances. We might, in practice, find that difficult to do when there is uncertainty about how to represent many of the flux terms, even those such as rainfalls and stream discharges that it might be feasible to measure. It might, indeed, at the present state of knowledge, only be possible to estimate partitioning between the required boundary fluxes and closure to a certain degree of approximation and uncertainty.

The functional requirement here is that the representation of each boundary flux for each process component should properly reflect the relationship between that flux and the changing state of the system (the *constitutive relationships* for a process). This immediately raises new possibilities about how to think about the process representations. We would expect the closure fluxes to be a hysteretic function of storage, for example, i.e. the fluxes would be different depending on the past sequence of wetting or drying of a control volume. In general, we would expect the discharge for a given storage to be lower

in wetting than in drying. I would encourage the reader to think about why there should be such an expectation, starting with a block of soil in a lysimeter and working up to hillslope and catchment scales. We also would expect the closure fluxes to be a function of spatial scale and geometry of the control volumes. The time scale of the response will be longer when the hillslopes or channel reaches are longer or the soil profile is deeper.

This type of hysteresis has been demonstrated in a number of field studies at different scales. Myrabø (1997), Ewen and Birkinshaw (2006) and Beven (2006b) have all demonstrated hysteretic relationships between storage and discharge in small catchments. Kendall *et al.* (1999) showed how the form of the hysteresis might vary with position in the catchment. Clockwise hysteresis loops between flow and groundwater levels were found in the riparian area in a small catchment in Vermont, while counter-clockwise hysteresis was apparent in the data from sites on the contributing hillslopes.

Hysteresis is not a new concept to hydrologists. It has been used within soil physics in representing the soil moisture characteristics which often show different behaviour in wetting and drying. This is often forgotten when soil moisture characteristic curves are used in hydrological models based on the Darcy–Richards equation, partly because of a limited amount of experimental information about hysteretic characteristics (though the GRIZZLY database of Haverkamp *et al.* (2002) has gathered together what information was available at that time). There are also multiple explanations of the causes of hysteresis and no real consensus about how it should be parameterised (see the discussions by O'Kane, 2005; Flynn *et al.*, 2005; and Beven, 2006b).

9.3 How are the REW Concepts Different from Other Hydrological Models?

In one sense nearly all previous rainfall–runoff models are implementations of the REW concepts in that they provide flux relationships that allow closure of balance equations, whether they are the lumped models of Chapter 4, the distributed models of Chapter 5 or the semi-distributed HRU models of Chapter 6. It would even be possible to analyse the flux relationships of such models to see what they imply about energy or momentum balances (even if they are not considered explicitly in the original model formulations). The difference is in the way in which the REW concepts require that the mass, energy and momentum balance equations are considered in an integrated and consistent way at the scale of a discrete partitioning of the landscape into REW elements. I consider the most important consequence to be the way in which this focuses attention on the scale dependence of the relationships between the boundary fluxes and the state of the system. This allows the theory of hydrological science to be comprehensively rethought in ways in which scale, heterogeneity and nonlinearity are more properly and explicitly reflected in the process representations and closures at the REW control volume. However, it has to be said that most implementations of the REW concepts to date have not considered this scale dependence in any way.

9.4 Implementation of the REW Concepts

There have been at least four implementations of the REW concepts reported in the literature: the REWv4.0 model (Fenicia *et al.*, 2005; Reggiani and Rientjes, 2005; Varado *et al.*, 2006); the CREW model (Lee *et al.*, 2007; Zehe *et al.*, 2006); the REWASH model (Zhang and Savenije, 2005); and the THModel (Tian *et al.*, 2006; Mou *et al.*, 2008). The latter extends the REW concepts to low temperature regions in which freeze–thaw processes have an important effect on the hydrology.

The closure schemes that have been used in these implementations of the REW concepts as hydrological models have largely ignored both hysteresis and scale issues. Instead, they have taken a path similar to models based on the Freeze and Harlan blueprint by assuming that small scale physics can be used at larger scales only with a change of effective parameters. This, however, gives an even greater approximation in the discrete framework of the REW than in the numerical solution of the partial differential equations of the Freeze and Harlan blueprint. The distributed nature of a direct numerical solution, at least, shows the effects of length scales and wetting and drying implicitly (and could be thought of as a form of REW model within which the REW elements are defined by the spatial discretisation of the numerical solution; there might also be the possiblity of transforming the orginal differential equations to boundary element form in cases where suitable continuum equations can be defined).

An interesting application in this respect is that of Zehe *et al.* (2006) and Lee *et al.* (2006). The aim of their papers was to provide constitutive relationships for predicting the response of ungauged basins within the REW framework (see Chapter 10). To do so, they looked initially at modelling a single small catchment, the Weiherbach catchment (3.6 km^2) in Germany. Erwin Zehe had already implemented a distributed model for this catchment (CATMOD) that produced reasonably accurate predictions of the discharge hydrograph and of the averaged soil moisture observations in the catchment (Zehe *et al.*, 2005). CATMOD is based on the Darcy–Richards equation, but with the option of additional macropores as local zones of high hydraulic conductivity in the solution domain. Zehe *et al.* (2006) then used this model as a dynamic upscaling tool to develop constitutive relationships for storage–discharge at the REW scale in the CREW model. They used the existing model to do so, as well as alternate representations of the hillslopes in terms of both soil types and macroporosity. They showed how the relationships could be derived from integrated values of storage and predicted discharges. The relationships showed some limited hysteresis in this case, but a non-hysteretic relationship could provide almost as good predictions. They allow that the limited range of soil moisture values simulated at this site might have limited the hysteresis in the response, so that this might not be a general conclusion. Lee *et al.* (2007) then take the relationships to provide closure in a REW model of the catchment, demonstrating predictive capability of equivalent accuracy to the orginal model.

In the application of the REW concepts to the Weiherbach catchment, the CREW model is emulating the CATMOD simulation rather than the real catchment. Zhang *et al.* (2006) have taken an alternative strategy, calibrating the effective parameter values of their REWASH model (where the closures are also based on small scale Darcian physics) directly against data from the Geer catchment in Belgium. In both cases, there is still a limited degree of accuracy in reproducing the behaviour of the real catchment that is similar to that achieved with other hydrological models calibrated against observations.

There may be two primary reasons why these new concepts might seem to perform no better than traditional models. One is that the level of performance that is achievable is limited by the uncertainties in the observations against which they are being calibrated or evaluated. As discussed in Chapter 7, rainfall–runoff models should not be expected to perform better than the quality of the data that is available to drive and evaluate them. The second is that the types of closure scheme that are being used in the first implementations of the REW concepts might still be improved, beyond the range of concepts that have been used in the rainfall–runoff models currently available.

The question is how this might be achieved in future research and practice. We cannot get around the fact that while experiments such as those in Figure 1.1 reveal the complexity of subsurface flow pathways, it is not generally possible to examine how water flows in the subsurface (and even how it might flow over the surface) everywhere in a landscape unit. It is clear, however, that to ignore such complexity and assume that small scale theory can simply be scaled up to the REW is not the correct response, even if it might appear to produce acceptable results in some cases. It would be better to try to account in a more general way for such complexity. It is not yet clear how this might be done, but two possibilities are considered here, one in which a scale-dependent hysteresis is inferrred from simplified theory and one which tries to represent the heterogeneity of flow pathways directly as distributions of velocities

(that will change with wetting and drying). Since velocities require a length scale in order to simulate the timing of boundary fluxes, the scale and geometry of the control volume are naturally taken into account (even if there is necessarily some uncertainty about the subsurface geometry and velocity distributions).

9.5 Inferring Scale-Dependent Hysteresis from Simplified Hydrological Theory

An analytical treatment of the scale-dependent hysteresis of hillslope scale runoff responses has been produced by Norbiato and Borga (2008). Their analysis was limited to horizontal kinematic subsurface fluxes on slopes of a certain range of shapes (assuming a parabolic width function and an exponential surface slope function) under conditions of homogeneous conductivity and uniform recharge (which also means that delays in the unsaturated zone can be neglected). For nine different shapes of slope their analysis represents the hysteresis in the storage–discharge relationship as a non-dimensional functional form (Figure 9.2). For each slope shape, the length of the slope and the uniform recharge rate control the scaled discharge from the slope.

Norbiato and Borga did not consider cases where the water table rises to the surface and allows a saturated contributing area to develop. They speculate about the hysteresis in the fast runoff response that would result in this case. This has been demonstrated by the work of Martina *et al.* (2011) who develop relationships between storage and saturated area for use in the lumped version of TOPKAPI (Ciarapica and Todini, 2002; Liu and Todini, 2002). These relationships are derived from the fully distributed version (see Section 6.5) in which the length scale is considered explicitly. The parameters of the lumped version

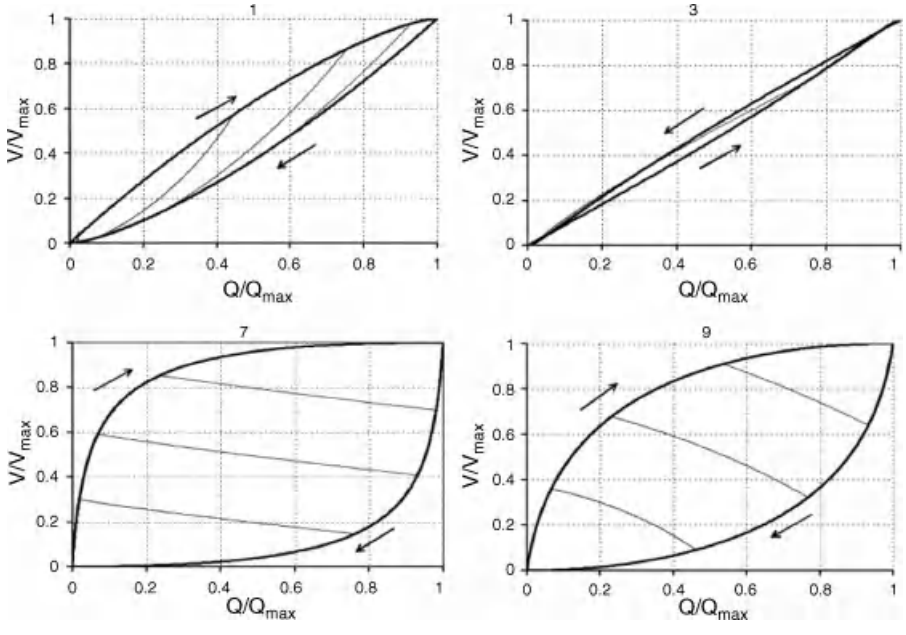


Figure 9.2 *Normalised storage ($r(t) = V(t)/V_{max}$)) versus flux ($u(t) = Q(L, t)/Q_{max}$) relationships for four hillslope forms: 1. divergent concave; 3. divergent convex; 7. convergent concave; and 9. convergent convex (after Norbiato and Borga, 2008, with kind permission of Elsevier).*

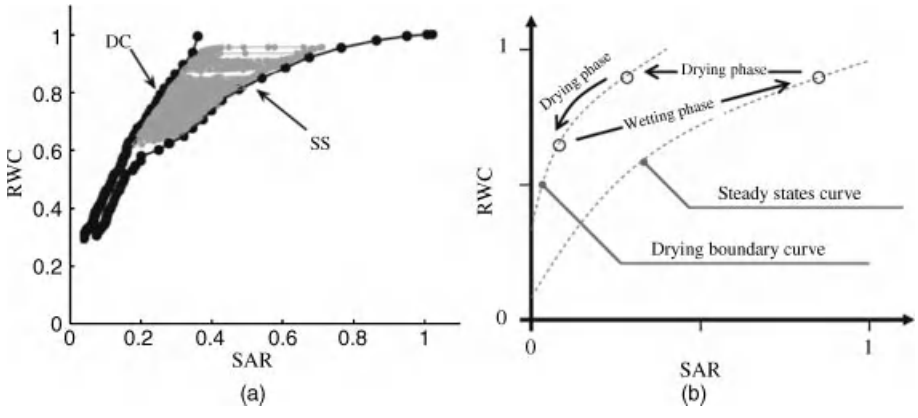


Figure 9.3 *(a) Saturated area ratio (SAR) against relative water content (RWC) for a five-year simulation period of the Reno catchment at the Calcara river section (grey dots); the steady state curve (SS) was obtained as a set of SAR–RWC values for different simulations relative to precipitation events of different intensity and after the equilibrium state has been reached; the drying curve (DC) represents the SAR and RWC values for the drying down transition phase only; (b) Schematic diagram showing how the hysteretic relationship is used in the lumped TOPKAPI model (after Martina et al., 2011, with kind permission of Elsevier).*

can then be related to the characteristics of the area being modelled, including slope lengths, hydraulic conductivity and porosity. Figure 9.3 shows the relationship between relative water content and saturated area derived for the whole of the Reno catchment in Italy (4000 km^2). Martina *et al.* also show that different subcatchments, with different characteristics, exhibit somewhat different hysteretic behaviour. Relationships were also developed for return flow as a function of saturated area.

These two studies do not represent a complete solution to the closure problem because of the simplifying assumptions that are made (constant hydraulic conductivities in relatively shallow flow pathways, etc.) but they point to the way in which the forms of better closure schemes mught be developed. They also do not attempt to predict the residence times of water in the system, which would be a useful additional attribute to include in a next generation of rainfall–runoff model.

9.6 Representing Water Fluxes by Particle Tracking

Another way of implementing a closure scheme based on velocity distributions is to treat the water as particle masses of different velocities (and therefore energy and momentum). The sample of particle velocities is then a representation of the assumed velocity distribution. As the system wets and dries, the particles can change state and velocity accordingly. There is some overlap here with stochastic storage models since the mean residence time of a water particle within a storage volume will be a function of both its point of origin and the *Lagrangian velocity* along its pathway to the outlet. Stochastic storage models have a long history in hydrology. This type of model was first implemented by Jay Bagley working at Stanford University at the same time as Norman Crawford was developing the Stanford Watershed Model. At the catchment scale, stochastic storage assumptions also underly the concepts of the geomorphological unit hydrograph (see Section 4.7.2). Many other conceptual storage models can be interpreted in this way. This approach is considered in more detail when we discuss residence time distribution models in Chapter 11.

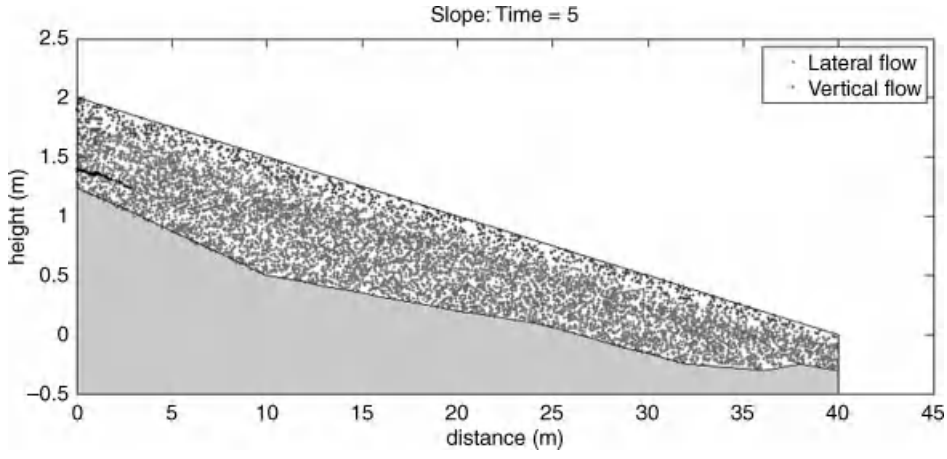


Figure 9.4 *Distribution of unsaturated zone (light grey), saturated zone (darker grey) and tracer (black) particles in the MIPs simulation of a tracer experiment at Gårdsjön, Sweden (after Davies et al., 2011, with kind permission of John Wiley and Sons).*

More explicit *particle tracking models* have been introduced by Beven *et al.* (1989) and in John Ewen's SAMP model (Ewen, 1996). The former considered steady state flows only; the latter, infiltration into a single soil profile, allowing for the effects of capillarity on particle movement. More recently, Davies *et al.* (2011) have suggested that such an approach might also be used in a more explicitly distributed model, tracking individual particle masses of water through a hillslope element. Their multiple interacting pathways (MIPs) model assigns Lagrangian velocities to particles from distributions according to whether they are in the unsaturated (vertical movement) or saturated (downslope movement) state. The distributions must integrate to match the specified saturated hydraulic conductivity of the soil, which in their application was a strongly nonlinear function of depth into the soil profile. In the initial application of the model, interactions between particles are ignored so that particle movements can be treated in parallel and bulk energy loss is assumed to be in equilibrium with the potential energy gained with loss of elevation (equivalent to saying that capillary forces can be ignored as having only a local effect and that the hydraulic gradient is equal to the slope angle). Figure 9.4 shows a snapshot of the particles on a representation of a hillslope at the Gårdsjön catchment in Sweden. More details of this modelling experiment are in Chapter 11, but Davies *et al.* (2011) demonstrated that the MIPs approach could provide good simulations of both flow and tracer concentrations at the site.

Having established that this model can reproduce the observed behaviour of the Gårdsjön slope, Davies and Beven (2010) also explored the effects of slope length scale on the nature of the storage–discharge relationship, the constitutive relationships of the REW framework. Figure 9.5 shows, in an extension of their original results, how the hydrographs and hysteresis in storage–discharge changes for slopes of different length and different antecedent conditions. This work demonstrates how other approaches to representing hillslope hydrology are possible, albeit as yet with simplifying kinematic assumptions.

Some of those assumptions will need to be relaxed in future. For example, in situations with deeper subsurface flow systems, the change in hydraulic gradient with wetting and drying will be much more important (although a kinematic assumption might still be a good approximation if the Lagrangian velocities scale well with the total head differences in longer slower pathways). Alternatively, there is a particle-based technique that allows the assessment of changes in pressure within the flow domain. This is "smoothed particle hydrodynamics" (SPH), first introduced by Gingold and Monaghan (1977), which

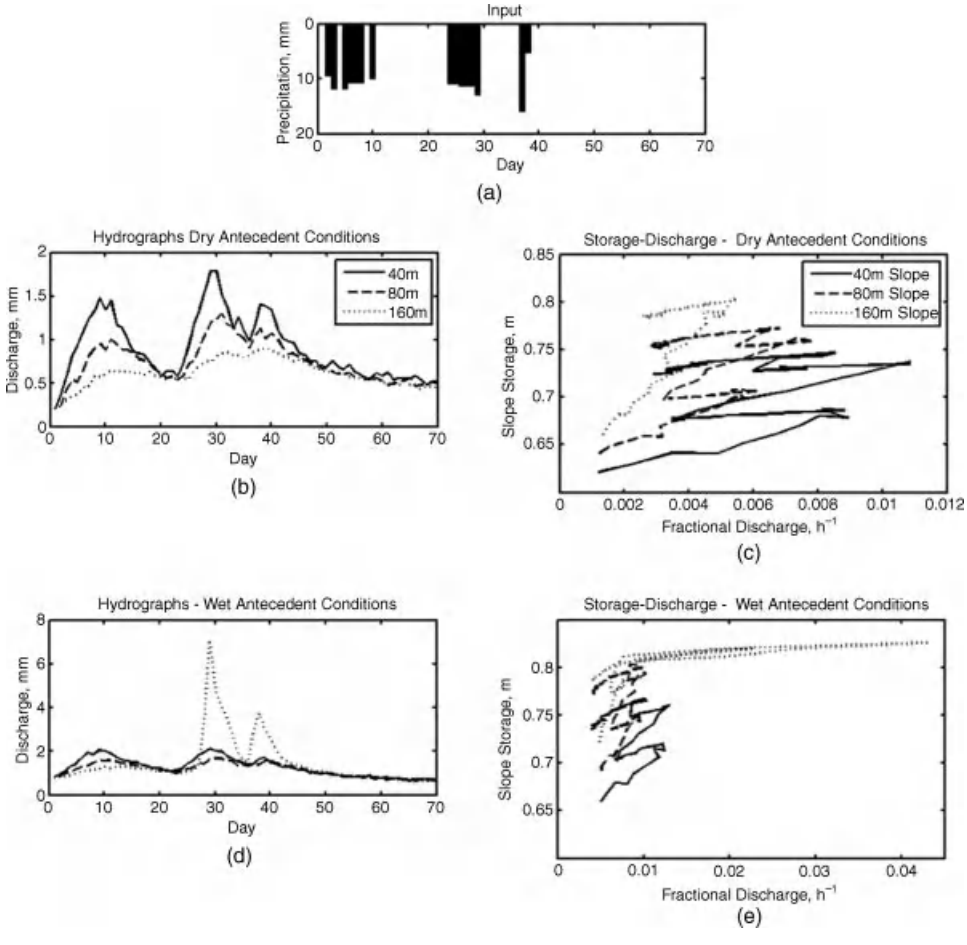


Figure 9.5 *(a) a series of rainfall events on slopes of 40, 80 and 160 m length: (b) hydrograph and (c) storage flux relationships under dry antecedent conditions; (d) hydrograph and (e) storage flux relationships under wet initial conditions; both sets of results were produced using the MIPs random particle tracking simulator on slopes of 2.86 degrees, with a constant soil depth of 1.5 m and a constant hydraulic conductivity at the soil surface of 50 m/day, declining exponentially with depth (calculations by Jessica Davies).*

has been increasingly used to represent the complexity of flows in a variety of situations, including film animations of breaking waves. SPH is computationally demanding in that it requires the interactions between neighbouring particles to be taken into account (which means that it is necessary to keep track of which particles are neighbours), but programming techniques are being developed to allow for this, including on the type of fast parallel processors used in graphics cards. SPH has been applied to free surface flows by Monaghan (1994), Rodriguez-Paz and Bonet (2005) and Roubtsova and Kahawita (2006); to debris flows by Rodriguez-Paz and Bonet (2004); and to dispersion in porous media by Zhu and Fox (2002) and Tartakovsky *et al.* (2007).

These particle-tracking approaches to representing fluxes of water in a heterogeneous flow domain are very flexible but there remains the issue of what velocity distributions should be used to represent the

fluxes in an arbitrary surface or subsurface flow domain. We would expect this to depend on the nature of the soil horizons, the spatial heterogeneity of the soil characteristics, the degree and connectivity of any macroporosity, and other complexities of the perceptual model for a particular site (see Section 1.4). Where there are observations of actual discharges then some calibration of the velocity distributions might be possible, but as with other parameterisations, we would expect there to be no unique calibration but rather some equifinality of representations. This will make it more difficult in estimating what velocity distribution characteristics might be suitable for an ungauged site, but the same problem arises in any modelling framework (see Chapter 10). Thus, it would be advantageous if other constraints on possible parameterisations could be imposed.

9.7 Catchments as Complex Adaptive Systems

There is an intellectual movement in hydrology that aspires to a grand unified theory of catchments as complex adaptive systems. This movement has its origins in the classic paper of Horton (1945) on the evolution of hillslope and catchment structures as a result of erosion and transport of sediments in surface runoff. The work was given impetus by Pete Eagleson in his book on dynamic hydrology and papers on the interaction of soil moisture and vegetation (Eagleson, 1978, 1982; Eagleson and Tellers, 1982) and later by Ignacio Rodriguez-Iturbe and many colleagues, work that has been summarised by Rodriguez-Iturbe and Rinaldo (1997) and Rodriguez-Iturbe and Porporato (2004). Modern interpretations of this theme, including taking account of the anthropomorphic effect on hydrological and ecological systems, have been expressed by, for example, Wagener *et al.* (2010) as the "future of hydrology". This is to evaluate catchment systems as complex adaptive systems, with co-evolving soil, vegetation and water components with or without significant human influence. A particular driving force in this effort has been to borrow concepts from nonlinear dynamics and complexity theory such as self-organisation, criticality, complex trajectories in phase space, etc. (e.g. Kumar, 2007).

I have long been sceptical of such an effort. This is not because I do not think of catchments as complex adaptive and co-evolving systems. Far from it, catchments exhibit their history both before and after significant human influence, only too obviously in many cases. But providing a meaningful unifying theory with predictive capability has two main issues. One is that the history of a catchment is unknowable; it is lost in the past. What we see at the present time is the result of millenia of different external forcings which might have had different *relaxation times*, that is the periods over which their impacts can be identified. In some cases, such as a storm event on a badland landscape, the relaxation time might be rather short, not much longer than the event itself. A significant flood might have a relatively short relaxation time in humid temperate landscapes (e.g. see the Anderson and Calver (1977) study on the 1952 Lynmouth flood in Devon, UK, or the different types of response to two major floods on the Plynlimon catchments described by Newson, 1980). In other cases, such as in a glaciated landscape, the relaxation time may be much longer. The flow pathways in the catchment may still be dominated by the results of periglacial and glacial processes of 10 000 years or more ago and have not necessarily changed much since. Some relaxation times may be even longer, such as the relict soils that are found in some landscapes, for example in the deep lateritic soils of western Australia which pre-date the Pleistocene (even if more recent events such as deforestation for cereal production has had a dramatic impact on the hydrology over large areas).

We cannot know very much detail about the patterns of external forcing that "evolved" the present landscape, nor about the "initial" conditions at the point of departure for evaluating that evolution. This problem has been recognised in geomorphology for decades. Bill Culling (1957) borrowed the word "equifinality" from Ludwig von Bertalanffy (1951, 1962) to describe the logical multiplicity of potential explanations of the evolution of landforms when their true initial conditions and history development are unknowable. Multiple theories of that development might be consistent with the landforms as observed

now. Other theories might be inconsistent, but that could just be because they have used the wrong initial or external forcing conditions over the period of evolution considered. The concept is very similar to equifinality of rainfall–runoff models in explaining catchment responses, though it should be pointed out that Culling (1987) later reconsidered the concept to intepret it as a vague and transient concept that will ultimately be subsumed into the well-defined apparatus of abstract dynamical systems. In this interpretation, he suggested that geomorphological systems are nonlinear and subject to random forcings of events of different magnitudes. Similarly to other nonlinear systems, they should be expected to show significant sensitivity to initial conditions and random perturbations. He distinguishes between equifinality *sensu strictu*, where a perturbed system will eventually return to its original form, and weaker forms of equifinality which imply only persistence of some property (see the discussion in Beven, 1996c).

A second issue is the uniqueness of catchment systems (see Beven, 2000). Even nearby catchments that appear to be similar in geology, soils, and vegetation show different hydrological responses (though they might be more similar than another nearby set of catchments that have developed on a different geology, with different soils and vegetation). Where the development of a catchment has been subjected to a catastrophic event (such as glaciation and the rather erratic deposition of glacial moraines), such differences might be marked. This uniqueness of place, arising from natural variability in initial conditions and the history of climatic forcings, is one of the issues that makes the ungauged catchment problem so hard in hydrology (see Chapter 10). It also makes it difficult to provide a unified theory in any meaningful way because, to be useful in practical applications, any such theory requires auxiliary conditions to deal with the uniqueness of catchment characteristics. This is also why rainfall–runoff modelling has had to rely so much on calibration against observed variables to either optimise or (much more appropriate) constrain the uncertainty in the model parameters that reflect those characteristics.

Many hydrologists have suggested that uniqueness of place is a negative concept in that science works initially by classification of similar entities, then by developing general theories to explain the generalities underlying the classes, then by testing those theories against observations. Emphasising uniqueness undermines the possibility of generalising in this way. There is certainly more yet to learn from the classification approach, especially in the prediction of ungauged basins (see Chapter 10). The problem is, as noted in previous chapters, there is that so much about catchment hydrology that is simply unknowable and about which we can only speculate. This is true both at the small scale – the limited possibilities for knowing about the subsurface characteristics of the different hillslopes in a catchment – and at the large scale – the limited possibilities of knowing about the forcing inputs in many large catchments around the world. These limitations are important in that they control the boundary conditions for the prediction we are really interested in. That is the **next** event in the catchment of interest, with all the particularities of both catchment and event, whether it be the next flood, the next drought, or some prediction of the future impact of land or climate change. That is a different kind of science: because we are interested in the particular, the general will only take us so far (Beven, 2002b). There is, however, a positive response to uniqueness of place that we develop in Chapter 12.

9.8 Optimality Constraints on Hydrological Responses

One approach that has been proposed for providing additional constraints for the next generation of hydrological models is to invoke optimality principles. The concept suggests that over a period of time an open system will evolve so as to optimise the use of the available energy subject to the prevailing constraints. This actually has a long history in hydrological research, instigated by the seminal series of papers on the interaction of soil moisture and natural vegetation by Pete Eagleson (1982, 2002; Eagleson and Tellers, 1982; Rodriguez-Iturbe and Porporato, 2004). There have now been a variety of applications of optimality principles to hydrological systems (see, for example, recent papers by Schymanski *et al.*,

2008, 2009; Wong, 2008; Kleidon *et al.*, 2010; and Zehe *et al.*, 2010) and it has been suggested that this should be a guiding principle in constraining the model representation of hydrological processes (Schaefli *et al.*, 2011).

There are a number of optimality principles that might be invoked. These include maximising the net primary productivity of the vegetation cover, maximising entropy production, minimising energy expenditure, maximising Helmholtz free energy dissipation, maximising water use, minimising water or oxygen stress, and maximising net carbon profit (see review by Paik and Kumar, 2010). Different principles might be appropriate over different time scales and there might be conflicts between different optimality constraints. The question that then arises is whether such concepts might be useful constraints on appropriate model structures in future, especially when there will be continuing uncertainty in the forcing data.

In fact, nearly all rainfall–runoff models implement one such principle already: the principle of mass balance. The input precipitation to a model is partitioned into discharge, evapotranspiration and change of storage such that, in the model, no water is lost or gained. A check on mass balance is commonly programmed into a model, because a modeller would be worried if mass was being lost or gained, for example as the result of an approximate numerical solution to a partial differential equation (this was a problem with my own early finite element model of hillslope hydrology Beven, 2001).

The problem that then arises is that the data available on the different elements of the water balance might be subject to sufficient uncertainties that the net mass balance should not actually be zero. This might arise if the monitoring network for precipitation was not adequate to always give accurate estimates of the inputs. There might be a non-stationary bias because the raingauge sites under-represent the rain and snow inputs in high elevation zones, or because they cannot adequately capture local convective raincells relative to synoptic frontal patterns. There might be inadequate rating curve information such that the total volume of discharge is not accurately estimated. Similarly estimates of actual evapotranspiration and storage changes (if available at all) are subject to uncertainties.

The result is that it can be difficult to be sure that the data driving the model are consistent with the water balance principle. This is one reason why a number of studies have tried to estimate rainfall and discharge errors as part of the calibration process (see Section 7.8). Similar issues arise in the application of energy and momentum balances and establishing optimality principles. Even if the vegetation is really acting optimally in all conditions (and it is not clear how much information would be required about water status, root growth, nutrient status, health of the plants, and other external forcings such as fire, land management, insect infestations, etc. to verify this), then the forcing data actually might not be consistent with the application of the principle. The propensity to change and sensitivity to new forcing events might also depend on the particular sequence of events (e.g. Beven, 1981b). However, that is not to say that this approach might not sometimes be valuable. For example, Hwang *et al.* (2009) show that, in an application of the RHESSys model, the rooting depth parameters that produce an optimal fit to observed discharges also maximise net primary productivity in fitting vegetation leaf area index data. Schymanski *et al.* (2008, 2009) start with the principle that plants maximise their net carbon profit and implement a "vegetation optimality model" (VOM) that constrains the number of parameters to be calibrated in hydrological modelling. They also show how this can be tested as a hypothesis, at least during the dry season when the vegetation is constrained by the availability of water. This type of hypothesis testing can be used within an uncertainty framework. Mitchell *et al.* (2011), for example, show how a forest growth model is not consistent with a set of observations for a site in Oregon, even allowing for uncertainty in measured fluxes.

In fact, a general principle for the hydrological modeller should be to invoke all prior knowledge and principles in trying to improve the representation of catchment processes. It may, however, be very difficult to verify the application of such principles in practical applications. Indeed, it might be the case that the imposition of a principle might constrain a model in a misleading way, in much the same way as an incorrect representation of a flow process (treating subsurface flows as purely Darcian; treating overland

flows as a uniform sheet flow; treating pollution transport as a simple advection–dispersion problem, etc.) should also be considered misleading. What this should teach us is that we need to be careful about evaluating process descriptions and constraining principles as hypotheses using appropriate but uncertain data. The issue in doing so is how far the information is commensurate with model predictions and whether uncertainties might mean that some periods of data might be disinformative in inferring a good model of the system, especially when the input data are poorly specified (Section 7.17; Beven *et al.*, 2008; Kuczera *et al.*, 2010a; Beven and Westerberg, 2011).

9.9 Key Points from Chapter 9

- A new generation of rainfall–runoff models is needed to replace the concepts of the Freeze and Harlan (1969) blueprint for physically based models.
- It is suggested that a suitable framework for a novel approach already exists in the representative elementary watershed (REW) concepts of Reggiani and others, which is based on equations of mass, energy and momentum balance in discrete elements of the landscape.
- These equations, however, give rise to a *closure problem*, because they involve fluxes of mass, energy and momentum across the boundaries of the elements. Since it is difficult to show by measurement that these balance equations hold for real elements of the landscape, there will necessarily be some uncertainty inherent in defining closure schemes.
- Implementations of the REW concepts to date are not satisfactory in that they do not generally recognise the hysteretic and scale dependence required of a closure scheme.
- Recent work that might form the basis for better closure schemes is reviewed, including the multiple interacting pathways (MIPs) modelling strategy that attempts to predict both flow and residence times of water in the system within the same framework.
- It might be possible to impose on the representations other constraints than the balance equations. Recent work has suggested that a local optimality principle for vegetation to maximise net carbon production given available energy and water might be useful in constraining the prediction of actual evapotranspiration.
- In general, it is useful to view catchments as complex adaptive systems subject to different external forcings that have variable *relaxation times*.

10
Beyond the Primer: Predictions in Ungauged Basins

PUB requires the development of new predictive approaches that are based on a deep "understanding" of hydrological functioning at multiple space–time scales. Indeed, PUB will herald a major change of paradigm in surface hydrology from the present one based on "calibration" to a new exciting one based on "understanding".

PUB Science Plan, 2003

Through its sharp focus on predictive uncertainty, PUB will therefore adopt and foster a self-critical approach to hydrologic predictions by addressing what is not known or not understood, and emphasizing the need for empirical exploration and explicit attempts to generate, falsify and validate new ideas and new forms of knowledge. Also, the focus on predictive uncertainty will enable hydrologists to go beyond quick fixes, to search for new and innovative solution approaches, and to seek knowledge and understanding of natural processes beyond the immediate problem-solving needs of construction, environmental management or regulation.

PUB Science Plan, 2003

10.1 The Ungauged Catchment Challenge

One of the great unsolved challenges in hydrology is the accurate simulation of a catchment without any observational data with which to calibrate a hydrological model, i.e. an ungauged basin (Wagener *et al.*, 2004; Vogel, 2005; Blöschl, 2005). This is also called the *regionalisation problem* in hydrological modelling. It is so much easier to get reasonable predictions if some discharge data are available to give an indication, at least, of the response of a catchment and to allow the calibration of model parameter values. Even if the hydrological modeller has experience of applying a particular model in many different

Rainfall–Runoff Modelling: The Primer, Second Edition. Keith Beven.

catchments, it is much more difficult if parameter values have to be estimated *a priori*, with no guarantee of equally good performance in prediction (of course if a catchment is really ungauged then there will be little reason to question the predictions, but researchers often play the game of pretending a catchment is ungauged to test different prediction methods as if that catchment was ungauged).

The ungauged catchment problem has a long history. Early methods were mostly based on regressions of model parameter values or runoff coefficients determined for gauged catchments against variables representing the characteristics of those catchments. Once the regression equations had been developed then they could be used to estimate the parameter values for the ungauged catchments of interest. This is a purely empirical approach to the regionalisation problem, but it would be nice to think that we might also be able to bring some hydrological science to bear on the problem. After all, a science that has to rely on calibration for every site that its theory is applied to cannot be said to be mature (even if there are very good reasons for these difficulties). This was recognised in the late 1990s when the International Association of Scientific Hydrologists (IAHS) started to develop plans for the predictions in ungauged basins (PUB) project. The PUB project has stimulated a lot of research activity, including attempts to develop methods based on the classification of catchment hydrological responses and the investigation of whether the problem could be solved by a minimum amount of gauging of the ungauged catchment. These approaches are discussed in this chapter.

10.2 The PUB Initiative

The prediction of ungauged basins (PUB) project is an initiative of the International Association of Hydrological Sciences (IAHS). Plans for the decade-long project were approved in 2001 and it started formally in 2003 (Sivapalan, 2003).The science plan for PUB (see www.iahs-pub.org) recognises that predicting the response of ungauged basins involves inherent uncertainties and that improving the predictions is primarily an exercise in constraining uncertainty, which may involve the collection of new types of data as well as improving the scientific basis for hydrology.

The aims of PUB are summarised as:

- to advance the ability of hydrologists worldwide to predict fluxes of water and associated constituents from ungauged basins, qualified by estimates of predictive uncertainty;
- to advance knowledge and understanding of climatic and landscape controls on hydrological processes at all scales, in order to constrain predictive uncertainty;
- to demonstrate the value of data for hydrological predictions and provide a rational basis for future data acquisitions, including alternative data sources, by quantifying the links between data and predictive uncertainty;
- to advance the scientific foundations of hydrology and provide a scientific basis for sustainable river basin management;
- to promote and encourage capacity-building activities in the development of appropriate scientific knowledge and technology to the areas and communities where it is needed most.

Two primary scientific targets were identified:

- Target 1: Examine and improve existing models in terms of their ability to predict in ungauged basins through appropriate measures of predictive uncertainty.
- Target 2: Develop new, innovative models to capture space–time variability of hydrological processes for making predictions in ungauged basins, with a concomitant reduction in predictive uncertainty.

The PUB decade has caused a significant rethink about how the ungauged basin problem should be addressed and the development of some new methodologies (Wagener and Montanari, 2011). They are outlined in Section 10.6.

10.3 The MOPEX Initiative

A recent international experiment has attempted to search for generalities in regionalisation approaches using a large database of long-term (10 years or more) catchment rainfall–runoff observations from different countries (although mostly daily time step data were used in this case). The primary aim of the model parameterisation experiment (MOPEX, see www.weather.gov/oh/mopex/mo_welcome.htm) is to develop techniques for the *a priori* estimation of the parameters used in land surface parameterisation schemes of atmospheric and hydrological models that can be applied widely in ungauged basins. A good overview of the project is provided by Duan *et al.* (2006). Particular attention was paid to having adequate precipitation data for the chosen basins, in terms of raingauge densities in a catchment.

Within the MOPEX project a number of different rainfall–runoff and land surface parameterisation models were applied to different groups of catchments. In general it was found that using parameter values estimated *a priori* none of the models could reproduce the observed discharges for all the catchments when treated as ungauged. Results were greatly improved if the model parameters were calibrated and tested in prediction using a split sample approach, although in the study reported in Duan *et al.* (2006) no uncertainty in the estimated parameter values was considered. This suggested that methods for the prior estimation of parameter values could be greatly improved given more hydrologically relevant information about catchment characteristics. It is worth noting that the performance of the rainfall–runoff models were mostly much better than the models that had been developed only as land surface parameterisations, and that the performance of a combined model output was generally better than the performance of any of the individual models.

A number of other studies have used the MOPEX data set as a test bed for estimating the responses of ungauged basins using different rainfall–runoff models and methodologies for estimating parameters (Ao *et al.*, 2006; Huang and Liang, 2006).

10.4 Ways of Making Predictions in Ungauged Basins

The choice of a way of making predictions in ungauged catchments in part depends on the purpose of a particular study. Estimates of discharges for ungauged catchments are required for a wide variety of purposes, including flood defence design, land planning, water resources, ecological, hydropower and other decision-relevant applications. Prediction of the impacts of future change also involve aspects of the ungauged basin problem since, by definition, future responses under different inputs or different land surface characteristics cannot be gauged (see also Chapter 8).

Regionalisation methods have been developed primarily for the estimation of the characteristics of flood frequency distributions, for estimation of flow duration curves, and for estimation of the parameters of hydrological models at ungauged sites. The first two are also, in essence, a question of estimating parameter values: in both frequency estimation and flow duration curve estimation the parameters are those of a frequency distribution. Recent work has started to look at the estimation of mean residence times for water in catchments (see also Chapter 11). There is also significant interest in estimating the parameters of *land surface parameterisations* (LSPs) of atmospheric models. The land surface hydrology can have an important effect on latent and sensible heat fluxes as a lower boundary condition for atmospheric models and, as noted in Section 8.9, has generally been represented rather poorly from a hydrological

perspective. This is also a problem of regionalisation since the land surface parameters at the atmospheric grid scales cannot be inferred from either direct measurement or calibration since most of the land surface is effectively ungauged. There is now a global network of several hundred observation sites for land surface fluxes, also including carbon dioxide fluxes (FLUXNET, see Appendix A) but these are local scale measurements that reflect fluxes in the effective fetch area around each site. So some regionalisation methodology is still required to provide parameters at large scales.

There is a relatively small number of common methods for regionalisation. These can be differentiated between those that make estimates at the catchment scale, and those that try to incorporate knowledge from geographical information systems about the distributed nature of catchment characteristics. The catchment scale methods generally use regression against catchment characteristics or define donor catchments or pooling groups using similarity measures based on summary measures of catchment characteristics (Vogel, 2005). The distributed methods generally relate model parameter values to overlays of soil, land use and topographic characteristics.

10.5 PUB as a Learning Process

The regionalisation problem would not be so difficult if it was possible to find a "similar" catchment to the ungauged site of interest in the same physioclimatic region. Unfortunately, even catchments that look superficially similar often exhibit rather different hydrological response characteristics. Catchments vary in their characteristics in so many ways (in area, in rainfall characteristics, in geology, in topography and geomorphological history, in soil characteristics, in vegetation and land management, in degree of urbanisation, ...) that close similarity should not be expected (especially since the sample of gauged catchments is generally very limited). However, it is worth noting that where the ungauged site of interest is within a catchment with one or more existing gauges, simply scaling the observations at those gauges will be difficult to beat, unless there are good physical reasons to suspect that this will not be successful because of a striking change in geology, urbanisation or strongly different rainfall inputs. This is the first approach recommended by the UK *Flood Estimation Handbook* (IH, 1999).

This is one argument for using a distributed parameters approach to regionalisation since the pattern of parameters can then reflect the spatial patterns of characteristics and processes within a catchment, at least partially. There remains a problem that we do not know enough about the nature of the processes at any point, particularly the subsurface processes, to be confident about whether the responses are being properly represented. This is because the small scale variability in hydrological responses is sufficiently complex that, even given the same rainfall regime, soil type, vegetation type, upslope contributing area, slope angle and aspect, it is difficult to be sure that the balance of surface and subsurface flow processes will be the same. This might because of fracturing in the bedrock, highly variable soil depths, or modifications by people in ditching and installing drainage. This is the problem of "uniqueness of place" (Beven, 2000). It is evident in the fairly common experience that some catchments appear to be "outliers" in regionalisation methods (e.g. Wagener and Wheater, 2006). It implies that we should expect local predictions to be uncertain, even if distributed parameters are used, because the use of generalised parameter values are expected to give only approximate results locally.

In extreme cases, of course, the generalised parameter values might give predictions that are (locally) quite wrong in some of the predicted catchments (e.g. Lamb and Kay, 2004). Where this is made known, however, action can be taken to try to improve the local model representations (Beven, 2007; see also Chapter 12). The philosophy that has developed over the course of the PUB decade has been to treat the ungauged basin problem as a learning process in which initial estimates of how a catchment responds are gradually improved, and uncertainty contrained, as new information becomes available (Sivapalan, 2003; Buytaert and Beven, 2009; Wagener and Montanari, 2011). Associated with this philosophy has been a change from an approach centred on model parameters to one based more on trying to estimate the

hydrological characteristics of a catchment directly. If model predictions are still required (as they often are) the estimates of the hydrological response characteristics can then be used to constrain suitable model parameters (see Section 10.8).

There is another approach to defining model parameters for an ungauged catchment within such a learning process. If an application justifies the cost, then it might be possible to provide some point discharge or other response measurements, or a short period of measurements, to improve the prior estimates of the hydrological response (see Section 10.10). This is consistent with the PUB approach of using any information available to constrain uncertainty in the predictions of the site of interest.

10.6 Regression of Model Parameters Against Catchment Characteristics

The basis for any regionalisation method based on direct estimation of model parameters is a sample of parameter values of interest determined for gauged catchments. Most regionalisation strategies then use a method of relating those parameter values to the physical characteristics of the gauged catchments. The methods are tested against other gauged catchments, treated as if they were ungauged. In the past, parameter values were generally fitted to the observations from the gauged catchments by optimisation. Thus, each gauged catchment would be associated with only one value of a particular parameter. This allowed the prediction problem to be set up as a statistical regression, with the parameter values treated as dependent variables and one or more catchment characteristic measures as independent variables. Generally, a form of multiple regression is used in which many different independent variables are tried and only those that make a significant contribution to explaining the variance of the dependent variable are retained. There are statistical tests that allow the significance of the contribution to be assessed, generally under an assumption that the residuals in fitting the dependent variable are independent and distributed as a normal random variate.

In the UK, this was the approach taken by the *Flood Studies Report* (NERC, 1975) and its later revision as the *Flood Estimation Handbook* (IH, 1999). Both these reports provided regression equations for flood frequency distributions and for rainfall–runoff model parameters (a model based on unit hydrograph concepts was used in both, see Kjeldsen, 2007). This type of regression method can also be used to estimate the parameters of rainfall–runoff models directly. This approach was first used in the UK with a unit hydrograph model by Eamonn Nash (1960), who was later involved in producing the *Flood Studies Report*. More recent studies of this type have been used in the USA (Abdulla and Lettenmaier, 1997a; Fernandez *et al.*, 2000); in Australia (Boughton and Chiew, 2007); in France (Oudin *et al.*, 2008); in Belgium (Heuvelmans *et al.*, 2006); in Germany (Hundecha *et al.*, 2008); in Austria (Merz and Blöschl, 2004); in Sweden (Seibert, 1999; Xu 1999); in South Africa (Kapangaziwiri and Hughes, 2008); in China (Xie *et al.*, 2007) and elsewhere for both flood frequency and hydrograph model parameters.

There are a number of problems with this approach. The first is the problem of optimisation of the model parameters at the gauged sites. As we have seen in Chapter 7, the optimised values of the parameters may not be robust to uncertainty in the calibration data, calibration period, performance measures, optimisation method or interactions between the parameters. The possibility of parameter interactions within a model structure is often ignored in this approach, with regressions against catchment characteristics being carried out independently for different parameters, without any allowance for uncertainty in the estimation of those parameters. Finally, although it is possible to use some hydrological reasoning in deciding which catchment characteristics might be related to different parameters, the regression method is purely statistical. As an example, Sefton and Howarth (1998) produced regression relationships for the parameters of the IHACRES model for catchments in England and Wales. One such relationship,

determined by multiple linear regression, takes the form:

$$\frac{1}{c} = 7.7UPLAND + 12.9DECID + 7.4TILLED + 9.0URBAN + 7.3OTHER - 0.4CLASSB - 0.8REHPMN \quad (10.1)$$

where c is the proportion of rainfall contributing to catchment storage; $UPLAND$ is the % area of heath, moor and bracken land use types; $DECID$ is the % area of deciduous and mixed forest land use types; $TILLED$ is the % area of arable crops; $OTHER$ is the % area of the remaining three land use types; $CLASSB$ is the % area of semi-permeable mineral soils (no groundwater) and $REHPMN$ is the mean relative humidity averaged over the period 1961–1991. The relationship was developed from calibrating the model to 60 different catchments in England and Wales. The correlation coefficient for this relationship was 0.61, implying significant uncertainty in estimating the value of $\frac{1}{c}$ for any individual catchment. Similar equations for other parameters of the model had better and worse fits to the data and two of the parameters showed a strong correlation across the 60 gauged catchments.

The independent variables in Equation (10.1) were chosen from a list of 10 morphometric variables, five soil descriptors, eight land use types and seven climate variables. The statistical fitting techniques allow the identification of independent variables that add a significant contribution to the explanation of the observed variations in the dependent variables. Any variables that do not provide a significant contribution are not included in the equation. The technique used to derive this type of equation is a linear regression analysis. Any nonlinear relationships must be achieved by transforming either the dependent or the independent variables. In one of the other predictive equations in the Sefton and Howarth paper, log transforms on all the variables were used before performing the analysis. This type of analysis is used widely in the estimation of model parameters or discharge characteristics within a region for ungauged catchments. What is often not recorded is what other transforms or independent variables were tried and rejected in favour of the final published equation. Variants on this approach include the identification of predictive equations for model parameters by optimisation over the full set of gauged catchments, rather than for individual catchments (Fernandez *et al.*, 2000; Hundecha and Bárdossy, 2004; Götzinger and Bárdossy, 2007; Samaniego *et al.*, 2010); the use of geostatistics to take account of spatial correlation in parameter values (Merz and Blöschl, 2004; Parajka *et al.*, 2005; Viviroli *et al.*, 2009); and the use of clustering methods or self-organising maps rather than regression for relating catchment characteristics to model parameters (Szolgay *et al.*, 2003; Rao and Srinivas, 2006; Di Prinzio *et al.*, 2011).

It would be nice if such relationships always made good hydrological sense. This is sometimes hard to extract from this type of regression equation. In Equation (10.1), one wonders what processes an additive linear function of the mean relative humidity variable is acting as a surrogate for? Although this is a widely used approach to regionalisation, we do not consider it further here. The PUB initiative was, at least in part, a recognition of the limitations of this approach. We make just one final comment which is to note that since these regression relationships are statistical in nature, then it should be possible to derive a standard error of estimate for each parameter value estimated (even if the covariance structure cannot be estimated if the regressions for each parameter are treated independently). Thus, at least the uncertainty in the estimates for the ungauged site arising from the regression could be assessed in the model predictions. This is rarely done (but see the work of Lamb and Kay (2004), Wagener and Wheater (2006), Kay *et al.* (2007) and Calver *et al.* (2009) for exceptions). It also does not account completely for different sources of uncertainty in the regionalisation process (see Jones and Kay (2007) for a more complete statistical treatment of the parameter generalisation problem) but it is at least a start when any decision depending on the modelling of the ungauged site might be sensitive to uncertainty.

10.7 Donor Catchment and Pooling Group Methods

While the regression equation approach to estimating catchment scale parameters has been widely used, the UK *Flood Estimation Handbook* (FEH) did not recommend it as the technique for regionalisation for most cases. This was a recognition of the high uncertainties (and occasional inaccuracies) associated with the regression estimates. Instead, where no nearby gauging station is available, an approach based on a *pooling group* of gauged catchments was preferred. The pooling group is chosen on the basis of proximity to the ungauged catchment of interest of potential donor gauged catchments. The concept is hydrological rather than simply statistical: catchments that are similar in their characteristics should be similar in their hydrological responses. The problem, of course, is that catchments can vary in so many ways that defining similarity is always going to be difficult and there are many different similarity measures that could be used. The FEH uses a database of physical characteristics of catchments (actually the same variables as used in the statistical regression approach) to define similarity as proximity in terms of a Euclidean distance measure in the space of these physical characteristics. A Euclidean distance in a k dimensional space between points i and j is defined as:

$$E = \sqrt{\sum_{k} \left(x_{j,k} - x_{i,k}\right)^2} \tag{10.2}$$

The Euclidean distance is often normalised by dividing the contribution on each dimension by the variance of the values x on that dimension. Other distance measures have also been used, such as the Mahalanobis distance which allows for the full covariance of values on different dimensions (Cunderlik and Burn, 2006). Some catchments are more similar than others, of course, and it is a somewhat subjective decision as to which catchments are included in the pooling group. It is often the case that catchments that are in the same physioclimatic region are closest to the study site but there may also be distant catchments that have similar characteristics and (we would expect) hydrological responses. The contribution of each donor catchment to the estimation of the parameters for the study site can be weighted, depending on the degree of similarity identified from the catchment characteristics.

This approach raises two major issues. The first is what sort of similarity measure to use; the second how to allow for uncertainty in the transfer of information from gauged donor sites to the study site. There are a variety of similarity and dissimilarity measures that have been used in the past (see review by Wagener *et al.*, 2007) but there is, as yet, no common agreement on what catchment characteristics should be used in defining similarity and what measures should be used. In fact, Oudin *et al.* (2010) question the basis for this approach, which depends on an assumption that hydrological similarity is equivalent to physical similarity (at least if the right physical characteristics are included in the analysis). They assessed hydrological similarity in terms of model parameters and physical similarity in terms of catchment characteristics and showed that there was only overlap between the two definitions for 60% of the 900 catchments in France and England included in the analysis. For the other catchments, the subsurface properties of the catchment, not well represented by the set of physical measures in the analysis, were a major control on the hydrological response. Reichl *et al.* (2009) also tried to optimise the weights on different physical characteristics in defining a distance measure. This gave some improvement over standard distance measures but did not give good predictions for all the test catchments, and they suggested that the weights were dependent on the sample of catchments used.

One response to this is to try to define groups of catchments or regions that might be considered to be hydrologically similar before carrying out a regionalisation exercise. This necessarily cuts down the number of potential donor catchments in each grouping but can mitigate the problem of including catchments with quite different rainfall regimes or geology and baseflows within the same regionalisation exercise. There has been quite a lot of research on trying to define homogeneous groups or regions of catchments, but the donor catchment or pooling group approaches to regionalisation centre this on the

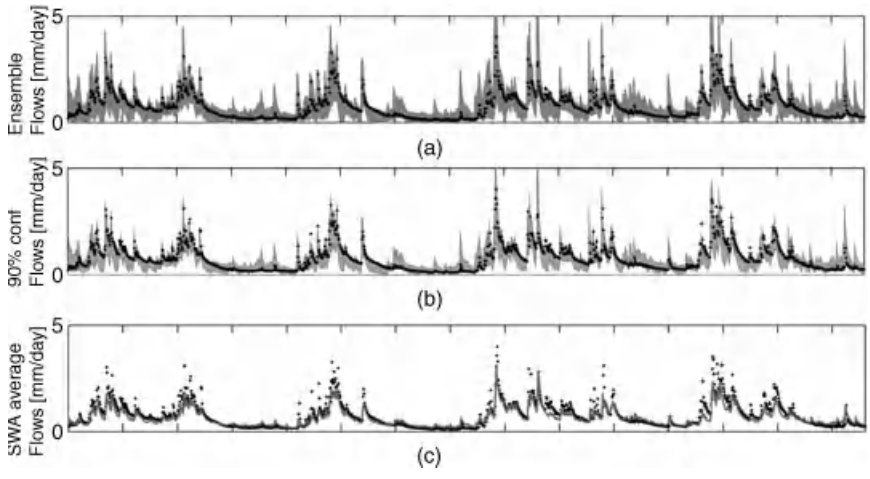


Figure 10.1 *Hydrograph predictions using the PDM model and parameters estimated by the ensemble method: (a) the 10 best parameter sets from the 10 most similar catchments; (b) 90th percentile of the same ensemble; (c) similarity weighted ensemble (after McIntyre et al., 2005, with kind permission of the American Geophysical Union).*

ungauged site of interest rather than the available gauged sites. Thus, for a new ungauged site, a group of gauged catchments is sought that are close in characteristics to that site; its "nearest neighbours" in terms of whatever measures are being used to assess similarity.

But we should expect the estimates of rainfall–runoff model parameters for the pooling group donor sites to be uncertain, and the transfer of information from those sites to the ungauged site should add to that uncertainty. McIntyre *et al.* (2005) have attempted to take these uncertainties into account by allowing for uncertainty in estimating parameter sets for the PDM model (Section 6.2) for the gauged catchments using the GLUE method of Section 7.10, then weighting the parameter sets from donor catchments according to similarity with respect to physical characteristics. Similarity was defined rather simply in this case in terms of the Euclidean distance in the space of the catchment area, standard annual average rainfall (SAAR) and an estimated baseflow index based on the fraction of HOST soil types in the catchment, as used in the *Flood Estimation Handbook* (IH, 1999). Log transforms of area and SAAR were used to reduce the skew in these variables. The similarity measure was then transformed into a weighting function that was applied to the behavioural parameter sets from each donor catchment in predicting the response of the ungauged catchment. In general, the method performed better than an estimation of model parameters based on regression against catchment characteristics especially during recession periods, but the full weighting method tended to underestimate the prediction uncertainty particularly for hydrograph peaks. An ensemble of the 10 best parameter sets from the 10 most similar catchments was better at bracketing the observed discharges for the ungauged catchments (Figure 10.1).

10.8 Direct Estimation of Hydrograph Characteristics for Constraining Model Parameters

Another approach to assessing the uncertainty associated with the regionalisation process is provided by Samaniego *et al.* (2010) using a new type of similarity measure based on copula distributions of discharge

data. This study is a nice lead in to the next approach to the regionalisation problem, since it makes use of the hydrograph characteristics at the gauged sites directly.

Given the limitations shown by the studies that attempt to regionalise model parameter values, recent work on regionalisation has tried to work to estimate the hydrograph characteristics of an ungauged catchment directly. This approach actually has a long history, but only in the specific circumstances, mentioned earlier, when the site of interest is not too far upstream or downstream from a stream gauge. Then, unless there are obvious differences between the sites, the recommended approach to regionalisation is simply to scale the observed hydrograph by catchment area. For such sites, this is the approach to beat (IH, 1999; Merz and Blöschl, 2004). The nearest gauge, however, may not actually be the best gauge to use (though see the work of Zhang and Chiew (2009) for an example where spatial proximity performed slightly better than physical similarity as defined by catchment characteristic measures). Archfield and Vogel (2010) suggest a method for deciding which nearby gauge might give the best results.

A more sophisticated version of this technique has been recently provided by Parada and Liang (2010). They treat the ungauged catchment discharge as a hidden variable to be estimated on the basis of uncertain observations (in this case, the discharges from two nearby catchments and predicted discharges for the study catchments using default parameter values in the VIC-3L model, see Box 2.2). The hidden variable is expected to be related to the observations by some arbitrary nonlinear transform. They then frame the problem as a recursive variational Bayesian Kalman Filter, using kernel functions to linearise the estimation problem. The procedure leads to better predictions in terms of bias, correlation and root mean square error than either the default model or the donor catchment discharges. Best results for low flows are obtained by using the $\log_{10}$ transformation of the observations. The method can also produce estimates of the uncertainty in the predicted discharges, though the authors do not show this in their hydrograph plots.

But the site of interest may not be close to a stream gauge, of course, so that it then becomes more difficult to estimate the hydrograph characteristics. Yadav *et al.* (2007) suggest that rather than trying to estimate model parameters for such sites, it might better to estimate hydrograph characteristics directly and then use those characteristics to constrain the uncertainty in using a model to represent the study catchment. The question then is what indices of hydrograph behaviour might most easily be extrapolated and which might be most useful to constrain model uncertainty. Such indices are used quite widely in hydroecological studies. Olden and Poff (2003) compiled 171 different indices from previous literature. Shamir *et al.* (2004) show how selected indices can be useful in hydrological model calibration since they can reflect different components of the hydrological response. Bárdossy (2007) used some simple indices of the discharge annual mean and variance to evaluate the predictions of model parameter sets from donor catchments when applied to ungauged sites.

Yadav *et al.* (2007) took this further in using 39 different hydrological indices (relating to aspects of high flow, average flows, low flows, rate of rise, event frequencies, timing and climate) on 30 UK gauged catchments chosen to reflect a range of areas, annual rainfalls, baseflow index and flow duration curves. They showed that there is redundancy of information between some characteristics which are highly correlated. They then developed (independent) step-wise regression equations to relate chosen hydrograph indices to the physical characteristics of the gauged catchments. Some 19 indices were then estimated for test catchments treated as ungauged, together with their estimation uncertainty. This gave a basis for evaluating runs of a rainfall–runoff model, in this case, the five-parameter HYMOD conceptual model. Monte Carlo realisations of the parameters were generated and the characteristics of the model outputs compared with the estimated hydrograph indices. High flow event frequency, runoff ratio and the slope of the flow duration curve provided strong regionalisable constraints while producing reliable predictions in the sense of bracketing the observed flows. Not all the regression equations, however, gave good predictions of the equivalent hydrograph indices for the test catchments; if too many constraints were used in testing the model realisations, all the models tried could be rejected. Zhang *et al.* (2008) extended this work to use a directed search algorithm that would be more efficient at finding model

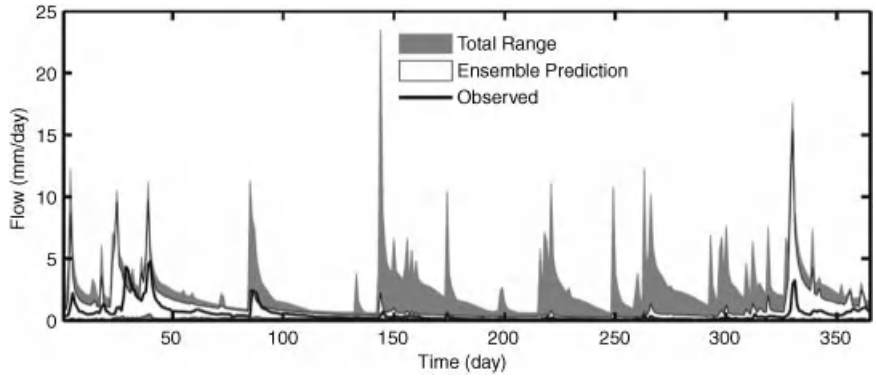


Figure 10.2 *Simulation results for the Harrowdown Old Mill catchment (194 km^2) treated as if ungauged with observed and ensemble prediction streamflows (the grey range is the unconstrained ensemble; the white range is the ensemble constrained by predicted hydrograph indices) (after Zhang et al., 2008, with kind permission of the American Geophysical Union).*

parameter sets that would satisfy multiple constraints. An example of their resulting ensemble of model predictions for a catchment treated as ungauged is shown in Figure 10.2.

The advantages of this approach are that it does not depend on a particular model structure, nor on the particular calibration methodology of the gauged catchments. It also focuses some attention on the uncertainty in the hydrograph indices when they are used in constraining the model predictions. This study might, however, have been too ambitious in its assumptions that single model structures for both the regression equations and for the hydrological model could be applied everywhere. It might be more interesting to see whether differences within a group of similar catchments, analogous to the pooling group concept, could be distinguished in this way. Bulygina *et al.* (2009) have, in fact, used an analogous approach to predict how afforestation and soil degradation might impact on hydrological indices and model parameters for a gauged catchment.

It is worth noting that there have also been attempts to regionalise complete flow duration curves for ungauged sites using different methods (Yu and Yang, 2000; Holmes *et al.*, 2002; Yu *et al.*, 2002; Castellerin *et al.*, 2007; Li *et al.*, 2010). These might also be used as a constraint on model parameterisations, though the uncertainty in the discharge estimates should be taken into account in doing so (e.g. Blazkova and Beven, 2009; Westerberg *et al.*, 2010b).

10.9 Comparing Regionalisation Methods for Model Parameters

With so many different methods available for the estimation of rainfall–runoff model parameters on ungauged catchments, it is not surprising that there have been a number of studies comparing different methods. These reveal, as noted above, that it is difficult to beat the method of scaling the measurements for an upstream or downstream gauge, for those special cases where such observations are available (Merz and Blöschl, 2004). More generally, however, the results of such comparisons have been equivocal (Parajka *et al.*, 2005; Kay *et al.*, 2006b; Zhang and Chiew, 2009), with some suggestion that a combination of methods might perform better than any individual method (e.g. Viviroli *et al.*, 2009). Oudin *et al.* (2008) note that none of the methods compared on a sample of 913 catchments in France performed well relative to full model calibration. Götzinger and Bárdossy (2007) report that the four methods in their study using a distributed form of the HBV model failed to perform well in subcatchments with karst geology

and those in which the flows were heavily modified or regulated. These are types of catchment that, elsewhere, have generally been excluded from regionalisation studies but which might be important in specific practical applications.

These results emphasise the real difficulty in addressing the ungauged catchment problem in hydrology. The final report of the PUB initiative has still to appear but, while considerable progress has been made, it will certainly not declare that the problem has been solved either by improved hydrological theorising or by common agreement on a regionalisation strategy. Approaches based on regionalisation of hydrograph indices perhaps show the greatest promise, particularly as they provide a means for allowing for uncertainty in the models that might be appropriate for an ungauged site. However, it might well be that the best near-term solution to the ungauged catchment problem is a short field campaign of discharge measurements (see Section 10.11).

10.10 HRUs and LSPs as Models of Ungauged Basins

There is a quite different approach to model parameterisation in ungauged catchments that has arisen out of the use of hydrological models as land surface parameterisations (LSPs) in global atmospheric models. The requirement is that the whole of the land surface of the globe must have some form of representation of the surface water and energy balances, albeit at the rather coarse resolution of global or regional atmospheric models. This includes all those large parts of the global land surface that are effectively ungauged at these scales. The parameters of the LSPs must therefore be derived in other ways, depending on what information might be available.

In some cases, models are provided with default parameter sets that have been derived from calibration exercises in some gauged catchments. This is the case for the VIC model of Box 2.2 (e.g. Abdulla and Lettenmaier, 1997b; Li *et al.*, 2009) and the SWAT model of Box 6.2. The default parameters depend on defining soil type and land management classes from other information, such as remote sensing imaging (e.g. Lakshmi, 2004). LSPs are generally used in a way that reflects the fractions of different classes (or "tiles") in a larger atmospheric model grid square. There is, therefore, some commonality with the hydrological response unit (HRU) models discussed in Section 6.6, albeit that the areas represented by particular parameter sets as homogeneous tend to be much larger when used as LSPs. In fact, the only recognition of spatial heterogeneity in hydrological responses recognised by operational LSPs is in terms of these fractional tiles. Thus, the representation of surface hydrology in operational LSP models can be considered totally inadequate (e.g. Beven, 2011).

More interesting is the question of whether this form of prior estimation of default model parameters can be successful in reproducing the response of ungauged catchments at finer HRU scales, especially when most such studies totally neglect the uncertainty inherent in estimating such parameters (see the discussion of pedotransfer functions in Box 5.5). Effectively, this is similar to a classification approach, where the classification is based on the soil and land use classes used to define the default parameters. Most studies have shown that some form of local conditioning results in improved performance over using default estimates of parameter values, even in ungauged catchments (Dornes *et al.*, 2008; Parada and Liang, 2010). Buytaert and Beven (2009) argue that this should be part of the learning process. They showed how prior estimates of parameter values taken from nearby catchments, with some (uncertain) speculation about the impacts of land use change, could provide surprisingly good predictions but that uncertainties could be significantly constrained as discharge observations became available for the new site.

10.11 Gauging the Ungauged Basin

One of the options in a learning process to improve the predictions for an ungauged site is always to take some discharge measurements at the site of interest. With the availability of the new generation of

floating acoustic doppler velocity measurement devices, the costs of such measurements have become smaller compared to more labour-intensive current metering. Even a small number of current metering estimates of discharge might be feasible for some applications if it can be shown that the measurements have value in reducing uncertainties for decision making.

This then raises the question of which measurements are the most cost-effective in the learning process, and how many measurements might be necessary to achieve a desired reduction in uncertainty. This remains an open question: there is almost no guidance in the hydrological modelling literature about the value of measurements in model identification, except for some vague suggestions about how long a discharge record is necessary to obtain an optimal model calibration (of the order of several years of data). If these suggestions are correct then clearly it is unlikely that funding to collect several years of data in an ungauged catchment would be made available, except perhaps in exceptional circumstances for high capital expenditure projects, such as dams. However, Juston *et al.* (2009) demonstrate that, if selected in an intelligent way, a small fraction of data points in a longer time series might contain almost all the information of the entire data series. Rode and Suhr (2007) applied a water quality model to the River Elbe and also found that a subset of the entire calibration data already provided good results. Rojas-Cerna *et al.* (2006) showed how a small number of measurements could be combined with prior estimates of model parameters derived from a regionalisation process to improve the predictions of ungauged sites. Binley and Beven (2003) also showed that a single set of geophysical measurements of a deep soil water profile contained most of the information content of 18 months of weekly measurements in conditioning a model of groundwater recharge. McIntyre and Wheater (2004) came to a similar conclusion for water quality modelling, providing that the data were chosen from storm periods. This is an indication of the potential value of limited observation data for conditioning the uncertainty in estimating the response of an ungauged basin.

Seibert and Beven (2009) have addressed this problem in an application to 11 catchments draining to Lake Mälaren in Sweden. Models were calibrated using different selections of daily discharge values for one year, and then used to predict the catchment response for a 10-year period. For each calibration, the best 100 models out of a sample of 10 000 runs were chosen based on the sum of squared errors. They showed that a relatively small number of randomly chosen observations (16–32 days) of data provided almost as much information as having a full year of data, but that a poor (random) choice of days might decrease the evaluation performance. Even fewer measurements might be needed when a hydrologically intelligent choice is made about when the measurements are taken, in this case (see Figure 10.3) either the six highest peaks in a two-month period (MAX6) or peak flow days followed by recession days (MAX1REC5, MAX2REC4). Interestingly, the ensemble mean prediction over the 100 models, performed better in evaluation than the single best model of the set for all the cases.

Others have also considered the use of limited streamflow measurements in constraining rainfall–runoff models for ungauged catchments including Rojas-Cerna *et al.* (2006) and Winsemius *et al.* (2009). The latter paper also suggests that "soft" information might be used where available (see also Seibert and McDonnell, 2002). There remains a logistical problem of collecting even a small number of discharge observations, especially in getting a team into the field to measure peak discharges (including the health and safety issues of making measurements at the very highest peaks). One strategy to make this easier might be to install a level sensor for a short period of time (e.g. during a single snowmelt season, in the case of Sweden) and use the small number of measurements to develop an approximate rating curve for a site. In doing so, however, it should be remembered that not only might the limited information in a discharge record constrain the accuracy of model predictions but more effort might be required in the estimation of the inputs to a model. Other, less obvious, sources of information for model calibration have been suggested, including satellite estimates of river width as an index of discharge (see Sun *et al.*, 2010). This is a technique that might benefit from the enhanced accuracy that would be available if the planned SWOT satellite is successfully deployed in the future. A number of studies have used remote

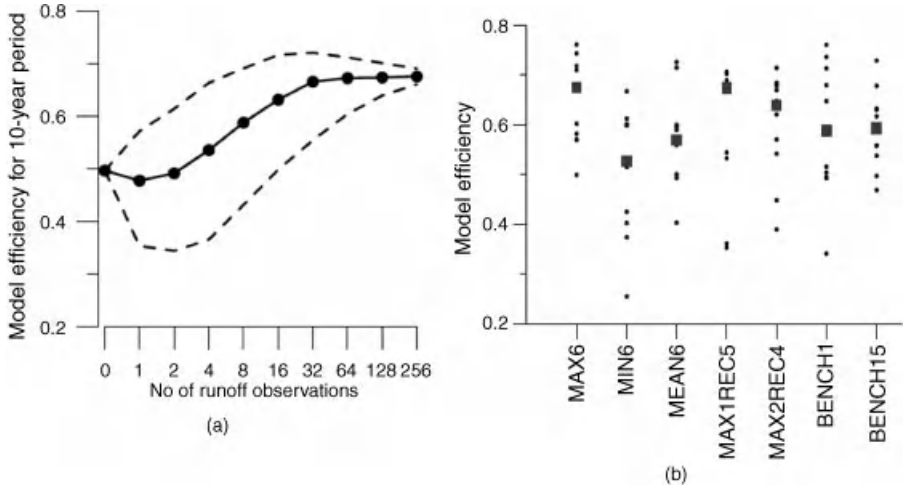


Figure 10.3 *(a) Model efficiencies for the entire 10-year period of the weighted ensemble mean where the ensemble has been selected based on n randomly chosen measurements during one year: the solid line and the circles represent the median over all years, catchments and random realisations of the selection of n days; the dashed lines show the medians of the percentiles (10% and 90%) for the different realisations of the selection of n days; (b) performance of different strategies to select six days of observations during one year for use in model calibration; black dots represent the median of 10 years for one catchment; squares represent the median of all catchments (after Seibert and Beven, 2009).*

sensing of flood inundation to condition the parameters of hydraulic models (e.g. Bates *et al.*, 2004; Pappenberger *et al.*, 2007a, 2007b; Di Baldassarre *et al.*, 2009).

10.12 Key Points from Chapter 10

- There has been significant progress in the problem of predicting the hydrological responses of ungauged catchments in the last decade, in part because of the IAHS PUB intiative and the MOPEX project. These have the aims of improving both the science base for hydrological modelling and the functionality of representing hydrology in data-sparse regions of the world in global modelling.
- Classical regionalisation approaches used in estimating model parameters for ungauged basins have been based on regression of fitted parameter values for gauged catchments against catchment characteristics. The limitations of this approach are now generally recognised, though it continues to be used. Statistical regressions can provide confidence limits for the estimated parameter values, but these are often ignored.
- Where data from a nearby upstream or downstream gauge are available, a simple area scaling of the observed hydrograph to the ungauged basin is the model to beat.
- Approaches based on a pooling group of donor catchments generally give (slightly) better results than regresssions against catchment characteristics but differences are often small.
- Recent work has suggested that a more effective approach to predicting the responses of ungauged basins is the regionalisation of hydrograph characteristics that are then used to constrain an ensemble of models. The ensemble approach allows the uncertainties in representing the ungauged catchment to be addressed.

- Direct estimation of model parameters based on soil, land use and other information (as used in defining land surface parameterisations of atmospheric circulation models) has not been shown to provide accurate simulations of ungauged basins. Where information is available for local conditioning of default parameter estimates, better results are usually obtained.
- One of the most effective ways of constraining prediction uncertainties for an ungauged basin might be to make a few discharge measurements or use a few measurements to constrain a rating curve for a level sensor. Other types of "soft" information might also be used in the conditioning process.

11

Beyond the Primer: Water Sources and Residence Times in Catchments

Every theory is based on physical concepts expressed through mathematical idealisations. They are introduced to give an adequate representation of the physical phenomena. No physical concept is adequately defined without the knowledge of its domain of validity.

Leon Rosenfeld (quoted in Prigogine, 1997)

L'augmentation du débit des eaux souterraines, constaté lors des crues, ne s'accompagne pas d'une variation de leur teneur en tritium . . . Tout se passe comme si la lame d'eau infiltrée chassait, plus vite, devant elle les eaux des épisodes pluvieux antérieurs.

Edith Crouzet *et al.*, 1970

11.1 Natural and Artificial Tracers

There are many purposes for which it is not sufficient just to predict the discharge from a catchment area. It would also be useful to have predictions of flow pathways and residence times of the water within the catchment especially when concerned about water quality, sediment and pollutant transport. It would also be nice to be able to have a more direct understanding of water sources and pathways when it is so difficult to observe the subsurface. As noted in Section 1.5, the use of *environmental tracer* data was fundamental in causing a re-evaluation of flow pathways in catchments by revealing that in many situations the bulk of the hydrograph was made of **pre-event** water displaced from storage rather than from precipitation in a storm. This meant that the simplistic Hortonian concept of the soil surface acting as a "separating surface" between "storm runoff" and infiltration was no longer tenable for many catchments (though it is worth noting that Robert Horton himself had a more complex view of infiltration and runoff processes (see Beven, 2004b). It is also worth noting that the concept of displacement is also not new, Horton was in correspondence with other hydrologists in the 1930s about ideas of "translationary flow", a phrase

Rainfall–Runoff Modelling: The Primer, Second Edition. Keith Beven.

that was later used by John Hewlett in discussing fast subsurface stormflow in catchments where surface runoff was observed to be rare (Beven, 2004c).

Analysis of residence times in catchments is fraught with difficulties of both observation and assumptions. This means that it has also proven rather difficult to make predictions that properly represent water pathways and residence time distributions. This is partly because of the limitations of the available tracers in hydrology. An ideal tracer would have a low background concentration, would be detectable in very low concentrations, and would not be subject to sorption, fractionation, volatilization or other chemical reactions during its residence time in the system so that it can be considered conservative in its behaviour. Ideally also, the plot, hillslope or catchment under study would be water-tight so that the tracer is not lost to pathways not being observed while input concentrations should be well defined and constant in space and time. These characteristics are difficult to meet completely. Many of the commonly used tracers are either not entirely conservative or, for the environmental tracers, have variations in space and time that make it difficult to assess the true mass balance for the tracer (even if the hydrological system is really water-tight).

This means that, as in the case of modelling hydrographs, inferences are often made on the basis of less than ideal observations and assumptions that are necessarily approximate. The nearest to the ideal case is probably when artificial applications of either hydrogen or oxygen isotopes (tritium, deuterium or ^{18}O) are made in short-term experiments (Rodhe *et al.*, 1996). This is because these isotopes can be added as part of the water molecule and therefore should track the water pathways directly. They are subject to fractionation due to freeze–thaw, evaporation and transpiration processes but, for short-term experiments where snow and ice are not an issue, this should not be a significant factor. More important is whether all the outputs of tracer are being measured or whether the tracer is retained in long-term *immobile* storage. Both result in not all the tracer that is input being measured, so that mass balance is apparently not preserved. It is rather typical in such experiments (and more so with other tracers subject to sorption or biogeochemical reactions) that only a low proportion of the input tracer can be accounted for in the output observations. The tracer (or solute) is then said to be *non-conservative*, though this might be just because not all the output fluxes or storage in the system have been detected, rather than because of biogeochemical processes. Even with the isotopes of the water molecule, interpretation of measurements may be difficult. It has been suggested by Brooks *et al.* (2010), for example, that the water extracted from soil storage by trees for evapotranspiration may be different in isotopic concentration to that in soil water discharge, even though no fractionation process would appear to be involved.

There have been relatively few comparative studies of using different tracers in the analysis of residence times but it is known that different tracers might reveal different aspects of the catchment response. In particular, the oxygen and deuterium isotopes cannot provide information about water ages over about four years, whereas tritium can identify older waters (Stewart *et al.*, 2010). Some groundwater systems, of course, contain waters that date back 25 000 years or more, having been replenished during the last Pleistocene glaciation. The use of such waters for water supply is effectively mining of waters that are not now being recharged. Other dating techniques (such as ^{14}C) can be used for such ages, but here we are concerned much more with the short-term rainfall–runoff processes for which oxygen and deuterium isotopes are useful. They are commonly assumed to be linearly related but might still provide somewhat different types of information (e.g. Lyon *et al.*, 2009).

In the remainder of this chapter, we look at the use of mixing models to infer sources of runoff and models for inferring distributions of *residence times* and *travel* (or *transit*) *times* in catchment systems (see Section 11.9). A residence time distribution summarises the lengths of time that water molecules have spent in the flow domain of interest (soil core, lysimeter, isolated plot, hillslope or catchment) but it will be seen that the residence time distribution for an increment of input might be quite different from the residence time distribution for an increment of output and both might be different from the residence time distribution for water in storage in the catchment at any time (e.g. Rinaldo *et al.*, 2011). In the general case we expect all of these distributions to vary over time as the catchment wets and dries. Thus, in identifying either water sources or residence times, the reader should be aware of the strong assumptions

that are necessary in the different analysis methods presented below. As ever, in using such techniques, start by evaluating the assumptions that are required against the perceptual model of the processes.

11.2 Advection and Dispersion in the Catchment System

The residence time of a water particle in the catchment will be related to both its origin on entry (in space and time) and to the net effect of all the different velocities experienced by that particle in its pathway to the catchment outlet. This might involve laminar flows in different pore sizes in the soil matrix; laminar, transitional or turbulent flows in macropores and over the soil surface; and turbulent flows in rills, gulleys or the stream channels. It might also involve periods of immobility or loss as evapotranspiration (see the perceptual model of Section 1.4). The description of the transport process of water and tracers through the system, for both surface and subsurface flow processes has traditionally been in terms of *advection* and *dispersion*. Advection is the movement associated with the mean velocity of the flow at any point; dispersion is the spread resulting from the distribution of velocities around the mean at that point.

The theory of advection and dispersion can be found in Box 11.1, including descriptions based on the advection–dispersion equation (ADE), and the aggregated dead zone (ADZ) model. In steady flows, both can be used in the form of linear transfer functions (i.e. similar to the unit hydrograph method for predicting hydrographs in Section 2.1). In many catchment systems, of course, we are not so much interested in steady flows (though it can be a good approximation, for example, in predicting transport in larger river channels during recession periods). What happens to the transport during hydrographs, and the consequent effects on residence time distributions, is much more interesting (e.g. Figure 11.1).

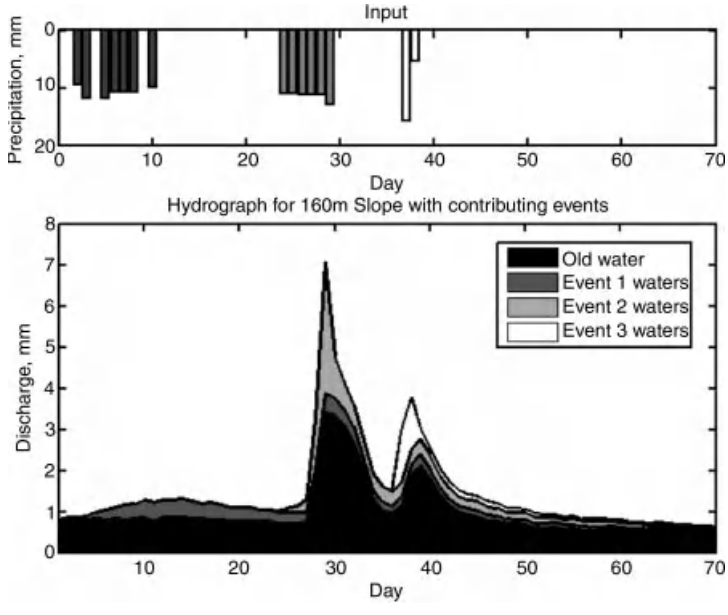


Figure 11.1 *Predicted contributions to individual hydrographs from different precipitation inputs (grey shading); this hypothetical simulation uses the MIPs model of Davies et al. (2011) on a 160 m hillslope with a soil depth of 1.5 m and a hydraulic conductivity profile similar to that found at the Gårdsjön catchment, Sweden (see Figure 11.8).*

In addition, it has been suggested that by combining hydrograph and environmental tracer information in calibrating catchment response models, the hydrologist might be able to get closer to having a catchment model that gets the right result for the right reasons. We return to this issue after discussing how environmental and artificial tracers can be used to throw light on the sources and residence time distributions of catchment discharges.

11.3 Simple Mixing Models

Tracers are also used to try to determine the sources of water in discharge hydrographs from hillslopes and catchments. To carry out tracer experiments using artificial additions of tracer at these scales is difficult, so it is generally necessary to resort to the use of naturally occurring substances or *environmental tracers* to do this. The use of environmental tracer information to infer the sources of water in the hydrograph started with the use of simple two-component mixing methods that distinguished between the contributions of water that is added during an event (**new** water) and water that is stored in the catchment prior to an event (**pre-event** or **old** water). The analysis depends on being able to assign characteristic concentrations to each of the two sources, and a simple mass balance of mass of water and tracer in catchment hydrograph (but not necessarily in the catchment as a whole).

Thus, at any time step, it is assumed that:

$$Q = Q_o + Q_n$$

$$QC = Q_o C_o + Q_n C_n$$

where Q is discharge, C is tracer concentration and the subscripts o and n refer to old and new water, respectively.

Given these two equations, the proportion of new water in the hydrograph can be calculated as:

$$\frac{Q_n}{Q} = \frac{(C - C_o)}{(C_n - C_o)} \tag{11.1}$$

This calculation can then be repeated for all time steps for which concentration measurements are available to provide a hydrograph separation based on water sources (e.g. the example shown in Figure 1.6). Note that for a well-defined separation there must be a distinct difference between C_o and C_n. If this is not the case then the denominator in the mixing equation is very small or zero and the separation cannot be made. Note also, in applying the equation at successive time steps it is necessary to assume that C_o and C_n are constant in space and time (at least for that particular storm).

It is known that this latter assumption is generally only very approximately true. New water concentrations often vary in space and time in a storm (McDonnell *et al.*, 1990; Harris *et al.*, 1995). In addition, for some tracers, such as silica, rain water might actually change its concentration as it interacts with the soil over rather short time scales. Old water concentrations are often taken to be the observed concentration in the hydrograph measured prior to a storm, but such concentrations represent a mix of pathways in space and there is no real reason why those concentrations should stay constant in either time or space during the forcing due to a storm event (and in Section 11.4, we see that the mix of waters from subsurface sources can change over a sequence of events).

Thus, this type of mixing model is a crude approximation. It has been an important approximation, however, in causing hydrologists to recognise the dominant contribution of subsurface runoff to the hydrograph in many catchments. However, it should not be assumed that this is always the case. Even in

the early paper of Sklash and Farvolden (1979), it was shown that samples taken from surface runoff at a particular location in a catchment were sometimes predominantly new water and sometimes old water and many catchment hydrographs might be consistently dominated by new water. It is possible to carry out an uncertainty analysis of the resulting hydrograph separations taking account of the uncertainties in the input information. We return to this in respect of the more general end member mixing analysis (Section 11.5).

11.4 Assessing Spatial Patterns of Incremental Discharge

This type of two-component mixing analysis can also be applied along discrete increments of a river to determine the increments of discharge in space. In this case, rather than old and new water components, in each reach of the channel the upstream and lateral contributions can be determined by applying Equation (11.1) sequentially in each reach. It is again necessary that the characteristic concentrations for the upstream and lateral contributions should be different. This can be achieved by adding an artifical tracer at the upstream end of the reach (Huff *et al.*, 1982) or making use of an environmental tracer (e.g. radon, as Genereux *et al.*, 1993, did). Some recent work has used water temperature, measured in a continuous downstream fibre-optic cable, as a form of tracer in a similar type of analysis based on a combination of mass and energy balances (Westhoff *et al.*, 2007), although this introduces additional terms into the energy balance that require a number of parameters to be estimated. Since those parameters are uncertain, the resulting estimates of lateral inflows are uncertain but, as yet, there has not been an adequate uncertainty analysis of this approach published.

Despite the uncertainties, however, this type of analysis has revealed some very strong spatial variability in the lateral inputs of discharge to stream channels under both high flow and low flow conditions. In the case of the Walker Branch catchment, studied by Huff *et al.* (1982) and Genereux *et al.* (1993), this variability appears to be related to dipping bedrock structures beneath the soil layers that are not manifest in the surface topography of the catchment.

11.5 End Member Mixing Analysis (EMMA)

The type of simple mixing analysis of Section 11.4 can be extended to the identification of multiple sources of water by the use of multiple tracers where the tracer concentrations can be used to distinguish those sources or "end members". This end member mixing analysis (EMMA) is most often used with two tracers to differentiate three components: event water, water from soil storage and water from longer-term groundwater storage. The three balance equations are then:

$$Q = Q_P + Q_{SW} + Q_{GW}$$

$$QC = Q_P C_P + Q_{SW} C_{SW} + Q_{GW} C_{GW}$$

$$QC' = Q_P C'_P + Q_{SW} C'_{SW} + Q_{GW} C'_{GW}$$

where Q is discharge, C and C' are the concentrations of two different tracers, and the subscripts P, SW and GW refer to precipitation, soil water and ground water respectively. Q, C and C' are measured in the stream and, if the concentrations of the three end members can also be assumed known and constant in time, there are only three unknowns, Q_P, Q_{SW} and Q_{GW}. The system can therefore be solved at every time step for which measurements of Q, C and C' are available.

Similar assumptions to the two-component case are necessary: that there is a distinct difference between the end member concentrations and that those concentrations can be assumed constant in time and space.

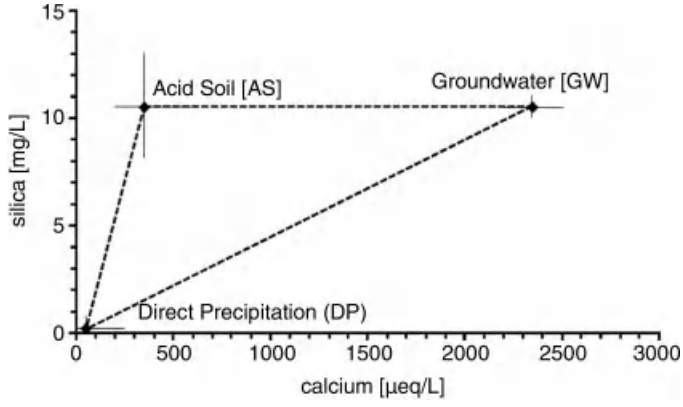


Figure 11.2 *Mixing diagram for direct precipitation, soil water and groundwater end members defined for silica and calcium in the Haute-Mentue catchment, Switzerland (after Iorgulescu et al., 2005, with kind permission of John Wiley and Sons).*

The first assumption can be checked by plotting the tracer concentrations for each end member on a mixing diagram similar to Figure 11.2, which shows soil water, rainfall and baseflow silica and calcium concentrations for the Haute-Mentue catchment in Switzerland. In this case, these two tracers separate the components quite nicely, since an acid soil (high in silica, low in calcium) has developed over a regolith of molasse bedrock (high in calcium, low in silica), while the rainfall should be low in both calcium and silica.

The analysis might still only be approximate, however. In particular, it might be expected that soil water concentrations would vary over time as the soil wets and dries, so that a set of samples might be needed prior to each event to improve the accuracy of the separation. There is also an issue about whether the soil concentrations (measured on samples that are usually obtained by imposing a suction) are representative of the mobile water stored in the soil that might contribute to the hydrograph. There is also the possibility that, when the event water interacts with the soil, it takes up silica into solution and rather quickly starts to appear as if it were soil water, thereby resulting in an overestimation of the soil water contribution and an underestimation of the event water contribution.

Similar issues arise in most EMMA applications and the validity of the assumptions of such an analysis should always be examined with some care. One way of doing so, as matter of good practice, is to evaluate the uncertainties in the resulting separations. Hooper *et al.* (1990) and Genereux (1998) used analytical forward uncertainty propagation to do this while Bazemore *et al.* (1994), Joerin *et al.* (2002) and Soulsby *et al.* (2003) have used Monte Carlo simulation.

EMMA can also be used with multiple chemical species contributing to the definition of the end members (Burns *et al.*, 2001; Hooper, 2003) and to test for the conservativity of pollutants in the system. Christophersen *et al.* (1990), using two source components, tested for deviations from the 1:1 mixing line between end members, while recently Jarvie *et al.* (2011) showed how EMMA can be extended for end members that are expected to vary with discharge.

11.6 On the Implications of Tracer Information for Hydrological Processes

The results of such mixing model calculations often suggest that a large part of the hydrograph is made up of water that has been stored in the catchment for a long time (see also Section 1.4). Kirchner

(2003; Kirchner *et al.* 2010) considers this to be a "paradox": how do many catchments store water for such long periods of time but so easily mobilise so much "old" water during storms.

In Section 5.5.3, the kinematic wave model for subsurface flow was used to explain how the displacement of subsurface water was a result of the difference between the velocity of flow and the wave speed or celerity of the response. Since the effective storage coefficient controlling the celerity is always less (and sometimes two orders of magnitude less) than the saturated porosity controlling the mean pore water velocity, the effect of an input propagates downslope faster than the mean pore water velocity, resulting in the displacement of stored water. An extreme case is that of a saturated pipe filled with soil. The effective storage coefficient is close to zero, so the celerity is close to infinite. Pouring water in at one end immediately causes a displacement at the other end, clearly of "old" water rather than "new" water, in this case.

This explanation is useful conceptually but the situation on real hillslopes is somewhat more complex because of the space and time variability of both unsaturated zone and saturated zone flows and the interaction between them, complicated by the possibility of a dynamic capillary fringe above the water table. In addition, both the unsaturated and saturated flows may involve a very wide range of pore water velocities, from preferential flows in (often discontinuous) macropores to relatively immobile storage in the soil matrix. Account also needs to be taken of the distributed nature of the inputs of new water to the system. Inputs near the base of the slope might contribute to the hydrograph at the same time as displacements of old water are taking place.

There is some evidence of this type of complex response from small-scale artificial tracer experiments carried out on undisturbed soil cores. Reeves *et al.* (1996), for example, looked at the case where a "return flow" (applied to the base of a core and labelled by a fluorobenzoate tracer) mixed with water stored in the core and water added from a sprinkler system above (labelled by a different fluorobenzoate tracer). The core had been taken from a valley bottom hollow and brought to saturation prior to the experiments and, at the end of the tracer runs, the return flow was switched to a methylene blue dye solution to stain the flow pathways of the return flow. The results showed that the pathways of the return flow were strongly preferential in the lower part of the core, while the mix of tracers in water overflowing from the surface of the saturated core suggested that some of the rainfall added was infiltrating into the soil and displacing stored water, despite the core being saturated. This suggests some rather local displacement circulations during such an event, even on a saturated surface. Local displacement in generating overland flow was also inferred from field evidence by Iorgulescu *et al.* (2007).

There is also the possibility of deep subsurface pathways contributing to the hydrograph (Loague *et al.*, 2005; Ebel *et al.*, 2008). Clearly, the inference of hydrograph sources, flow pathways and residence times of water in a catchment remains complex (McDonnell *et al.*, 2010).

11.7 Case Study: End Member Mixing with Routing

The Haute-Mentue catchment data described in Section 11.6 have been used in a slightly more sophisticated modelling study of source identification by Iorgulescu *et al.* (2005, 2007) that has also tried to allow for routing times in the catchment for the different sources. The catchment is in Switzerland, on the plateau area north of Lausanne, with a strong differentiation of the chemistry of the soil and bedrock (Figure 11.2). Six subcatchments were monitored; the results in this case study come from the Bois-Vuacoz (24 ha). Isotope and chemical concentrations were sampled in rainfall and streamflow during a sequence of events in a wetting-up period at the end of the summer. Thus it was expected that the storages in the system would be dynamic and that the effective concentrations for the different sources would change over time.

Thus, a model was proposed that included the partitioning of inputs into a direct runoff component (DP), a soil water component (AS) and a groundwater component (GW). The partitioning was made a nonlinear function of the storage in the soil. Stationary transfer functions with fast and slow components were used to route all three components to the stream. The concentrations in each component are assumed to evolve

only slowly and are treated as uncertain parameters within the ranges indicated by Figure 11.2. This is equivalent to saying that the inputs in this period are causing a displacement of soil water and groundwater from storages that are large relative to the inputs. Even this very simple model of three components has 17 parameters. The model was calibrated within the GLUE framework, using constraints on matching the stream discharges, silica and calcium concentrations based on the residual errors (Iorgulescu *et al.*, 2005). To search such a high-dimensional model space, two billion realisations of the parameters were run and evaluated. Some 216 were accepted as behavioural (see Figure 11.3). The model can be interpreted to illustrate how the contribution of the different components changes as a function of relative storage in the system (Figure 11.4). The nonlinear nature of the response is clear, with a steep rise in the contribution of the soil water component as the system wets up. Interestingly the transfer functions of precipitation and soil water are inferred as being very similar. This implies either that the precipitation component has a relatively slow pathway or that it is taking up some silica and appearing as part of the soil water contribution. Probably both are actually occurring.

The constant source concentration assumption in this analysis means that nothing can be inferred about residence times and effective storages in the system. The transfer functions in the model are related to the hydrograph response, not to the actual travel times of water through the system. The isotope observations, however, can be used to take the analysis a step further. Thus, Iorgulescu *et al.* (2007) used

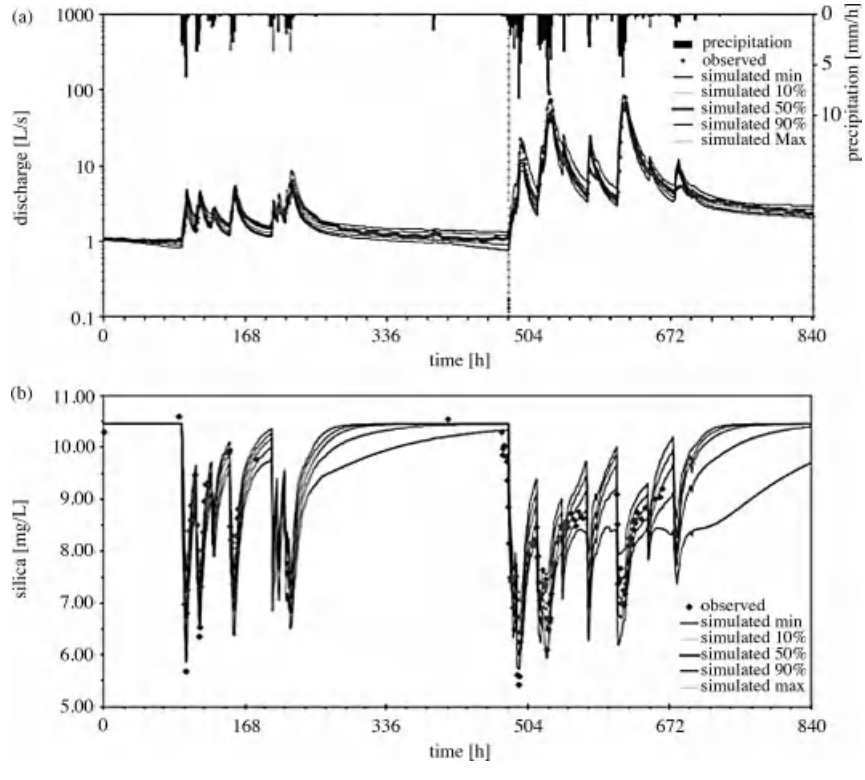


Figure 11.3 *Observed and predicted discharges, silica and calcium concentrations for the Bois-Vuacoz subcatchment; lines represent the range and quantiles of the predictions from the 216 behavioural models (after Iorgulescu et al., 2005, with kind permission of John Wiley and Sons).*

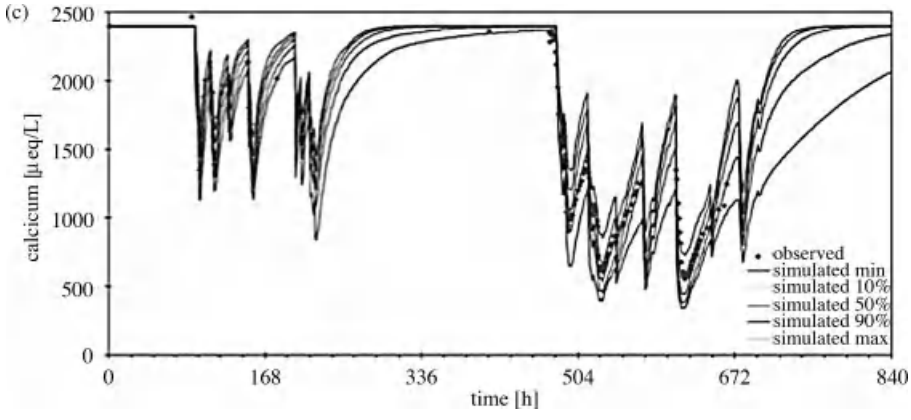


Figure 11.3 *(continued)*

the source separations of the 216 behavioural models to see what storage and mixing assumptions would be necessary to reproduce the isotope concentrations in the stream. A number of different hypotheses were considered. The most successful assumed an effective mixing volume for precipitation and soil water, and a rate constant for the precipitation water entering the soil to take up silica. The effective mixing volumes allow for the difference between celerities and velocities. They do not necessarily give an indication of the total storage associated with each component and, for the GW component, the time scale of the measurements is too short even to estimate an effective mixing volume.

The results from this model in matching the stream ^{18}O concentrations is shown in Figure 11.5. The resulting dynamic effective mixing volumes are shown in Figure 11.6 where it can be seen that this changes only slowly for the AS soil water component but much more quickly for the DP precipitation component. The uncertainties in both estimates are quite high, despite the good match to the stream concentrations shown in Figure 11.5. This study gives a good illustration of the nonlinearity and nonstationarity of

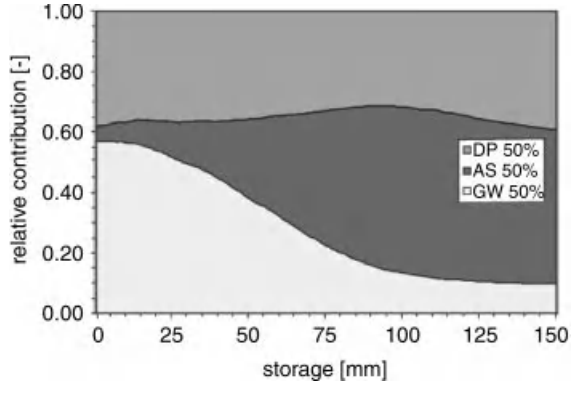


Figure 11.4 *Changing contributions of the different components of stream flow in the Bois-Vuacoz subcatchment as a function of storage; results are the median estimates over 216 behavioural models (after Iorgulescu et al., 2005, with kind permission of John Wiley and Sons).*

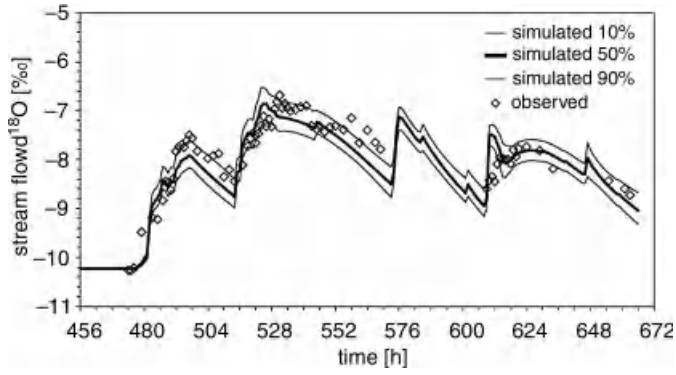


Figure 11.5 *Prediction of stream ^{18}O isotope concentrations for the Bois-Vuacoz subcatchment (from Iorgulescu et al., 2007, with kind permission of the American Geophysical Union).*

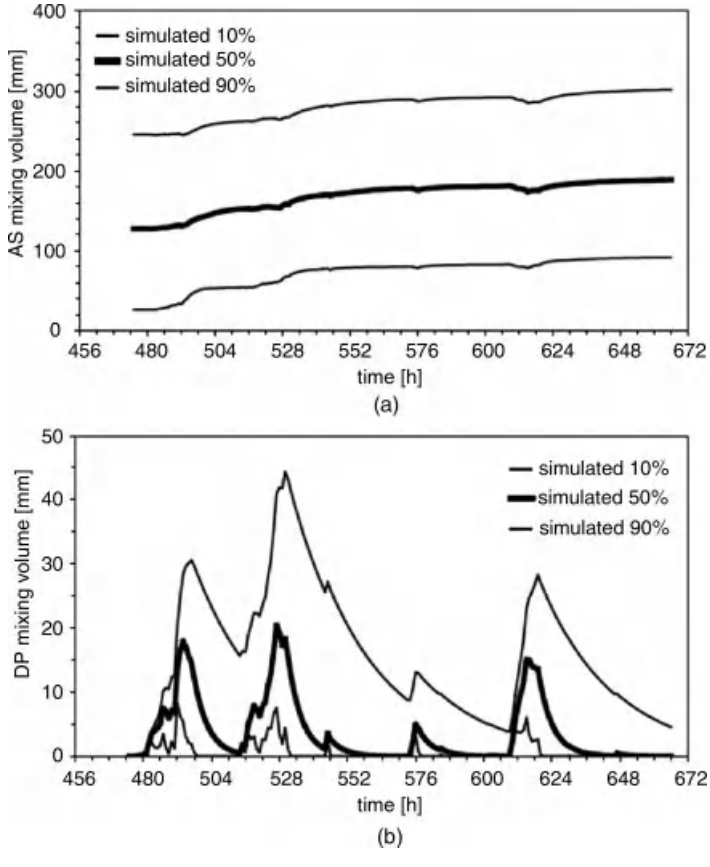


Figure 11.6 *Inferred dynamic storages for (a) soil water and (b) direct precipitation from modelling isotope concentrations in the Bois-Vuacoz subcatchment (from Iorgulescu et al., 2007, with kind permission of the American Geophysical Union).*

mixing processes in catchments, suggesting that analyses of residence time distributions based on steady state assumptions over the longer term (see Section 11.8) should be interpreted with some care.

11.8 Residence Time Distribution Models

The types of mixing model outlined above have concentrated on short time scale contributions to storm hydrographs. It is not necessary in those analyses to estimate how long the stored pre-event water has been in the catchment. To do so, it is necessary to have measurements of concentrations of inputs and outputs over much longer periods of time and to assume some model for the full residence time distribution of water in the catchment while making some similar assumptions about the spatial homogeneity of input concentrations and (in the simplest cases) temporal stationarity of the residence time distribution. As with the simple mixing models and end member mixing analysis methods decribed above, it is necessary to have some differences between the concentrations in the inputs and outputs to make the analysis tractable. In general, because of the damping effects of storage in a catchment the variability in the input concentrations are greater than the variability in the output concentrations. Some signal in the output concentration is still necessary, however. If all the input variability is damped out, it will not be possible to infer anything about the residence times in the catchment except that they are very long relative to the time scale of variability in the inputs and that the storage in the catchment must be very large relative to the mass flux.

For the isotopes of oxygen and hydrogen/deuterium, a structured signal in the input concentrations happens naturally because of the fractionation that takes place when condensation takes place in the atmosphere. Thus, the heavier isotopes are preferentially released in the formation of precipitation. The fractionation process depends on temperature. The lower the temperature, the heavier the isotopic concentration in precipitation. In humid temperate climates, this results in a general seasonal pattern in the isotopic concentrations in precipitation, with generally heavier concentrations in winter. However, individual events can vary significantly from this general seasonal trend and there can also be marked concentration variations within events in both time and space. In other climate regimes, the seasonal patterns can also be more complex, such as the monsoonal effects reported by Xie *et al.* (2011).

It is possible to provide a general framework for the analysis of residence time distributions from input precipitation to output discharges within a catchment (or representative element watershed, see Chapter 9) storage.

The mass balance equation for the storage can be written as:
for the water

$$\frac{dS(t)}{dt} = P(t) - E(t) - Q(t) \tag{11.2}$$

and for an arbitrary tracer

$$\frac{dM(t)}{dt} = P(t)C_P(t) - E(t)C_E(t) - Q(t)C_Q(t) - R(t) \tag{11.3}$$

where $S(t)$ is storage at time t, $P(t)$ is precipitation input, $E(t)$ is evapotranspiration output, $Q(t)$ is discharge output, M is mass of tracer, C is tracer concentration, and $R(t)$ is a source or sink term for mass of tracer. The transport of tracer within the storage volume can then be thought of most simply as a form of time variable convolution so that

$$Q(t) = \int_0^t P(\tau) f(t - \tau|t) d\tau \tag{11.4}$$

and $f(t - \tau|t)$ is the residence time distribution for inputs reaching the stream outlet which is expected, in this general case, to vary with time as the catchment wets and dries and therefore conditional on t.

It is important to note that this is a different convolution to the unit hydrograph convolution of $h()$ in Section 2.2 and transfer functions of Box 4.1. Here, the convolution represents the travel times of inputs to the outlet rather than the hydrograph response. It is therefore representing the integrated effect of velocities in the system rather than the (faster) celerities that control the hydrograph response. Following Botter *et al.* (2010), we can write a similar general function for the loss of water to evapotranspiration as:

$$E(t) = \int_0^t P(\tau) g(t - \tau|t) d\tau \tag{11.5}$$

where $g(t - \tau|t)$ is the residence time distribution for inputs to be lost from the system, again conditional on t as evapotranspiration is also expected to vary with wetting and drying and the energy available to drive the evapotranspiration process which varies with time on both daily and seasonal time scales. Both $f(t - \tau|t)$ and $g(t - \tau|t)$ can also be thought of as conditional probability distributions for particles of water entering the system at time τ to reach the outlet at time t.

The difficulty in estimating these residence time distributions from input and output fluxes and concentrations is then readily apparent. We have no *a priori* knowledge of the form of the distributions $f()$ and $g()$, nor how they vary with time. It is also known that, even in simple cases where the distributions can be assumed constant and linear, estimation from time series of inputs and outputs (a "deconvolution" operation) is known to be poorly posed and highly sensitive to errors in the data series. In the case of catchments, not only are the residence time distributions expected to vary with time, they may not be linear, and observations of both the inputs and output fluxes and concentrations may be subject to significant uncertainty. In one sense, therefore, the identification of $f()$ and $g()$ is an impossible problem. However, significant insights into how the catchment system is responding can be obtained if some assumptions are made about the nature of the distributions in order to simplify the identification. More detail on the estimation of residence time distributions is given in Box 11.2.

11.8.1 Which Residence Time Distribution?

As defined above, a residence time distribution is a summary of the times that water molecules have spent within the flow domain of interest. There are, however, a number of different residence time distributions that are of interest (Rinaldo *et al.* 2011). Consider a simple watertight catchment system in which the rainfall inputs become either discharge or evapotranspiration. Then a given increment of rainfall will have a *residence time* (or *travel* (*transit*) *time*) distribution in the system that is different for discharge and evapotranspiration fluxes. A given increment of discharge will have a residence time (or travel (transit) time) distribution that (except for some very special circumstances) is different from the residence time distribution for rainfall increments (and the same is true for an increment of evapotranspiration). These distributions will also be different from the residence time distribution for water stored in the system at any time (travel (transit) time distributions are generally used only for input or output fluxes and not for storage). Thus, in referring to the residence time distribution for a catchment, it is necessary to be clear about what is actually meant. A theorem of Niemi (1977) can be used to define the relationship between these different residence times analytically under certain simplifying assumptions (Botter *et al.*, 2010; Rinaldo *et al.*, 2011). In the general time variable case, if the fraction of rainfall reaching the discharge outlet at time t is θ_t then

$$f'(t - \tau|t) Q_t = P_t \theta_t f(t - \tau|t) \tag{11.6}$$

where $f'(t - \tau|t)$ is the residence time distribution for water making up the increment of the discharge at time t. The input and output flux increment residence time distributions can only be the same if the system is completely dry before the increment of rainfall input (so that no displacement of pre-event water forms the discharge) or if the system is at steady state (with linear partitioning of inputs between discharge and evapotranspiration (θ_t constant) such that all increments of input have the same travel

time distribution through the system). This is not generally the case, so that input flux residence time or travel time distributions are different from output flux residence time or travel time distributions; they are different from the residence time distributions of storage in the catchment; and they all change over time. As we see in the following sections, it is often a convenient assumption that residence time distributions do not change over time but this is, in fact, an assumption difficult to justify.

11.8.2 Introducing Some Simplifying Assumptions

Where long time series of input and output fluxes and concentrations are available we might be able to ignore the variations in residence time distributions over time and look only at the dominant filtering of the input time variability in producing the outputs by assuming that the distribution $f()$ is a constant. The analysis is further simplified by assuming that the mean flux rate through the storage is a constant. This means that variations in flux can be ignored so that only the input and output concentration time series are needed to determine a residence time distribution. It has been argued that this might be an adequate simplification if the variation in storage in the catchment over time is small relative to the total volume of active storage (McGuire and McDonnell, 2006), but this is also an argument from quasi-steady-state principles. It may not be a compelling argument when small changes in storage result in very large changes in the fluxes and velocities of discharge and tracer. The importance of this has perhaps been neglected in the past because, as well as the convenience in simplifying the analysis, there have been very few datasets that have adequately characterised the mass flux of tracer during storm periods.

The identification of a residence time distribution is simpler still if we can further assume a parametric form for the residence time distribution. Then only a small number of parameters need to be estimated. A number of different distributions have been used in the literature including the exponential and gamma distributions (see Box 11.2). From a fitted distribution, characteristics such as the mean travel time of the tracer in the system can be estimated. Although the value derived is affected by the simplifying assumptions that have been made, it may still be a useful comparative measure across catchments, giving an indication of the nature of the hydrological response. Indeed, it may be possible to get a good estimate of these mean travel times from catchment characteristics (McGuire *et al.*, 2005; Soulsby *et al.*, 2010; Lyon *et al.*, 2010). This can give some idea of the residence times to be expected in ungauged catchments, albeit with significant uncertainty (Figure 11.7).

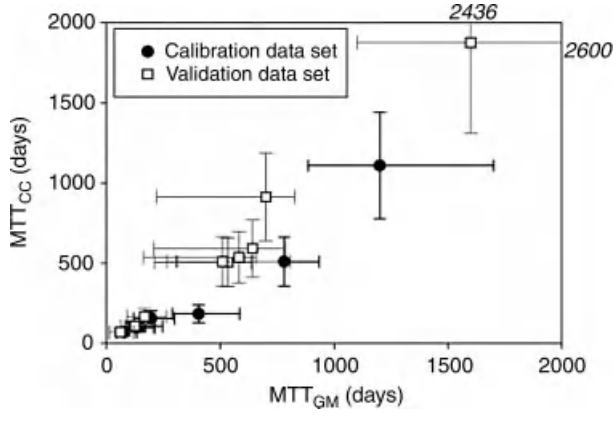


Figure 11.7 *A comparison of mean travel times estimated from observations (MTT_{GM}) and those estimated from catchment characteristics (MTT_{CC}) for catchments in Scotland (from Soulsby et al., 2010, with kind permission of John Wiley and Sons). Error bars represent 5 and 95% percentiles.*

11.8.3 Fractal Residence Time Distributions

An interesting development in the analysis of transit and residence time distributions originates with the work of Kirchner *et al.* (2000, 2001), which was based on the analysis of three years of daily chloride data in rainfalls and discharges collected at the Plynlimon experimental catchments in Wales. Their analysis showed that the simplest assumption of an exponential residence time distribution in this catchment was not a good one, and that a much better fit to the data could be obtained by using a two-parameter gamma distribution (see Box 11.2). The gamma distribution can take on a wide range of forms; with shape parameter values <1, there is greater weight in the tails of the distribution. In an analysis of 22 catchments using a spectral filtering method, Godsey *et al.* (2010) suggest that the majority of those catchments have shape factors significantly less than 1 (see also Kirchner *et al.*, 2010). Four catchments with the highest values of the shape parameter (>0.65), and therefore closest to exponential distribution (with shape factor = 1), all had catchment lakes, which might provide more complete mixing behaviour.

There are a number of interpretations of this type of heavy-tailed behaviour. One is that the residence time distributions are "fractal" (Kirchner *et al.*, 2000), which means that there is a not a single representative mean residence time but a distribution of storages with different mean residence times. Another is that rather than a lumped storage in the system, it is behaving more like a storage with inputs distributed in space that are then advected and dispersed according to the ADE (Box 11.1; Kirchner *et al.*, 2001). A catchment does, of course, have distributed inputs over its hillslopes, though the advection and dispersion processes are likely to be much more complex than the simple ADE representation and may involve multiple storages and heavy-tailed velocity distributions (Lindgren *et al.*, 2004; Davies *et al.*, 2011).

11.8.4 Issues in the Interpretation of Residence Time Distributions

The Kirchner *et al.* (2000, 2001) studies raise a number of interesting issues:

- how the form of numerical analysis chosen can affect the apparent residence times in the catchment (see also Kirchner, 2005);
- that there may be multiple sets of assumptions and consequent explanations of the type of filtering of an input that results from the dynamics of the catchment;
- that how far we can distinguish between those dynamics might depend on the uncertainties in sampling the signal, both in the frequency and representativity of the sampling and in the length of the sampling period. Very long residence times might be poorly represented by short periods of record (and might depend on the environmental tracer used: see Lyon *et al.*, 2009; Kirchner *et al.*, 2010; Stewart *et al.*, 2010).

The original chloride data from Plynlimon used by Kirchner *et al.* (2000) was three years of daily data in a catchment that responds very rapidly to rainfall. The daily time series might therefore not be an adequate sample of the mass flux of chloride out of the catchment. Some of evidence of this was suggested by the study of Page *et al.* (2007) which attempted to simulate the same data using a catchment scale model based on TOPMODEL (Box 6.1). The catchment model also requires an estimate of chloride inputs and the data available suggested that there was a gross imbalance of about 30% in chloride, with the inputs in rainfall being much lower than the outputs (it is possible that more detailed sampling of the chloride in discharge might have made this imbalance worse). The Plynlimon catchments lie in an area where there can be dry deposition of chloride in westerly airstreams from the Atlantic and the Irish Sea and scavenging of chloride from low clouds that often cover the upland forest and grassland. Page *et al.* (2007) produced a model of the inputs to correct for the mass imbalance based on the meteorological conditions, but were only partially successful in reproducing the chloride concentrations in discharge. In fact, analysing the simulated concentrations using similar spectral filtering methods to Kirchner *et al.*

(2000) on the observations showed that the catchment model based on TOPMODEL did not have the right sort of filtering of the signal to reproduce the catchment behaviour. It was getting reasonably correct results (given the uncertainties in the inputs) but not, it seems, for the right reasons. This might be because of the particular mixing assumptions incorporated for this study but also because of the lack of a space dimension in the TOPMODEL process representations, which Kirchner *et al.* (2001) suggested might be important in explaining the mixing characteristics inferred from the chemistry (see also the discussion in Chapter 9).

However, it is becoming increasingly possible to obtain more frequent measurements of concentrations and this may gradually become less of an issue in the future (e.g. for environmental isotopes, see the work of Berman *et al.*, 2009). It might then become more feasible to relax more of the simplifying assumptions and consider the time variability in the residence time distributions. Botter *et al.* (2010) and Rinaldo *et al.* (2011) have produced a framework that does allow the time variability in the residence time distributions to be predicted as the catchment wets and dries. It remains to be seen how well this dynamic residence time distribution formulation can reproduce the behaviour of different catchments. Earlier work on transient residence time distributions was reported by Niemi (1977), Zuber (1986) and Foussereau *et al.* (2001).

11.8.5 Residence Time Distributions and Rainfall–Runoff Model Calibration

There has been some discussion in the literature about how the use of environmental tracer information might help in calibrating rainfall–runoff models so that we do get the right results for the right reasons. There have also been a number of model applications that have been checked against both flow and environmental tracer concentrations (Vaché and McDonnell, 2006; McGuire *et al.*, 2007; the case studies in Sections 11.7 and 11.9). The application of these models is also aimed at understanding the "paradox" of Kirchner (2003) although, as noted earlier, this is perhaps not such a paradox after all.

However, nearly all the models used in simulating both flow and tracer concentrations might be getting reasonable results for the wrong reasons. The storage models often use a complete mixing assumption for the local concentrations and an instantaneous response of discharge to changes in storage regardless of the length scale of the calculation element. With these assumptions, such models cannot adequately represent the difference between the pressure wave celerity and the distribution of flow velocities (see Chapter 9). They are relying on uncontrolled numerical dispersion interacting with the mixing assumptions to get the right sorts of answer. Even distributed models, such as that of Sayama and McDonnell (2009), which include some routing between elements, but with explicit time stepping, could be subject to solution error if their numerical algorithms are not properly implemented. In that case, the tracer dispersion also relies on numerical dispersion. We should be wary of inferences derived from models with numerics of this type (Clark and Kavetski, 2010; Kavetski and Clark, 2010). The processes and storages involved in the hydrograph response and tracer response are certainly closely linked but they are different.

11.9 Case Study: Predicting Tracer Transport at the Gårdsjön Catchment, Sweden

An interesting alternative to this storage-based approach has recently been suggested by Davies *et al.* (2011), although it is still very much in its early stages of testing (see also Section 9.6). This methodology represents the water and associated tracer as discrete particles that move through the flow domain. Such particle tracking models have been used widely in the past to predict the transport of tracer and solutes in a specified velocity field, determined by a flow model, particularly in groundwater modelling where there is interest in the movement of solutes over long periods of time. The USGS MODFLOW groundwater

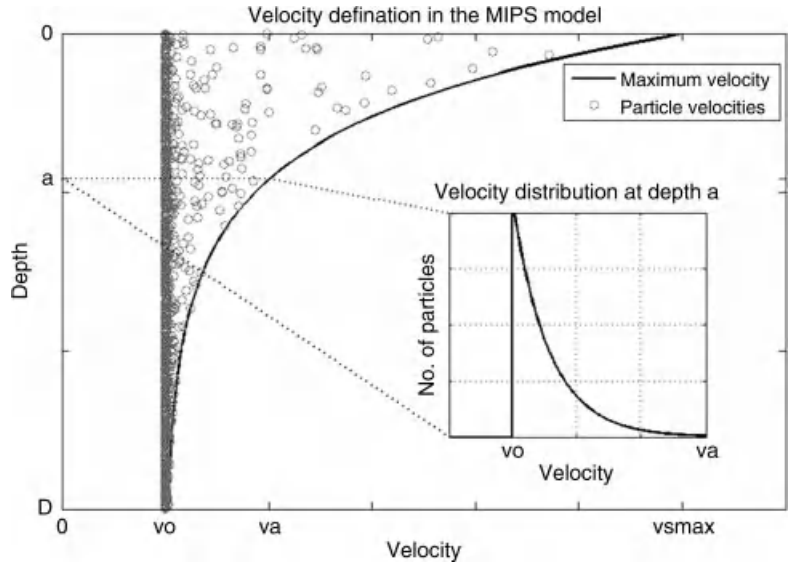


Figure 11.8 *A representation of particle velocities chosen from an exponential distribution for each layer in a way consistent with the nonlinear profile of hydraulic conductivity at Gårdsjön (after Davies et al., 2011, with kind permission of John Wiley and Sons).*

package has a particle tracking component to predict transport (see Appendix A). Often these particles are advected only with the local mean pore water velocity, but dispersion can also be added by a random velcoity component chosen from a specified distribution around the mean.

The multiple interacting pathways (MIPs) model of Davies *et al.* (2011) is different in that the particle formulation is chosen to represent both the flow and the transport. The representation of celerity effects must therefore result from the movement of the water particles as the slope wets and dries. The initial implementation of the model assumes that flow in the unsaturated zone is vertical and in the saturated zone is parallel to the surface slope (the kinematic assumption). This was illustrated in Figure 9.4. The step equation for each particle at each time step in the saturated zone can then be expressed as:

$$x_{i,t+t_s} = x_{i,t} + v \tan \beta t_s \tag{11.7}$$

where t_s is the time step, x is horizontal distance, v is a randomly chosen Lagrangian velocity defined for a unit hydraulic gradient and β is the local slope angle. The choice of velocity distribution is therefore an important consideration in setting up the model. It was found for the case at Gårdsjön that this could be done in a way compatible with the field information on the strongly nonlinear profile of hydraulic conductivity. The velocity distribution at each level in the profile should integrate to be consistent with the conductivity at that level. By assuming an exponential velocity distribution for each level and that the increase in conductivity is a result of an increase in larger pores, functional relationships were developed between porosity, layer conductivity and velocity distribution (Figure 11.8).

Applying these relationships without further calibration resulted in the predicted hydrograph for the Gårdsjön slope in Figure 11.9a. Reproducing the tracer concentrations proved to be much more difficult, but thinking directly about velocities within the MIPs framework turned out to be valuable. The tracer was input to the system at a depth of 60 cm. 96% of the tracer was recovered at the outlet, so there was little uncertainty about the tracer mass balance in this case. If it was allowed to continue to move at the 60 cm

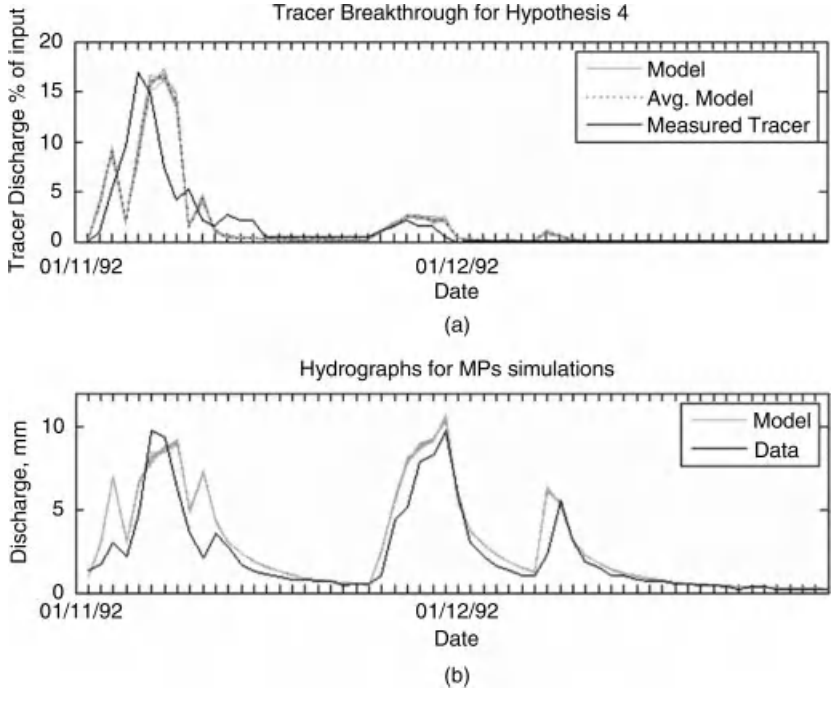


Figure 11.9 *Observed and predicted flow and tracer concentrations for the Trace B experiment at Gårdsjön, Sweden; the first three hypotheses about the tracer movement were rejected as inconsistent with the observations; the hydrograph simulation was based only on field data and the assumed velocity distribution of Figure (11.8) without calibration (after Davies et al., 2011, with kind permission of John Wiley and Sons).*

depth, however, it arrived at the outlet much too slowly. There was some evidence from piezometers that there was some upwards return flow into the soil though the bedrock. There was also evidence of tracer higher in the soil. The results of Figure 11.9b are based on a hypothesis that there was some initial mixing around the tubes where the tracer was injected and that lateral inflows from the sideslopes were supplied to the base of the slope forcing the tracer into the higher, more conductive layers. Simulating the second peak of tracer that is the result of displacement during the second spell of rain was the result of tracer being held in the upper layers of the soil when the simulated water table fell during the dry spell between events. It is then remobilised when the water table rises again so that particles that have been moving vertically start to move towards the outlet again. The hydrograph response suggests that the celerities, which depend on the rate of rise and fall of the water table, are also being quite well simulated.

11.10 Implications for Water Quality Models

A full review of water quality modelling is beyond the scope of this text, and there are existing texts that review the field (e.g. Ji, 2008). Clearly, however, there are important implications for water quality modelling of being able to identify the sources of water in the hydrograph and the full residence time distributions of water in a catchment, even if there is necessarily some uncertainty associated with such estimates.

To date, there have been very few studies that have attempted to model both flow and water quality in a totally consistent way. To do so requires that the different dynamics of celerities and the full distribution of pore water velocities would need to be properly represented (as well as all the other biogeochemical processes that affect water quality in the soil, groundwater and stream channels with their own uncertainties in process representations and effective parameter values). Most water quality models at the catchment scale rely on mixing in conceptual storage elements in some way, with no attempt to represent flow pathways (understandably so, given the complexity and time variability of flow pathways).

This is, however, a promising direction for future work and pointers in that direction are provided by the work of Botter *et al.* (2009, 2010) and Duffy (2010). Other techniques in this chapter, such as the end member mixing analysis (Section 11.5) can also be used in water quality contexts (see, for example, Jarvie *et al.* (2011) and references therein).

11.11 Key Points from Chapter 11

- The use of artificial and environmental tracers to understand and improve the respresentation of hydrological processes is a valuable tool and has led to new insights into catchment behaviour.
- The "paradox" of how catchments can store water for such long periods of time but so easily mobilise so much "old" water during storms is, in part at least, explained by the different storage mechanisms involved in the wave celerities that control the hydrograph response and the pore water velocities that control tracer transport.
- The assumptions required in analysing sources of water and residence time distributions represent significant simplifications, so that inferences may be subject to multiple sources of uncertainty. Uncertainty may also be introduced by the limitations of the sampling of concentrations in inputs and discharges. Consideration of the uncertainties associated with the analysis should be a matter of good practice.
- Collection of tracer concentrations at shorter time scales, including oxygen and hydrogen isotopes, is becoming cheaper and easier. This should mean that assumptions of linearity and stationarity in residence time distributions can be relaxed. Theoretical constructs that allow time varying residence time distributions still need to be properly tested.
- The use of both flow and tracer data to calibrate hydrological models needs to take account of the fact that the effective storages involved in the hydrograph and tracer responses are different. Thus, using tracer observations requires additional parameters to be identified, which means that the uncertainties in the discharge predictions may not necessarily be significantly reduced.

Box 11.1 Representing Advection and Dispersion

The concentrations of any solute or tracer moving through the catchment system with the flow of water are subject to *advection* and *dispersion*. Advection is controlled by the local mean velocity of the flow, dispersion by the distribution of velocities around the mean value. Similar theory is used to describe advection and dispersion for both surface and subsurface flows.

By analogy with molecular diffusion of concentration gradients in a solution, dispersion is often described by the product of the gradient of concentration of a tracer or solute in the water, C, and a scaling constant called the dispersion coefficient, D. This is a simple representation of the expectation that the steeper the concentration gradient, then the greater the net dispersive flux of solute. If there is no concentration gradient (the tracer is completely and homogeneously mixed in the water), there is no net dispersive flux even if the water is moving with a velocity distribution. This first-order gradient relationship is here known as Fick's Law:

$$J_x = -D\frac{dC}{dx} \tag{B11.1.1}$$

where J_x is the mass flux in direction x. The dimensions of J_x are $[ML^{-1}T^{-1}]$ and of C are $[ML^{-3}]$, so that D must have the dimensions $[L^2T^{-1}]$. Similar equations are used for heat transport in response to a gradient of temperature and, as shown in Chapter 5, for the transport of water in response to a gradient of hydraulic potential as Darcy's law.

If Fick's law is combined with a mass balance equation for the tracer or solute (this can be derived using the control volume approach explained in Box 2.3):

$$\frac{\partial C}{\partial t} = -\frac{\partial J}{\partial x} - \bar{v}\frac{\partial C}{\partial x} - R(t) \tag{B11.1.2}$$

where $\bar{v}$ is the mean flow velocity and $R(t)$ is a general source or sink term that will be zero for a fully conservative tracer. If D can be assumed constant and $R(t)$ can be assumed to be proportional to C, then subsituting for J from Equation (B11.1.1) gives

$$\frac{\partial C}{\partial t} = D\frac{\partial^2 C}{\partial x^2} - \bar{v}\frac{\partial C}{\partial x} - \alpha C \tag{B11.1.3}$$

This is the one dimensional "advection–dispersion equation" (ADE). The first term on the RHS represents the dispersive flux of solute; the second term, the advective flux; and the last term, a loss proportional to concentration. The ADE is a partial differential equation in which the dispersion coefficient D is, in general, different for different flow directions. This is often represented in two or three dimensions, again for both surface and subsurface flows, in terms of the "longitudinal dispersion coefficient" in the direction of the mean flow and the "transverse dispersion coefficient" orthogonal to the direction of mean flow.

As a quick exercise, show how the ADE can be derived using the control volume method of Box 2.3.

It can be shown that, for a conservative tracer, the ADE has an analytical solution for an impulse input. For the one-dimensional case of only longitudinal dispersion in the direction of the mean velocity and $\alpha = 0$, then:

$$C(x, t) = \frac{M}{2A\,(\pi Dt)^{0.5}} \exp\left[-\frac{(x - \bar{v}t)^2}{4Dt}\right] \tag{B11.1.4}$$

where M is the mass of solute added, A is the cross-sectional area involved in the flow and π is the constant pi (3.141592 ...). This is the same form as that used to describe the normal or Gaussian distribution in statistics (Figure B11.1.1). It implies that the impulse input will gradually disperse as it moves through the system with a Gaussian distribution in the space

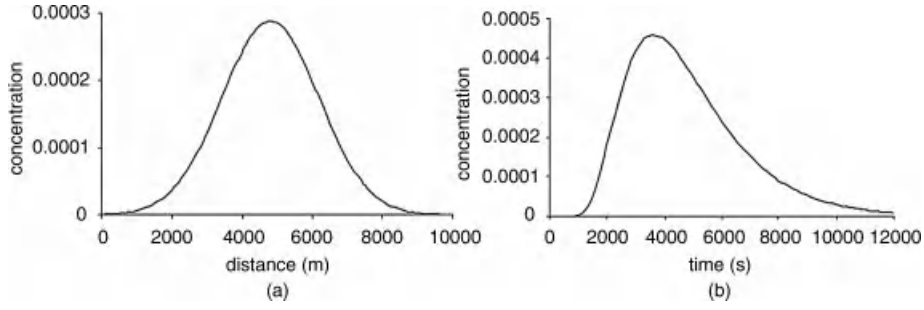


Figure B11.1.1 *Solution of the advection–dispersion equation (Equation (B11.1.4)): (a) plotted against distance in the flow direction; (b) plotted against time for a single point in space.*

coordinate in the direction of flow. The further the travel of the tracer through the system, the lower the peak concentration and the greater the spread but the cloud of tracer will continue to have the form of a Gaussian distribution if the theory holds. It is also a linear solution, so that twice the mass input in the pulse will result in a curve of the same form with twice the concentration. For steady flow conditions, it can therefore be used as a linear transfer function in the same way as the unit hydrograph is used in predicting runoff generation in Section 2.2.

This form of this solution is quite revealing in that it implies that the dispersion is the result of each tracer particle sampling all the various velocities in the distribution of the mean velocity over time. With a large number of particles, and regardless of the underlying form of the actual velocity, then the law of large numbers suggests that as all the velocities are sampled then the resulting distribution will tend asymptotically towards a normal distribution. Before that is the case, in the very early stages of the dispersion, the shape of the concentration curve will be skewed in space because particles of tracer that start with fast velocities will not have had time to sample the slow velocities. How long it will take for this initial mixing to be complete depends on the structure of the flow domain and the accessibility of all the possible pathways. The length scale of the initial mixing process is called the "mixing length". Equation (B11.1.4) is strictly valid only after this initial phase of mixing has been reached. Tracer experiments have revealed that, in both surface and subsurface systems, the mixing length can be long, sometimes very long, such that the dispersion coefficient that is calculated from the results of an experiment can depend on the length scale of the experiment (e.g. the work of Gelhar *et al.* (1992) in groundwater systems, Beven *et al.* (1993) in soils and Green *et al.* (1994) in rivers). Thus the skewness of the initial mixing phase can persist for long distances and long times. However, in applying the ADE for prediction, it is often assumed that the initial mixing is complete and the dispersion coefficient is constant or a simple function of velocity.

In fact, it is somewhat rare to measure the dispersion of a tracer as distributed in space at a given time. Much more often the time distribution is measured as the cloud passes a particular point in space. This depends on making only sequential measurements at a point rather than simulataneous measurements at many points so is much easier. The resulting concentration curve should then be expected to be skewed in time (even if it had the symmetric Gaussian distribution in space), since particles arriving later will have been subjected to more dispersion than those in the early part of the cloud. Because of incomplete mixing, however, that skewness will be greater and will persist for longer, so that concentration curves remain skewed with long tails.

The ADE can be modified to help simulate these skewed distributions by introducing an exchange with an immobile store. In soils and groundwaters, this is known as the

"mobile–immobile model" (van Genuchten *et al.*, 1989) and in rivers as the "transient storage model" (Bencala and Walters, 1983). The general form is:

$$\frac{\partial C}{\partial t} = D\frac{\partial^2 C}{\partial x^2} - \bar{v}\frac{\partial C}{\partial x} - \alpha\,(C - C_s) \tag{B11.1.5}$$

$$\beta\frac{dC_s}{dt} = \alpha\,(C - C_s) \tag{B11.1.6}$$

where C_s is the concentration in the immobile or transient storage, α is the mass exchange coefficient between the main flow and the storage zone and β is the ratio between the volumes of the storage zone and the main flow. This also has a linear analytical solution for predictions beyond the mixing length (see, for example, Wörman, 2000; De Smedt *et al.*, 2005).

The ADE can also be solved numerically, using similar methods to those described in Box 5.3. Care must be taken in this case, however, so as not to introduce too much *numerical dispersion* into the solution due to the numerical approximations to the partial differentials. This might cause unrealistic dispersion in the predictions of transport relative to that expected for a given value of the dispersion coefficient. It can also induce oscillations in the solution in the vicinity of sharp concentration gradients. There is an extensive literature about the solution of such equations, which generally require higher order numerical approximations to be used (see, for example, Pinder and Gray, 1977). Fortunately, with the greater computer power now available, this is not such a problem and in real applications the problem of defining the appropriate sources and boundary conditions for the transport may still be much more significant.

There is an alternative dispersion model that has been derived directly from data analysis of tracer concentrations in rivers and soils. This is the aggregated dead zone (ADZ) model (Beer and Young, 1983; Wallis *et al.*, 1989; Green *et al.*, 1994). It is also a linear model and takes the form of the simple transfer functions described in Section 4.3 and Box 4.1. In the ADZ model, advection is treated as a pure time delay and dispersion by routing through one or more linear stores. It generally produces very good fits to available tracer data, including the long tails that are often a feature of observed tracer curves. If both input and output concentrations to the flow domain can be considered well mixed, for example in a river reach where tracer has been added some way upstream of the input to the reach, then only a single linear store (a first-order transfer function) should be necessary. In the intial stages of mixing, or where there is mixing with a high volume storage, a higher order model might be required.

The dispersion in the ADZ model is controlled by the mean residence time of the linear store (or stores). In the first-order case, under steady flow conditions, this has a very simple intepretation. The ADZ model is best understood in terms of simple "lag and route" concepts. In any transport problem, there is a lag between the time at which the solute concentration starts to rise at the start of a reach and when it starts to rise at the end of the reach. This, in the ADZ methodology, is called the advective time delay, τ. It is expected to decrease with increasing discharge in the reach. The dispersion of the input plume in the reach is modelled as a linear transfer function. In the simplest first-order case, the transfer function is equivalent to a linear store with a mean residence time, T. The mean travel time in the reach is then $(\tau + T)$. Given a steady discharge, Q, the volume of this effective mixing store is then $V_e = Q * T$. The total volume in the reach involved in the transport process is $V * [\tau + T]$. The ratio V_e/V is called the "dispersive fraction" (*DF*). It can also be defined in terms of the adective travel time and mean residence time as $T/(\tau + T)$.

In some analyses of river tracing experiments it has been shown that a higher order transfer function can give a slightly better fit to the observed tracer experiments (e.g. Wallis *et al.*, 1989; Young and Wallis, 1993). This suggests that there might be different mechanisms affecting the dispersion with rather longer mean residence times than in the bulk flow and its

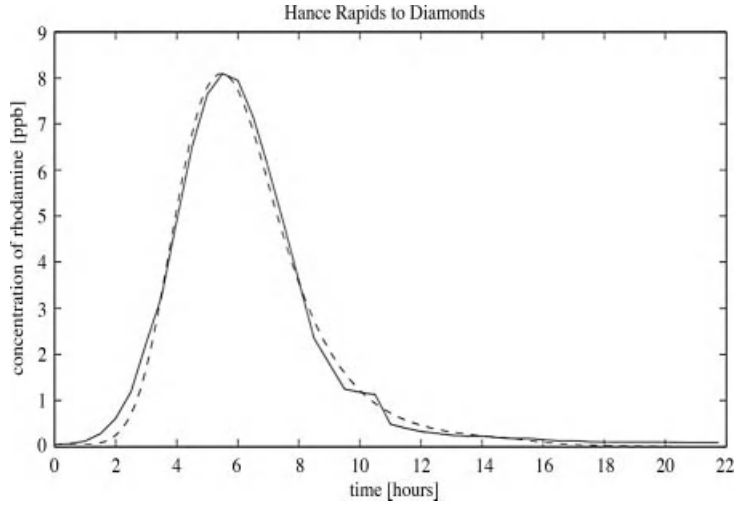


Figure B11.1.2 *Fit of the aggregated dead zone model to a tracer experiment for the Colorado River for the Hance Rapids to Diamond reach (238.8 km).*

associated dead zones. This might be due to exchanges into larger dead zones or perhaps mixing with hyporheic zone waters in the bed and banks. It is worth noting that, as shown by Reynolds *et al.* (1991), the lateral velocity shear at the boundary between a dead zone and the main flow can induce locally efficient mixing. However, once solute or pollutant is transferred into such a storage, mixing is much slower and the time scale of pollutant retention much longer. This contributes to the heavy tails of observed tracer concentrations. Examples of fitting a simple first-order ADZ model to tracer data from the Colorado River is shown in Figure B11.1.2.

Both the ADE and ADZ models are difficult to use in predicting catchment scale transport of tracers or solutes. This is because of the complexity of the flow pathways and the possibility that at the soil profile and hillslope scale, mixing is still in its initial stages so that it is difficult to justify the assumption of the ADE that mixing is beyond the mixing length (or at least to be able to estimate the dispersion parameters required at a given scale). Despite this it has been used in some practical applications; quite widely in the prediction of transport and dispersion (most hydraulic routing packages have a transport and dispersion option that uses the mean velocities calculated for the flow); and in groundwater transport (again, many groundwater modelling packages have an option to predict dispersion based on the ADE, such as MT3D for MODFLOW, see Appendix A).

The ADE model has also been applied at the catchment scale in an interesting way, by simplifying the subsurface flow pathways into a series of one-dimensional *stream tubes* (e.g. Simic and Destouni, 1999; Persson and Destouni, 2009). The steady mean velocities are then calculated in each stream tube, taking account of the changing cross-sectional area of each stream tube. The ADE is then solved for the one-dimensional transport along each stream tube. The delivery of all the stream tubes to the channels in the catchment provides a prediction of the complete transport on the hillslopes, given some specified sources of the solute of interest. Biogeochemical reactions can be added to these calculations for non-conservative solutes.

Box 11.2 Analysing Residence Times in Catchment Systems

The analysis of residence times of water in catchment systems has received much attention in the last decade, in part because of the improvements of measurement techniques for environmental tracers in terms of both cost and reliability, and in part because of the additional information provided about flow sources and pathways in understanding how the catchment works. The presentation here builds upon a number of review papers where more detail may be found (see particularly Kendall and McDonnell, 1998; Ozyurt and Bayari, 2005a, 2005b; McGuire and McDonnell, 2006).

As discussed in the main text of Chapter 11, difficulties of adequate observations and the strong simplifying assumptions required in the analysis of residence times remain. The interpretation of the results must bear these limitations in mind. Sources of uncertainty in the analysis should be considered as a matter of good practice.

The first and most important requirement for a successful analysis is that there needs to be a clear signal, a difference in concentrations between inputs and outputs. This is quite obvious for the simple mixing models of Section 11.3, for which the calculation cannot be carried out if there is no difference in concentration between input and output. The same holds for longer term input and output signals. In this case, certainly for the isotopes of the water molecule, we expect the variability in the isotope signal in the input to be smoothed out in the output signal. The catchment acts as a low-pass filter for the higher frequency variations in the input. However, if there is no variation at all in the output, we also cannot infer much about the processes since the input variability has then been integrated out completely. We can say that there must be a very large effective storage in the catchment relative to the input fluxes but the mean residence time of the storage would be indeterminate. Thus, the ideal case is to have some structured variability in the input that can still be distinguished as a signal in the output after the smoothing in the catchment.

B11.2.1 Forms of Residence Time Distribution

Analysis of residence time distributions has classically been based on treating the output concentrations as a convolution of the input concentrations and the residence time distribution such that:

$$C_Q(t) = \int_0^t C_P(\tau) f(t-\tau) d\tau \tag{B11.2.1}$$

where $C_Q(t)$ is the sequence of output concentrations, $C_P(t)$ is the sequence of input concentrations and $f(t)$ is the residence time distribution. Note that this form is not as general as Equation (B11.1.2) in that it assumes linearity and stationarity of the residence time distribution in relating the inputs to the outputs.

Given an identification of the residence time distribution $f(t)$, the mean residence time $\bar{T}$ can then be calculated as:

$$\bar{T} = \frac{\int_0^\infty t f(t) dt}{\int_0^\infty f(t) dt} \tag{B11.2.2}$$

The form of residence time distribution of water in a catchment is not known *a priori*. In one of the first analyses of this type, Eriksson (1971) argued for an exponential distribution:

$$f(t) = \frac{1}{K} \exp(-t/K) \tag{B11.2.3}$$

where the parameter K has units of time and is equal to the mean residence time. As shown in Box 4.1, this is the output from a simple well-mixed linear store, or what is known in the

chemical engineering literature as a "continuous stirred tank reactor" (CSTR). This has been used widely as a model of the residence time distribution since.

A more general form, which includes the exponential distribution as a special case, but which involves more parameters, is the gamma distribution:

$$f(t) = \left(\frac{t}{K}\right)^{N-1} \frac{\exp(-t/K)}{K\Gamma(N)} \qquad \text{(B11.2.4)}$$

where $\Gamma(N)$ is the gamma function ($\Gamma(N) = (N-1)!$ for integer values of N). This form was also used as a functional form for the unit hydrograph (Equation 2.2) but the parameters for $f()$ for predicting residence time distributions are not the same as $h()$ for predicting hydrographs. The general linear transfer function models of Box 4.1 could also be used (Young and Wallis, 1993; Young, 2011a; see also Amin and Campana, 1996). The way in which the gamma distribution changes shape for different values of the parameters K and N is shown in Figure B11.2.1. A third, time shift, parameter can also be introduced.

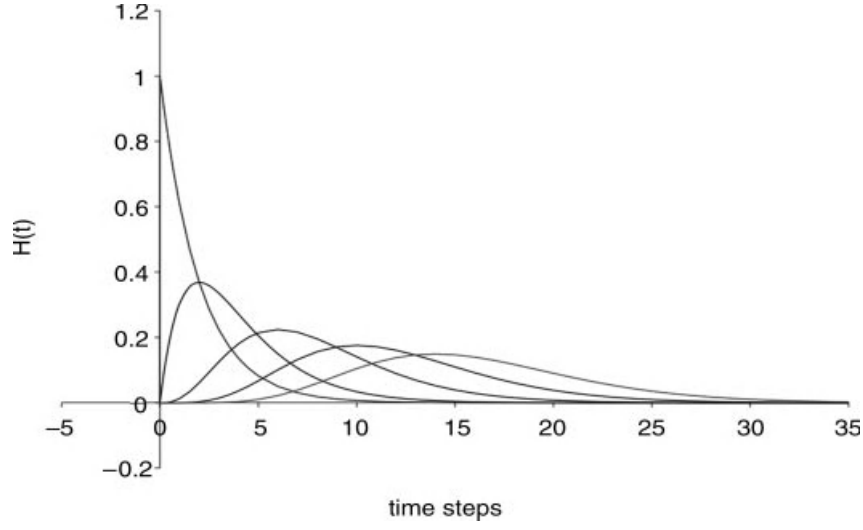


Figure B11.2.1 *Forms of the gamma distribution for different values of the parameters K and N.*

A number of packages exist to fit residence time distributions to tracer concentration observations (e.g. Maloszewski and Zuber, 1993; Ozyurt and Bayari, 2005a; the TFM package (Appendix A) could also be used to fit ADZ models). A particularly strong assumption in fitting a distribution to time series of inputs and outputs is that the residence time distribution used should be linear and stationary. Linearity means that there is a constant scaling of the integrated inputs and outputs; stationarity means that the distribution is treated as constant in time. These assumptions simplify the fitting process (only a small number of parameters need to be fitted once a distributional form has been chosen) but are not true. A number of studies have compared the use of different distributional forms to data under these assumptions but there have been few studies that have evaluated the effect of the linearity and stationarity assumptions on the inferred residence times in the catchment.

It is well known from the theory of linear systems that the fitting of convolution kernels, such as the residence time distribution $f(t)$ of interest here, to noisy data can be improved by the inherent filtering involved in the use of Fourier or Laplace transforms (for applications in this area see e.g. Eriksson, 1971; Duffy and Gelhar, 1985; Kirchner, 2005). The type of

robust identification algorithm used in the CAPTAIN package of the TFM modelling package (Appendix A) are also based on a filtering approach, in that case applied recursively (Young, 2011a).

The Fourier transform transforms the original observations into frequency amplitudes. The use of standard Fourier-transform-based algorithms usually requires regularly spaced observations with little or no missing data. This is not generally the case for catchment input and output tracer concentrations. Kirchner (2005), however, presents an algorithm that can be used with irregular spaced data and investigates the difficulties introduced by various factors including the "aliasing" of the input and output signals.

B11.2.2 Introducing Time Variability

The steady state flow assumption and constant residence time distribution assumption greatly simplify the identification problem and many of the published studies of catchment residence time distributions are based on these assumptions. However, given the concern in much of the rest of this book with trying to reproduce the short time scale dynamics of catchment response, these assumptions seem to be a rather extreme simplification. It greatly simplifies the identification problem, but clearly we need to be wary about the interpretation of the results. It is therefore worth considering whether at least some of the assumptions can be relaxed.

The first simple modification that can be made is to allow for the variability in discharge as the catchment wets and dries. This can be done very approximately by carrying out the analysis in terms of a time scale transformed by discharge flux rather than time (Rodhe *et al.*, 1996; Simic and Destouni, 1999) so that

$$t_Q = \frac{\int_{t_o}^{t} Q(t)dt}{\bar{Q}} \tag{B11.2.5}$$

where $\bar{Q}$ is the mean discharge over the period. This is a very useful approach where there is a strong seasonal dependence of flow, such as in catchments dominated by snowmelt runoff. Snowmelt also tends to affect the isotopic concentration of inputs to the system. Studies have shown that the earliest melt can be relatively light in ^{18}O and deuterium, with progressive enrichment over time. This has an effect on the inferred residence time distribution (Lyon *et al.*, 2009).

It would be better still if both inputs and outputs could be transformed to normalised cumulative mass flux scales so that

$$M_P = \frac{\int_{t_o}^{t} P(t)C_P(t)dt}{\int_{t_o}^{T} P(t)C_P(t)dt} \tag{B11.2.6}$$

$$M_Q = \frac{\int_{t_o}^{t} Q(t)C_Q(t)dt}{\int_{t_o}^{T} Q(t)C_Q(t)dt} \tag{B11.2.7}$$

where T is the length of the complete time period considered, but this immediately makes two issues more readily apparent. The first is that normalising the inputs and outputs in this way obscures any mass imbalance between inputs and outputs. The second is that the concentrations may not have been sampled frequently enough (particularly in the early days of isotope analysis) to properly characterise the mass flux of tracer during storm periods. This is still a problem with some of today's studies – the catchment studies of Soulsby *et al.* (2000) were based on weekly to fortnightly samples of isotope concentrations; even the Kirchner *et al.* (2000, 2001) studies at Plynlimon were based on daily samples of chloride as a tracer, but in a catchment where the hydrograph generally peaks in two to four hours. Even the seven-hourly

samples that have been collected at Plynlimon since 2009 do not give a complete picture of the variations in chemistry during storm periods. In addition, where studies of the variability in tracer concentrations in the inputs have been made, distinct variability in the measured concentrations in space and time has been found (e.g. McGuire *et al.*, 2005). Thus, most analyses are not based on a full characterisation of inputs and outputs.

The limitations of the assumption of a residence time distribution that is linear and stationary are increasingly recognised. Theoretical frameworks that allow for the time variability of the residence time distribution are being developed (e.g. Duffy, 2010; Botter *et al.*, 2010; Rinaldo *et al.*, 2011) and the prediction of tracer concentrations is being integrated with rainfall–runoff models (Davies *et al.*, 2011). By making hypothetical simulations, it is relatively easy to show that the changes in the residence time distribution, and hence in the inferred mean residence time, can be very significant as the catchment wets and dries. These models involve additional parameters, however, and have not to date been adequately tested against observational data which may not (at least as yet) contain sufficient information to estimate all the parameters required. It may indeed, be difficult to reproduce, or even to understand, the short time scale variations in concentrations that have been observed in some experiments (Page *et al.*, 2007).

B11.2.3 Tracer Residence Times and Hydraulic Turnover Times

The mean residence time of a tracer derived from $f()$ represents the effective turnover time for water in the catchment system. We can also derive a hydraulic turnover time from the $h()$ unit hydrograph in Section 2.2. These will be different (even if they are parameterised within the same family of distributions, such as the gamma). It might seem surprising that they should be expected to be different if a tracer is conservative. A conservative tracer is, after all, expected to follow the velocities of the water in the various storages in the catchment. However, there are several reasons why the tracer mean residence time and effective storage might differ from the hydraulic mean turnover time. One is the difference between wave celerities and velocities discussed earlier in this chapter. The implication of the wave celerity being faster than the mean pore velocity is that the effective storage for the hydrograph response is different from that for a tracer response (see the case studies in Sections 11.7 and 11.9). In addition, if the mean residence time is determined from a distribution fitted using the simplifying assumptions discussed above, it will be subject to all the uncertainties inherent in the fitting process.

A second reason is that both the shorter and the longer residence times of water in catchments might not be adequately represented by the available observations: the shorter because of the sampling intervals of tracer concentrations missing the dynamics during storm periods, the longer because of limited time series of data or choice of tracer. Thus, what appears in the residence time distribution may not sample all the storages in the catchment, either because it is relatively immobile (and may not interact with a mobile tracer except over very long time periods) or because some of the fast flow pathways are simply missed by the sampling interval.

12

Beyond the Primer: Hypotheses, Measurements and Models of Everywhere

. . . although co-production involved both academic Scientists and local members becoming scientists, there were still some roles, of retained expertise, that were left intact. The academic scientists, for instance, had to code the model and handle some of the large datasets required by the model. Local scientists had detailed knowledge of particular places and issues that a Scientist could not possibly bring to their work. The difference was that, in theory, any of these elements could be subject to wider interrogation and the means by which this interrogation happened was the practice of doing flood risk science and not only discussing or debating scientific reports.

Stuart Lane *et al.*, 2011

Scientists have no monopoly on wisdom where this kind of trans-science is involved: they will have to accommodate to the will of the public and its representatives. The republic of trans-science, bordering as it does on both the political republic and the republic of science, can be neither as pure as the latter nor as undisciplined as the former. The most science can do is to inject some intellectual discipline into the republic of trans-science; politics in an open society will surely keep it democratic.

Alvin Weinberg, 1972

12.1 Model Choice in Rainfall–Runoff Modelling as Hypothesis Testing

It should be apparent from the previous chapters that a very wide range of models are available to any rainfall–runoff modelling application without any clear basis for making a choice between them.

Rainfall–Runoff Modelling: The Primer, Second Edition. Keith Beven.

Recall the criteria that were established in Chapter 1 for model choice. These may be summarised as follows:

- Is a model readily available or could one be made available if the investment of time (and money!) appeared to be worthwhile?
- Does the model predict the variables required by the aims of the particular project?
- Are the assumptions made by the model likely to be limiting in terms of what you know about the response of the catchment you are interested in?
- Can all the inputs required by the model, for specification of the flow domain, for specification of the boundary and initial conditions and for specification of the parameters values, be provided within the time and cost constraints of the project?

At this stage it should be apparent that these criteria essentially provide the basis for model rejection and, as pointed out way back in Chapter 1, it is all too easy to reject all the available models due to inadequate assumptions or infeasible demands for input data. This is not very helpful: in very many projects, the hydrologist is still required to make predictions of what might be expected in terms of flood peaks, reservoir inflows or other variables under different conditions. Thus one or more models will need to be retained.

This is, however, where the idea of conditioning of models becomes very important. We can, to some extent, overcome some of the limitations of the available models by conditioning their predictions on any available observations or prior knowledge about the catchment of interest. Traditionally, this has been done by the calibration or optimisation of parameter values: it is suggested in Chapter 7 that a more general strategy of conditioning within an uncertainty framework is a much more satisfactory approach for the future. In this framework, **any** model that predicts a variable of interest is a potentially useful predictor, until there is evidence (or a justifiable opinion) to reject it. The value of conditioning in this way is that the model or models retained must be consistent with the available data (at least to some level of acceptability), otherwise they would have been rejected. This gives some basis for belief in the predictions when those models are used to extrapolate or predict responses for other conditions.

Try to picture the modelling process as a form of mapping of a particular (unique) catchment or element of a catchment into a model space with dimensions of different model structures and parameter values. Chapter 7 shows how this mapping can be done using conditioning based on likelihood measures or fuzzy weights. For many reasons, primarily associated with the limitations of model structure and parameter estimation, the mapping is necessarily approximate. The mapping can be done for both gauged catchments and ungauged catchments based on whatever prior information or observations are available, but is likely to be more exact where observations of hydrological response are available and more approximate for the ungauged catchment case. There is, however, the potential to refine this mapping by the collection of more data and this approach to model evaluation also has implications for the types of data collection exercise that might be used as the basis for a rainfall–runoff modelling study.

In fact, as discussed in Section 7.17, the process of model evaluation can be set up in terms of hypothesis testing. Given a set of models for a catchment that have survived the initial selection and evaluation procedure, what hypotheses can be tested by the collection of data to allow some of those models to be rejected? It would seem that some types of data might be more valuable than others in testing models and in model rejection. For example Lamb *et al.* (1998b) considered the use of water table information in predicting the responses of the small Saeternbekken catchment in Norway (see Section 6.4). Their results highlighted the problem of using spatially distributed observations in model evaluation in that local water table responses depend strongly on the local transmissivity and storage characteristics of the soil. Thus it is unlikely, given the heterogeneity of soil characteristics to be expected in a catchment, that even if the model structure were correct, an effective transmissivity and storage coefficient calibrated for a catchment would predict local variations in water table depth accurately. Local parameter values

would be needed to make accurate local predictions. Essentially their conclusion was that knowledge of local water tables greatly improved the prediction of local water tables (although even then there were significant local anomalies) but did not help much in constraining the uncertainties in discharge predictions as estimated using the GLUE methodology.

The difficulty of making measurements remains an issue here. It is much easier (and very much cheaper) to make measurements at, or very close to, the surface. The types of measurements available for the investigation of subsurface flows are much more limited and we still do not have any adequate non-destructive way of assessing preferential flow pathways, whether they be due to natural piping or through mechanically induced cracks to mole drains in an agricultural field. Our perceptual model might allow for the possible importance of such processes but, if their effects must be inferred rather than measured, it is difficult to define an appropriate model description or condition a model on the basis of perceptions alone.

A view of model calibration as a process of hypothesis testing and model rejection would seem to be a positive one in moving from a situation in which the limitations of a simple optimisation approach in the face of parameter identifiability problems have become increasingly clear. It also brings together the variety of methods described in Chapter 7 in which multiple criteria are used in the model evaluation, from Pareto optimal sets to GLUE. Each has its own way of deciding on which models should be rejected and what weighting coefficients should be used in projecting the predictions of the retained models. In the Pareto optimal set method, all models that are not part of the set are rejected and each retained parameter set is given equal weight. In formal Bayes methods, the likelihoods are based on a statistical model of the residuals. In GLUE, different ways of combining informal likelihood measures can be chosen and the weights are based on the current likelihood value associated with each set.

12.2 The Value of Prior Information

This type of conditioning framework does, however, allow for the use of prior information in the conditioning process. There are many types of prior information that might be used, for example prior degrees of belief in different types of model structure; prior estimates of ranges of parameter values that might be appropriate for different vegetation, soil or rock types; and the perceptual model of how a particular catchment might respond to rainfall. Such information can constrain, in a sensible way, the range of modelling possibilities to be considered in an application as a result of previous experience.

The use of prior information requires making choices and, in particular, rejecting possibilities. As stressed earlier, such choices need to be made with care. There is no shortage of examples in the literature of prior choices of inappropriate models, particularly the use of models based on infiltration excess concepts in applications where this may not be the dominant runoff mechanism. It might still be possible to calibrate or condition the model parameters to reproduce the observed discharges but should a model based on inappropriate concepts be used in extrapolation?

The fact that calibration can bring nearly all rainfall–runoff models into line with an observed discharge record can therefore be a problem as well as an advantage, since in many cases we do not know if an inappropriate model is being used. There is, for example, a very natural tendency in making prior choices about model structures for every model developer to give a prior weight of one to his or her model and a prior weight of zero to all other models. This is, however, not a necessary choice: multiple models may be included, as additional dimensions of the model space, within the modelling framework suggested here.

There is also a natural tendency to give a greater prior weight to models that are considered to be physically based relative to those that are more conceptual or data based in formulation. physically based models should, in principle, reflect our understanding of hydrological systems more closely and therefore be more robust in extrapolation to other conditions. This is only true, however, if such models truly reflect

the nature of the catchment processes and it is not clear whether we have really reached this stage, given the limitations of current physically based process descriptions in really representing the nonlinearities of hydrological processes at the scales of hillslopes and catchments and the difficulties in estimating effective values of the model parameters (see Chapter 9).

12.3 Models as Hypotheses

In testing models as hypotheses, these types of uncertainty in the modelling process mean that there is always the possibility of accepting a poor model when it should be rejected (false positive or Type I error) or rejecting a good model when it should be accepted (false negative or Type II error). In statistical hypothesis testing, we generally choose to test a hypothesis only with a certain probability of being wrong (e.g. at the 5% level). It is difficult to carry this over to simulation models applied to places and data sets that are unique in space and time with only single realisations of any observational errors (see Beven, 1981). Except in special circumstances, we cannot replicate sets of observations so any such 5% error criterion would have to be assessed in some different way.

Consider each of the possibilities for being wrong. We would wish to avoid Type I errors in testing because the performance of models falsely retained for use in prediction might lead to false inferences and poor decision making. So how to avoid Type I errors? The primary reason why a Type I error might occur is because there is enough uncertainty in the inputs to the model and the observational data used in evaluation that whatever performance indicator is used to make a decision about model acceptability, it cannot differentiate between good and poor models. This might be because of a particular or peculiar sequence of observational errors, including "rogue" observations. To avoid Type I errors, we need to be careful about the commensurability of observed and model variables and try to ensure that only periods of good quality data are used in calibration.

It is perhaps more important still to avoid Type II errors. If we do retain a poor model for use in prediction, then hopefully further evaluation in the future might reveal that it gives poor predictions and therefore our choice can be later refined. But we really would not want to eliminate a good model just because of input errors. However, again because of particular or peculiar sequences of observational errors, it may be difficult to differentiate between models that are good in prediction from those that are not. This might also be the reason why a model that performs well in calibration does less well in testing, purely because of different input error characteristics. Again, it is necessary to be careful about commensurability of observed and model variables and ensure that only periods of good quality data are used in calibration or validation.

So is there any possibility of distinguishing between Type I and Type II errors in this type of model application? Some error is inevitable and almost certainly not reducible to statistical noise while the characteristics of different sources of error in calibration might well be different from those in prediction. It therefore remains difficult to avoid both Type I and Type II errors until there is a good case for rejection when a model performs poorly but we have belief that the observational data are adequate. The conclusion therefore is that we should use a rejectionist approach to hypothesis testing while trying to avoid Type II errors by avoiding periods of low quality or "rogue" observations (see Section 7.17).

This then suggests that, to achieve some objectivity in hypothesis testing, the criteria for model rejection need to be set independently of any model run. One way of doing so is the "limits of acceptability" approach suggested in Section 7.10. This is a form of possibilistic rather than probabilistic model evaluation. To avoid rejecting a good model because of poor input data, the limits of acceptability should be set so as to reflect the potential effects of input, boundary condition, and observation error and should be set prior to running the model.

The set of models which consistently fail such limits of acceptability tests should be rejected as hypotheses, even if that means rejecting all the models tried. This is, of course, a positive result because

of the Popperian asymmetry of learning by falsification: i.e. we learn more from falsification of a hypothesis than from one that cannot yet be rejected on the basis of the available observations (although Popper did allow that there might be degrees of verisimilitude and falsification). If a model as hypothesis satisfies some (relaxed) limits of acceptability then there is no guarantee that it is getting the right results for the right reasons. If it is not, then it may still fail in prediction (although such models are necessarily still our best bet in making predictions). We can only continue to test for rejection as new data become available.

On the other hand, if a model fails the limits of acceptability (and we are happy that we have taken adequate account of uncertainties in the input data so as to avoid Type II errors) then we have eliminated a model from the prediction process. If all the models tried fail such a test, then there might be information in the periods or types of failure that would suggest how a model might be improved, resulting in a positive learning process, albeit with work still to be done.

Given these (rather serious) issues, the problem is how to define when a model is acceptable, conditional on some imperfect evaluation data. This might be different for scientific research and for practical applications. For scientific research, hypothesis testing (ideally) should be posed in terms of falsifiability as discussed above. But hydrology is also a practical science and, for practical applications, what is needed of a model is that it should provide predictions of future outcomes that lead to good management decisions (or, at least, do not lead to poor management decisions). Any conditioning or evaluation process given some imperfect data is only a means to that end. We want to avoid Type I errors but the potential for epistemic errors means that there will be limitations as to how far success in prediction can be guaranteed by success in calibration or conditioning. Such errors are a form of knowledge uncertainty (and, when referring to errors that could not be known beforehand, are sometimes called Type III errors). The assumption that error characteristics in prediction are the same as in calibration is a convenience when we actually expect nonstationarity of sources of uncertainty.

These are some of the issues that underlie the testing of models as hypotheses about system functioning. The dialogue between experimentalists and modellers now needs to proceed to the stage of agreeing what might constitute appropriate limits of acceptability or other principles on which adequate hypothesis tests might be based. A key question, therefore, is just how we can define an adequate hypothesis test that would, in fact, identify models that were getting the right result for the right reasons (Klemeš, 2000; Kirchner, 2006; Beven, 2010) or, indeed, pose the right sorts of question (see Sivapalan, 2009). We have mostly been trained to consider that hypothesis testing should normally involve a statistical analysis and there is certainly no shortage of statistical techniques for model choice and model testing. The use of such methods, however, implies that the sources of uncertainty that affect model performance can be considered as if they are random. The problem in hydrological modelling is that there are uncertainties arising from lack of knowledge or understanding that it might be misleading to treat as if they were random. In Chapter 7, these two types of error were introduced as *aleatory* and *epistemic* errors.

The recognition of epistemic errors has two important implications. One is that it cannot be assured that the nature of errors in prediction are the same as in calibration. Of course, we have no idea what the nature of such errors might be in prediction so it is usually only possible to assume that they are, in some sense, similar to the calibration period. But it does have a second implication in that the value of the information in calibration might be less than that implied by calculating the type of formal statistical likelihoods shown in Box 7.1. In particular, some residual errors in calibration might be *disinformative* about what constitutes a good hypothesis of the system and consequently introduce bias into any calibration process (see, for example, Beven *et al.*, 2008; Beven and Westerberg, 2011; Section 7.17).

It may not be sufficient to test model performance purely on the basis of the output observations from a catchment system. It would be more rigorous to also test whether the model can predict the changing internal states of the catchment adequately. This raises three further problems in hypothesis testing. The first is the possibility of *incommensurability* between observed state variables at a certain scale in space and time and the equivalent model state variables, often predicted at different space and time scales

(Beven, 1989, 2006a). The comparison cannot be done at all, of course, if the model state variables have no measureable physical equivalents.

A second, related, problem is the use of spatially distributed parameters in a model. It is difficult to use most observations of internal catchment states in model evaluation without making use of a distributed model. But then, as noted above, in order to predict a local response correctly, distributed parameter fields might be needed, which then introduce many additional degrees of freedom. If, on the other hand, global values of parameters (such as hydraulic conductivity or transmissivity) are used, then we should expect only approximate predictions of local variables nearly everywhere, where the global parameters cannot reproduce the local responses (e.g. Lamb *et al.*, 1997; Blazkova *et al.*, 2002a; Vazquez *et al.*, 2009).

A third problem is that any prediction requires the knowledge of future boundary conditions. We have already noted that the error characteristics of observations used in calibration might be different from those in prediction, and there are very many reported hydrological modelling studies where poorer performance in prediction, relative to calibration, has been accepted because there is an expectation that this will be the case. Certainly, in the case of groundwater models, poor performance in post-audits could often be assigned to poor specification of future boundary conditions (see the review by Konikow and Bredehoeft, 1992).

12.4 Models of Everywhere

We have seen how detailed distributed models of large catchments and even of whole countries are becoming more and more computationally feasible. We have also seen how many of the available distributed models depend on very approximate representations of hydrological processes. In implementing such models of everywhere, however, it is a hydrological modelling aphorism that every catchment is unique, making regionalisation or the prediction of the responses of ungauged catchments difficult (see Chapter 10). That is for the water: it is even more so for water quality and ecological variables. So, given this uniqueness, is it even useful to think in terms of models of everywhere and everything in catchment management when such models will inevitably be wrong in some places or some of the time?

Looking ahead to the availability of models of everywhere, Beven (2007) argues that it will be useful, and even necessary, to think in these terms. It will change the nature of the modelling process, from one in which general model structures are used in particular catchment applications to one in which modelling becomes a learning process about places. In particular, if a model is obviously wrong in its predictions about a place, then this will be an important driver to do better. This has already been seen in Denmark, where the National Water Resources Model is already in its fourth generation (in almost as many years) because it was deemed to be wrong in its implementation in some parts of the groundwater system (Henriksen *et al.*, 2008). Every successive generation should be an improvement. The uncertainties in the modelling process will not, of course, disappear (particular with respect to future boundary conditions) but they may be gradually constrained. If, in the words of George Box, all models are wrong but some might be useful, then we would hope that models of everywhere would become increasingly useful to the management process as the representation of processes in particular places is improved.

This learning process about place is a way of doing science in complex open adaptive systems. Model hypotheses can be tested within the limitations of the uncertainties in available data and either survive locally or be rejected. As new data become available, further tests can be carried out as part of the learning process. If the models survive some agreed testing process then they can be retained for use in prediction. Uncertainty might mean that multiple models survive. We can therefore only generally say that those models that have survived a testing process up to now are the best we have available for prediction, subject to future testing as new information becomes available.

Treating the modelling problem as a learning process implies a modelling system that needs to be flexible, with the possibility of replacing a component or components to suit local conditions, or replacing

an existing model structure with a more complete or more satisfactory representation of the local processes (perhaps including next generation components such as those discussed in Chapter 9). Some recent work on flexible modelling systems has already been reported, notably the FLEX and FUSE systems cited in Sections 2.4 and 4.6.4. Branger *et al.* (2010) discuss past work in providing distributed modelling systems and introduce the LIQUID modelling system. There has also been work on standards for linking different model components, such as the OpenMI initiative (see Section 3.11) or the Python-based catchment modelling framework (CMF) (Kraft *et al.*, 2011). Some sources of software for linking model components are given in Appendix A.

12.5 Guidelines for Good Practice

One way forward in this situation is to agree upon assumptions by consensus of the parties involved, both those setting up the model and those who will use the model results. The advantage in such an approach comes from the use of a simple, transparent decision process as a communication tool with users and stakeholders. If decisions about different sources of uncertainty have to be agreed upon (or at least be open to scrutiny and discussion) then a greater understanding will develop on both sides about the uncertainties essential to making a particular decision. The resulting assumptions might well be quite wrong but this might only become apparent in hindsight when reviewing the process. Because of the nature of epistemic uncertainties, some sources might also be left out of the analysis but, again, this might only be evident in hindsight. The essence of such a consensus would be not to knowingly underestimate the potential uncertainties in making a decision.

Clearly, however, we can use experience to reach consensus, experience that might be encapsulated in sets of rules or guidelines for good practice (Beven and Alcock, 2011). Such guidelines can set out the decisions needed in considering sources of uncertainty for different types of application and provide advice on how they have been handled previously. Those decisions can provide a useful structure for interaction with stakeholders and users, serving to structure the translationary discourse for communication of uncertainty advocated by Faulkner *et al.* (2007).

Agreeing on assumptions about different sources of uncertainty is a heuristic approach to allow for uncertainty in model predictions. Any resulting assessment of uncertainties in model predictions that might be used in decision making will necessarily be approximate since we cannot be sure that all sources of uncertainty have been considered, nor if those that have been considered are properly represented. In fact, just like the model structures themselves, we will be pretty certain that we do not know how to properly represent different types of uncertainty. However, the very process of defining and debating the assumptions within some guidelines for good practice produces an agreed-upon working tool. As a heuristic process, it is implicit that the assumptions should be evaluated and refined in the future as more information about system responses becomes available. This is all part of the learning process.

Applying the guidelines will produce a range, possibly a wide range, of potential outcomes (or else, where the model predictions can be evaluated, possibly a conclusion that all the models tried can be rejected and decisions have to be made in some other way). Consideration of these outcomes in decision making should reveal the range of conditions under which a potential future decision might not satisfy the decision criteria. This is already a more robust heuristic than relying on some "best estimate" prediction. Ideally, a decision would be taken that satisfies the decision criteria over all potential outcomes, at reasonable cost, without compromising future decisions. By considering the uncertainty in the predictions more explicitly, the possibility for failure of a decision, conditional on the potential outcomes, can be assessed directly and judgements made as to whether the resulting risk is acceptable or not (see also Beven, 2011). Such a judgement is likely to be highly dependent on the context, particularly where extremes in the potential outcomes might involve catastrophic failures.

Traditionally, of course, engineers and others have had agreed-upon heuristics for dealing with uncertainties (factors of safety, freeboard, etc.) that err in the direction of more robust design. While involving significant subjectivity in the choice of appropriate values, that has not stopped them from being incorporated into standards and codes of practice (and, without doubt, preventing many engineering failures). The types of guideline for good practice argued for here represents a formal extension of this approach.

Does it matter to robustness in decision making that the underlying model structure, or the assumptions about the relevant sources of uncertainty might be quite wrong? This would suggest that, for whatever reason, we have not (yet) been able to detect a Type I error in choosing a model representation. So we would not therefore have a good reason to know that the model is wrong - until some information came along to question that conclusion (although in some cases, such as climate change projections, we are already very well aware of deficiencies in reproducing past observations). This might involve the collection of more observations that reveal the deficiencies of the model; it might be that an evaluation of the predictions of potential future outcomes does not seem to produce sensible results; it could be that specific experiments are carried out with a view to testing a model as hypothesis about how a particular part of the system functions. In either case, a continuing review of the heuristic assumptions on which the analysis is based will be justified as part of an adaptive management strategy. If neither case is evident, then we have no evidence to question the assumptions.

12.6 Models of Everywhere and Stakeholder Involvement

Once models of everywhere are available however it will very soon become apparent that modelling is a learning process about places and all their idiosyncracies (Beven, 2007). This is because there will be a continuing local feedback process about how well a model is performing, particularly if visualisations of model predictions are made available to stakeholders who have local knowledge. The stakeholders might be local authority or agency employees, they might be farmers, they might be environmentalists. They might be interested in flood levels, nutrient levels, or patterns of erosion and sediment deposition. They will all be interested in checking whether what is being predicted for their place is consistent with their local knowledge and experience. If it is not, they will say so – requiring some action by those who set up and run the model (as in the example of the Danish National Water Resources Model noted in Section 12.5).

An interesting strategy within such a framework is then to involve local stakeholders in setting up the model. The type of decision framework embodied in the guidelines for good practice is a way of providing a structure for thinking about local issues relevant to the problem under study. A practical example of this approach has been demonstrated in a project involving a variety of local stakeholders contributing to improved flood risk management for the town of Pickering in Yorkshire, UK (Lane *et al.*, 2011). The stakeholders were directly involved, through the local Ryedale Flood Research Group, in formulating a model of the runoff generation, flood storage areas and mitigation strategies in the catchment. The result was a learning process for both the scientists and the stakeholders involved (see the quotation from Stuart Lane at the head of this chapter). A similar approach has been taken in addressing water quality problems in Sweden (see Olsson and Berg, 2005; Olsson and Andersson, 2007).

Some people have gone further and suggested that many environmental problems are so complex, multi-faceted and uncertain that decisions should only made by such a process of discussion amongst the relevant stakeholders. This view was stimulated by the book by Funtowicz and Ravetz (1990) which introduced the concept of "post-normal science", although the idea that some problems might be too difficult to address using normal scientific methods goes back much further (e.g. Weinberg's (1972) concept of "trans-scientific problems"). More recent expressions of similar concepts include the "tangled thickets" of Wimsatt (2007). Examples of the use of such an approach in water management are provided, for example, by Pahl-Wostl (2002, 2007). It is clear that the question of how far model predictions can

provide useful information and evidence in complex environmental systems is one that still needs to clarified (e.g. Beven, 2002b; Harris and Heathwaite, 2005; Harris, 2007).

It is difficult to be sure about predictions of water flows in catchments, and water management problems involve more than just water flows. In satisfying the requirements of the EU Water Framework Directive, chemical and ecological processes also need to be considered. These additional dimensions have their own (even greater) uncertainties, as do the economic costs and impacts of management decisions. These are indeed complex, multi-faceted problems but I believe that modelling still has an important role to play in structuring the nature of a problem and in assessing the effects of a local change on the wider catchment environment. This again is where the learning process within the models of everywhere framework will serve to bring together scientists and stakeholders in solving management issues.

12.7 Models of Everywhere and Information

A learning process requires information. Post-normal science allows that the relevant information might be purely qualitative, based on the practical experience of both scientists and stakeholders. However, for the scientist, there is also an interest in what data might be of value in hypothesis testing and constraining the uncertainty in predictions. It was noted in Section 7.18 that there is little guidance in the literature about the value of different types of data and normally, of course, we make use of whatever data we can get hold of (while remembering to be wary of disinformation resulting from epistemic errors, see Section 7.17).

As noted in the Preface, the last paragraph of the first edition of this book read:

> The future in rainfall–runoff modelling is therefore one of uncertainty: but this then implies a further question as to how best to constrain that uncertainty. The obvious answer is by conditioning on data, making special measurements where time, money and the importance of a particular application allow. It is entirely appropriate that this introduction to available rainfall–runoff modelling techniques should end with this focus on the value of field data.

Ten years on, we know a little bit more about the interaction between model uncertainty and measurements. It has been the subject of a number of different workshops and sessions in major conferences. I have suggested that one of the most productive frameworks for such an interaction might be in testing quantitative hydrological models as hypotheses about how a hydrological system is functioning. Field data, however, are not often collected in experimental catchments with this aim in mind, but rather with the aim of extending our perceptual understanding about how catchments work. The quantitative modelling often comes later, as an approximation.

The models of everywhere concept can potentially change this way of doing things, since the effect of collecting additional data and information should be cumulative in the learning process. It is also possible to be pro-active about gathering additional information in specific places. We would want that process to be cost effective, which implies learning more about the value of different types of data collection. However, the availability of a model of everywhere allows an experimental design to be tested and costed. This is a form of what in statistical experimental design is sometimes called *pre-posterior prior analysis*. It involves using the current model of the system as a prior estimator of the different observations that might be collected. The benefit of actually collecting those values in constraining uncertainty in the model predictions can then be assessed, allowing for the potential observation uncertainties. The cost of different observation strategies can also be assessed, and the cost–benefit evaluated prior to actually making the measurements. A similar process could also be used in choosing critical observations that might best differentiate between models that differ in their predictions. Such a critical measurement might reveal

that none of the models give a good prediction but, as in any testing of a theory or model in science, this is actually the most valuable and interesting result in learning how to represent the processes.

Again, we are ending with a focus on the value of observation to modelling rainfall–runoff processes but in a way that demonstrates that there has been significant progress in the last 10 years in understanding some of the issues involved. The models of everywhere concept would seem to be a useful and positive way ahead. It should change hydrological modelling practice in the future.

12.8 Some Final Questions

This is likely to be the last book I publish before retiring as a hydrological modeller. When I started my research career, I had thought to understand the evolution of landscape (having been heavily influenced by Mike Kirkby as an undergraduate student). But to do that I reasoned that it would be really important to get the water flows right. So my PhD thesis was concerned with modelling water flows on hillslopes. The physically based finite element model I developed was not a great success (the story is told in Beven, 2001) but the problems I encountered were the basis for the rest of a research career. It is clear from the discussions in this book, however, that not all the problems have been solved and it seems pertinent to end with some questions that need to be addressed in the future.

In Chapter 9, I suggested a framework for developing a new generation of rainfall–runoff models and in this chapter I have tried to convey some really positive ways in which rainfall–runoff modelling might develop in the future. It is, however, difficult to predict the future of rainfall–runoff modelling. There are no predictive techniques for methodological advances. Hydrological science is currently in a period of gradual development typical of "normal" science in the sense of Thomas Kuhn. It was Kuhn who suggested that science often progresses by "paradigm shifts" interspersed by periods of normal science. Computer advances have certainly made it easier to prepare GIS databases of spatially distributed inputs, to present space and time variable model outputs and to calibrate and carry out sensitivity analyses of more complex models and parameter spaces, but there have been no major paradigm shifts in approach unless we count a shift from purely deterministic simulation to stochastic formulations (but of much the same model structures and without much in the way of additional supporting data).

If such a paradigm shift is to be achieved, then we need advances that respond to the following questions:

- Can we provide some guidance on the value of different types of data for hypothesis testing and constraining the uncertainty in rainfall–runoff models, particularly where there might be little in the way of discharge observations available?
- What novel measurement techniques can be envisaged and implemented in the future that will be effective in hypothesis testing and constraining the uncertainty in rainfall–runoff models?
- What new process representations can be supported by the observations in different types of catchment?
- Can we make remote sensing a central part of defining and evaluating the dynamic behaviour of rainfall–runoff modelling? The difficulty of sensing the subsurface and the need for an interpretative model in deriving hydrologically useful information from remote-sensing images will generally limit the utility of such data for rainfall–runoff modelling.
- Can we find some ways of dealing with epistemic uncertainties in model definition, evaluation and prediction that does not involve pretending that they are equivalent to statistical uncertainties? In particular, can we find ways of distinguishing between information and disinformation (or some way of eliminating the latter by improved observational methods) for hypothesis testing and constraining the uncertainty in rainfall–runoff models?

- Can local stakeholders be involved in defining modelling strategies for local conditions in a way that leads to more realistic rainfall–runoff models (a socially "bottom-up" approach within the models of everywhere framework)?

Finally, a last word about the equifinality of models in representing rainfall–runoff processes. A major theme of this book has been the apparent equivalence in performance of different model structures or different sets of parameters within a model structure. This has been referred to as "equifinality" to suggest that it is a generic problem, not simply a difficulty of finding the "right" model. But equifinality smacks of relativism, something that is not very acceptable to many scientists, including hydrological scientists. There should, after all, be an answer to aim for: a single, proper description of the catchment system that would be generally applicable and recognised as acceptable for predictive purposes. Perhaps such a description will develop in the future but even if such a model were available, applying it would not necessarily avoid the equifinality problem unless techniques for the direct measurement of effective parameters were evolved at the same time. The unique nature of every catchment system will always require the estimation of parameter values to reflect that uniqueness. Without direct measurement some form of conditioning on observed responses would be required and, with what would undoubtedly be a complex model, some equifinality of different parameter sets will undoubtedly ensue.

Personally, however, I see this as a positive concept for the very reason that it leads naturally to a hypothesis testing framework and to a focus on observational data in testing hypotheses. If multiple models can be found that are consistent with available observations, allowing for the uncertainty in those observations, then they can be treated as hypotheses to be tested. That is a methodology that we can call scientific, even if we have to develop new testing methods to deal with epistemic uncertainties. The nature of catchment systems and hydrological data is such that there might still be a plurality of models that are consistent with the data . . . but that is the nature of the prediction problem. Looking into the future, the limitations of observation techniques for subsurface flow pathways and processes, in particular, are likely to constrain how well we can predict rainfall–runoff processes. That is what makes rainfall–runoff modelling still challenging and will do so for some considerable time to come. Good luck!

Appendix A
Web Resources for Software and Data

This appendix gives information about sources of data and modelling software that can be downloaded from the web, organised by chapter. All sources were checked at the time of completing this edition but addresses and content might change over time. With a few exceptions where software is a commercial product, the sources can be freely downloaded. Some require registration as a user or are restricted for use in non-commercial activities in education and research. No particular recommendation is intended by mention here. I have not been a direct user of most of the data sets or packages mentioned.

Chapter 3 Data for Rainfall–Runoff Modelling

ASTER DEM The 30 m ASTER global topographic data (and other NASA satellite products) can be retrieved from the NASA Warehouse Inventory Search Tool (WIST) at https://wist.echo.nasa.gov/wist/api/imswelcome/.

Climate Change Projections UKCP09 projections of change in various climate parameters for the UK are derived from an ensemble of regional climate model projections and expressed in terms of 5, 50 and 95% probabilities of change relative to the 1961–1990 control period for the ensemble. The UKCP09 projections can be found at http://ukclimateprojections.defra.gov.uk/.

The EU Ensembles project produced climate projections at various resolutions, from the global scale to regional scales, using an ensemble of climate models. The data are available for research purposes at www.ensembles-eu.org.

In the USA, maps of climate projections produced by NCAR and statistical downscaling to finer resolution can be found at www.gisclimatechange.org.

CUAHSI The US-based Consortium of Universities for the Advancement of Hydrologic Science Inc. has been developing open source standards for exchange of data in its WaterML and CUAHSI Hydrologic Information System (CUAHSI-HIS) that are compatible with standards set by the Open Geospatial Consortium (OGC). CUAHSI can be found at www.cuahsi.org.

Rainfall–Runoff Modelling: The Primer, Second Edition. Keith Beven.

INTAMAP The Interoperability and Mapping project is developing a web-based framework for the real-time mapping of critical environmental variables using geostatistical methods. It can be used for exchanging data, carrying out statistical analysis, and visualising results. INTAMAP can be found at www.intamap.org.

DTM-Analysis Topographic analysis software to calculate the TOPMODEL topographic index using a multiple flow direction algorithm (see the list for Chapter 6).

FLUXNET acts as a repository for more than 500 sites around the globe making observations from micrometeorological tower sites that use eddy covariance methods to measure the exchanges of carbon dioxide (CO_2), water vapour and energy between terrestrial ecosystems and the atmosphere. Details of the sites and available data can be found at www.fluxnet.ornl.gov/fluxnet/index.cfm.

Global Runoff Data Centre (GRDC) The World Meteorological Global Runoff Database, hosted at the Bundesanstalt für Gewässerkunde (BfG) in Germany, contains river discharge data collected at daily or monthly intervals from more than 7700 stations in 156 countries (a total of around 310 000 station-years with an average record of 40 years). The GRDC provides discharge data and data products for non-commercial applications at http://www.bafg.de/GRDC/EN/Home/homepage_node.html.

HydroSHEDS DEM The Hydrological data and maps based on SHuttle Elevation Derivatives at multiple Scales (HydroSHEDS) is a mapping product maintained by the US Geological Survey that offers a suite of geo-referenced data sets (vector and raster) at various scales, including river networks, watershed boundaries, drainage directions and flow accumulations at regional and gobal scales. HydroSHEDS is based on high-resolution elevation data obtained during a space shuttle flight for NASA's Shuttle Radar Topography Mission (SRTM). Data are available at 1 arc second (about 30 m) resolution in the conterminous US and has been averaged to 3 arc seconds elswhere. Vectorised river network data are available at the same site. The data are known to contain some voids and artifacts. HydroSHEDS can be found at http://hydrosheds.cr.usgs.gov/index.php.

TARDEM Topographic analysis software available from www.crwr.utexas.edu/gis/gishydro99/uwrl/tardem.html.

UK hydrological data The UK National River Flow Archive, hosted by the Centre for Ecology and Hydrology allows daily river discharge data and catchment averaged rainfalls for some 200 stations to be downloaded and data from other stations to be ordered at www.ceh.ac.uk/data/nrfa/.

US hydrological data In the USA, river discharge data are available from the USGS (http://waterdata.usgs.gov/nwis) and climate data from NOAA (www.ncdc.noaa.gov/oa/climate/climatedata.html). The NOAA National Climate Data Centre also allows some global climate datasets to be accessed. Data for several hundred US and other basins are also available by way of the MOPEX project (see Section 10.3) at ftp://hydrology.nws.noaa.gov/pub/gcip/mopex/US_Data/. The Land Surface Hydrology Research Group at the University of Washington maintains databases of both meteorological and hydrological data at www.hydro.washington.edu/Lettenmaier/DataInfo.html.

Chapter 4 Data-Based Models

IHACRES The PC IHACRES software, along with its 94-page user guide, which includes tutorials and other documents, is available to download, free of charge from www.ceh.ac.uk/products/software/CEHSoftware-PC-IHACRES.htm.

RRMT Versions of the PDM, IHACRES, TFM and other conceptual model structures can also be found in the Imperial College Rainfall–Runoff Modelling Toolbox for MATLAB (see www3.imperial.ac.uk/ewre/research/software/toolkit).

TFM is a program for the analysis of rainfall-discharge catchment data based on Transfer Function Model concepts, similar to those used in the data-based mechanistic (DBM) modelling approach of Section 4.3. It is available at www.lec.lancs.ac.uk/research/catchment_and_aquatic_processes/software.php

Much of the DBM modelling functionality has also been incorporated in the CAPTAIN toolbox for MATLAB. CAPTAIN was developed by Peter Young and colleagues at Lancaster University and provides a variety of tools for non-stationary time series analysis, system identification, signal processing and forecasting using unobserved components models, time variable parameter models, state-dependent parameter models and multiple-input transfer function models. CAPTAIN is available at www.es.lancs.ac.uk/cres/captain.

Chapter 5 Distributed Models

Dynamic Flood Routing Models There are a number of commercial packages that solve the 1D or 2D depth averaged St Venant equations for flow routing in channels. A freely available 1D package is the Hydrologic Engineering Center's River Analysis System (HEC-RAS). It was the first of HEC's Next Generation (NexGen) software packages and can be used for steady flow and unsteady flow simulations; sediment transport/mobile bed computation; and water temperature modeling. It is available at www.hec.usace.army.mil/software/hec-ras/.

HYDRUS Early versions of HYDRUS 1D and 2D are available for free download at http://igwmc.mines.edu/software/igwmcsoft/. More recent versions are only available commercially.

KINEROS The KINEROS overland flow and sediment transport distributed model can be downloaded from www.tucson.ars.ag.gov/kineros/.

MODFLOW The USGS groundwater simulation package, MODFLOW, has a variety of options, including linked distributed rainfall–runoff predictions, particle tracking transport calculations and model calibration options. A number of companies have also packaged MODFLOW within a user-friendly interface offered as a commercial product. MODFLOW can be found at http://water.usgs.gov/nrp/gwsoftware/modflow.html.

Chapter 6 Distribution Function and Semi-Distributed Models

Green Kenue Green Kenue from Environment Canada embodies an HRU-based hydrological model within a GIS framework. It is available at www.nrc-cnrc.gc.ca/eng/ibp/chc/software/kenue/green-kenue.html.

SWAT The official Soil and Water Assessment Tool (SWAT) web site contains full details of various versions of the model (including versions where the model is embedded within a GIS framework), downloads, training courses, publications, etc. SWAT is available at http://swatmodel.tamu.edu/.

TOPKAPI A commercially available version is provided by Progea at www.progea.net/prodotti.php?p=TOPKAPI&c=Software&lin=inglese.

TOPMODEL Since 1974, many variants of TOPMODEL have been developed at Leeds, Lancaster and elsewhere but never a "definitive" version. This has been quite intentional. TOPMODEL is not intended to be a traditional model package but is more a collection of concepts that can be used where appropriate. It is up to the user to verify that the assumptions made are appropriate. The version of the program discussed in Box 6.1 is best suited to catchments with shallow soils and moderate topography which do not suffer from excessively long dry periods. Ideally, predicted contributing areas should be checked against what actually happens in the catchment (at least qualitatively).

There are a number of sources of TOPMODEL Software. The Lancaster University demonstration version available at www.lec.lancs.ac.uk/research/catchment_and_aquatic_processes/software.php was described in some detail in the first edition of this book. This program is intended as a demonstration version of TOPMODEL for Windows (32 bit only) and has been developed from versions used for teaching purposes. It includes an option for making Monte Carlo runs of the model that link to the Windows GLUE software (see software for Chapter 7). A Windows digital terrain analysis package, DTM-Analysis, is also available. The DTM-ANALYSIS program is used to derive a distribution of $ln(a/tanB)$ values from a regular raster grid of elevations for any catchment or subcatchment using the multiple direction flow algorithm of Quinn *et al.* (1995). Output from the program is a histogram of the distribution of the $ln(a/tanB)$ values and a map file of $ln(a/tanB)$ values that can be used for map output in the TOPMODEL program.

A version of the code for TOPMODEL is available as an R package through CRAN sites (see http://rwiki.sciviews.org/doku.php?id=packages:cran:topmodel). It is also part of a wider R package called R-Hydro that is currently under development by Wouter Buytaert and Dominik Reusser (see http://r-forge.r-project.org/projects/r-hydro/).

Chapter 7 Sensitivity Analysis, Model Calibration and Uncertainty Estimation

A more detailed guide to software for random number generation and uncertainty estimation methods can be found at www.uncertain-future.org.uk.

GLUE Lancaster University provides a Windows (32 bit only) demonstration GLUE package as a teaching aid. It provides tools for sensitivity analysis and uncertainty estimation using the results of Monte Carlo simulations. It is available from www.lec.lancs.ac.uk/research/catchment_and_aquatic_processes/software.php.

A MATLAB version of GLUE has been developed by Marco Ratto at the EU Joint Research Centre, Ispra, Italy. This allows for more parameters and more Monte Carlo runs than the Windows demonstration version and can be downloaded from http://eemc.jrc.ec.europa.eu/Software-GLUEWIN.htm.

DYNIA, with GLUE, is included in the Imperial College Monte Carlo toolbox for MATLAB that can be found at www3.imperial.ac.uk/ewre/research/software/toolkit.

PEST is a suite of software for parameter estimation, sensitivity analysis and uncertainty estimation, developed by John Doherty *et al.* It is applicable to a wide range of models including highly parameterised distributed models. It has a variety of regularisation techniques to reduce the parameter dimensions. The main inversion technique is based on a weighted least squares approach. It is available from www.pesthomepage.org/Home.php.

UncertML OGC-compatible standards for communicating uncertainty in variables and model results are being developed under the UncertML project. See www.uncertml.org.

Chapter 8 Changing Risk

FEWS The Delft Flood Early Warning System (Delft-FEWS) is available under license from http://public.deltares.nl/display/FEWSDOC/Home.

Chapter 12 Models of Everywhere

CMF The catchment modelling framework is a Python-based set of modelling tools for linking hydrological model components and data. It can be downloaded from www.uni-giessen.de/ gh1961/.

LIQUID The LIQUID modelling platform is a flexible system for linking distributed modelling components, available from www.hydrowide.com/liquid/current/.

OpenMI The OpenMI association (see www.openmi.org/reloaded/) is developing open source software designed to provide common interfaces between models and data on a time step by time step basis during simulations. The standards support distributed model components with different grid sizes.

RFortran RFortran provides code for linking R routines to Fortran code. It has been used for processing the results of the BATEA code (see Thyer *et al.*, 2011). It is available from www.rfortran.org.

Appendix B
Glossary of Terms

Actual Evapotranspiration The rate of evapotranspiration from a surface or vegetation canopy to the atmosphere under the prevailing meteorological conditions and water availability [Section 3.3; Box 3.1]

Advection Transport with the velocity in a flow pathway. Often used to refer to transport with the mean velocity of the flow within some control volume (see also Dispersion) [Sections 4.3, 11.2; Boxes 3.1, 11.1]

Aerodynamic Resistance Scaling parameter for sensible and latent heat fluxes in Penman–Monteith equation [Section 3.3; Box 3.1]

Aleatory Uncertainty Uncertainty due to random variations that can be described by a statistical model [Chapter 7, Section 12.3]

Antecedent Conditions The state of wetness of a catchment prior to an event or period of simulation [Section 1.4]

Aquiclude A layer of soil or rock that is impermeable to water [Section 5.1.1]

Atmospheric Demand The rate of potential evapotranspiration for given atmospheric conditions of temperature, humidity and wind speed without any limit due to the availability of water [Section 3.3]

Autocorrelated Errors A time series of model residuals that are not independent at each time step, i.e. that exhibit statistical correlation at one or more time steps apart (see also Heteroscedastic Errors) [Section 7.3; Box 7.1]

Automatic Optimisation Calibration of model parameters using a computer algorithm to maximise or minimise the value of an objective function [Sections 7.1, 7.4]

Baseflow That part of the discharge hydrograph that would continue after an event if there were no further rainfall. Sometimes taken to be equivalent to a subsurface flow contribution to stream discharge, in which case a response during a storm might be expected, but environmental tracer measurements suggest that this is not good usage of the term since subsurface flow may be the dominant contribution to the hydrograph in many storms [Section 2.2]

Baseflow Separation A procedure associated with use of the unit hydrograph to separate the hydrograph into "storm runoff" and "baseflow" components. Many different methods are available, mostly without any firm basis [Box 4.2]

Rainfall–Runoff Modelling: The Primer, Second Edition. Keith Beven.

Basis Functions Interpolation functions used in representing the variation of a predicted variable within each element of a finite element solution [Box 5.3]

Bayes Equation Equation for calculating a posterior probability given a prior probability and a likelihood function. Used in the GLUE methodology to calculate posterior model likelihood weights from subjective prior weights and a likelihood measure chosen for model evaluation [Section 7.7.4; Box 7.2]

Behavioural Simulation A simulation that gives an acceptable reproduction of any observations available for model evaluation. Simulations that are not acceptable are nonbehavioural [Sections 7.2.2, 7.7]

Big Leaf Assumption The representation of a vegetation canopy in predicting evapotranspiration as if it was a uniform surface [Box 3.1]

Black Box Model A model that relates an input to a predicted output by a mathematical function or functions without any attempt to describe the processes controlling the response of the system [Sections 1.1, 4.1]

Blind Validation Evaluation of a model using parameter values estimated before the modeller has seen any output data [Section 5.4]

Boundary Conditions Constraints and values of variables required to run a model for a particular flow domain and time period. May include input variables, such as rainfall and temperatures, or constraints, such as specifying a fixed head (Dirichlet boundary condition) or impermeable boundary (Neumann boundary condition) or specified flux rate (Cauchy boundary condition) [Sections 1.3, 5.1]

Calibration The process of adjusting parameter values of a model to obtain a better fit between observed and predicted variables. May be done manually or using an automatic calibration algorithm [Section 1.8, Chapter 7]

Canopy Resistance An effective resistance to the transport of water vapour from leaf stomata to the atmosphere [Section 3.3]

Celerity or Wave Speed. The speed with which a disturbance of pressure propagates through the flow domain. Important in the explanation of the large "old water" component of storm runoff in many catchments. May vary greatly with different processes and with catchment wetness [Sections 1.5, 5.5; Box 5.7]

Closure Problem The problem of estimating boundary fluxes for a discrete calculation unit in part of the flow domain when the appropriate relationships should be state and scale dependent [Section 9.2]

Cloud Computing The use of methods and data storage available via the Internet without the user having to worry about where computations are made or information is stored. These are becoming increasingly sophisticated and can provide elements that can be integrated with hydrological modelling (e.g. the use of GoogleMaps and GoogleView in visualisation of model results) [Section 3.11]

Complementarity Approach A method for the prediction of actual evapotranspiration based on the idea that the greater the actual evapotranspiration rate (and therefore the ambient humidity), the lower will be a measurement of evaporation from a free water surface or evaporation pan [Section 3.3]

Conceptual Model A hydrological model defined in the form of mathematical equations. A simplification of a perceptual model [Section 1.3]

Constitutive Relationships Relationships between the current state of volume in the flow domain and the boundary fluxes required to close mass, energy and momentum balance equations for that volume. Should be expected to be hysteretic (dependent on the history of wetting and drying) and scale dependent [Section 9.2.2]

Contributing area A term used in a variety of ways in hydrology. Most often it refers to the part of the catchment contributing either surface or subsurface storm runoff to the hydrograph [Section 1.4]

Data Assimilation The process of using observational data to update model predictions (see also Real-Time Forecasting and Updating) [Section 5.6]

Data-Based Mechanistic (DBM) Modelling An approach to modelling input–output systems using linear transfer functions, usually coupled with a nonlinear filter on the inputs. The structure of the model is identified from the available observations [Section 4.1, Chapter 6]

Degree-Day Method A method for predicting snowmelt as proportional to the difference between mean daily temperature and a threshold value [Section 3.4]

Depression Storage Water in excess of the infiltration capacity of the soil that is retained in surface hollows before significant downslope overland flow occurs. May later infiltrate into the soil after rainfall has ceased [Section 1.4]

Deterministic Model A model that, with a set of initial and boundary conditions, has only one possible outcome or prediction [Section 1.7]

Diffusivity The produce of unsaturated hydraulic conductivity and the gradient of the curve relating capillary potential to soil moisture content [Section 5.1.1; Box 5.1]

Disinformation Used to describe observational data that might be used in calibrating a model but which is inconsistent as a result of measurement or other errors and therefore might not be informative in identifying a good model of the system [Section 1.8]

Dispersion Variation in the transport of water or other quantity as a result of variation in flow velocity within some control volume. This may be a result of variations in velocity within a particular pathway or of variations across multiple pathways as a result of heterogeneity in the characteristics of different pathways. The latter is often referred to as macrodispersion (see also Advection) [Sections 4.3, 11.2; Box 11.1]

Distributed Model A model that predicts values of state variables varying in space (and normally time) [Section 1.7]

Double Mass Curve A plot of the cumulative volumes associated with two measurement stations (either rainfalls or discharges) [Section 3.2]

Dynamic contributing area The area generating surface runoff that will tend to expand during a storm [Section 1.4]

Eddy Correlation Method A technique for measuring actual evapotranspiration and sensible heat fluxes by integrating the rapid fluctuations in humidity and temperature associated with turbulent eddies in the lower boundary layer [Section 3.3.3]

Effective Rainfall A part of the storm rainfall inputs to a catchment that is equivalent in volume to the "storm runoff" part of the hydrograph (but note that the storm runoff may not be all water from that event) [Section 1.3, 2.2]

Effective Storage Capacity The difference between the current moisture in the soil immediately above the water table and saturation [Section 1.5]

Ensemble Predictions A set of simulations based on running multiple models with different parameter sets or initial conditions. Used in weather forecasting, climate change predictions and uncertainty estimation [Sections 7.10, 8.2, 8.9]

Environmental Tracer A tracer for water flow through a hillslope or catchment system that occurs naturally in the environment. Oxygen and hydrogen isotopes are the most commonly used environmental tracers because they are part of the water molecule, but other tracers are also used [Sections 11.2, 11.3, 11.8]

Ephemeral Streams Streams that are frequently dry between storm periods [Section 1.4]

Epistemic Uncertainty Uncertainty that is due to lack of knowledge in describing processes or defining input and boundary conditions or to changes in the system over time [Chapter 7, Section 12.3]

Equifinality The concept that there may be many models of a catchment that are acceptably consistent with the observations available [Section 1.8, 7.10]

ESMA Model The "explicit soil moisture accounting" model (sometimes called a conceptual model). A hydrological model made up of a series of storage elements, with simple equations to control transfers

between the elements. Mostly applied at the lumped scale, but some models use ESMA components to represent distributed hydrological response units [Section 2.4]

Evaluation see Validation

Explicit Solution The independent calculation of predicted variables at a time step given values of the variables at the previous time step (see also Implicit solution) [Section 5.1.1; Box 5.3]

Field Capacity An imprecisely defined variable normally expressed as the water content of the soil when it is allowed to drain from saturation until rapid drainage has ceased (see Soil Moisture Deficit) [Box 6.1]

Finite Difference The approximate representation of a time or space differential in terms of variables separated by discrete increments in time or space [Box 5.3]

Finite Element Method The approximate representation of time or space differentials in terms of integrals of simple interpolation functions involving variables defined at nodes of an irregular discretisation of the flow domain into elements [Box 5.3]

Fuzzy logic A system of logical rules involving variables associated with a continuous fuzzy measure (normally in the range 0 to 1) rather than the binary measure (right/wrong, 0 or 1) of traditional logic. Rules are available for operations such as addition and multiplication of fuzzy measures and for variables grouped in fuzzy sets. Such rules can be used to reflect imperfect knowledge of how a variable will respond in different circumstances [Sections 1.7, 5.2.2]

Gain A multiplier applied to a transfer function to scale inputs to outputs in a linear systems analysis. May be made adaptive in real-time forecasting [Box 8.1]

Geomorphological Unit Hydrograph A unit hydrograph derived from the structural relationships of the catchment geomorphology, in particular the branching structure of the channel network [Sections 2.3, 4.7.2]

Global optimum A set of parameter values that gives the best fit possible to a set of observations [Sections 1.8, 7.2]

Head An expression of pressure as energy per unit weight commonly used in hydrology and hydraulics since it has units of length [Section 5.1.1]

Heteroscedastic Errors A time series of model residuals that exhibit a changing variance over a simulation period (see also Autocorrelated Errors) [Section 7.3; Box 7.1]

Hortonian model Runoff production by an infiltration excess mechanism. Named after Robert E. Horton [Section 1.4, Box 1.1, see also Partial Area Model]

Hydrological Response Unit A parcel of the land surface defined in terms of its soil, vegetation and topographic characteristics [Sections 1.7, 2.6, 3.8, 6.6]

Hyperparameter Used to describe the parameters of a structural model of the rainfall–runoff model residuals in statistical calibration of the rainfall–runoff model parameters (see also Model Inadequacy Function). Hyperparameters are estimated as part of the calibration process [Box 7.1]

Hysteretic function A function between two hydrological variables that is different under wetting and drying conditions. Might refer to storage–discharge at the catchment scale or soil moisture-capillary potential at the point scale. Extreme wetting and drying values are called the primary wetting and drying curves; when there is a sequence of wetting and drying periods, the function might also follow intermediate secondary or scanning curves

Implicit Solution The simultaneous solution of predicted variables at a time step given values of the variables at the previous time step, usually by an iterative method (see also Explicit solution) [Section 5.1.1; Box 5.3]

Importance Sampling Methods for sampling the model space with a view to obtaining samples with a density that depends on the local value of the sampled model output. Where the output is a likelihood in model calibration, the sample density should reflect the local magnitude of the likelihood [Section 7.10.2]

Incommensurate Used here to refer to variables or parameters with the same name that refer to different quantities because of a change in scale [Sections 1.8, 7.7]

Infiltration Capacity The limiting rate at which a soil surface can absorb rainfall – depends on factors such as antecedent moisture content, volume of infiltrated water, presence of macropores and surface crusting [Section 1.4; Box 5.2]

Infiltration Excess Runoff generated by rainfall intensities exceeding the infiltration capacity of the soil surface. May be used either at the local point scale within a catchment (when the surface runoff may infiltrate further downslope) or at the catchment scale to represent that part of the storm hydrograph generated by an infiltration excess mechanism [Section 1.4]

Initial Conditions Values of storage or pressure variables required to initialise a model at the start of a simulation period [Section 5.1]

Instrumentalism Approach to science based on the pragmatic concept that empirical adequacy is sufficient to justify a model or theory. In rainfall–runoff modelling, this is equivalent to saying that a model that gives a good fit to observations in calibration should be useful in prediction, regardless of its theoretical underpinning [Section 6.7]

Interception Rainfall that is stored on a vegetation canopy and later evaporated back to the atmosphere [Section 3.3.2; Box 3.2]

Inverse Method Model calibration by adjusting parameter values to reduce the differences between observations and predicted variables [Section 5.1.1]

Lagrangian velocity The average velocity taken along a flow pathway by a molecule of water or solute or particle of sediment subject to time and space variability in local flow velocities [Section 9.6]

Land Surface Parameterisation Hydrological model used to calculate water and energy fluxes from the land surface to the atmosphere in atmospheric circulation models [Sections 2.4, 10.4]

Lead Time The required time for a forecast ahead of the current time in real-time flood forecasting [Section 8.1]

Learning set A set of observed data used in the calibration of a neural net model [Section 4.6]

Likelihood Measure A quantitative measure of the acceptability of a particular model or parameter set in reproducing the hydrological response being modelled [Section 7.7; Boxes 7.1, 7.2]

Linearity A model is linear if the outputs are in direct proportion to the inputs [Section 2.2; Boxes 2.1, 4.1]

Linear Store A model component in which the output is directly proportional to the current storage value. The basic building block of the general linear transfer function model and the Nash cascade [Section 2.3; Box 4.1]

Local Optimum A local peak in the parameter response surface where a set of parameter values gives a better fit to the observations than all parameter sets around it, but not as good a fit as the global optimum [Section 7.2]

Lumped Model A model that treats the whole of a catchment as a single accounting unit and predicts only values of variables averaged over the catchment area [Sections 1.5, 1.7]

Macropores Large pores in the soil that may form important pathways for infiltration and redistribution of water bypassing the soil matrix as a preferential flow. May result from soil cracking and ped formation, root channels and animal burrows [Section 1.4]

Model Inadequacy Function A mathematical function (usually simple in form) designed to correct for non-random structure in a series of model residuals that would otherwise lead to biased inference in model calibration. The simplest correction might be a constant bias or a linear trend [Section 7.3]

Monte Carlo simulation Simulation involving multiple runs of a model using different randomly chosen sets of parameter values [Sections 7.5, 7.6, 7.7; Box 7.3]

Network Width Function A histogram of the number of reaches in the channel network at a given distance from the catchment outlet. Can be used as the basis for both linear and nonlinear flow routing algorithms [Section 2.3, 4.7.1]

Nomogram An empirical method for estimating runoff by using a series of graphs [Section 2.1]

Nonbehavioural Simulation see Behavioural Simulation

Nonlinear A model is nonlinear if the outputs are not in direct proportion to the inputs but may vary with intensity or volume of the inputs or with antecedent conditions [Box 2.1]

Nonparametric Method A method of estimating distributions without making any assumptions about the mathematical form of the distribution [Section 7.2.2]

Nonstationarity A model in which the parameters are expected to change over time [Box 2.1]

Numerical Dispersion The effect of numerical approximation in the solution of differential equations in smoothing the pattern of a predicted variable (see Finite Difference, Finite Element Method) [Section 5.5.5; Box 11.1]

Objective Function A measure of how well a simulation fits the available observations [Sections 1.8, 7.3; Box 7.1]

Optimisation The process of finding a parameter set that gives the best fit of a model to the observations available. May be done manually or using an automatic calibration algorithm [Sections 1.8, 7.4]

Overland Flow Downslope flow of water on the surface of the soil in excess of the infiltration capacity and depression storage capacity of the surface [Section 1.4]

Parameter A constant that must before defined before running a model simulation [Sections 1.6, 1.8]

Parameter Space A space defined by the ranges of feasible model parameters, with one dimension for each parameter [Sections 1.8, 7.2]

Parsimony The concept, sometimes known as Occam's razor, that a model should be no more complex than necessary to predict the observations sufficiently accurately to be useful [Box 4.1]

Partial Area Model Runoff production (by an infiltration excess mechanism) only over part of the hillslopes (the partial area) in a catchment [Section 1.4]

Particle Tracking Models The representation of water, solute mass, or sediments as a large number of particles that are tracked through a hydrological system according to some rules used to represent the flow processes [Sections 9.6, 11.12]

Pedotransfer Function A function for predicting soil hydraulic parameters from knowledge of soil texture and other more easily measured variables [Sections 3.8, 5.1.1; Box 5.5]

Perceptual Model A qualitative description of the processes thought to be controlling the hydrological response of an area [Sections 1.3, 1.4]

Phreatophytes Plants whose roots extract water from below the water table [Section 1.4]

Pooling Group A group of gauged catchments defined as being similar to an ungauged catchment and used to estimate its hydrological response [Section 10.7]

Post-audit Analysis An evaluation of predictions made of the future behaviour of a system, once the period of predictions has actually occurred [Section 8.8]

Potential Evapotranspiration Rate of evapotranspiration from a surface or vegetation canopy with no limitation due to water availability (see also atmospheric demand) [Section 3.3; Box 3.1]

Preferential flow Local concentrations of flow in the soil that may be due to the effects of macropores, local variations in hydraulic properties or fingering of a wetting front moving into the soil profile. May lead to rapid and deep infiltration of water bypassing much of the soil matrix (see also macropores) [Section 1.3]

Pre-posterior prior analysis The concept of using predicted variables from a model (with uncertainty) to help design a programme of observations that would lead to the most effective constraint on posterior predictive uncertainty [Section 12.7]

Procedural Model A model represented as a computer program. May be an exact or an approximate solution of the equations defining the conceptual model of a system [Section 1.3]

Raster Digital Elevation Model A gridded set of elevation values at regular spacing [Section 3.7]

Rational Method An empirical method, first used in the 19th century, for predicting peak discharges based on catchment area and a measure of average rainfall [Section 2.1]

Realisation effect The effect that short periods of observations or stochastic model outputs might lead to different inferences when different periods or different realisations from the stochastic model are used [Section 8.7]

Real-Time Forecasting and Updating Operational forecasting of flows during an event, usually to predict the possibility of flooding, often with adaptive updating of model parameters based on the errors between observed and predicted variables (see also Lead Time) [Sections 4.8, 8.4; Box 8.1]

Regionalisation The process of estimating the response of an ungauged catchment using information transferred from gauged catchments [Chapter 10]

Relaxation time The time scale over which the impact of a particular event can be distinguished [Section 9.7]

Residence Time Distributions The distribution of times that particles have been within a hillslope, channel network or catchment system. There are at least three different uses: to indicate the distribution of times that it takes an increment of input to get to the outlet; to indicate the distribution of times it has taken for the water in an increment of discharge to arrive at the outlet; and to indicate the distribution of times that water has been stored in the system at any given moment. These are all different and should be differentiated (see also Travel Time Distributions) [Section 11.8–11.10]

Response Surface The surface defined by the values of an objective function (or other model output variable) as it changes with changes in parameter values. May be thought of conceptually as a surface with "peaks" and "troughs" in the multidimensional space defined by the parameter dimensions, where the peaks represent good fits to the observations and the troughs represent poor fits to the observations (see also Parameter Space) [Sections 1.8, 7.2]

Riparian Area The part of the catchment that is immediately adjacent to the stream and is often the most important source of runoff, both surface and subsurface [Section 1.4]

Runoff see Overland Flow, Storm Runoff, Surface Runoff, Subsurface Stormflow.

Runoff Coefficient The proportion of the rainfall volume in a storm that appears in the storm hydrograph. The value will depend on how the storm runoff component of the hydrograph is defined [Section 2.2]

Runoff Generation The process mechanisms that contribute water to the hydrograph in a channel for a rainfall or snowmelt event. May include both surface and subsurface processes. May also include contributions from both the event inputs and displacement of pre-event stored water.

Runoff Routing The translation of surface runoff and subsurface stormflow to a point of interest, usually the catchment outlet, taking account of surface, subsurface and channel flow velocities [Sections 1.6, 4.4, 5.5, 5.6, 6.1]

Saturation Excess Runoff generated by rainfall onto saturated soil, even though the rainfall intensity may not exceed the normal infiltration capacity of the soil. May be used either at the local point scale within a catchment (when the surface runoff may infiltrate further downslope) or at the catchment scale to represent that part of the storm hydrograph generated by a saturation excess mechanism [Section 1.4]

Similar Media A method of scaling the soil moisture characteristics of heterogeneous soils by making assumptions about the structure of the media (e.g. that the geometry of the soil matrix is identical, different only in length scale between different samples) [Box 5.4]

Slope-Area Method Method of measuring a peak discharge after a flood event using one of the uniform flow equations by estimating the cross-sectional area, water surface slope and roughness coefficient at a site [Section 3.2]

Snow Course Transect where regular measurements of snow depth and density are made [Section 3.1]

Soil Moisture Deficit (SMD) A state variable used in many hydrological models as an expression of soil water storage. SMD is zero when the soil is at field capacity and gets larger as the soil dries out. It is usually expressed in units of depth of water [Sections 1.4, 3.1]

Soil Moisture Characteristics Curves or functions relating soil moisture content to unsaturated hydraulic conductivity and capillary potential [Section 5.1.1; Box 5.2]

Specific Moisture Capacity The gradient of the curve relating unsaturated soil moisture to capillary potential [Section 5.1.1; Box 5.1]

State Variable A variable in a model that is part of the solution of the model equations and varies with time during a simulation but which is not a flux or exchange of mass. May include storage and pressure variables, depending on the definition of the model [Section 5.8]

Stemflow Rainfall that reaches the ground via the stems of plants [Section 1.4; Box 3.2]

Stochastic A model is stochastic if, for a given set of intial and boundary conditions, it may have a range of possible outcomes, often with each outcome associated with an estimated probability [Section 1.7]

Storm Profile Time series of rainfall intensities during a storm [Section 3.1]

Storm Runoff There are many conflicting definitions of storm runoff. Here it is that part of the stream hydrograph due to a rainfall event over and above the discharge that would have occurred without a rainfall event and may involve both surface and subsurface flow processes, and both event and "old" water contributions [Sections 1.4, 1.5, 1.6]

Streamline A line following the direction of flow of water in a stream, aquifer or other flow domain [Section 3.7]

Stream Tube The cross-sectional area bounded by surfaces following streamlines in a stream, aquifer or other flow domain [Section 3.7; Box 11.1]

Sublimation Direct loss of water from a snowpack to the atmosphere by evaporation [Section 3.1]

Subsurface Stormflow The contribution to the stream hydrograph generated by purely subsurface flow processes [Section 1.4]

Superposition The addition of linear model responses to construct a total response [Section 2.2; Box 2.1]

Surface Runoff The contribution to the stream hydrograph from overland flow [Section 1.4]

Tessellation The discretisation of space into a spatial grid or network of elements [Section 3.7]

Throughfall Rainfall that reaches the ground either directly or indirectly by dripping from the leaves of plants [Section 1.4; Box 3.2]

Throughflow Often used for rapid near-surface downslope subsurface flow in the soil profile [Section 1.4]

Time Compression Assumption Treating the volume of infiltrated water during an event as if it had infiltrated at the infiltration capacity of the soil in order to calculate an equivalent time to ponding [Box 5.2]

Time to Ponding The time taken during a rainfall event to bring the soil surface just to saturation [Box 5.2]

Top-down Modelling Using the available observational data to suggest appropriate model structures, usually applied at the catchment scale [Chapter 4]

Transfer Function A representation of the output from a system due to a unit input [Section 2.2, Chapter 4]

Transit Time Distributions, Travel Time Distributions The distribution of times taken by water to move through a hillslope, channel network or catchment. Used to indicate both the transit or travel times for an increment of input to reach an outlet and the transit or travel times for an increment of discharge to reach the outlet. These are different and should be properly differentiated (see also Residence Time Distributions) [Section 11.8–11.10]

Triangular Irregular Network A way of representing topography by a network of triangles between points of known elevation [Section 3.7]

Uniform Flow An open channel flow or overland flow in which the slope of the water surface is equal to the bed slope such that energy loss due to frictional shear stress is exactly matched by the potential energy gain as the water moves downhill [Section 5.2; Box 5.6]

Unit Hydrograph The "storm runoff" response from a unit of "effective rainfall" [Sections 2.2, 2.3, 4.3]

Validation A process of evaluation of models to confirm that they are acceptable representations of a system. Philosphers of science have some problems with the concept of validation (see Section 1.8) and it may be better to use "evaluation" or "confirmation" rather than "validation" (which, from its Latin root, implies a degree of truth in the model) [Sections 1.8, 5.3, 10.5]

Vector Digital Elevation Model A set of irregularly spaced elevation points defining contours of equal elevation [Section 3.7]

Wave speed see Celerity.

References

Abbott M. B., Bathurst J. C., Cunge J. A., O'Connell P. E. and Rasmussen J. An introduction to the European Hydrological System – Système Hydrologique Européen (SHE): 1. History and philosophy of a physically-based, distributed modelling system. *Journal of Hydrology*, 87:45–59, 1986a.

Abbott M. B., Bathurst J. C., Cunge J. A., O'Connell P. E. and Rasmussen J. An introduction to the European Hydrological System – Système Hydrologique Européen, (SHE): 2. Structure of a physically-based, distributed modelling system. *Journal of Hydrology*, 87:61–77, 1986b.

Abbott M. B. and Refsgaard J. C. *Distributed Hydrological Modelling*. Kluwer Academic, Dordrecht, 1996.

Abdulkabir M. O., Beven K. J. and Reeves A. D. Flow separation in undisturbed soil using multiple anionic tracers (3) unsteady core scale infiltration experiments. *Hydrological Processes*, 10:1467–1482, 1996.

Abdulla F. A. and Lettenmaier D. P. Application of regional parameter estimation schemes to simulate the water balance of a large continental river. *Journal of Hydrology*, 197:258–285, 1997a.

Abdulla F. A. and Lettenmaier D. P. Development of regional parameter estimation equations for a macroscale hydrologic model. *Journal of Hydrology*, 197:230–257, 1997b.

Abdulla F. A., Lettenmaier D. P., Wood E. F. and Smith J. Application of a macroscale hydrologic model to the water balance of the Arkansas Red River basin. *Journal of Geophysical Research*, 101:7449–7460, 1996.

Abebe A. J., Solomatine D. P. and Venneker R. G. W. Application of adaptive fuzzy rule-based models for reconstruction of missing precipitation events. *Journal of Hydrology*, 45:425–436, 2004.

Abeliuk R. and Wheater H. S. Parameter identification of solute transport models for unsaturated soils. *Journal of Hydrology*, 117:199–224, 1990.

Abrahart R. J. and See L. Neural network modelling of non-linear hydrological relationship. *Hydrology and Earth System Sciences*, 11:1563–1597, 2007.

Acutis M. and Donatelli M. SOILPAR 2.00: Software to estimate soil hydrological parameters and functions. *European Journal of Agronomy*, 18:373–377, 2003.

Ahsan M. and O'Connor K. M. A simple non-linear rainfall–runoff model with a variable gain factor. *Journal of Hydrology*, 155:151–183, 1994.

Akay O., Fox G. A. and Šimůnek J. Numerical simulation of flow dynamics during macropore–subsurface drain interactions using HYDRUS. *Vadose Zone Journal*, 7:909–918, 2008.

Al-Wagdany A. S. and Rao A. R. Correlation of the velocity parameter of three geomorphological instantaneous unit hydrograph models. *Hydrological Processes*, 12:651–659, 1998.

Alley R. B., Martozke J., Nordhaus W. D., Overpeck J. T., Peteet D. M., Pielke Jr R. A., Pierrehumbert R. T., Rhines P. B., Stocker T. F., Talleyt L. D. and Wallace J. M. Abrupt climate change. *Science*, 299:2005–2010, DOI: 10.1126/science.1081056, 2003.

Alvisi S., Mascellani G., Franchini M. and Bárdossy A. Water level forecasting through fuzzy logic and neural network approaches. *Hydrology and Earth System Sciences*, 10:1–17, 2006.

Ambroise B., Beven K. J. and Freer J. Towards a generalisation of the TOPMODEL concepts: Topographic indices of hydrologic similarity. *Water Resources Research*, 32:2135–2145, 1996.

Rainfall–Runoff Modelling: The Primer, Second Edition. Keith Beven.

Ambroise B., Freer J. and Beven K. J. Application of a generalised TOPMODEL to the small Ringelbach catchment, Vosges, France. *Water Resources Research*, 32:2147–2159, 1996b.

Amerman C. R. The use of unit-source watershed data for runoff prediction. *Water Resources Research*, 1:499–507, 1965.

Amin I. E. and Campana M. E. A general lumped parameter model for the interpretation of tracer data and transit time calculation in hydrologic systems. *Journal of Hydrology*, 179(1–4):1–21, 1996.

Amorocho J. and Brandstetter A. Determination of nonlinear functional response functions in rainfall–runoff processes. *Water Resources Research*, 7:1087–1101, 1971.

Anagnostou E. N., Anagnostou M. N., Krajewski W. F., Kruger A., Robinson B. A. and Miriovsky J. High-resolution rainfall estimation from X-band polarimetric radar measurements. *Journal of Hydrometeorology*, 5:110–128, 2004.

Anderson E. A. Development and testing of snowpack energy balance equations. *Water Resources Research*, 3:19–38, 1968.

Anderson M. G., editor. *Encyclopaedia of Hydrological Sciences*. John Wiley & Sons, Chichester, UK, 2005. 5 volumes.

Anderson M. G. and Burt T. P., editors. *Process Studies in Hillslope Hydrology*. John Wiley & Sons, 1990.

Anderson M. G. and Calver A. On the persistence of landscape features formed by a large flood. *Transactions of the Institute of British Geographers*, 2:243–254, 1977.

Andreadis K. M. and Lettenmaier D. P. Assimilating remotely sensed snow observations into a macroscale hydrology model. *Advances in Water Resources*, 29:872–886, 2006.

Andréassian V., Lerat J., Loumagne C., Mathevet T., Michel C., Oudin L. and Perrin C. What is really undermining hydrologic science today? *Hydrological Processes*, 21:2819–2822, 2007.

Ao T., Ishidaira H., Takeuchi K., Kiem A., Yoshitani J., Fukami K. and Magome J. Relating BTOPMC model parameters to physical features of MOPEX basins. *Journal of Hydrology*, 320:84–102, 2006.

Arabi M., Govindaraju R. S. and Hantush M. M. A probabilistic approach for analysis of uncertainty in evaluation of watershed management practices. *Journal of Hydrology*, 333:459–471, 2007.

Archer D. R. The use of flow variability analysis to assess the impact of land use change on the paired Plynlimon catchments, mid-Wales. *Journal of Hydrology*, 347:487–496, 2007.

Archfield S. A. and Vogel R. M. Map correlation method: Selection of a reference streamgage to estimate daily streamflow at ungaged catchments. *Water Resources Research*, 46:W10513, DOI:10.1029/2009WR008481, 2010.

Arnaud-Fassetta G., Ballais J.-L., Beghin E., Jorda M., Meffre J.-C., Provensal M., Roditis J.-C. and Suanez S. La crue de l'Ouvèze à Vaison-la-Romaine (22 septembre 1992): Ses effets morphodynamiques, sa place dans la fonctionnement d'un géosystème anthropisé. *Revue de Géomorphologie Dynamique*, 42:34–48, 1993.

Arnell N. W. A simple water balance model for the simulation of streamflow over a large geographic domain. *Journal of Hydrology*, 217:314–335, 1999.

Arnold J. G. and Allen P. M. Automated methods for estimating baseflow and groundwater recharge from streamflow records. *Journal of the American Water Resources Association*, 35:411–424, 1999.

Arnold J. G., Potter K. N., King K. W. and Allen P. M. Estimation of soil cracking and the effect on surface runoff in a Texas Blackland Prairie watershed. *Hydrological Processes*, 19:589–603, 2005.

Arnold J. G., Srinavasan R., Muttiah R. S. and Williams J. R. Large area hydrologic modeling and assessment: Part 1. Model development. *Journal of the American Water Resources Association*, 34:73–89, 1998.

Arnold J. G. and Williams J. R. SWRRB – a watershed scale model for soil and water resources management. In Singh V. P. , editor, *Computer Models of Watershed Hydrology*, pages 847–908. Water Resource Publications, Highlands Ranch, CO, 1995.

Aronica G., Bates P. D. and Horritt M. S. Assessing the uncertainty in distributed model predictions using observed binary pattern information within GLUE. *Hydrological Processes*, 16:2001–2016, 2002.

Aronica G., Hankin B. G. and Beven K. J. Uncertainty and equifinality in calibrating distributed roughness coefficients in a flood propagation model with limited data. *Advances in Water Resources*, 22:349–365, 1998.

ASCE. Task Committee on the application of ANN in hydrology. Artificial neural networks in hydrology: I Preliminary concepts. *ASCE Journal of Hydrologic Engineering*, 5:115–123, 2000a.

ASCE. Task Committee on the application of ANN in hydrology. Artificial neural networks in hydrology: II Hydrological applications. *ASCE Journal of Hydrologic Engineering*, 5:124–137, 2000b.

Bales J. and Betson R. P. The curve number as a hydrologic index. In Singh V. P., editor, *Rainfall–Runoff Relationship*, pages 371–386. Water Resource Publications, 1982.

Band L. E., Patterson P., Nemani R. and Running S. W. Forest ecosystem processes at the watershed scale: Incorporating hillslope hydrology. *Agricultural and Forest Meteorology*, 63:93–126, 1993.

Band L. E., Peterson D. L., Running S. W., Coughlan J. C., Lammers R., Dungan J. and Nemani R. Forest ecosystem processes at the watershed scale: Basis for distributed simulation. *Ecological Modelling*, 56:171–196, 1991.

Bárdossy A. *Encyclopaedia of Hydrological Sciences*, pages 2007–2014. John Wiley & Sons, Chichester, UK, 2005.

Bárdossy A. Calibration of hydrological model parameters for ungauged catchments. *Hydrology and Earth System Sciences*, 11:703–710, 2007.

Bárdossy A., Bronstert A. and Merz B. 1-dimensional, 2-dimensional and 3-dimensional modelling of water movement in the unsaturated soil matrix using a fuzzy approach. *Advances in Water Resources*, 18(4):237–251, 1995.

Bárdossy A. and Singh S. K. Robust estimation of hydrological model parameters. *Hydrology and Earth System Sciences*, 12:1273–1283, 2008.

Barling R. D., Moore I. D. and Grayson R. B. A quasi-dynamic wetness index for characterizing the spatial distribution of zones of surface saturation and soil water content. *Water Resources Research*, 30(4):1029–1044, 1994.

Bartholmes J. C., Thielen J., Ramos M. H. and Gentilini S. The European Flood Alert System EFAS – Part 2: Statistical skill assessment of probabilistic and deterministic operational forecasts. *Hydrology and Earth System Sciences*, 13:141–153, 2009.

Bastiaanssen W. G. M., Hoekman D. H. and Roebeling D. *A methodology for the assessment of surface resistance and soil water storage variability at mesoscale based on remote sensing measurements*. The International Association of Hydrological Sciences Special Publication 2, Wallingford, UK, 1994.

Bastiaanssen W. G. M., Menenti M., Feddes R. A. and Holtslag A. A. M. A remote sensing surface energy balance algorithm for land (SEBAL): 1. Formulation. *Journal of Hydrology*, 212-213:198–212, 1998.

Bates P. D., Anderson M. G., Baird L., Walling D. E. and Simm D. Modelling flood plain flows using a two-dimensional finite element model. *Earth Surface Processes and Landforms*, 17:575–588, 1992.

Bates P. D., Anderson M. G. and Hervouet J.-M. An initial comparison of 2-dimensional finite element codes for river flood simulation. *Proc. Inst. Civ. Eng., Water, Maritime, Energy*, 112:238–248, 1995.

Bates P. D. and De Roo A. P. J. A simple raster-based model for flood inundation simulation. *Journal of Hydrology*, 236(1–2):54–77, 2000.

Bates P. D., Horritt M. S., Aronica G. and Beven K. J. Bayesian updating of flood inundation likelihoods conditioned on flood extent data. *Hydrological Processes*, 18:3347–3370, 2004.

Bates P. D., Horritt M. S. and Fewtrell T. J. A simple inertial formulation of the shallow water equations for efficient two dimensional flood inundation modelling. *Journal of Hydrology*, 387:33–45, 2010.

Bathurst J. C. Physically-based distributed modelling of an upland catchment using the Système Hydrologique Européen. *Journal of Hydrology*, 87:79–102, 1986a.

Bathurst J. C. Sensitivity analysis of the Système Hydrologique Européen for an upland catchment. *Journal of Hydrology*, 87:103–123, 1986b.

Bathurst J. C. and Cooley K. R. Use of the SHE hydrological modelling system to investigate basin response to snowmelt at Reynolds Creek, Idaho. *Journal of Hydrology*, 175:181–211, 1996.

Bathurst J. C., Ewen J., Parkin G., O'Connell P. E. and Cooper J. D. Validation of catchment models for predicting land-use and climate change impacts: 3. Blind validation for internal and outlet responses. *Journal of Hydrology*, 287:74–94, 2004.

Bathurst J. C., Wicks J. M. and O'Connell P. E. The SHE/SHESED basin scale water flow and sediment transport modelling system. In Singh V. P., editor, *Computer Models of Watershed Hydrology*, pages 563–594. Water Resource Publications, Highlands Ranch, CO, 1995.

Bayes T. An essay towards solving a problem in the doctrine of chances. *Philosophical Transactions of the Royal Society London*, 53:370–418, 1763.

Bazemore D. E., Eshleman K. N. and Hollenbeck K. The role of soil water in stormflow generation in a forested headwater catchment: Synthesis of natural tracer and hydrometric evidence. *Journal of Hydrology*, 162:47–75, 1994.

Beasley D. B., Huggins L. F. and Monke E. J. ANSWERS: A model for watershed planning. *Trans. Am. Soc. Agr. Engrs.*, 23:938–944, 1980.

Beaumont M., Cornuet J.-M., Marin J.-M. and Robert C. Adaptive approximate Bayesian computation. *Biometrika*, 96:983–990, 2009.

Beaumont M. A., Zhang W. and Balding D. J. Approximate Bayesian computation in population genetics. *Genetics*, 162:2025–2035, 2002.

Beck M. B. Water quality modeling: A review of the analysis of uncertainty. *Water Resources Research*, 23:1393–1442, 1987.

Becker A. and Braun P. Disaggregation, aggregation and spatial scaling in hydrological modelling. *Journal of Hydrology*, 217:239–252, 1999.

Beer T. and Young P. C. Longitudinal dispersion in natural streams. *ASCE Journal of Environmental Engineering*, 109:1049, DOI:10.1061/(ASCE)0733–9372(1983)109:5(1049), 1983.

Bell V. A., Kay A. L., Jones D. A. and Moore R. J. Development of a high resolution grid-based hydrological model for use with RCM outputs. *Hydrology and Earth System Sciences*, 11:532–549, 2007b.

Bell V. A., Kay A. L., Jones R. G. and Moore R. J. Use of a gridbased hydrological model and regional climate model outputs to assess changing flood risk. *International Journal of Climatology*, 27:1657–1671, 2007a.

Bell V. A., Kay A. L., Jones R. G., Moore R. J. and Reynard N. S. Use of soil data in a grid-based hydrological model to estimate spatial variation in changing flood risk across the UK. *Journal of Hydrology*, 377:335–350, 2009.

Bell V. A. and Moore R. J. A grid-based distributed flood forecasting model for use with weather radar data: 1. Formulation. *Hydrology and Earth System Sciences*, 2:265–281, 1998a.

Bell V. A. and Moore R. J. A grid-based distributed flood forecasting model for use with weather radar data: 2. Case studies. *Hydrology and Earth System Sciences*, 2:283–298, 1998b.

Bencala K. E. and Walters R. A. Simulation of solute transport in a mountain pool and riffle stream: A transient storage model. *Water Resources Research*, 19:718–724, 1983.

Beran M. Hydrograph prediction: How much skill? *Hydrology and Earth System Sciences*, 3:305–307, 1999.

Bergstrom S. The development of a snow routine for the HBV-2 model. *Nordic Hydrology*, 6:73–92, 1975.

Berkamp G. A hierarchical view of the interactions of runoff and infiltration with vegetation and microtopography in semiarid shrublands. *Catena*, 33:201–220, 1998.

Berman E. S. F., Gupta M., Gabrielli C., Garland T. and McDonnell J. J. High-frequency field-deployable isotope analyzer for hydrological applications. *Water Resources Research*, 45:W10201, DOI:10.1029/2009WR008265, 2009.

Bernier P. Y. Variable source areas and stormflow generation: An update of the concept and simulation effort. *Journal of Hydrology*, 79:195–213, 1985.

Berod D. D., Singh V. P. and Musy A. A geomorphologic kinematic-wave (GKW) model for estimation of floods from small alpine watersheds. *Hydrological Processes*, 13:1391–1416, 1999.

Von Bertalanffy L. An outline of general systems theory. *British Journal for the Philosophy of Science*, 1:134–165, 1951.

Von Bertalanffy L. *General Systems Theory*. Brazilier, New York, 1962.

Berthet L., Andréassian V., Perrin C. and Javelle P. How crucial is it to account for the antecedent moisture conditions in flood forecasting? Comparison of event-based and continuous approaches on 178 catchments. *Hydrology and Earth System Sciences*, 13:819–831, 2009.

Betson R. P. What is watershed runoff? *Journal of Geophysical Research*, 69:1541–51, 1964.

Beven K. J. Hillslope hydrographs by the finite element method. *Earth Surface Processes and Landforms*, 2:13–28, 1977.

Beven K. J. A sensitivity analysis of the Penman–Monteith actual evapotranspiration estimates. *Journal of Hydrology*, 44:169–190, 1979a.

Beven K. J. On the generalised kinematic routing method. *Water Resources Research*, 15:1238–1242, 1979b.

Beven K. J. Kinematic subsurface stormflow. *Water Resources Research*, 17(5):1419–1424, 1981a.
Beven K. J. The effect of ordering on the geomorphic effectiveness of hydrologic events. In *Erosion and Sediment Transport in Pacific Rim Steeplands*, IAHS Publ. No. 132, pages 510–526, IAHS Press: Wallingford, UK, 1981b.
Beven K. J. Infiltration into a class of vertically non-uniform soils. *Journal of Hydrology*, 29:425–434, 1984.
Beven K. J. Hillslope runoff processes and flood frequency characteristics. In Abrahams A. D., editor, *Hillslope Processes*, pages 187–202. Allen and Unwin, Boston, 1986a.
Beven K. J. Runoff production and flood frequency in catchments of order *n*: An alternative approach. In Gupta V. K., Rodriguez-Iturbe I. and Wood E. F., editors, *Scale Problems in Hydrology*, pages 107–131. Reidel, Dordrecht, 1986b.
Beven K. J. Towards the use of catchment geomorphology in flood frequency predictions. *Earth Surface Processes and Landforms*, 12:69–82, 1987.
Beven K. J. Changing ideas in hydrology: The case of physically-based models. *Journal of Hydrology*, 105:157–172, 1989.
Beven K. J. Spatially distributed modelling: Conceptual approach to runoff prediction. In Bowles D. S. and O'Connell P. E., editors, *Recent Advances in the Modelling of Hydrologic Systems*, pages 373–387. Kluwer Academic, 1991a.
Beven K. J. Hydrograph separation? In *Proceedings of the BHS 3rd National Hydrology Symposium*, Southampton, 1991b.
Beven K. J. Prophecy, reality and uncertainty in distributed hydrological modelling. *Advances in Water Resources*, 16:41–51, 1993.
Beven K. J. Linking parameters across scales: Sub-grid parameterisations and scale dependent hydrological models. *Hydrological Processes*, 9:507–525, 1995.
Beven K. J. Equifinality and uncertainty in geomorphological modelling, in Rhoads B. L. and Thorn C. E., editors, *The Scientific Nature of Geomorphology*, pages 289–313. John Wiley & Sons, Chichester, UK, 1996.
Beven K. J. A discussion of distributed modelling. In Abbott M. B. and Refsgaard J. C., editors, *Distributed Hydrological Modelling*, pages 255–278. Kluwer, Dordrecht, 1996a.
Beven K. J. The limits of splitting: Hydrology. *Science of the Total Environment*, 183:89–97, 1996b.
Beven K. J. *Distributed Hydrological Modelling: Applications of the TOPMODEL Concept*. John Wiley & Sons, Chichester, UK, 1997.
Beven K. J. Uniqueness of place and process representations in hydrological modelling. *Hydrology and Earth System Sciences*, 4(2):203–213, 2000.
Beven K. J. Dalton Medal Lecture: How far can we go in distributed hydrological modelling? *Hydrology and Earth System Sciences*, 5(1):1–12, 2001.
Beven K. J. Towards an alternative blueprint for a physically-based digitally simulated hydrologic response modelling system. *Hydrological Processes*, 16(2):189–206, 2002a.
Beven K. J. Towards a coherent philosophy for environmental modelling. *Proceedings of the Royal Society London*, A458:2465–2484, 2002b.
Beven K. J. Runoff production in semi-arid areas, in Kirkby M. J. and Bull L., editors, *Dryland Rivers*, pages 57–105. John Wiley & Sons, Chichester, UK, 2002c.
Beven K. J. Surface runoff at the Horton Hydrologic Laboratory (or not?). *Journal of Hydrology*, 293:219–234, 2004a.
Beven K. J. Robert Horton's perceptual model of infiltration. *Hydrological Processes*, 18:3447–3460, 2004b.
Beven K. J. Robert Horton and abrupt rises of groundwater. *Hydrological Processes*, 18:3687–3696, 2004c.
Beven K. J. On the concept of model structural error. *Water Science and Technology*, 52:165–175, 2005.
Beven K. J. A manifesto for the equifinality thesis. *Journal of Hydrology*, 320:18–36, 2006a.
Beven K. J. The holy grail of scientific hydrology: Qt = H(S R) A as closure. *Hydrology and Earth System Sciences*, 10:609–618, 2006b.
Beven K. J. On undermining the science? *Hydrological Processes*, 20:3141–3146, 2006c.
Beven K. J., editor. *Benchmark Papers in Hydrology: Streamflow generation processes*. IAHS Press, Wallingford, UK, 2006d.

Beven K. J. Working towards integrated environmental models of everywhere: Uncertainty, data, and modelling as a learning process. *Hydrology and Earth System Sciences*, 11:460–467, 2007.

Beven K. J. On doing better hydrological science. *Hydrological Processes*, 22:3549–3553, 2008.

Beven K. J. *Environmental Modelling: An Uncertain Future?* Routledge, London, UK, 2009.

Beven K. J. Comment on "Equifinality of formal (DREAM) and informal (GLUE) Bayesian approaches in hydrologic modeling?" by Vrugt J. A., ter Braak C. J. F., Gupta H. V. and Robinson B. A. *Stochastic Environmental Research and Risk Assessment*, 23:1059–1060, 2009b.

Beven K. J. Preferential flows and travel time distributions: Defining adequate hypothesis tests for hydrological process models. *Hydrological Processes*, 24:1537–1547, 2010.

Beven K. J. I believe in climate change but how precautionary do we need to be in planning for the future? *Hydrological Processes*, 25:1517–1520, 2011.

Beven K. J. and Alcock R. Modelling everything everywhere: A new approach to decision making for water management under uncertainty. *Freshwater Biology*, DOI:10.1111/j.1365–2427.2011.02592.x, 2011.

Beven K. J. and Binley A. M. The future of distributed models: Model calibration and uncertainty prediction. *Hydrological Processes*, 6(3):279–298, 1992.

Beven K. J., Calver A. and Morris E. M. The Institute of Hydrology Distributed Model. Technical Report Report No. 98, Institute of Hydrology, Wallingford, UK, 1987.

Beven K. J. and Clarke R. T. On the variation of infiltration into a homogeneous soil matrix containing a population of macropores. *Water Resources Research*, 22:383–388, 1986.

Beven K. J. and Freer J. A dynamic TOPMODEL. *Hydrological Processes*, 15(10):1993–2011, 2001.

Beven K. J. and Germann P. F. Water flow in soil macropores, II. A combined flow model. *Journal of Soil Science*, 32:15–29, 1981.

Beven K. J. and Germann P. F. Macropores and water flow in soils. *Water Resources Research*, 18:1311–1325, 1982a.

Beven K. J., Henderson D. E. and Reeves A. D. Dispersion parameters for undisturbed partially saturated soil. *Journal of Hydrology*, 143:19–44, 1993.

Beven K. J., Hornberger G. M. and Germann P. F. Hillslope hydrology: A multiple interacting pathways model. In *Proceedings of the British Hydrological Society Second National Hydrology Symposium*, pages 1.1–1.8, Wallingford, UK, 1989. British Hydrological Society.

Beven K. J. and Kirkby M. J. A physically based, variable contributing area model of basin hydrology. *Hydrological Sciences Bulletin*, 24(1):43–69, 1979.

Beven K. J., Kirkby M. J., Schoffield N. and Tagg A. Testing a physically-based flood forecasting model (TOPMODEL) for three UK catchments. *Journal of Hydrology*, 69:119–143, 1984.

Beven K. J., Lamb R., Quinn P. F., Romanowicz R. and Freer J. TOPMODEL. In Singh V. P., editor, *Computer Models of Watershed Hydrology*, pages 627–668. Water Resources Publications, Highlands Ranch, CO, 1995.

Beven K. J., Smith P. J. and Freer J. E. So just why would a modeller choose to be incoherent? *Journal of Hydrology*, 354:15–32, 2008a.

Beven, K. J., Smith P. J. and Wood A. On the colour and spin of epistemic error (and what we might do about it), *Hydrology and Earth System Sciences*, 15:3123–3133, 2011.

Beven K. J., Warren R. and Zaoui J. SHE: Towards a methodology for physically-based distributed forecasting in hydrology. *Int. Assoc. Sci. Hydrol. Publ. No.*, 129:133–137, 1980.

Beven K. J. and Westerberg I. On red herrings and real herrings: Disinformation and information in hydrological inference. *Hydrological Processes*, 25:1676–1680, 2011.

Beven K. J. and Wood E. F. Catchment geomorphology and the dynamics of runoff contributing areas. *Journal of Hydrology*, 65:139–158, 1983.

Beven K. J. and Wood E. F. Flow routing and the hydrological response of channel networks. In Beven K. J. and Kirkby M. J., editors, *Channel Network Hydrology*. John Wiley & Sons, Chichester, UK, 1993.

Beven K. J., Young P., Romanowicz R., O'Connell P. E., Ewen J., O'Donnell G., Holman I., Posthumus H., Morris J., Hollis J., Rose S., Lamb R. and Archer D. Analysis of historical data sets to look for impacts of land use and management change on flood generation. Final Report FD2120, Defra, London, 2008c.

Beven K. J. and Young P. C. Comment on "Bayesian recursive parameter estimation for hydrologic models" by Thiemann M., Trosset M., Gupta H. and Sorooshian S. *Water Resources Research*, 39:doi: W01116, 10.1029/2001WR001183, 2003.

Beven K. J., Young P. C., Leedal D. T. and Romanowicz R. Computationally efficient flood water level prediction (with uncertainty). In Samuels P., Huntingdon S., Allsop W. and Harrop J., editors, *Flood Risk Management: Research and Practice*. CRC Press, Balkema, 2008b.

Bingeman A. K., Kouwen N. and Soulis E. D. Validation of the hydrological processes in a hydrological model. *ASCE Journal of Hydrologic Engineering*, 11(5):451–463, 2006.

Binley A. and Beven K. J. Vadose zone model uncertainty as conditioned on geophysical data. *Ground Water*, 41(2):119–127, 2003.

Binley A. M. and Beven K. J. Three dimensional modelling of hillslope hydrology. *Hydrological Processes*, 6:347–359, 1992.

Binley A. M., Beven K. J., Calver A. and Watts L. G. Changing responses in hydrology: Assessing the uncertainty in physically based model predictions. *Water Resources Research*, 27(6):1253–1261, 1991.

Binley A. M., Beven K. J. and Elgy J. A physically-based model of heterogeneous hillslopes: II. Effective hydraulic conductivities. *Water Resources Research*, 25(6):1227–1233, 1989.

Binley A. M., Elgy J. and Beven K. A physically based model of heterogeneous hillslopes: I. Runoff production. *Water Resources Research*, 25(6):1219–1226, 1989.

Birkett C. M., Mertes L. A. K., Dunne T., Costa M. and Jasinski J. Altimetric remote sensing of the Amazon: Application of satellite radar altimetry. *Journal of Geophysical Research*, 107(D20):8059, DOI:10.1029/2001JD000609, 2002.

Blackie J. R. and Eeles C. W. O. Lumped catchment models. In Anderson M. G. and Burt T. P. editors, *Hydrological Forecasting*, pages 311–346, Chichester, UK, 1985. John Wiley & Sons.

Blasone R.-S., Vrugt J. A., Madsen H., Rosbjerg D., Robinson M. A. and Zyvoloski G. A. Generalized likelihood uncertainty estimation (GLUE) using adaptive Markov chain Monte Carlo sampling. *Advances in Water Resources*, 31:630–648, 2008.

Blazkova S. and Beven K. J. Flood frequency prediction for data limited catchments in the Czech Republic using a stochastic rainfall model and TOPMODEL. *Journal of Hydrology*, 195:256–278, 1997.

Blazkova S. and Beven K. J. Flood frequency estimation by continuous simulation for an ungauged catchment with fuzzy possibility uncertainty estimation. *Water Resources Research*, 292(1–4):153–172, 2000.

Blazkova S. and Beven K. J. Flood frequency estimation by continuous simulation of subcatchment rainfalls and discharges with the aim of improving dam safety assessment in a large basin in the Czech Republic. *Journal of Hydrology*, 292:153–172, 2004.

Blazkova S. and Beven K. J. A limits of acceptability approach to model evaluation and uncertainty estimation in flood frequency estimation by continuous simulation: Skalka catchment, Czech Republic. *Water Resources Research*, 45:W00B16, DOI:10.1029/2007WR006726, 2009.

Blazkova S., Beven K. J. and Kulasova A. On constraining TOPMODEL hydrograph simulations using partial saturated area information. *Hydrological Processes*, 16(2):441–458, 2002b.

Blazkova S., Beven K. J., Tacheci P. and Kulasova A. Testing the distributed water table predictions of TOPMODEL (allowing for uncertainty in model calibration): The death of TOPMODEL? *Water Resources Research*, 38(11):W01257, 10.1029/2001WR000912, 2002a.

Blöschl G. Rainfall-runoff modelling of ungauged catchments *in M. G. Anderson, editor, Encyclopaedia of Hydrological Sciences*, pages 2061–2080. John Wiley & Sons, Chichester, UK, 2005.

Blöschl G., Kirnbauer R. and Gutnecht D. Distributed snowmelt simulations in an alpine catchment 1. Model evaluation on the basis of snow cover patterns. *Water Resources Research*, 27:3171–3179, 1991.

Blume T., Zehe E. and Bronstert A. Use of soil moisture dynamics and patterns at different spatio-temporal scales for the investigation of subsurface flow processes. *Hydrology and Earth System Sciences*, 13:1215–1234, 2009.

Bonan G. B., Oleson K. W., Vertenstein M., Levis S., Zeng X., Dai Y., Dickinson R. E. and Yang Z.-L. The land surface climatology of the Community Land Model coupled to the NCAR Community Climate Model. *Journal of Climate*, 15:3123–3149, 2002.

Bonell M. Runoff generation in tropical forests. In Bonell M. and Bruijnzeel L. A., editors, *Forests, Water and People in the Humid Tropics: Past, present and future hydrological research for integrated land and water management*, pages 314–406. Cambridge University Press, Cambridge, UK, 2004.

Boorman D. B., Hollis J. M. and Lilly A. Hydrology of soil types: A hydrologically-based classification of the soils of the United Kingdom. Technical Report 126, Institute of Hydrology, Wallingford, UK, 1995.

Borah D. K., Prasad S. N. and Alonso C. V. Kinematic wave routing incorporating shock fitting. *Water Resources Research*, 16:529–541, 1980.

Botter G., Bertuzzo E. and Rinaldo A. Transport in the hydrologic response: Travel time distributions, soil moisture dynamics, and the old water paradox. *Water Resources Research*, 46:W03514, DOI:10.1029/2009WR008371, 2010.

Botter G., Milan E., Bertuzzo E., Zanardo S., Marani M. and Rinaldo A. Inferences from catchment-scale tracer circulation experiments. *Journal of Hydrology*, 369:368–380, 2009.

Boughton W. *Rainfall–Runoff Modelling with the AWBM*. EA Books, Crows Nest, NSW, 2011.

Boughton W. and Chiew F. Estimating runoff in ungauged catchments from rainfall, PET and the AWBM model. *Environmental Modelling & Software*, 22:476–487, 2007.

Box G. E. P. and Cox D. R. An analysis of transformations. *Journal of the Royal Statistical Society*, B26:211–243, Discussion 244–252, 1964.

Box G. E. P. and Jenkins G. M. *Time Series Analysis, Forecasting and Control*. Holden-Day, San Fransisco, 1970.

Boyle D. P., Gupta H. V. and Sorooshian S. Toward improved calibration of hydrologic models: Combining the strengths of manual and automatic methods. *Water Resources Research*, 36:3663–3674, 2000.

Bradbrook K. JFLOW: A multiscale two-dimensional dynamic flood model. *CIWEM Water and Environment Journal*, 20:79–86, 2006.

Brakensiek D. L. and Rawls W. J. Soil containing rock fragments: Effects on infiltration. *Catena*, 23:99–110, 1994.

Branger F., Braud I., Debionne S., Viallet P., Dehotin J., Henine H., Nedelec Y. and Anquetin S. Towards multi-scale integrated hydrological models using the LIQUID framework: Overview of the concepts and first application examples. *Environmental Modelling & Software*, 25:1672–1681, 2010.

Bras R. E. *Hydrology*. Addison-Wesley, 1990.

Brath A., Montanari A. and Moretti G. Assessing the effect on flood frequency of land use change via hydrological simulation (with uncertainty). *Journal of Hydrology*, 324:141–153, 2006.

Brath A., Montanari A. and Toth E. Analysis of the effects of different scenarios of historical data availability on the calibration of a spatially-distributed rainfall–runoff model. *Journal of Hydrology*, 291:271–288, 2004.

Braud I., De Condappa D., Soria J. M., Haverkamp R., Angulo-Jaramillo R., Galle S. and Vauclin M. Use of scaled forms of the infiltration equation for the estimation of unsaturated soil hydraulics properties (the Beerkan method). *European Journal of Soil Science*, 56:361–374, 2005.

Braun L. N. and Lang H. Simulations of snowmelt runoff in lowland and lower alpine regions of Switzerland. In *Modelling Snowmelt-Induced Processes*, pages 125–140, Wallingford, UK, 1986. Int. Assoc. Sci. Hydrol., Publ. No. 155.

Bray M. and Han D. Identification of support vector machines for runoff modelling. *Journal of Hydroinformatics*, 6:265–280, 2004.

Brazier R. E., Beven K. J., Freer J. and Rowan J. S. Equifinality and uncertainty in physically-based soil erosion models: Application of the GLUE methodology to WEPP – the water erosion prediction project – for sites in the UK and USA. *Earth Surface Processes and Landforms*, 25:825–845, 2000.

Breiman L. Random forests. *Machine Learning*, 45:5–32, 2001.

Breiman L., Friedman J., Olshen R. and Stone C. *Classification and Regression Trees*. Wadsworth, Belmont, CA, 1984.

Bronstert A. Capabilities and limitations of distributed hillslope hydrological modelling. *Hydrological Processes*, 13:21–48, 1999.

Bronstert A., Bárdossy A., Bismuth C., Buiteveld H., Disse M., Engel H., Fritsch U., Hundecha Y., Lammersen R., Niehoff D. and Ritter N. Multi-scale modelling of land-use change and river training effects on floods in the Rhine basin. *River Research and Applications*, 23:1102–1125, 2007.

Bronstert A. and Plate E. Modelling of runoff generation and soil moisture dynamics for hillslopes and micro-catchments. *Journal of Hydrology*, 198:177–195, 1997.

Brooks J. R., Barnard H. R., Coulombe R. and McDonnell J. J. Ecohydrologic separation of water between trees and streams in a Mediterranean climate. *Nature Geoscience*, 3:100–104, 2010.

Brooks R. H. and Corey A. T. Hydraulic properties of porous media. Technical Report Hydrology Paper No. 3, Colorado State University, Fort Collins, CO, 1964.

Brun E., Durand Y. and Martin E. A meteorological estimation of relevant parameters for snow models. *Annals of Glaciology*, 18, 1993.

Brutsaert W. and Nieber J. L. Regionalized drought flow hydrographs from a mature glaciated plateau. *Water Resources Research*, 13(3):637–643, 1977.

Brutsaert W. and Sugito M. Application of self-preservation in the diurnal evolution of the surface energy budget to determine daily evaporation. *Journal of Geophysical Research*, 97(D17):18377–18382, 1992.

Bulygina N. and Gupta H. Estimating the uncertain mathematical structure of a water balance model via Bayesian data assimilation. *Water Resources Research*, 45:W00B13, DOI:10.1029/2007WR006749, 2009.

Bulygina N. and Gupta H. How Bayesian data assimilation can be used to estimate the mathematical structure of a model. *Stochastic Environmental Research and Risk Assessment*, 24:925–937, 2010.

Bulygina N., McIntyre N. and Wheater H. Conditioning rainfall–runoff model parameters for ungauged catchments and land management impacts analysis. *Hydrology and Earth System Sciences*, 13:893–904, 2009.

Burdine N. T. Relative permeability calculations from pore size distribution data. *Transactions of the American Institute of Mining and Metallurgical Engineers*, 198:71–78, 1953.

Burges S. J., Wigmosta M. S. and Meena J. M. Hydrological effects of land-use change in a zero-order catchment. *ASCE Journal of Hydrologic Engineering*, 3:86–97, 1998.

Burnash R. J. C. The NWS river forecast system: Catchment modeling. In Singh V. P., editor, *Computer Models of Watershed Hydrology*, pages 311–366, Highlands Ranch, CO, 1995. Water Resources Publications.

Burns D. A., McDonnell J. J., Hooper R. P., Peters N. E., Freer J. E., Kendall C. and Beven K. J. Quantifying contributions to storm runoff through end-member mixing analysis and hydrologic measurements at the Panola mountain research watershed (Georgia, USA). *Hydrological Processes*, 15:1903–1924, 2001.

Butts M. B., Payne J. T., Kristensen M. and Madsen H. An evaluation of the impact of model structure on hydrological modelling uncertainty for streamflow prediction. *Journal of Hydrology*, 298:242–266, 2004.

Buytaert W. and Beven K. J. Regionalisation as a learning process. *Water Resources Research*, 45:W11419, DOI:10.1029/2008WR007359, 2009.

Buytaert W., Reusser D., Krause S. and Renaud J.-P. Why can't we do better than TOPMODEL? *Hydrological Processes*, 22:4175–4179, 2008.

Calcagno G., Mendicino G., Monacelli G., Senatore A. and Versace P. *Methods and Tools for Drought Analysis and Management*, pages 125–147, DOI: 10.1007/978–1–4020–5924–7_7. Springer-Verlag, Berlin, 2007.

Calder I. R. A model of transpiration and interception loss from a spruce forest in Plynlimon, central Wales. *Journal of Hydrology*, 33:247–265, 1977.

Calder I. R. A stochastic model of rainfall interception. *Journal of Hydrology*, 89:65–71, 1986.

Calder I. R. *Evaporation in the Uplands*. John Wiley & Sons, Chichester, UK, 1990.

Calder I. R. Dependence of rainfall interception on drop size: 1. Development of the two-layer stochastic model. *Journal of Hydrology*, 185:363–378, 1996.

Calder I. R., Hall R. L., Rosier P. T. W., Bastable H. G. and Prasanna K. T. Dependence of rainfall interception on drop size: 2. Experimental determination of the wetting functions and two-layer stochastic model parameters for five tropical tree species. *Journal of Hydrology*, 185:379–388, 1996.

Calder I. R., Harding R. J. and Rosier P. T. W. An objective assessment of soil moisture deficit models. *Journal of Hydrology*, 60:329–355, 1983.

Calver A. Calibration, sensitivity and validation of a physically-based rainfall–runoff model. *Journal of Hydrology*, 103:103–115, 1988.

Calver A. and Cammeraat L. H. Testing a physically-based runoff model against field observations on a Luxembourg hillslope. *Catena*, 20:273–288, 1993.

Calver A. and Lamb R. Flood frequency estimation using continuous rainfall–runoff modelling. *Physics and Chemistry of the Earth*, 20:479–483, 1996.

Calver A., Stewart E. and Goodsell G. Comparative analysis of statistical and catchment modelling approaches to river flood frequency estimation. *Journal of Flood Risk Management*, 2:24–31, 2009.

Calver A. and Wood W. L. The Institute of Hydrology Distributed Model. In Singh V. P., editor, *Computer Models of Watershed Hydrology*, chapter 17, pages 595–626. Water Resources Publications, 1995.

Cameron D., Beven K. J. and Tawn J. Modelling extreme rainfalls using a modified random pulse Bartlett–Lewis stochastic rainfall model (with uncertainty). *Advances in Water Resources*, 24:203–211, 2001.

Cameron D., Kneale P. and See L. An evaluation of a traditional and a neural net modelling approach to flood forecasting for an upland catchment. *Hydrological Processes*, 16:1033–1046, 2002.

Cameron D. S., Beven K. J., Tawn J., Blazkova S. and Naden P. Flood frequency estimation by continuous simulation for a gauged upland catchment (with uncertainty). *Journal of Hydrology*, 219:169–187, 1999.

Campolo M., Andreussi P. and Soldati A. River flood forecasting with a neural network model. *Water Resources Research*, 35:1191–1197, 1999.

Campolo M., Soldati A. and Andreussi P. Artificial neural network approach to flood forecasting in the River Arno. *Hydrological Sciences Journal*, 48:381–398, 2003.

Canfield H. E. and Goodrich D. C. Differentiating the impact of parameter lumping from the impact of geometric simplification in modeling runoff and erosion. *Hydrological Processes*, 20:17–35, 2006.

Cappus P. Bassin versant expérimental d'Alrance: études des lois de l'écoulement. Application au calcul et à la prévision des débits. *La Houille Blanche*, A:493–520, 1960.

Carsel R. F., Mulkey L. A., Lorber M. N. and Bakin L. B. The pesticide root zone model (PRZM): A procedure for evaluating pesticide leaching threats to groundwater. *Ecological Modelling*, 30:49–69, 1985.

Castellarin A., Camorani G. and Brath A. Predicting annual and long-term flow-duration curves in ungauged basins. *Advances in Water Resources*, 30:937–953, 2007.

Chapman T. G. Common unitgraphs for sets of runoff events: Part 2. Comparisons and inferences for rainfall loss models. *Hydrological Processes*, 10:783–792, 1996.

Chappell N. A. Soil pipe distribution and hydrological functioning within the humid tropics: A synthesis. *Hydrological Processes*, 24:1567–1581, 2010.

Chappell N. A., Tych W., Chotai A., Bidin K., Sinun W. and Chiew T. H. Barumodel: Combined data based mechanistic models of runoff response in a managed rainforest catchment. *Forest Ecology and Management*, 224:58–80, 2006.

Charbeneau R. J. Kinematic models for soil moisture and solute transport. *Water Resources Research*, 20:699–706, 1984.

Chaubey I., Haan C. T., Grunwald S. and Salisbury J. M. Uncertainty in the model parameters due to spatial variability of rainfall. *Journal of Hydrology*, 220:48–61, 1999.

Chehbouni A., Kerr Y. H., Watts C., Hartogensis O., Goodrich D., Scott R., Schieldge J., Lee K., Shuttleworth W. J., Dedieu G. and De Bruin H. A. R. Estimation of area-average sensible heat flux using a large-aperture scintillometer during the semi-arid land–surface–atmosphere (SALSA) experiment. *Water Resources Research*, 35:2505–2511, 1999.

Chibanga R., Berlamont J. and Vandewalle J. Modelling and forecasting of hydrological variables using artificial neural networks: The Kafue River sub-basin. *Journal of Hydrology*, 48:363–379, 2003.

Chiew F. H. S., Stewardson M. J. and McMahon T. A. Comparison of six rainfall–runoff modelling approaches. *Journal of Hydrology*, 147:1–36, 1993.

Chirico G. B., Grayson R. B. and Western A. W. On the computation of the quasi-dynamic wetness index with multiple-flow-direction algorithms. *Water Resources Research*, 39:W01115, DOI:10.1029/2002WR001754, 2003.

Choi H. T. and Beven K. J. Multi-period and multi-criteria model conditioning to reduce prediction uncertainty in distributed rainfall–runoff modelling within GLUE framework. *Journal of Hydrology*, 332(3–4):316–336, 2007.

Chow V. T. *Handbook of Applied Hydrology*. McGraw-Hill, New York, 1964.

Chow V. T. and Kulandaiswamy V. C. General hydrologic systems model. *ASCE Journal of the Hydraulics Division*, 97 (HY6):791–804, 1971.

Christensen N. S., Wood A. W., Voisin N., Lettenmaier D. P. and Palmer R. N. The effects of climate change on the hydrology and water resources of the Colorado River basin. *Climatic Change*, 62:337–363, 2004.

Christiaens K. and Feyen J. Constraining soil hydraulic parameter and output uncertainty of the distributed hydrological MIKE SHE model using the GLUE framework. *Hydrological Processes*, 16:373–391, 2002.

Christophersen N., Neal C., Hooper R. P., Vogt R. D. and Andersen S. Modelling streamwater chemistry as a mixture of soilwater end-members: A step towards second-generation acidification models. *Journal of Hydrology*, 116:307–320, 1990.

Chutha P. and Dooge J. C. I. The shape parameters of the geomorphologic unit hydrograph. *Journal of Hydrology*, 117:81–97, 1990.

Ciarapica L. *TOPKAPI: Un Modello Afflussi–Deflussi Distribuito Applicabile Dalla Scala Di Versante Alla Scala Di Bacino*. PhD thesis, Università di Bologna, Italy, 1998.

Ciarapica L. and Todini E. TOPKAPI: A model for the representation of the rainfall–runoff process at different scales. *Hydrological Processes*, 16:207–229, 2002.

Clapp R. B. and Hornberger G. M. Empirical equations for some soil hydraulic properties. *Water Resources Research*, 14(4):601–604, 1978.

Clapp R. B., Hornberger G. M. and Cosby B. J. Estimating spatial variation in soil moisture with a simple dynamic model. *Water Resources Research*, 19:739–745, 1983.

Clapp R. B., Timmins S. P. and Huston M. A. Visualising the surface hydrodynamics of a forested watershed. In Russell T. F. *et al.*, editor, *Proc. IX Internat. Conf. On Computational Methods in Water Resources, V.2 Mathematical Models of Water Resources*. Elsevier, New York, 1992.

Clark C. O. Storage and unit hydrograph. *Transactions of the American Society of Civil Engineers*, 110:1416–1446, 1945.

Clark M. P. and Kavetski D. Ancient numerical daemons of conceptual hydrological modeling: 1. Fidelity and efficiency of time stepping schemes. *Water Resources Research*, 46:W10510, DOI:10.1029/2009WR008894, 2010.

Clark M. P., Slater A. G., Rupp D. E., Woods R. A., Vrugt J. A., Gupta H. V., Wagener T. and Hay L. E. Framework for Understanding Structural Errors (FUSE): A modular framework to diagnose differences between hydrological models. *Water Resources Research*, 44:W00B02, DOI:10.1029/2007WR006735, 2008.

Clarke R. T. A review of some mathematical models used in hydrology, with observations on their calibration and use. *Journal of Hydrology*, 19:1–20, 1973.

Clarke R. T. *Stochastic Processes for Water Scientists: Developments and applications*. John Wiley & Sons, Chichester, UK, 1998.

Clausnitzer V., Hopmans J. H. and Nielsen D. R. Simultaneous scaling of soil water retention and hydraulic conductivity curves. *Water Resources Research*, 28:19–31, 1992.

Climent-Soler D., Holman I. P. and Archer D. R. Application of flow variability analysis to identify impacts of agricultural land-use change on the River Axe, southwest England. *Hydrology Research*, 40:380–393, 2009.

Cloke H. L. and Pappenberger F. Ensemble flood forecasting: A review. *Journal of Hydrology*, 375:613–626, 2009.

Cluckie I. D. Hydrological forecasting: Real time. In Young P. C., editor, *Encyclopaedia of Environmental Systems*, pages 291–298. Pergamon Press, Oxford, UK, 1993.

Cobby D. M., Mason D. C., Horritt M. S. and Bates P. D. Two-dimensional hydraulic flood modelling using a finite-element mesh decomposed according to vegetation and topographic features derived from airborne scanning laser altimetry. *Hydrological Processes*, 17:1979–2000, 2003.

Colbeck S. C. Water flow through snow overlying an impermeable boundary. *Water Resources Research*, 10(1):119–123, 1974.

Cole S. J. and Moore R. J. Hydrological modelling using raingauge- and radar-based estimators of areal rainfall. *Journal of Hydrology*, 358(3-4):159–181, 2008.

Connolly R. D., Silburn D. M. and Ciesiolka C. A. A. Distributed parameter hydrology model (Answers) applied to a range of catchment scales using rainfall simulator data: III. Application to a spatially complex catchment. *Journal of Hydrology*, 193:183–203, 1997.

Cooper D. M. and Wood E. F. Parameter estimation of multiple input-output time series models: Application to rainfall–runoff processes. *Water Resources Research*, 18:1352–1364, 1982.

Corradini C., Melone F. and Smith R. E. A unified model for infiltration and redistribution during complex rainfall patterns. *Journal of Hydrology*, 192:104–124, 1997.

Corradini C., Morbidelli R. and Melone F. On the interaction between infiltration and Hortonian runoff. *Journal of Hydrology*, 204:52–67, 1998.

Cosby B. J., Hornberger G. M., Clapp R. B. and Ginn T. R. A statistical exploration of the relationship of soil moisture characteristics to the physical properties of soils. *Water Resources Research*, 20:682–690, 1984.

Cowpertwait P. S. P., O'Connell P. E., Metcalf A. V. and Mawdsley J. Stochastic point process modelling of rainfall: I. Single-site fitting and validation. *Journal of Hydrology*, 175:17–46, 1996a.

Cowpertwait P. S. P., O'Connell P. E., Metcalf A. V. and Mawdsley J. Stochastic point process modelling of rainfall: II. Regionalisation and disaggregation. *Journal of Hydrology*, 175:47–65, 1996b.

Crago R. D. and Brutsaert W. Daytime evaporation and the self-preservation of the evaporation fraction and the Bowen ratio. *Journal of Hydrology*, 178:241–255, 1996.

Crago R. D. and Crowley R. Complementary relationships for near-instantaneous evaporation. *Journal of Hydrology*, 300:199–211, 2005.

Cranmer A., Kouwen N. and Mousavi S. F. Proving WATFLOOD: Modelling the non-linearities of hydrologic response to storm intensities. *Canadian Journal of Civil Engineering*, 28:837–855, 2001.

Crawford N. H. and Linsley R. K. Digital simulation in hydrology: Stanford Watershed Model IV. *Technical Report*, Dept. of Civil Engineering, Stanford University, CA, 39, 1966.

Cristianini N. and Shawe-Taylor J. *An Introduction to Support Vector Machines*. Cambridge University Press, Cambridge, UK, 2000.

Croke B. F. W., Andrews F., Jakeman A. J., Cuddy S. M. and Luddy A. IHACRES Classic Plus: A redesign of the IHACRES rainfall–runoff model. *Environmental Modelling & Software*, 21:426–427, 2006.

Croke B. F. W. and Jakeman A. J. A catchment moisture deficit module for the IHACRES rainfall–runoff. *Environmental Modelling & Software*, 19:1–5, 2004.

Crouzet E., Hubert P., Olive P. and Siwertz E. Le tritium dans les mesures d'hydrologie de surface. Détermination expérimentale du coefficient de ruisellement. *Journal of Hydrology*, 11:217–229, 1970.

Culling W. E. H. Multicyle streams and the equilibrium theory of grade. *J. Geology*, 65:259–274, 1957.

Culling W. E. H. Equifinality: Modern approaches to dynamical systems and their potential for geographical thought. *Transactions of the Institute of British Geographers*, 12:57–72, 1987.

Cunderlik J. M. and Burn D. H. Switching the pooling similarity distances: Mahalanobis for Euclidean. *Water Resources Research*, 42:W03409, DOI:10.1029/2005WR004245, 2006.

Cunge J. A. On the subject of a flood routing method (Muskingum method). *Journal of Hydraulic Research*, 7:205–230, 1969.

Cunnane C. Statistical distributions for flood frequency analysis. *Operational Hydrology Report 33*, World Meteorological Organisation, Geneva, 1989.

Darcy H. *Les Fontaines de la Ville de Dijon*. Dalmont, Paris, 1856.

Davie T. J. A. Modelling the influence of afforestation on hillslope storm runoff. In Anderson M. G. and Brooks S. M., editors, *Advances in Hillslope Processes, V*, pages 149–184, John Wiley & Sons, Chichester, UK, 1996.

Davies J. and Beven K. J. Scaling implications of a new multiple interacting pathways hillslope model. In *Proceedings of BHS International Hydrology Symposium 2010*. British Hydrological Society, 2010.

Davies J., Beven K. J., Nyberg L. and Rodhe A. A discrete particle representation of hillslope hydrology: Hypothesis testing in reproducing a tracer experiment at Gårdsjön, Sweden. *Hydrological Processes*, 25:DOI: 10.1002/hyp.8085, 2011.

Davis L. *Handbook of Genetic Algorithms*. Van Nostrand Reinhold, New York, 1991.

Dawdy D. R. and O'Donnell T. Mathematical models of catchment behaviour. *ASCE Journal of the Hydraulics Division*, 91(HY4):123–127, 1965.

Dawes W. R., Zhang L., Hatton T. J., Reece P. H., Beale G. T. H. and Packer I. Evaluation of a distributed parameter ecohydrological model (TOPOG-IRM) on a small cropping rotation catchment. *Journal of Hydrology*, 191:64–86, 1997.

Dawson C. and Wilby R. An artificial neural network approach to rainfall–runoff modelling. *Journal of Hydrology*, 43:47, 1998.

De Bruin H. A. R., van Den Hurk B. J. J. M. and Kohsiek W. The scintillation method tested over a dry vineyard area. *Boundary Layer Meteorol.*, 76:25–40, 1995.

De Roo A. P. J., Wesseling C. G. and van Deursen W. P. A. Physically based river basin modelling within a GIS: The LISFLOOD model. *Hydrological Processes*, 14:1981–1992, 2000.

Dean S., Freer J. E., Beven K. J., Wade A. J. and Butterfield D. Uncertainty assessment of a process-based integrated catchment model of phosphorus (INCA-P). *Stochastic Environmental Research and Risk Assessment*, 23:991–1010, DOI 10.1007/s00477–008–0273–z, 2009.

Di Baldassarre G. and Montanari A. Uncertainty in river discharge observations: A quantitative analysis. *Hydrology and Earth System Sciences*, 13:913–921, 2009.

Di Baldassarre G., Schumann G. and Bates P. D. A technique for the calibration of hydraulic models using uncertain satellite observations of flood extent. *Journal of Hydrology*, 367:276–282, 2009.

Dibike Y. B., Solomatine D. and Abbott M. B. On the encapsulation of numerical-hydraulic models in artificial neural network. *Journal of Hydraulic Research*, 37:147–161, 1999.

Dietrich W. E., Reiss R., Hsu M-L. and Montgomery D. R. A process-based model for colluvial soil depth and shallow landsliding using digitial elevation data. *Hydrological Processes*, 9:383–400, 1995.

Dietterick B. C., Lynch J. A. and Corbett E. S. A calibration procedure using TOPMODEL to determine suitability for evaluating potential climate change effects on water yield. *Journal of the American Water Resources Association*, 35:457–468, 1999.

Dinçer T., Payne B. R., Florkowski T., Martinec J. and Tongiorgi E. G. E. I. Snowmelt runoff from measurements of tritium and oxygen-18. *Water Resources Research*, 6:110–124, 1970.

Diskin M. H. and Boneh A. Determination of optimal kernels for second-order stationary surface runoff systems. *Water Resources Research*, 9:311–325, 1973.

Dixon R. M. and Linden D. R. Soil air pressure and water infiltration under border irrigation. *Proceedings of the Soil Science Society of America*, 36:948–953, 1972.

Doe W. W., Saghafian B. and Julien P. Y. Land-use impact on watershed response: The integration of two-dimensional hydrological modelling and geographical information systems. *Hydrological Processes*, 10:1503–1511, 1996.

Dolman A. J. Summer and winter rainfall interception in an oak forest. Predictions with an analytical and a numerical simulation model. *Journal of Hydrology*, 90:1–9, 1987.

Donigian A. S. D., Bicknell B. R. and Imhoff J. C. Hydrological simulation program – fortran. In Singh V. P., editor, *Computer Models of Watershed Hydrology*, pages 395–442. Water Resource Publications, Highlands Ranch, CO, 1995.

Donnelly-Makowecki L. M. and Moore R. D. Hierarchical testing of three rainfall–runoff models in small forested catchments. *Journal of Hydrology*, 219:136–152, 1999.

Dooge J. C. I. The rational method for estimating flood peaks. *Engineering*, 184:311–313, 374–377, 1957.

Dooge J. C. I. A general theory of the unit hydrograph. *Journal of Geophysical Research*, 64:241–256, 1959.

Dooge J. C. I. and O'Kane P. *Deterministic Methods in Systems Hydrology*. Balkema A. A., Lisse, NL, 2003.

Dornes P. F., Tolson B. A., Davison B., Pietroniro A., Pomeroy J. W. and Marsh P. Regionalisation of land surface hydrological model parameters in subarctic and arctic environments. *Physics and Chemistry of the Earth*, 33:1081–1089, 2008.

Dottori F., Martina M. L. V. and Todini E. A dynamic rating curve approach to indirect discharge measurement. *Hydrology and Earth System Sciences*, 13:847–863, 2009.

Dou C., Woldt W., Bogardi I. and Dahab M. Steady state groundwater flow simulations with imprecise parameters. *Water Resources Research*, 31:2709–2719, 1995.

Downer C. W. and Ogden F. L. GSSHA: Model to simulate diverse stream flow producing processes. *ASCE Journal of Hydrologic Engineering*, 9:161–174, 2004.

Downer C. W., Ogden F. L., Martin W. and Harmon R. S. Theory, development, and applicability of the surface water hydrologic model CASC2D. *Hydrological Processes*, 16:255–275, 2002.

Downer C. W., Ogden F. L., Niedzialek J. M. and Liu S. *Gridded Surface/Subsurface Hydrologic Analysis (GSSHA) Model: A model for simulating diverse streamflow producing processes*, pages 131–159. Taylor and Francis Group, CRC Press, 2005.

Draper D. Assessment and propagation of model uncertainty. *Journal of the Royal Statistical Society Series B*, 57(1):45–97, 1995.

Duan Q., Schaake J., Andréassian V., Franks S. and Goteti G. *et al.* The model parameter estimation experiment (MOPEX): An overview of science strategy and major results from the second and third workshops. *Journal of Hydrology*, 320:3–17, 2006.

Duan Q., Sorooshian S. and Gupta V. Effective and efficient global optimisation for conceptual rainfall–runoff models. *Water Resources Research*, 28:1015–1031, 1992.

Duan Q., Gupta V. K. and Sorooshian S. Shuffled complex evolution approach for effective and efficient global minimization. *J. Optimization Theory and Applications*, 76:501–521, 1993.

Duband D., Obled C. and Rodriguez J.-Y. Unit hydrograph revisited: An alternative approach to UH and effective precipitation identification. *Journal of Hydrology*, 150:115–149, 1993.

Duffy C. J. A two-state integral-balance model for soil moisture and groundwater dynamics in complex terrain. *Water Resources Research*, 32:2421–2434, 1996.

Duffy C. J. Dynamical modelling of concentration–age–discharge in watersheds. *Hydrological Processes*, 24:1711–1718, 2010.

Duffy C. J. and Gelhar L. W. Frequency domain approach to water quality modeling in groundwater: Theory. *Water Resources Research*, 21:1175–1184, 1985.

Dumenil L. and Todini E. A rainfall–runoff scheme for use in the Hamburg climate model. In O'Kane J. P., editor, *Advances in Theoretical Hydrology: A Tribute to James Dooge*, pages 129–157. Elsevier, Amsterdam, 1992.

Dunn S. M. and Ferrier R. C. Natural flow in managed catchments: A case study of a modelling approach. *Water Research*, 33:621–630, 1999.

Dunn S. M., Mackay R., Adams R. and Oglethorpe D. R. The hydrological component of the NELUP decision-support system: An appraisal. *Journal of Hydrology*, 177:213–235, 1996.

Dunne T. Field studies of hillslope flow processes. In Kirkby M. J., editor, *Hillslope Hydrology*, pages 227–293. John Wiley & Sons, 1978.

Dunne T. and Black R. D. An experimental investigation of runoff production in permeable soils. *Water Resources Research*, 6:478–490, 1970.

Dunne T., Price J. G. and Colbeck S. C. The generation of runoff from subarctic snowpacks. *Water Resources Research*, 12(4):677–685, 1976.

Dunne T., Zhang W. and Aubrey B. F. Effects of rainfall, vegetation and microtopography on infiltration and runoff. *Water Resources Research*, 27:2271–2286, 1991.

Durand P., Robson A. and Neal C. Modelling the hydrology of sub-Mediterranean montane catchments (Mont Lozère, France), using TOPMODEL: Initial results. *Journal of Hydrology*, 139:1–14, 1992.

Eagleson P. Climate, soil, and vegetation, 6: Dynamics of the annual water balance. *Water Resources Research*, 14:749–764, 1978.

Eagleson P. S. *Dynamic Hydrology*. McGraw-Hill, New York, 1970.

Eagleson P. S. Dynamics of flood frequency. *Water Resources Research*, 8:878–898, 1972.

Eagleson P. S. Ecological optimality in water-limited natural soil vegetation systems, 1: Theory and hypothesis. *Water Resources Research*, 18:325–340, 1982.

Eagleson P. S. *Ecohydrology: Darwinian expression of vegetation form and function*. Cambridge University Press, Cambridge, MA., 2002.

Eagleson P. S. and Tellers T. E. Ecological optimality in waterlimited natural soil-vegetation systems, 2: Tests and applications. *Water Resources Research*, 18:341–354, 1982.

Easton Z. M., Fuka D. R., Walter M. T., Cowan D. M., Schneiderman E. M. and Steenhuis T. S. Re-conceptualizing the soil and water assessment tool (SWAT) model to predict runoff from variable source areas. *Journal of Hydrology*, 348(3-4):279–291, 2008.

Easton Z. M., Fuka D. R., White E. D., Collick A. S., Biruk Ashagre B., McCartney M., Awulachew S. B., Ahmed A. A. and Steenhuis T. S. A multi basin SWAT model analysis of runoff and sedimentation in the Blue Nile, Ethiopia. *Hydrology and Earth System Sciences*, 14:1827–1841, 2010.

Ebel B. A. and Loague K. Physics-based hydrologic-response simulation: Seeing through the fog of equifinality. *Hydrological Processes*, 20:2887–2900, 2006.

Ebel B. A. and Loague K. Rapid simulated hydrologic response with the variably saturated near surface. *Hydrological Processes*, 22:464–471, 2008.

Ebel B. A., Loague K., Montgomery D. R. and Dietrich W. E. Physics-based continuous simulation of long-term near-surface hydrologic response for the Coos Bay experimental catchment. *Water Resources Research*, 44:W07417. DOI: 10.1029/2007WR006442, 2008.

Ebel B. A., Loague K., VanderKwaak J. E., Dietrich W. E., Montgomery D. R., Torres R. and Anderson S. P. Near-surface hydrologic response for a steep, unchanneled catchment near Coos Bay, Oregon: 2. Comprehensive physics-based simulations. *American Journal of Science*, 307:709–748, 2007.

Editjatno N. D. O., Yang X., Makhlouf Z. and Michel C. GR3J: A daily watershed model with three free parameters. *Journal of Hydrology*, 44:263–277, 1999.

Ehlers W. Observations on earthworm channels and infiltration on tilled and untilled loess soil. *Soil Science*, 119:242–249, 1975.

Endreny T. A. and Wood E. F. Representing elevation uncertainty in runoff modelling and flowpath mapping. *Hydrological Processes*, 15:2223–2236, 2001.

Engeland K., Xu C. Y. and Gottschalk L. Assessing uncertainties in a conceptual water balance model using Bayesian methodology. *Journal of Hydrology*, 50:45–63, 2005.

Engman E. T. and Rogowski A. S. A partial area model for storm flow synthesis. *Water Resources Research*, 10(3):464–472, 1974.

Erichsen B. and Myrabø S. Studies of the relationship between soil moisture and topography in a small catchment. In Gambolati G., editor, *Proc. 8th International Conference on Computational Methods in Water Resources, Venice*, 1990.

Eriksson E. Compartment model and reservoir theory. *Annual Review of Ecology and Systematics*, 2:67–84, 1971.

Espino A., Mallants D., Vanclooster M. and Feyen J. Cautionary notes on the use of pedotransfer functions for estimating soil hydraulic properties. *Agricultural Water Management*, 29:235–253, 1995.

Ewen J. 'SAMP' model for water and solute movement in unsaturated porous media involving thermodynamic subsystems and moving packets: 2. Design and application. *Journal of Hydrology*, 182:195–207, 1996.

Ewen J. and Birkinshaw S. J. Lumped hysteretic model for subsurface stormflow developed using downward approach. *Hydrological Processes*, 21(11):1496–1505, 2007.

Ewen J. and Parkin G. Validation of catchment models for predicting land-use and climate change impacts: 1. Method. *Journal of Hydrology*, 175:583–594, 1996.

Ewen J., Parkin G. and O'Connell P. E. Shetran: Distributed river basin flow and transport modelling system. *ASCE Journal of Hydrologic Engineering*, 5:250–258, 2000.

Faeh A. O., Scherrer S. and Naef F. A combined field and numerical approach to investigate flow processes in natural macroporous soils under extreme precipitation. *Hydrology and Earth System Sciences*, 1(4):787–800, 1997.

Fagre D. B., Comanor P. L., White J. D., Hauer F. R. and Running S. W. Watershed responses to climate change at glacier national park. *Journal of the American Water Resources Association*, 33:755–765, 1997.

Famiglietti J. S., Wood E. F., Sivapalan M. and Thongs D. J. A catchment scale water balance model for FIFE. *Journal of Geophysical Research*, 97(D17):18997–19007, 1992.

Farmer D. L., Sivapalan M. and Jothityangkoon C. Climate, soil and vegetation controls on variability of water balance in temperate and semi-arid landscapes: Downward approach to hydrological analysis. *Water Resources Research*, 39:W01035, DOI: 10.1029/2001 WR000328, 2003.

Faulkner H., Parker D., Green C. and Beven K. J. Developing a translational discourse to communicate uncertainty in flood risk between science and the practitioner. *Ambio*, 16(7):692–703, 2007.

Faurès J.-M., Goodrich D. C., Woolhiser D. A. and Sorooshian S. Impact of small-scale rainfall variability on runoff modelling. *Journal of Hydrology*, 173:309–326, 1995.

Federer C. A., Vörösmarty C. and Fekete B. Intercomparison of methods for calculating potential; evaporation in regional and global water balance models. *Water Resources Research*, 32:2315–2321, 1996.

Feldman A. D. HEC-1 flood hydrograph package. In Singh V. P., editor, *Computer Models of Watershed Hydrology*, pages 119–150, Water Resource Publications: Highland Ranch, CO, 1995.

Fenicia F., McDonnell J. J. and Savenije H. H. G. Learning from model improvement: On the contribution of complementary data to process understanding. *Water Resources Research*, 44:W06419, DOI:10.1029/2007WR006386, 2008b.

Fenicia F., Savenije H. H. G. and Avdeeva Y. Anomaly in the rainfall–runoff behaviour of the Meuse catchment: Climate, land-use, or land-use management? *Hydrology and Earth System Sciences*, 13:1727–1737, 2009.

Fenicia F., Savenije H. H. G., Matgen P. and Pfister L. A comparison of alternative multiobjective calibration strategies for hydrological modeling. *Water Resources Research*, 43:W03434, DOI:10.1029/2006WR005098, 2007b.

Fenicia F., Savenije H. H. G., Matgen P. and Pfister L. Understanding catchment behavior through stepwise model concept improvement. *Water Resources Research*, 44:W01402, DOI:10.1029/2006WR005563, 2008a.

Fenicia F., Solomatine D. P., Savenije H. H. G. and Matgen P. Soft combination of local models in a multi-objective framework. *Hydrology and Earth System Sciences*, 11:1797–1809, 2007a.

Fenicia F., Zhang G. P., Rientjes T., Hoffmann L., Pfister L. and Savenije H. H. G. Numerical simulations of runoff generation with surface water-groundwater interactions in the Alzette River alluvial plain (Luxembourg). *Physics and Chemistry of the Earth, Parts A/B/C*, 30(4–5):277–284, 2005.

Fernandez W., Vogel R. M. and Sankarasubramanian A. Regional calibration of a watershed model. *Journal of Hydrology*, 25:689–707, 2000.

Fernando D. A. K. and Jayawardena A. W. Runoff forecasting using RBF networks with OLS algorithm. *ASCE Journal of Hydrologic Engineering*, 3:203, 1998.

Feyen L., Vázquez R., Christiaens K., Sels O. and Feyen J. Application of a distributed physically-based hydrological model to a medium size catchment. *Hydrology and Earth System Sciences*, 4:47–63, 2000.

Feyen L., Vrugt J. A., O'Nuallain B., van der Knijff J. and De Roo A. Parameter optimisation and uncertainty assessment for large-scale streamflow simulation with the LISFLOOD model. *Journal of Hydrology*, 332:276–289, 2007.

Fisher J. and Beven K. J. Modelling of streamflow at Slapton Wood using TOPMODEL within an uncertainty estimation framework. *Field Studies*, 8:577–584, 1995.

Fleming G. *Computer Simulation Techniques in Hydrology*. Environmental Science Series, Elsevier, 1975.

Fleming G. and Franz D. D. Flood frequency estimating techniques for small watersheds. *Journal of the Hydraulics Division*, 97(9):1441–1460, 1971.

Flügel W.-A. Delineating hydrological response units by geographical information system analyses for regional hydrological modelling using PRMS/MMS in the drainage basin of the River Bröl, Germany. *Hydrological Processes*, 9:423–436, 1995.

Flury M., Flühler H., Jury W. A. and Leuenberger J. Susceptibility of soils to preferential flow of water: A field study. *Water Resources Research*, 30(7):1945–1954, 1994.

Flynn D., McNamara H., O'Kane P. and Pokrovskii A. *The Science of Hysteresis*, pages 681–736. Elsevier, Amsterdam, 2005.

Fontaine T. A. Rainfall–runoff model accuracy for an extreme flood. *ASCE Journal of Hydraulic Engineering*, 121(4):365–374, 1995.

Forrest S. Genetic algorithms: Principles of natural selection applied to computation. *Science*, 261:872–878, 1993.

Fortin J. P., Turcotte R., Massicotte S., Moussa R. and Fitzback J. A distributed watershed model compatible with remote sensing and GIS data: 1. Description of the model. *ASCE Journal of Hydrologic Engineering*, 6:91–99, 2001.

Foussereau X., Graham W., Aakpoji A., Destouni G. and Rao P. S. C. Solute transport through a heterogeneous coupled vadose-saturated zone system with temporally random rainfall. *Water Resources Research*, 37:1577–1588, 2001.

Fowler H. J. and Wilby R. L. Detecting changes in seasonal precipitation extremes using regional climate model projections: Implications for managing fluvial flood risk. *Water Resources Research*, 46:W03525, DOI:10.1029/2008WR007636, 2010.

Fox D. M., Le Bissonnais Y. and Bruand A. The effect of ponding depth on infiltration in a crusted surface depression. *Catena*, 32:87–100, 1998.

Franchini M. and Galeati G. Comparing several genetic algorithm schemes for the calibration of conceptual rainfall–runoff models. *Journal of Hydrology*, 42:357, 1997.

Franchini M., Helmlinger K. R., Foufoula-Georgiou E. and Todini E. Stochastic storm transposition coupled with rainfall–runoff modelling for estimation of exceedence probability of design floods. *Journal of Hydrology*, 175:511–532, 1996.

Franchini M. and Pacciani M. Comparative analysis of several conceptual rainfall–runoff models. *Journal of Hydrology*, 122:161–219, 1991.

Franks S. W. and Beven K. J. Bayesian estimation of uncertainty in land surface-atmosphere flux predictions. *Journal of Geophysical Research*, 102(D20):23991–23999, 1997a.

Franks S. W. and Beven K. J. Estimation of evapotranspiration at the landscape scale: A fuzzy disaggregation approach. *Water Resources Research*, 33:2929–2938, 1997b.

Franks S. W., Gineste P., Beven K. J. and Merot P. On constraining the predictions of a distributed model: The incorporation of fuzzy estimates of saturated areas into the calibration process. *Water Resources Research*, 34:787–797, 1998.

Fread D. L. Technique for implicit flood routing in rivers with tributaries. *Water Resources Research*, 9:918–926, 1973.

Fread D. L. Channel routing. In Anderson M. G. and Burt T. P., editors, *Hydrological Forecasting*, pages 437–503, John Wiley & Sons: Chichester, UK, 1985.

Freer J. *Uncertainty and Calibration of Conceptual Rainfall–Runoff Models*. PhD thesis, Lancaster University, Lancaster, UK, 1998.

Freer J., Beven K. J. and Ambroise B. Bayesian estimation of uncertainty in runoff prediction and the value of data: An application of the GLUE approach. *Water Resources Research*, 32:2161–2173, 1996.

Freer J., McDonnell J., Beven K. J., Brammer D., Burns D., Hooper R. P. and Kendal C. Topographic controls on subsurface stormflow at the hillslope scale for two hydrologically distinct small catchments. *Hydrological Processes*, 11:1347–1352, 1997.

Freeze R. A. Role of subsurface flow in generating surface runoff: 2. Upstream source areas. *Water Resources Research*, 8(5):1272–1283, 1972.

Freeze R. A. and Harlan R. L. Blueprint for a physically-based, digitally-simulated hydrologic response model. *Journal of Hydrology*, 9:237–258, 1969.

Funtowicz S. O. and Ravetz J. R. *Uncertainty and Quality in Science for Policy*. Kluwer Academic, Dordrecht, 1990.

Gallagher M. and Doherty J. Parameter estimation and uncertainty analysis for a watershed model. *Environmental Modelling & Software*, 22:1000–1020, 2007.

Gallart G., Latron J., Llorens P. and Beven K. J. Using internal catchment information to reduce the uncertainty of discharge and baseflow prediction. *Advances in Water Resources*, 30(4):808–823, 2007.

Gandolfi C., Bishetti G. B. and Whelan M. J. A simple triangular approximation of the area function for the calculation of network hydrological response. *Hydrological Processes*, 13:2639–2654, 1999.

Gash J. H. C. An analytical model of rainfall interception by forests. *Quarterly Journal of the Royal Meteorological Society*, 105:43–55, 1979.

Gash J. H. C. and Morton A. J. An application of the Rutter Model to the estimation of the interception loss from Thetford Forest. *Journal of Hydrology*, 38:49–58, 1978.

Gash J. H. C., Wright I. R. and Lloyd C. R. Comparative estimates of interception loss from three coniferous forests in Great Britain. *Journal of Hydrology*, 48:89–105, 1980.

Gassman P. W., Arnold J. G., Srinivasan R. and Reyes M. The worldwide use of the SWAT model: Technological driver, networking impacts, and simulation trends. *Transactions of the ASABE*, 50(4):1211–1, 2010.

Gassman P. W., Reyes M. R., Green C. H. and Arnold J. G. The Soil and Water Assessment Tool: Historical development, applications and future research directions. *Transactions of the ASABE*, 50:1211–1250, 2007.

Gaume E. and Gosset R. Over-parameterisation, a major obstacle to the use of artificial neural networks in hydrology? *Hydrology and Earth System Sciences*, 7:693–706, 2003.

Gelhar L. W., Welty C. and Rehfeldt K. R. A critical review of data on field-scale dispersion in aquifers. *Water Resources Research*, 28:1955–1974, 1992.

Genereux D. Quantifying uncertainty in tracer-based hydrograph separations. *Water Resources Research*, 34:915–919, 1998.

Genereux D. P., Hemond H. F. and Mulholland P. J. Spatial and temporal variability in streamflow generation on the west fork of Walker Branch Watershed. *Journal of Hydrology*, 142:137–166, 1993.

Gerke H. H., Germann P. F. and Nieber J. Preferential and unstable flow: From the pore to the catchment scale. *Vadose Zone Journal*, 9:207–212, 2010.

Germann P. and Hensel D. Poiseuille flow geometry inferred from velocities of wetting fronts in soils. *Vadose Zone Journal*, 5:867–876, 2006.

Germann P. F. Kinematic wave approach to infiltration and drainage into and from soil macropores. *Transactions of the ASABE*, 28:745–749, 1985.

Germann P. F. Macropores and hydrologic hillslope processes. In Anderson M. G. and Burt T. P., editors, *Process Studies in Hillslope Hydrology*, pages 327–364. John Wiley & Sons, Chichester, UK, 1990.

Germann P. F., Helbling A. and Vadilonga T. Rivulet approach to rates of preferential infiltration. *Vadose Zone Journal*, 6:207–220, 2007.

Gingold R. A. and Monaghan J. J. Smoothed particle hydrodynamics: Theory and application to non-spherical stars. *Monthly Notices of the Royal Astronomical Society*, 181:375–389, 1977.

Godsey S. E., Aas W., Clair T. A., de Wit H. A., Fernandez I. J., Kahl J. S., Malcolm I. A., Neal C., Neal M., Nelson S. J., Norton S. A., Palucis M. C., Skjelkvåle B. L., Soulsby C., Tetzlaff D. and Kirchner J. W. Generality of fractal 1/f scaling in catchment tracer time series, and its implications for catchment travel time distributions. *Hydrological Processes*, 24:1660–1671, 2010.

Gong L., Widen-Nilsson E., Halldin S. and Xu C.-Y. Large-scale runoff routing with an aggregated network-response function. *Journal of Hydrology*, 368:237–250, 2009.

Goodrich D. C., Lane L. J., Shillito R. M., Miller S. N., Syed K. H. and Woolhiser D. A. Linearity of basin response as a function of scale in a semiarid watershed. *Water Resources Research*, 33:2951–2965, 1997.

Goodrich D. C., Schmugge T. J., Jackson T. J., Unkrich C. L., Keefer T. O., Parry R., Bach L. B. and Amer S. A. Runoff simulation sensitivity to remotely sensed initial soil moisture content. *Water Resources Research*, 30:1393–1405, 1994.

Goodrich D. C., Woolhiser D. A. and Keefer T. O. Kinematic routing using finite elements on a triangular irregular network. *Water Resources Research*, 27:995–1003, 1991.

Götzinger J. and Bárdossy A. Comparison of four regionalisation methods for a distributed hydrological model. *Journal of Hydrology*, 333:374–384, 2007.

Götzinger J. and Bárdossy A. Generic error model for calibration and uncertainty estimation of hydrological models. *Water Resources Research*, 44:W00B07, DOI:10.1029/2007WR006691, 2008.

Grayson R. B., Blöschl G. and Moore I. D. Distributed parameter hydrologic modelling using vector elevation data: THALES and TAPES-C. In Singh V. P., editor, *Computer Models of Watershed Hydrology*, pages 669–696. Water Resource Publications, Highlands Ranch, CO, 1995.

Grayson R. B., Moore I. D. and McMahon T. A. Physically-based hydrologic modelling: 1. A terrain-based model for investigative purposes. *Water Resources Research*, 28:2639–2658, 1992a.

Grayson R. B., Moore I. D. and McMahon T. A. Physically-based hydrologic modelling: 2. Is the concept realistic? *Water Resources Research*, 28:2659–2666, 1992b.

Green C. H., Tomer M. D., Di Luzio M. and Arnold J. G. Hydrologic evaluation of the Soil and Water Assessment Tool for a large tile-drained watershed in Iowa. *Transactions of the ASABE*, 49:413–422, 2006.

Green H. M., Beven K. J., Buckley K. and Young P. C. *Mixing and Transport in the Environment*, pages 113–140. John Wiley & Sons, Chichester, UK, 1994.

Green W. H. and Ampt G. Studies of soil physics: Part 1. The flow of air and water through soils. *Journal of Agricultural Science*, 4:1–24, 1911.

Gregersen J. B., Gijsbers P. J. A. and Westen S. J. P. OpenMI: Open Modelling Interface. *Journal of Hydroinformatics*, 9:175–191, 2007.

Grove M., Harbor J. and Engel B. Composite vs. distributed curve numbers: Effects on estimates of storm runoff depths. *Journal of the American Water Resources Association*, 34:1015–1023, 1998.

Guinot V., Cappelaere B., Delenne C. and Ruelland D. Towards improved criteria for hydrological model calibration: Theoretical analysis of distance- and weak form-based functions. *Journal of Hydrology*, 401(1–2):1–13, 2011.

Güntner A. Improvement of global hydrological models using GRACE data. *Surveys in Geophysics*, 29:375–397, 2008.

Güntner A., Uhlenbrook S., Seibert J. and Leibundgut Ch. Multi-criterial validation of TOPMODEL in a mountainous catchment. *Hydrological Processes*, 13:1603–1620, 1999.

Gupta H. V., Sorooshian S. and Yapo P. O. Toward improved calibration of hydrologic models: Multiple and noncommensurable measures of information. *Water Resources Research*, 34:751–763, 1998.

Gupta H. V., Sorooshian S. and Yapo P. O. Status of automatic calibration for hydrologic models: Comparison with multilevel expert calibration. *ASCE Journal of Hydrologic Engineering*, 4:135, 1999.

Gupta S. C. and Larson W. E. Estimating soil water retention characteristics from particle size distribution, organic matter percent and bulk density. *Water Resources Research*, 15:1633–1635, 1979.

Gupta V. K. and Sorooshian S. Uniqueness and observability of conceptual rainfall–runoff model parameters: The percolation process examined. *Water Resources Research*, 19:269–276, 1983.

Gupta V. K. and Sorooshian S. The relationship between data and the precision of parameter estimates of hydrologic models. *Journal of Hydrology*, 81:57–77, 1985.

Gurtz J., Baltensweiler A. and Lang H. Spatially distributed hydrotope-based modelling of evapotranspiration and runoff in mountainous basins. *Hydrological Processes*, 13:2751–2768, 1999.

Hahn G. J. and Meeker W. Q. *Statistical Intervals: A guide for practitioners*. John Wiley & Sons, New York, 1991.

Hall F. R. Contributions of Robert E. Horton. In *History of Geophysics*, volume 3 of *AGU Monograph on the History of Geophysics*, pages 113–117. American Geophysical Union, Washington, DC, 1987.

Hall J., O'Connell P. E. and Ewen J. On not undermining the science: Coherence, validation and expertise. *Hydrological Processes*, 21:985–988, 2007.

Hall J. W. Probabilistic climate scenarios may misrepresent uncertainty and lead to bad adaptation decisions. *Hydrological Processes*, 21:1127–1129, 2007.

Hall J. W., Tarantola S., Bates P. D. and Horritt M. S. Distributed sensitivity analysis of flood inundation model calibration. *ASCE Journal of Hydrologic Engineering*, 131:117–126, 2005.

Hall R. L., Calder I. R., Gunawardena E. R. N. and Rosier P. T. W. Dependence of rainfall interception on drop size: 3. Implementation and comparative performance of the stochastic model using data from a tropical site in Sri Lanka. *Journal of Hydrology*, 185:389–407, 1996.

Halpern J. Y. *Reasoning about uncertainty*. MIT Press, Cambridge, MA., 2003.

Hamon W. R. Estimating potential evapo-transpiration. *ASCE Journal of the Hydraulics Division*, 87:107–120, 1961.

Han D., Kwong T. and Li S. Uncertainties in real-time flood forecasting with neural networks. *Hydrological Processes*, 21:223–228, 2007.

Hansen J. R., Refsgaard J. C., Hansen S. and Ernstsen V. Problems with heterogeneity in physically based agricultural catchment models. *Journal of Hydrology*, 342:1–16, 2007.

Hapuarachchi H. A. P., Wang Q. J. and Pagano T. C. A review of advances in flash flood forecasting. *Hydrological Processes*, 25:DOI: 10.1002/hyp.8040, 2011.

Harlin J. and Kung C.-S. Parameter uncertainty and simulation of design floods in Sweden. *Journal of Hydrology*, 137:209–230, 1992.

Harmel R. D., Cooper R. J., Slade R. M., Haney R. L. and Arnold J. G. Cumulative uncertainty in measured streamflow and water quality data for small watersheds. *Transactions of the ASABE*, 49:689–701, 2006.

Harris D., McDonnell J. J. and Rodhe A. Hydrograph separation using continuous open system isotope mixing. *Water Resources Research*, 31:DOI:10.1029/94WR01966, 1995.

Harris G. *Seeking Sustainability in an Age of Complexity*. Cambridge University Press, Cambridge, UK, 2007.

Harris G. P. and Heathwaite A. L. Inadmissible evidence: Knowledge and prediction in land and riverscapes. *Journal of Hydrology*, 304:3–19, 2005.

Hartman M. D., Baron J. S., Lammers R. B., Cline D. W., Band L. E., Liston G. E. and Tague C. Simulations of snow distribution and hydrology in a mountain basin. *Water Resources Research*, 35:1587–1603, 1999.

Hassanizadeh S. M. Derivation of basic equations of mass transport in porous media: Part 2. Generalized Darcy's law and Fick's law. *Advances in Water Resources*, 9:207–222, 1986.

Haverkamp R., Reggiani P., Ross P. J. and Parlange J.-Y. *Environmental Mechanics: Water Mass and Energy Transfer in the Biosphere*, pages 213–246. American Geophysical Union, Washington, DC, 2002.

He Z., Wu W. and Wang S. S. Y. Coupled finite-volume model for 2D surface and 3D subsurface flows. *ASCE Journal of Hydrologic Engineering*, 13:835–845, 2008.

Helton J. C. and Oberkampf W. L. Alternative representations of epistemic uncertainty. *Reliability Engineering and System Safety*, 85:1–10, 2004.

Henderson F. M. and Wooding R. A. Overland flow and groundwater flow from a steady rainfall of finite duration. *Journal of Geophysical Research*, 69:1531–1540, 1964.

Henriksen H. J., Troldborg L., Hojberg A. L. and Refsgaard J. C. Assessment of exploitable groundwater resources of Denmark by use of ensemble resource indicators and a numerical groundwater-surface water model. *Journal of Hydrology*, 348:224–240, 2008.

Heppner C. S. and Loague K. Characterizing long-term hydrologic-response and sediment-transport for the R-5 catchment. *Journal of Environmental Quality*, 37:2181–2191, 2007.

Heppner C. S., Loague K. and VanderKwaak J. E. Long-term InHM simulations of hydrologic response and sediment transport for the R-5 catchment. *Earth Surface Processes and Landforms*, 32:1273–1292, 2007.

Herr H. D. and Krzysztofowicz R. Bayesian ensemble forecast of river stages and ensemble size requirements. *Journal of Hydrology*, 387:151–164, 2010.

Herschy R. W. *Streamflow Measurement*. Chapman and Hall, London, 1995.

Heuvelmans G., Muys B. and Jan Feyen. Regionalisation of the parameters of a hydrological model: Comparison of linear regression models with artificial neural nets. *Journal of Hydrology*, 319:245–265, 2006.

Hewlett J. D. Comments on letters relating to "Role of subsurface flow in generating surface runoff: 2. Upstream source areas" by Allen Freeze R. *Water Resources Research*, 10:605–607, 1974.

Hewlett J. D. and Hibbert A. R. Factors affecting the response of small watersheds to precipitaion in humid areas. In Sopper W. E. and Lull H. W., editors, *Forest Hydrology*, pages 275–290. Pergamon, 1967.

Hill M. C. and Tiedeman C. R. *Effective groundwater model calibration*. John Wiley & Sons, Hoboken, NJ, 2007.

Hincapié I. and Germann P. Abstraction from infiltrating water content waves during weak viscous flow. *Vadose Zone Journal*, 8:891–901, 2009.

Hjelmfelt A. T. and Amerman C. R. The mathematical basin model of Merrill Bernard. *The International Association of Hydrological Sciences Publication No.*, 130:343–349, 1980.

Hjelmfelt A. T., Kramer L. A. and Burwell R. E. Curve numbers as random variables. In Singh V. P., editor, *Rainfall–Runoff Relationships*, pages 365–370. Water Resource Publications, 1982.

Hock R. Temperature index melt modelling in mountain areas. *Journal of Hydrology*, 282:104–115, 2003.

Holländer H. M., Blume T., Bormann H., Buytaert W., Chirico G. B., Exbrayat J. F., Gustafsson D., Hölzel H., Kraft P., Stamm C., Stoll S., Blöschl G. and Flühler H. Comparative predictions of discharge from an artificial catchment (Chicken Creek) using sparse data. *Hydrology and Earth System Sciences*, 13:2069–2094, 2009.

Hollenbeck K. J. and Jensen K. H. Experimental evidence of randomness and nonuniqueness in unsaturated outflow experiments designed for hydraulic parameter estimation. *Water Resources Research*, 34:595–602, 1998.

Holmes M. G. R., Young A. R., Gustard A. and Grew R. A region of influence approach to predicting flow duration curves within ungauged catchments. *Hydrology and Earth System Sciences*, 6:721–731, 2002.

Holtan H. N., England C. B., Lawless G. P. and Schumaker G. A. Moisture-tension data for selected soils on experimental watersheds. Technical Report ARS 41-144, USDA ARS, Agricultural Research Service, Beltsville, MD, 1968.

Holwill C. J. and Stewart J. B. Spatial variability of evapotranspiration derived from aircraft and ground based data. *Journal of Geophysical Research*, 97(D17):18673–18680, 1992.

Hooper R. P. Diagnostic tools for mixing models of stream water chemistry. *Water Resources Research*, 39:W01055, DOI:10.1029/2002WR001528, 2003.

Hooper R. P., Christophersen N. and Peters N. E. Modelling streamwater chemistry as a mixture of soil water end members: An application to the Panola mountain catchment, Georgia, USA. *Journal of Hydrology*, 116:321–343, 1990.

Hornberger G. M., Beven K. J., Cosby B. J. and Sappington D. E. Shenandoah watershed study: Calibration of a topography-based, variable contributing area hydrological model to a small forested catchment. *Water Resources Research*, 21:1841–1850, 1985.

Hornberger G. M. and Spear R. C. An approach to the preliminary analysis of environmental systems. *Journal of Environmental Management*, 12:7–18, 1981.

Horritt M. S. and Bates P. D. Predicting floodplain inundation: Raster-based modelling versus the finite-element approach. *Hydrological Processes*, 15:825–84, 2001.

Horritt M. S. and Bates P. D. Evaluation of 1D and 2D numerical models for predicting river flood inundation. *Journal of Hydrology*, 268:87–99, 2002.

Horritt M. S., Bates P. D., Fewtrell T. J., Mason D. C. and Wilson M. D. Modelling the hydraulics of the Carlisle 2005 flood event. *Proc. Inst. Civ. Eng., Water Management*, 163:273–281, 2010.

Horritt M. S., Di Baldassarre G., Bates P. D. and Brath A. Comparing the performance of a 2-D finite element and a 2-D finite volume model of floodplain inundation using airborne SAR imagery. *Hydrological Processes*, 21:2745–2759, 2007.

Horritt M. S., Mason D. C., Cobby D. M., Davenport I. J. and Bates P. D. Waterline mapping in flooded vegetation from airborne SAR imagery. *Remote Sensing of Environment*, 85:271–281, 2003.

Horton R. E. The role of infiltration in the hydrologic cycle. *Transactions of the American Geophysical Union*, 14:446–460, 1933.

Horton R. E. *Surface Runoff Phenomena: Part 1. Analysis of the Hydrograph*. Horton Hydrologic Laboratory Publication 101. Edward Bros., Ann Arbor, MI., 1935.

Horton R. E. Maximum ground-water levels. *Transactions of the American Geophysical Union*, 17:344–357, 1936.

Horton R. E. Determination of infiltration-capacity for large drainage basins. *Transactions, American Geophysical Union*, 18:371–385, 1937.

Horton R. E. The interpretation and application of runoff experiments with reference to soil erosion. *Proceedings of the Soil Science Society of America*, 3:340–349, 1938.

Horton R. E. An approach to the physical interpretation of infiltration capacity. *Proceedings of the Soil Science Society of America*, 5:399–417, 1940.

Horton R. E. Remarks on hydrologic terminology. *Transactions of the American Geophysical Union*, 23:479–482, 1942.

Horton R. E. Erosional development of streams and their drainage basins: Hydrophysical to quantitative morphology. *Bull. Geol. Soc. Am.*, 56:275–370, 1945.

Hosking J. R. M. and Clarke R. T. Rainfall–runoff relations derived from the probability theory of storage. *Water Resources Research*, 26:1455–1463, 1990.

Hosking J. R. M. and Wallis J. R. *Regional frequency analysis: An approach based on L-moments*. Cambridge University Press, 1997.

Hottelet C., Braun L., Leibundgut C. and Rieg A. Simulation of snowpack and discharge in an alpine karst basin. *The International Association of Hydrological Sciences Publication No.*, 218:249–260, 1993.

Houghton-Carr H. A. Assessment criteria for simple conceptual daily rainfall–runoff models. *Journal of Hydrology*, 44:237–263, 1999.

Houser P. R., Shuttleworth W. J., Famiglietti J. S., Gupta H. V., Syed K. H. and Goodrich D. C. Integration of soil moisture remote sensing and hydrologic modelling using data assimilation. *Water Resources Research*, 34:3405–3420, 1998.

Howson C. and Urbach P. *Scientific Reasoning: The Bayesian approach*. Open Court, Chicago, 2nd edition edition, 1993.

Hsu K.-L., Moradkhani H. and Sorooshian S. A sequential Bayesian approach for hydrologic model selection and prediction. *Water Resources Research*, 45:W00B12, DOI:10.1029/2008WR006824, 2009.

Huang M. and Liang X. On the assessment of the impact of reducing parameters and identification of parameter uncertainties for a hydrologic model with applications to ungauged basins. *Journal of Hydrology*, 320:37–61, 2006.

Huard D. and Mailhot A. Calibration of hydrological model GR2M using Bayesian uncertainty analysis. *Water Resources Research*, 44:W02424, DOI:10.1029/2007WR005949, 2008.

Huff D. D., O'Neill R. V., Emmanuel W. R., Elwood J. W. and Newbold J. D. Flow variability and hillslope hydrology. *Earth Surface Processes and Landforms*, 7:91–94, 1982.

Huggins L. F. and Monke E. J. A mathematical model for simulating the hydrologic response of a watershed. *Water Resources Research*, 4:529–539, 1968.

Hundecha Y. and Bárdossy A. Modeling of the effects of land use changes on the runoff generation of a river basin through parameter regionalization of a watershed model. *Journal of Hydrology*, 292:281–295, 2004.

Hundecha Y., Bárdossy A. and Theisen H.-W. Development of a fuzzy logic based rainfall–runoff model. *Hydrological Sciences Journal*, 46:363–377, 2001.

Hundecha Y., Ouarda T. B. M. J. and Bárdossy A. Regional estimation of parameters of a rainfall–runoff model at ungauged watersheds using the "spatial" structures of the parameters within a canonical physiographic-climatic space. *Water Resources Research*, 44:W01427, DOI:10.1029/2006WR005439, 2008.

Hunter N. M., Bates P. D., Horritt M. S. and Wilson M. D. Improved simulation of flood flows using storage cell models. *Proceedings of the ICE, Water Management*, 159:9–18, 2006.

Hunter N. M., Bates P. D., Neelz S., Pender G., Villanueva I., Wright N. G., Liang D., Falconer R. A., Lin B., Waller S., Crossley A. J. and Mason D. C. Benchmarking 2D hydraulic models for urban flood simulations. *Proc. Inst. Civ. Eng., Water Management*, 161:13–30, 2008.

Hursh C. R. and Brater E. F. Separating storm-hydographs from small drainage-areas into surface- and subsurface-flow. *Transactions, AGU*, 22:863–870, 1941.

Hwang T., Band L. and Hales T. C. Ecosystem processes at the watershed scale: Extending optimality theory from plot to catchment. *Water Resources Research*, 45:W11425, DOI:10.1029/2009wr007775, 2009.

Ibbitt R. P. and O'Donnell T. Fitting methods for conceptual catchment models. *ASCE Journal of the Hydraulics Division*, 97:1331–1342, 1971.

Ibbitt R. P. and O'Donnell T. Designing conceptual catchment models for automatic fitting methods. In *Mathematical Models in Hydrology*, The International Association of Hydrological Sciences Publication No. 101. 461–475, Wallingford, UK, 1974.

IH. *Flood Estimation Handbook (5 volumes)*. Institute of Hydrology, Wallingford, UK, 1999.

Iman R. L. and Conover W. J. A distribution-free approach to inducing rank correlation among input variables. *Communications in Statistics, Series B*, 11:311–334, 1982.

Imbeaux E. La Durance: Régime, crues et inondations. *Annales des Ponts et Chaussées, Mémoires et Documents*, 7e séries(Tome III):5–200, 1892.

Ingber A L. Simulated annealing: Practice versus theory. *J. Math. Comput. Modelling*, 18:29–57, 1993.

Iorgulescu I. and Beven K. J. Non-parametric direct mapping of rainfall–runoff relationships: An alternative approach to data analysis and modelling. *Water Resources Research*, 40:W08403, 10.1029/2004WR003094, 2004.

Iorgulescu I., Beven K. J. and Musy A. Data-based modelling of runoff and chemical tracer concentrations in the Haute-Mentue (Switzerland) research catchment. *Hydrological Processes*, 19:2557–2574, 2005.

Iorgulescu I., Beven K. J. and Musy A. Flow, mixing, and displacement in using a data-based hydrochemical model to predict conservative tracer data. *Water Resources Research*, 43:W03401, DOI:10.1029/2005WR004019, 2007.

Iorgulescu I. and Musy A. Generalisation of TOPMODEL for a power law transmissivity profile. *Hydrological Processes*, 11, 1997.

Ivanov V. Y., Bras R. L. and Vivoni E. R. Vegetation-hydrology dynamics in complex terrain of semiarid areas: 1. A mechanistic approach to modeling dynamic feedbacks. *Water Resources Research*, 44:W03429, DOI:10.1029/2006WR005588, 2008.

Ivanov V. Y., Vivoni E. R., Bras R. L. and Entekhabi D. Catchment hydrologic response with a fully distributed triangulated irregular network model. *Water Resources Research*, 40:W11102, DOI:10.1029/2004WR003218, 2004.

Jackson B. M., Wheater H. S., Mcintyre N. R., Chell J., Francis O. J., Frogbrook Z., Marshall M., Reynolds B. and Solloway I. The impact of upland land management on flooding: Insights from a multiscale experimental and modelling programme. *Journal of Flood Risk Management*, 1:71–80, 2008.

Jacquin A. and Shamseldin A. Y. Development of a possibilistic method for the evaluation of predictive uncertainty in rainfall–runoff modeling. *Water Resources Research*, 43:W04425, DOI:10.1029/2006WR005072, 2007.

Jain A., Sudheer K. P. and Srinivasulu S. Identification of physical processes inherent in artificial neural network rainfall runoff models. *Hydrological Processes*, 18:571–581, 2004.

Jain S. K., Storm B., Bathurst J. C., Refsgaard J. C. and Singh R. D. Application of the SHE to catchments in India: Part 2. Field experiments and simulation studies with the SHE on the Kolar subcatchment of the Narmada River. *Journal of Hydrology*, 140:25–47, 1992.

Jakeman A. J., Chen T. H., Post D. A., Hornberger G. M., Littlewood I. G. and Whitehead P. G. Assessing uncertainties in hydrological response to climate at large scale. In Wilkinson W. B. editor, *Macroscale Modelling of the Hydrosphere*, IAHS Publ. No. 214:37–47, Wallingford, UK, 1993b.

Jakeman A. J., Green T. R., Beavis S. G., Zhang L., Dietrich C. R. and Crapper P. F. Modelling upland and instream erosion, sediment and phosphorus transport in a large catchment. *Hydrological Processes*, 13:745–752, 1999.

Jakeman A. J. and Hornberger G. M. How much complexity is warranted in a rainfall–runoff model? *Water Resources Research*, 29(8):2637–2649, 1993.

Jakeman A. J., Littlewood I. G. and Whitehead P. G. Computation of the instantaneous unit hydrograph and identifiable component flows with application to two small upland catchments. *Journal of Hydrology*, 117:275–300, 1990.

Jakeman A. J., Littlewood I. G. and Whitehead P. G. An assessment of the dynamic response characteristics of streamflow in the Balquhidder Catchments. *Journal of Hydrology*, 145:337–355, 1993a.

James A. L., McDonnell J. J., Tromp van Meerveld I. and Peters N. E. Gypsies in the palace: Experimentalist's view on the use of 3-D physics-based simulation of hillslope hydrological response. *Hydrological Processes*, 24:3878–3893, 2010.

Jarvie H. P., Neal C., Withers P. J. A., Baker D. B., Richards R. P. and Sharpley A. N. Quantifying phosphorus retention and release in rivers and watersheds using extended end-member mixing analysis (E-EMMA). *Journal of Environmental Quality*, 40:492–504, 2011.

Jarvis A. J., Young P. C., Leedal D. T. and Chotai A. A robust sequential CO_2 emissions strategy based on optimal control of atmospheric CO_2 concentrations. *Climatic Change*, pages DOI 10.1007/s10584–007–9298–4, 2007.

Jarvis N. J., Jansson P.-E. and Dik P. E. Modelling water and solute transport in macroporous soil: 1. Model description and sensitivity analysis. *Journal of Soil Science*, 42:59–70, 1991.

Jarvis P. J. The interpretation of the variation in leaf water potential and stomatal conductance found in canopies in the field. *Philosophical Transactions of the Royal Society London B*, 273:593–610, 1976.

Jaynes D. B. *Soil Water Hysteresis: Models and Implications*, chapter 4, pages 93–126. John Wiley & Sons, Chichester, UK, 1990.

Jenny P., Lee S. H. and Tchelepi H. A. Multi-scale finite-volume method for elliptic problems in subsurface flow simulation. *Journal of Computational Physics*, 187:47–67, 2003.

Ji Z.-G. *Hydrodynamics and Water Quality*. Wiley-Interscience, Hoboken, NJ, 2008.

Joerin C., Beven K. J., Iorgulescu I. and Musy A. Uncertainty in hydrograph separations based on geochemical mixing models. *Journal of Hydrology*, 255:90–106, 2002.

Jones D. A. and Kay A. L. Uncertainty analysis for estimating flood frequencies for ungauged catchments using rainfall–runoff models. *Advances in Water Resources*, 30:1190–1204, 2007.

Jones J. A. A. Soil piping and catchment response. *Hydrological Processes*, 24:1548–1566, 2010.

Jones N. L., Wright S. G. and Maidment D. R. Watershed delineation with triangle-based terrain models. *ASCE Journal of the Hydraulics Division*, 110:1232–1251, 1990.

Jordan J.-P. Spatial and temporal variability of stormflow generation processes on a Swiss catchment. *Journal of Hydrology*, 357–382:153, 1994.

Jothityangkoon C., Sivapalan M. and Farmer D. L. Process controls of water balance variability in a large semi-arid catchment: Downward approach to hydrological model development. *Journal of Hydrology*, 254:174–198, 2001.

Juston J., Seibert J. and Johansson P. O. Temporal sampling strategies and uncertainty in calibrating a conceptual hydrological model for a small boreal catchment. *Hydrological Processes*, 23:3093–3109, 2009.

Kafle H. K. and Yamaguchi Y. Effects of topography on the spatial distribution of evapotranspiration over a complex terrain using two-source energy balance model with ASTER data. *Hydrological Processes*, 23:2295–2306, 2009.

Kalma J. D., McVicar T. R. and Matthew McCabe F. Estimating land surface evaporation: A review of methods using remotely sensed surface temperature data. *Surveys in Geophysics*, 29:421–469, DOI: 10.1007/s10712–008–9037–z, 2008.

Kannan N., White S. M., Worrall F. and Whelan M. J. Hydrological modelling of a small catchment using SWAT-2000 – ensuring correct flow partitioning for contaminant modelling. *Journal of Hydrology*, 334:64–72, 2007.

Kapangaziwiri E. and Hughes D. A. Towards revised physically based parameter estimation methods for the Pitman monthly rainfall–runoff model. *Water SA*, 34:183–192, 2008.

Kavetski D. and Clark M. P. Ancient numerical daemons of conceptual hydrological modeling: 2. Impact of time stepping schemes on model analysis and prediction. *Water Resources Research*, 46:W10511, DOI:10.1029/2009WR008896, 2010.

Kavetski D. and Kuczera G. Model smoothing strategies to remove microscale discontinuities and spurious secondary optima in objective functions in hydrological calibration. *Water Resources Research*, 43:W03411, DOI:10.1029/2006WR005195, 2007.

Kavetski D., Kuczera G. and Franks S. W. Calibration of conceptual hydrological models revisited: 1. Overcoming numerical artefacts. *Journal of Hydrology*, 320:173–186, 2006.

Kay A. L., Jones D. A., Crooks S. M., Calver A. and Reynard N. S. A comparison of three approaches to spatial generalization of rainfall–runoff models. *Hydrological Processes*, 20:3953–3973, 2006b.

Kay A. L., Jones D. A., Crooks S. M., Kjeldsen T. R. and Fung C. F. An investigation of site-similarity approaches to generalisation of a rainfall–runoff model. *Hydrology and Earth System Sciences*, 11:500–515, 2007.

Kelly K. S. and Krzysztofowicz R. A bivariate meta-Gaussian density for use in hydrology. *Stochastic Hydrology and Hydraulics*, 11:17–31, 1997.

Kelly K. S. and Krzysztofowicz R. Precipitation uncertainty processor for probabilistic river stage forecasting. *Water Resources Research*, 36:2643–2653, 2000.

Kendall C. and McDonnell J. J., editors. *Isotope Tracers in Catchment Hydrology*. Elsevier Science B. V., Amsterdam, 1998.

Kendall K. A., Shanley J. B. and McDonnell J. J. A hydrometric and geochemical approach to test the transmissivity feedback hypothesis during snowmelt. *Journal of Hydrology*, 219:188–205, 1999.

Kennedy M. C. and O'Hagan A. Bayesian calibration of mathematical models. *Journal of the Royal Statistical Society*, D63:425–450, 2001.

Keulegan G. H. Spatially varied discharge over a sloping plane. *Transactions of the American Geophysical Union*, 26:821–824, 1945.

Khan M. S. and Coulibaly P. Bayesian neural network for rainfall–runoff modeling. *Water Resources Research*, 42:W07409, DOI:10.1029/2005WR003971, 2006.

Kiem A. S., Franks S. W. and Kuczera G. Multi-decadal variability of flood risk. *Geophysical Research Letters*, 30:1035, DOI:10.1029/2002GL015992, 2003.

Kim J. G., Lee G. H., Lee J. S., Chon C. M., Kim T. H. and Ha K. Infiltration pattern in a regolith-fractured bedrock profile: Field observation of a dye stain pattern. *Hydrological Processes*, 20:241–250, 2006.

Kirchner J. W. A double paradox in catchment hydrology and geochemistry. *Hydrological Processes*, 17:871–874, 2003.

Kirchner J. W. Aliasing in 1/f noise spectra: Origins, consequences, and remedies. *Physical Review*, E 71:066110, 2005.

Kirchner J. W. Getting the right answers for the right reasons: Linking measurements, analyses, and models to advance the science of hydrology. *Water Resources Research*, 42:42, W03S04, DOI:10.1029/2005WR004362, 2006.

Kirchner J. W. Catchments as simple dynamical systems: Catchment characterization, rainfall–runoff modeling, and doing hydrology backward. *Water Resources Research*, 45:W02429, 2009.

Kirchner J. W., Feng X. and Neal C. Fractal stream chemistry and its implications for contaminant transport in catchments. *Nature*, 403(6769):524–527, 2000.

Kirchner J. W., Feng X. and Neal C. Catchment-scale advection and dispersion as a mechanism for fractal scaling in stream tracer concentrations. *Journal of Hydrology*, 254:82–101, 2001.

Kirchner J. W., Tetzlaff D. and Soulsby C. Comparing chloride and water isotopes as hydrological tracers in two Scottish catchments. *Hydrological Processes*, 24:1631–1645, 2010.

Kirkby M. J. Hydrograph modelling strategies. In Peel R., Chisholm M. and Haggett P. editors, *Processes in Physical and Human Geography*, chapter 3, pages 69–90. Heinemann, London, 1975.

Kirkby M. J. Tests of the random network model and its application to basin hydrology. *Earth Surface Processes and Landforms*, 1:197–212, 1976.

Kirkby M. J. Implications for sediment transport. In Kirkby M. J., editor, *Hillslope Hydrology*. John Wiley & Sons, Chichester, UK, 1978.

Kirkby M. J. Hillslope runoff processes and models. *Journal of Hydrology*, 100:315–340, 1988.

Kirkby M. J. *The Scientific Nature of Geomorphology: Proceedings of the 27th Binghamton Symposium in Geomorphology*, pages 257–272. John Wiley & Sons, New York, 1996.

Kirkby M. J. Topmodel: A personal view. *Hydrological Processes*, 11(9):1087–1097, 1997.

Kite G. W. Scaling of input data for hydrologic modelling. *Water Resources Research*, 31:2769–2781, 1995.
Kjeldsen T. R. The revitalised FSR/FEH rainfall–runoff method. Flood estimation handbook supplementary report no. 1, Centre for Ecology & Hydrology, Wallingford, UK, 2007.
Klaus J. and Zehe E. Modelling rapid flow response of a tile-drained field site using a 2D physically based model: Assessment of 'equifinal' model setups. *Hydrological Processes*, 24:1595–1609, 2010.
Kleidon A., Malhi Y. and Cox P. M. Maximum entropy production in environmental and ecological systems: Introduction. *Philosophical Transactions of the Royal Society London B*, 365:1297–1302, 2010.
Klemeš V. *Common sense and other heresies: Selected papers on hydrology and water resources engineering*. Canadian Water Resources Association, Cambridge, Ontario, 2000.
Klemeš V. Operational testing of hydrologic simulation models. *Journal of Hydrology*, 31:13–24, 1986.
Klepper O., Scholten H. and van de Kamer J. P. G. Prediction uncertainty in an ecological model of the Oosterschelde Estuary. *Journal of Forecasting*, 10:191–209, 1991.
Knight D. W., McGahey C., Lamb R. and Samuels P. *Practical Channel Hydraulics: Roughness, Conveyance and Afflux*. CRC Press, 2010.
Knight D. W., Yuen K. W. H. and Al-Hamid A. A. I. Boundary shear stress distributions in open channel flow. In Beven K. J., Chatwin P. C., and Millbank J. H., editors, *Mixing and Transport in the Environment*, pages 51–88, John Wiley & Sons, Chichester, UK, 1994.
Knight F. H. *Risk, Uncertainty and Profit*. Hart, Schaffner and Marx (reprinted by Beard Books, Washinton, DC, 2002), 1921.
Knisel W. G. and Williams J. R. Hydrology components of CREAMS and GLEAMS. In Singh V. P., editor, *Computer Models of Watershed Hydrology*, pages 1069–1114, Water Resource Publications: Highlands Ranch, CO, 1995.
Kokkonen T. S., Jakeman A. J., Young P. C. and Koivusalo H. J. Predicting daily flows in ungauged catchments: Model regionalization from catchment descriptors at the Coweeta Hydrologic Laboratory, North Carolina. *Hydrological Processes*, 17:2219–2238, 2003.
Konikow L. F. and Bredehoeft J. D. Groundwater models cannot be validated. *Advances in Water Resources*, 15:75–83, 1992.
Kool J. B., Parker J. C. and van Genuchten M. T. Parameter estimation for unsaturated flow and transport models: A review. *Journal of Hydrology*, 91:255–293, 1987.
Kothyari U. C. and Singh V. P. A multiple-input single-output model for flow forecasting. *Journal of Hydrology*, 220:12–26, 1999.
Kraft P., Vaché K. B., Frede H.-G. and Breuer L. A hydrological programming language extension for integrated catchment models. *Environmental Modelling & Software*, 26:828–830, DOI: 10.1016/j.envsoft.2010.12.009, 2011.
Krueger T., Freer J., Quinton J. N., Macleod C. J. A., Bilotta G. S., Brazier R. E., Butler P. and Haygarth P. M. Ensemble evaluation of hydrological model hypotheses. *Water Resources Research*, 46:W07516, DOI: 10.1029/2009WR007845, 2010.
Krzysztofowicz R. Bayesian theory of probabilistic forecasting via deterministic hydrologic model. *Water Resources Research*, 35:2739–2750, 1999.
Krzysztofowicz R. The case for probabilistic forecasting in hydrology. *Journal of Hydrology*, 249:2–9, DOI:10.1016/S0022–1694(01)00420–6, 2001a.
Krzysztofowicz R. Integrator of uncertainties for probabilistic river stage forecasting: Precipitation-dependent model. *Journal of Hydrology*, 249:69–85, 2001b.
Krzysztofowicz R. Bayesian system for probabilistic river stage forecasting. *Journal of Hydrology*, 268:16–40, 2002.
Krzysztofowicz R. and Herr H. D. Hydrologic uncertainty processor for probabilistic river stage forecasting: Precipitation-dependent model. *Journal of Hydrology*, 249:46–68, 2001.
Krzysztofowicz R. and Kelly K. S. Hydrologic uncertainty processor for probabilistic river stage forecasting. *Water Resources Research*, 36:3265–3277, 2000.
Kuczera G. Improved parameter inference in catchment models: 1. Evaluating parameter uncertainty. *Water Resources Research*, 19:1151–1162, 1983a.
Kuczera G. Improved parameter inference in catchment models: 2. Combining different kinds of hydrologic data and testing their compatibility. *Water Resources Research*, 19:1163–1172, 1983b.

Kuczera G. Efficient subspace probabilistic parameter optimisation for catchment models. *Water Resources Research*, 33(1):177–185, 1997.

Kuczera G., Kavetski D., Franks S. and Thyer M. Towards a Bayesian total error analysis of conceptual rainfall–runoff models: Characterising model error using storm-dependent parameters. *Journal of Hydrology*, 331:161–177, 2006.

Kuczera G. and Parent E. Monte Carlo assessment of parameter uncertainty in conceptual catchment models: The Metropolis algorithm. *Journal of Hydrology*, 211:69–85, 1998.

Kuczera G., Kavetski D., Renard B. and Thyer M. A limited-memory acceleration strategy for MCMC sampling in hierarchical Bayesian calibration of hydrological models. *Water Resources Research*, 46:W07602, DOI:10.1029/2009WR008985, 2010b.

Kuczera G., Renard B., Thyer M. and Kavetski D. There are no hydrological monsters, just models and observations with large uncertainties! *Hydrological Sciences Journal*, 55:980–991, 2010a.

Kull D. W. and Feldman A. D. Evolution of Clark's unit graph method to spatially distributed runoff. *ASCE Journal of Hydrologic Engineering*, 3:9, 1998.

Kumar P. Variability, feedback, and cooperative process dynamics: Elements of a unifying hydrologic theory. *Geogr. Compass*, 1:1338–1360, DOI:10.1111/j.1749–8198.2007.00068.x, 2007.

Kumara M., Duffy C. J. and Salvage K. M. A second-order accurate, finite-volume-based, integrated hydrologic modeling (FIHM) framework for simulation of surface and subsurface flow. *Vadose Zone Journal*, 8:873–890, 2009.

Kundzewicz Z. W., Graczyk D., Maurer T., Pińskwar I., Radziejewski M., Svensson C. and Szwed M. Trend detection in river flow series: 1. Annual maximum flow. *Journal of Hydrology*, 50(5):797–810, 2005.

Kuo W.-L., Steenhuis T. S., McCulloch C. E., Mohler C. L., Weinstein D. A., DeGloria S. D. and Swaney D. P. Effect of grid size on runoff and soil moisture for a variable-source-area hydrology model. *Water Resources Research*, 35:3419–3428, 1999.

Kustas W. P. and Goodrich D. C. Preface to the Monsoon'90 issue. *Water Resources Research*, 30:1211–1225, 1994.

Kustas W. P., Rango A. and Uijlenhoet R. A simple energy budget algorithm for the snowmelt runoff model. *Water Resources Research*, 30:1515–1527, 1994.

Lakshmi V. The role of satellite remote sensing in the prediction of ungauged basins. *Hydrological Processes*, 18:1029–1034, 2004.

Lamb R. Calibration of a conceptual rainfall–runoff model for flood frequency estimation by continuous simulation. *Water Resources Research*, 35:3103–3114, 1999

Lamb R. and Beven K. J. Using interactive recession curve analysis to specify a general catchment storage model. *Hydrology and Earth System Sciences*, 1:101–113, 1997.

Lamb R., Beven K. J. and Myrabo S. Discharge and water table predictions using a generalised TOPMODEL formulation. *Hydrological Processes*, 11:1145–1168, 1997.

Lamb R., Beven K. J. and Myrabo S. A generalised topographic-soils hydrological index. In Richards K. S., Lane S. N. and Chandler J. H., editors, *Landform Monitoring, Modelling and Analysis*, pages 263–278. John Wiley & Sons, Chichester, UK, 1998a.

Lamb R., Beven K. J. and Myrabo S. Use of spatially distributed water table observations to constrain uncertainty in a rainfall–runoff model. *Advances in Water Resources*, 22:305–317, 1998b.

Lamb R., Crossley A. and Waller S. A fast two-dimensional floodplain inundation model. *Proceedings of the ICE, Water Management*, 162:363–370, 2009.

Lamb R. and Kay A. L. Confidence intervals for a spatially generalised continuous simulation flood frequency model for Great Britain. *Water Resources Research*, 40:W07501, 2004.

Lambert A. O. A comprehensive rainfall–runoff model for an upland catchment area. *Journal of the Institute of Water Engineering*, 23:231–238, 1969.

Lambert A. O. Catchment models based on ISO functions. *Journal of the Institute of Water Engineering*, 26:413–422, 1972.

Lane S., Odoni N., Landström C., Whatmore S. J., Ward N. and Bradley S. Doing flood risk science differently: An experiment in radical scientific method. *Transactions of the Institute of British Geographers*, 36:15–36, 2011.

Lane S. J. Hydraulic modelling in hydrology and geomorphology: A review of high resolution approaches. *Hydrological Processes*, 12:1131–1150, 1998.

Lane S. N., Bradbrook K. F., Richards K. S., Biron P. A. and Roy A. G. The application of computational fluid dynamics to natural river channels: Three-dimensional versus two-dimensional approaches. *Geomorphology*, 29:1–20, 1999.

Lange J., Leibundgut C., Greenbaum N. and Schick A. P. A noncalibrated rainfall–runoff model for large, arid catchments. *Water Resources Research*, 35:2161–2172, 1999.

Lassabatère L., Angulo-Jaramillo R., Soria Ugalde J. M., Cuenca R., Braud I. and Haverkamp R. Beerkan estimation of soil transfer parameters through infiltration experiments – BEST. *Soil Science Society of America Journal*, 70:521–532, 2006.

Leavesley G. H., Markstrom S. L., Restrepo P. J. and Viger R. J. A modular approach for addressing model design, scale, and parameter. *Hydrological Processes*, 16:173–187, DOI:10.1002/hyp.344, 2002.

Leavesley G. H. and Stannard L. G. The precipitation-runoff modelling system (PRMS). In Singh V. P., editor, *Computer Models of Watershed Hydrology*, pages 281–310. Water Resource Publications, Highlands Ranch, CO, 1995.

LeDrew E. A diagnostic examination of a complementary relationship between actual and potential evapotranspiration. *Journal of Applied Meteorology*, 18:495–501, 1979.

Lee H., Zehe E. and Sivapalan M. Predictions of rainfall–runoff response and soil moisture dynamics in a microscale catchment using the CREW model. *Hydrology and Earth System Sciences*, 11:819–849, 2007.

Leedal D. T., Beven K. J., Young P. C. and Romanowicz R. Data assimilation and adaptive real-time forecasting of water levels in the River Eden catchment, UK. In Samuels P., Huntingdon S., Allsop W. and Harrop J., editors, *Flood Risk Management: Research and Practice*. CRC Press, Balkema, 2008.

Leedal D. T., Neal J., Beven K., Young P. and Bates P. Visualization approaches for communicating real-time flood forecasting level and inundation information. *Journal of Flood Risk Management*, 3:140–150, 2010.

Lees M., Young P. C., Ferguson S., Beven K. J. and Burns J. An adaptive flood warning scheme for the River Nith at Dumfries. In White W. R. and Watts J., editors, *River Flood Hydraulics*, pages 65–75. John Wiley & Sons, Chichester, UK, 1994.

Leijnse H., Uijlenhoet R. and Stricker J. N. M. Rainfall measurement using radio links from cellular communication networks. *Water Resources Research*, 43:W03201, 2007.

Leijnse H., Uijlenhoet R. and Stricker J. N. M. Microwave link rainfall estimation: Effects of link length and frequency, temporal sampling, power resolution, and wet antenna attenuation. *Advances in Water Resources*, 31:1481–1493, 2008.

Leith N. A. and Chandler R. E. A framework for interpreting climate model outputs. *Journal of the Royal Statistical Society C (Applied Statistics)*, 59:279–296, 2010.

Lek S., Dimopoulos I., Derraz M. and El Ghachtoul Y. Modélisation de la relation pluie-débit à l'aide des réseaux de neurones artificiels. *Revue des Sciences de l'Eau*, 3:319–331, 1996.

LeVeque R. *Finite Volume Methods for Hyperbolic Problems*. Cambridge University Press, Cambridge, UK, 2002.

Lhomme J. P. and Giulioni L. Comments on some articles about the complementary relationship. *Journal of Hydrology*, 323:1–3, 2006.

Li H., Zhang Y., Chiew F. H. S. and Xu S. Predicting runoff in ungauged catchments by using Xinanjiang model with MODIS leaf area index. *Journal of Hydrology*, 370:155–162, 2009.

Li M., Shao Q., Zhang L. and Chiew F. H. S. A new regionalization approach and its application to predict flow duration curve in ungauged basins. *Journal of Hydrology*, 389:137–145, 2010.

Li R. M., Simons D. B. and Stevens M. A. Nonlinear kinematic wave approximation for water routing. *Water Resources Research*, 11:245–252, 1975.

Liang G. C., O'Connor K. M. and Kachroo R. K. A multiple-input, single-output, variable gain-factor model. *Journal of Hydrology*, 155:185–198, 1994a.

Liang X., Lettenmaier D. P., Wood E. F. and Burges S. J. A simple hydrologically-based model of land surface water and energy fluxes for general circulation models. *Journal of Geophysical Research*, 99:14415–14428, 1994b.

Liang X., Wood E. F. and Lettenmaier D. P. Surface soil moisture parameterisation of the VIC-2L model: Evaluation and modifications. *Global Planetary Change*, 13:195–206, 1996.

Liang X. and Xie Z. A new surface runoff parameterization with subgrid-scale soil heterogeneity for land surface models. *Advances in Water Resources*, 24:1173–1193, 2001.

Lighthill M. H. and Whitham G. B. On kinematic waves: 1. Flood movement in long rivers. *Proceedings of the Royal Society London Series A*, 229:281–316, 1955.

Lin Z. and Radcliffe D. E. Automatic calibration and predictive uncertainty analysis of a semidistributed watershed model. *Vadose Zone Journal*, 5:248–260, 2006.

Linden D. R. and Dixon R. M. Infiltration and water table effects of soil air pressure under border irrigation. *Proceedings of the Soil Science Society of America*, 37:95–98, 1973.

Lindgren G. A., Destouni G. and Miller A. V. Solute transport through the integrated groundwater-stream system of a catchment. *Water Resources Research*, 40:W03511. DOI:10.1029/ 2003WR00276., 2004.

Lindström G., Pers C., Rosberg J., Strömqvist J. and Arheimer B. Development and testing of the HYPE (hydrological predictions for the environment) water quality model for different spatial scales. *Hydrology Research*, 41:295–319, 2010.

Linsley R. K., Kohler M. A. and Paulhus J. L. H. *Applied Hydrology*. McGraw-Hill, New York, 1949.

Liston G. E., Sud Y. C., and Wood E. F. Evaluating GCM land surface hydrology parameterisations by computing discharges using a runoff routing model: Application to the Mississippi basin. *Journal of Applied Meteorology*, 33:394–405, 1994.

Littlewood I. G. Improved unit hydrograph characterisation of the daily flow regime (including low flows) for the River Teifi, Wales: Towards better rainfall-streamflow models for regionalisation. *Hydrology and Earth System Sciences*, 6:899–911, 2002.

Liu Y., Freer J. E., Beven K. J. and Matgen P. Towards a limits of acceptability approach to the calibration of hydrological models: Extending observation error. *Journal of Hydrology*, 367:93–103, DOI:10.1016/ j.jhydrol.2009.01.016, 2009.

Liu Y. and Gupta H. Uncertainty in hydrologic modeling: Towards an integrated data assimilation framework. *Water Resources Research*, 43:W07401, DOI:10.1029/2006WR005756, 2007.

Liu Z., Martina M. L. V. and Todini E. Flood forecasting using a fully distributed model: Application of the TOPKAPI model to the Upper Xixian catchment. *Hydrology and Earth System Sciences*, 9:347–364, 2005.

Liu Z. and Todini E. Towards a comprehensive physically-based rainfall–runoff model. *Hydrology and Earth System Sciences*, 6:859–881, 2002.

Llorens P., Poch R., Latron J. and Gallart F. Rainfall interception by a *pinus sylvestris* forest patch overgrown in a Mediterranean mountainous abandoned area: 1. Monitoring design and results down to the event scale. *Journal of Hydrology*, 199:331–345, 1997.

Lloyd C. R., Gash J. H. C., Shuttleworth W. J. and de O. Marques F. A. The measurement and modelling of rainfall interception by Amazonian rain forest. *Agricultural and Forest Meteorology*, 43:277–294, 1988.

Loague K. R-5 revisited: 2. Reevaluation of a quasi-physically based rainfall–runoff model with supplemental information. *Water Resources Research*, 26:973–987, 1990.

Loague K. Soil water content at R-5: Part 2. Impact of antecedent conditions on rainfall–runoff simulations. *Journal of Hydrology*, 139:253–262, 1992.

Loague K., editor. *Benchmark Papers in Hydrology: Rainfall–Runoff Modelling*. IAHS Press, Wallingford, UK, 2010.

Loague K., Gander G. E., VanderKwaak J. E., Abrams R. H. and Kyriakidis P. C. Simulating hydrologic response for the R-5 catchment: A never ending story. *Journal of Floodplain Management*, 1:57–83, 2000.

Loague K., Heppner C. S., Abrams R. H., Carr A. E., VanderKwaak J. E. and Ebel B. A. Further testing of the integrated hydrology model (InHM): Event-based simulations for a small rangeland catchment located near Chickasha, Oklahoma. *Hydrological Processes*, 19:1375–1395, 2005.

Loague K. and VanderKwaak J. E. Simulating hydrological response for the R-5 catchment: Comparison of two models and the impact of the roads. *Hydrological Processes*, 16:1015–1032, 2002.

Loague K. and VanderKwaak J. E. Physics-based hydrological response simulation: Platinum bridge, 1958 Edsel, or useful tool? *Hydrological Processes*, 18:2949–2956, 2004.

Loague K. M. and Freeze R. A. A comparison of rainfall–runoff modelling techniques on small upland catchments. *Water Resources Research*, 21:229–248, 1985.

Loague K. M. and Gander G. A. R-5 revisited: Spatial variability of infiltration on a small rangeland watershed. *Water Resources Research*, 26:957–971, 1990.

Loague K. M. and Kyriakidis P. C. Spatial and temporal variability in the R-5 infiltration data set: Déjà vu and rainfall–runoff simulations. *Water Resources Research*, 33:2883–2896, 1997.

Lohmann D., Lettenmaier D. P., Liang X., Wood E. F. and Boone A. *et al*. The project for intercomparison of land-surface parameterisation schemes (PILPS) Phase 2c: Red-Arkansas River basin experiment: 3. Spatial and temporal analysis of water fluxes. *Global Planetary Change*, 19:161–198, 1998c.

Lohmann D., Raschke E., Nijssen B. and Lettenmaier D. P. Regional scale hydrology: I. Formulation of the VIC-2L model coupled to a routing model. *Journal of Hydrology*, 43:131–141, 1998a.

Lohmann D., Raschke E., Nijssen B. and Lettenmaier D. P. Regional scale hydrology: II. Application of the VIC-2L model to the Weser River, Germany. *Journal of Hydrology*, 43:143, 1998b.

Lopes V. L. and Canfield H. E. Effects of watershed representation on runoff and sediment-yield modeling. *Journal of the American Water Resources Association*, 40:311–320, 2004.

Lu J., Sun G., McNulty S. G. and Amatya D. M. A comparison of six potential evapotranspiration methods for regional use in the southeastern United States. *Journal of the American Water Resources Association*, pages 621–633, 2005.

Luxmoore R. J. and Sharma M. L. Runoff responses to soil heterogeneity: Experimental and simulation comparisons for two contrasting watersheds. *Water Resources Research*, 16:675–684, 1980.

Lyon S. W., Desilets S. L. E. and Troch P. A. A tale of two isotopes: Differences in hydrograph separation for a runoff event when using δD versus δ^{18}O. *Hydrological Processes*, 23:2095–2101, 2009.

Lyon S. W., Laudon H., Seibert J., Mörth M., Tetzlaff D. and Bishop K. H. Controls on snowmelt water mean transit times in northern boreal catchments. *Hydrological Processes*, 24:1672–1684, 2010.

Lyon S. W., Walter M. T., Gerard-Marchant P. and Steenhuis T. S. Using a topographic index to distribute variable source area runoff predicted with the SCS curve-number equation. *Hydrological Processes*, 18:2757–2771, 2004.

Maidment D. R. *Arc Hydro: GIS for Water Resources*. MediaCompass LLC, Woodstock, GA, 2002.

Maki M., Iwanami K., Misumi R., Park S.-G., Moriwaki H., Maruyama K.-I., Watabe I., Lee D.-I., Jang M., Kim H.-K., Bringi V. N. and Uyeda H. Semi-operational rainfall observations with X-band multi-parameter radar. *Atmospheric Science Letters*, 6:12–18, 2005.

Maloszewski P. and Zuber A. Principles and practice of calibration and validation of mathematical models for the interpretation of environmental tracer data. *Advances in Water Resources*, 16:173–190, 1993.

Manning L. J., Hall J. W., Fowler H. J., Kilsby C. G. and Tebaldi C. Using probabilistic climate change information from a multimodel ensemble for water resources assessment. *Water Resources Research*, 45:W11411, DOI:10.1029/2007WR006674, 2009.

Mantovan P. and Todini E. Hydrological forecasting uncertainty assessment: Incoherence of the GLUE methodology. *Journal of Hydrology*, 330:368–381, 2006.

Marks D. and Dozier J. Climate and energy exchange at the snow surface in the alpine region of the Sierra Nevada: 2. Snow cover energy balance. *Water Resources Research*, 28:3043–3054, 1992.

Markstrom S., Niswonger R., Regan R., Prudic D. E. and Barlow P. GSFLOW: Coupled ground-water and surface-water flow model based on the integration of the precipitation-runoff modeling system (PRMS) and the modular ground-water flow model (MODFLOW-2005). Techniques and Methods 6-D1, US Geological Survey, Reston, VA, 2008.

Marshall M. R., Francis O. J., Frogbrook Z. L., Jackson B. M., McIntyre N., Reynolds B., Solloway I. and Wheater H. S. The impact of upland land management on flooding: Results from an improved pasture hillslope. *Hydrological Processes*, 23:464–475, 2009.

Martina M. L. V., Todini E. and Liu Z. Preserving the dominant physical processes in a lumped hydrological model. *Journal of Hydrology*, 399:121–131, 2011.

Martinec J. and Rango A. Areal distribution of snow water equivalent evaluated by snow cover monitoring. *Water Resources Research*, 17:1480–1488, 1981.

Matrosov S. Y., Clark K. A., Martner B. E. and Tokay A. X-band polarimetric radar measurements of rainfall. *Journal of Applied Meteorology*, 41:941–952, 2010.

Matsumoto M. and Nishimura T. Mersenne Twister: A 623-dimensionally equidistributed uniform pseudo-random number generator. *ACM Transactions on Modeling and Computer Simulation*, 8:3, 1998.

Mayes W. M., Walsh C. L., Bathurst J. C., Kilsby C. G., Quinn P. F., Wilkinson M. E., Daugherty A. J. and O'Connell P. E. Monitoring a flood event in a densely instrumented catchment, the upper Eden, Cumbria, UK. *CIWEM Water and Environment Journal*, 20:217–226, 2006.

McAneny K. J., Green A. E. and Astill M. S. Large aperture scintillometry: The homogeneous case. *Agricultural and Forest Meteorology*, 76:149–162, 1995.

McBratney A. B., Minasny B., Cattle S. R. and Vervoort R. W. From pedotransfer functions to soil inference systems. *Geoderma*, 109:41–73, 2002.

McCuen R. H. The role of sensitivity analysis in hydrologic modelling. *Journal of Hydrology*, 18:37–53, 1973.

McCuen R. H. *A Guide to Hydrologic Analysis Using SCS Methods*. Prentice-Hall, Englewood Cliffs, NJ, 1982.

McCuen R. H., Knight Z. and Cutter A. G. Evaluation of the Nash–Sutcliffe efficiency index. *ASCE Journal of Hydrologic Engineering*, 11:597–603, 2006.

McDonnell J. J., Bonell M., Stewart M. and Pearce A. Deuterium variations in storm rainfall: Implications for stream hydrograph separation. *Water Resources Research*, 26:DOI:10.1029/WR026i003p00455, 1990.

McDonnell J. J., Freer J., Hooper R., Kendall C., Burns D., Beven K. J. and Peters J. New method developed for studying flow on hillslopes. *EOS, Transactions AGU*, 77(47):465/472, 1996.

McDonnell J. J., McGuire K., Aggarwal P., Beven K. J., Biondi D., Destouni G., Dunn S., James A., Kirchner J. and Kraft, P. *et al.*, How old is streamwater? Open questions in catchment transit time conceptualization, modelling and analysis. *Hydrological Processes*, 24:1745–1754, 2010.

McGlynn B. L., McDonnell J. J. and Brammer D. D. A review of the evolving perceptual model of hillslope pathways at the Maimai catchment. *Journal of Hydrology*, 257:1–26, 2002.

McGuire K. J. and McDonnell J. J. A review and evaluation of catchment transit time modeling. *Journal of Hydrology*, 330:543–563,DOI: 10.1026/j.hydrol.2006.04.020, 2006.

McGuire K. J., McDonnell J. J., Weiler M., Kendall C., Welker J. M., McGlynn B. L. and Seibert J. The role of topography on catchment-scale water residence time. *Water Resources Research*, 41:W05002. DOI:10.1029/2004WR00365., 2005.

McGuire K. J., Weiler M. and McDonnell J. J. Integrating tracer experiments with modeling to assess runoff processes and water transit times. *Advances in Water Resources*, 30:824–837, 2007.

McIntyre N., Lee H., Wheater H., Young A. and Wagener T. Ensemble predictions of runoff in ungauged catchments. *Water Resources Research*, 41:W12434, DOI:10.1029/2005WR004289, 2005.

McIntyre N. and Marshall M. Identification of rural land management signals in runoff response. *Hydrological Processes*, 24:3521–3534, 2010.

McIntyre N., Young P., Orellana B., Marshall M., Reynolds B. and Wheater H. Identification of nonlinearity in rainfall-flow response using data-based mechanistic modeling. *Water Resources Research*, 47:W03515, DOI:10.1029/2010WR009851, 2011.

McIntyre N. R. and Wheater H. S. Calibration of an in-river phosphorus model: Prior evaluation of data needs and model uncertainty purchase. *Journal of Hydrology*, 290:100–116, 2004.

McMillan H. and Clark M. Rainfall–runoff model calibration using informal likelihood measures within a Markov Chain Monte Carlo sampling scheme. *Water Resources Research*, 45:W04418, DOI:10.1029/2008WR007288, 2009.

McMillan H., Freer J., Pappenberger F., Krueger T. and Clark M. Impacts of uncertain river flow data on rainfall–runoff model calibration and discharge predictions. *Hydrological Processes*, 24:1270–1284, 2010.

Melching C. S. Reliability estimation. In Singh V. P., editor, *Computer Models of Watershed Hydrology*, chapter 3. Water Resources Publications, Highlands Ranch, CO, 1995.

Meng H., Green T. R., Salas J. D. and Ahuja L. R. Development and testing of a terrain-based hydrologic model for spatial Hortonian infiltration and runoff/on. *Environmental Modelling & Software*, 23(6):794–812, 2008.

Merz R. and Blöschl G. Regionalization of catchment model parameters. *Journal of Hydrology*, 287:95–123, 2004.

Mesa O. J. and Mifflin E. R. On the relative role of hillslope and network geometry in hydrologic response. In Gupta V. K., Rodriguez-Iturbe I. and Wood E. F., editors, *Scale Problems in Hydrology*, Pages 1–17, Reidel, Dordrecht, 1986.

Michaud J. and Sorooshian S. Comparison of simple versus complex distributed runoff models on a midsized semiarid watershed. *Water Resources Research*, 30:593–606, 1994.

Michel C., Andréassian V. and Perrin C. Soil conservation service curve number method: How to mend a wrong soil moisture accounting procedure? *Water Resources Research*, 41:W02011, DOI:10.1029/2004WR003191, 2005.

Miller E. E. and Miller R. D. Physical theory for capillary flow phenomena. *Journal of Applied Physics*, 27:324–332, 1956.

Milly P. C. D. A refinement of the combination equations for evaporation. *Surveys in Geophysics*, 12:145–154, 1991.

Milly P. C. D., Betancourt J., Falkenmark M., Hirsch R. M., Kundzewicz Z. W., Lettenmaier D. P. and Stouffer R. J. Stationarity is dead: Whither water management? *Science*, 319:573–574, 2008.

Minns A. W. and Hall M. J. Artificial neural networks as rainfall–runoff models. *Journal of Hydrology*, 41:399–417, 1996.

Minshall N. E. Predicting storm runoff on small experimental watersheds. *ASCE Journal of the Hydraulics Division*, 86 (HY8):17–38, 1960.

Mishra S. K. and Parker J. C. Parameter estimation for coupled unsaturated flow and transport. *Water Resources Research*, 25:385–396, 1989.

Mishra S. K. and Singh V. P. Another look at SCS-CN method. *ASCE Journal of Hydrologic Engineering*, 4:257–264, 1999.

Mitchell K. M. and DeWalle D. R. Application of the snowmelt runoff model using multiple-parameter landscape zones on the Towanda Creek basin, Pennsylvania. *Journal of the American Water Resources Association*, 34:335–346, 1998.

Mitchell S., Beven K. J., Freer J. and Law B. Processes influencing model-data mismatch in drought-stressed, fire-disturbed, eddy flux sites. *Journal of Geophysical Research*, 116, 2011.

Mitchell S., Freer J. and Beven K. J. Multiple sources of predictive uncertainty in modeled estimates of net ecosystem CO_2 exchange. *Ecological Modelling*, 220:3259–3270, DOI:10.1016/j.ecolmodel.2009.08.021, 2009.

Mockus V. *Estimation of Total (and Peak Rates of) Surface Runoff for Individual Storms: Exhibit A, Appendix B, Interim Survey Report, Grand (Neosho) River Watershed*. US Department of Agriculture, Washington DC, 1949.

Mohanty B. P. Scaling hydraulic properties of a macroporous soil. *Water Resources Research*, 35:1927–1931, 1999.

Mohanty B. P., Bowman R. S., Hendrickx J. M. H., Simunek J. and van Genuchten M. T. Preferential transport of nitrate to a tile drain in an intermittent-flood-irrigated field: Model development and experimental evaluation. *Water Resources Research*, 34:1061–1076, 1998.

Mohanty B. P. et al. Piecewise hydraulic functions for preferential flow. *Water Resources Research*, 33:2049, 1997.

Monaghan J. J. Simulating free surface flows with SPH. *Journal of Computational Physics*, 110:399–406, 1994.

Montanari A. What do we mean by 'uncertainty'? The need for a consistent wording about uncertainty assessment in hydrology. *Hydrological Processes*, 21:841–845, 2007.

Montanari A. and Brath A. A stochastic approach for assessing the uncertainty of rainfall–runoff simulations. *Water Resources Research*, 40:W01106, DOI:10.1029/2003WR002540, 2004.

Montanari A. and Grossi G. Estimating the uncertainty of hydrological forecasts: A statistical approach. *Water Resources Research*, 44:W00B08, DOI:10.1029/ 2008WR006897, 2008.

Montanari A. and Toth E. Calibration of hydrological models in the spectral domain: An opportunity for scarcely gauged basins? *Water Resources Research*, 43:W05434, DOI:10.1029/2006WR005184, 2007.

Monteith J. Accommodation between transpiring vegetation and the convective boundary layer. *Journal of Hydrology*, 166:251–263, 1995a.

Monteith J. A reinterpretation of stomatal responses to humidity. *Plant, Cell and Environment*, 18:357–364, 1995b.

Monteith J. L. Evaporation and environment. In *The State and Movement of Water in Living Organisms*, Proc. 19th Symposium, Soc. Exper. Biol., pages 205–233, Cambridge University Press, London, 1965.

Montgomery D. R., Dietrich W. E. and Heffner J. T. Piezometric response in shallow bedrock at CB1: Implication for runoff generation and landsliding. *Water Resources Research*, 38:W01274, DOI:10.1029/2002WR001429, 2002.

Moore I. D., Mackay S. M., Wallbrink P. J., Burch G. J. and O'Loughlin E. M. Hydrologic characteristics and modelling of a small forested catchment in southeastern New South Wales. *Journal of Hydrology*, 83:307–335, 1986.

Moore R. J. The probability-distributed principle and runoff production at point and basin scales. *Hydrological Sciences Journal*, 30:273–297, 1985.

Moore R. J. Real-time flood forecasting systems: Perspectives and prospects. In Casali R. and Margottini C., editors, *Floods and Landslides: Integral Risk Assessment*, pages 147–189. Springer, Berlin, 1999.

Moore R. J., Bell V. A., Austin R. M. and Harding R. J. Methods for snowmelt forecasting in upland Britain. *Hydrology and Earth System Sciences*, 3:233–246, 1999.

Moore R. J. and Clarke R. T. A distribution function approach to rainfall–runoff modelling. *Water Resources Research*, 17:1367–1382, 1981.

Moore R. J. and Clarke R. T. A distribution function approach to modelling basin sediment yield. *Journal of Hydrology*, 65:239–257, 1983.

Moore R. J., Cole S. J., Bell V. A. and Jones D. A. Issues in flood forecasting: Ungauged basins, extreme floods and uncertainty. In *Frontiers in flood forecasting*, IAHS Publ. No. 305:103–122. IAHS Press: Wallingford, U.K., 2006.

Moore R. J., Jones D. A., Bird P. B. and Cottingham M. C. A basin-wide flow forecasting system for real-time flood warning, river control and water management. In White W. R., editor, *International Conference on River Flood Hydraulics*, pages 21–30, John Wiley & Sons: Chichester, UK, 1990.

Moore R. J., Jones D. A., Black K. B., Austin R. M., Carrington D. S., Tinninon M. and Akhondi A. RFFS and HYRAD: Integrated systems for rainfall and river flow forecasting in real-time and their application in Yorkshire. In *Proc. BHS National Meeting on Analytical Techniques for the Development and Operations Planning of Water Resource and Supply Systems, BHS Occasional Paper No. 4*, 1994.

Moore R. J., Jones D. A., Cox D. R. and Isham V. S. Design of the HYREX raingauge network. *Hydrology and Earth System Sciences*, 4(4):521–530, 2000.

Morel-Seytoux H. J. and Khanji J. Derivation of an equation of infiltration. *Water Resources Research*, 10:795–800, 1974.

Morris E. M. Physics-based models of snow. In Bowles D. S. and O P. E. 'Connell, editors, *Recent Advances in the Modelling of Hydrological Systems*, pages 85–112. Kluwer Academic, Dordrecht, 1991.

Morton F. I. Estimating evapotranspiration from potential evaporation: Practicality of an iconoclastic approach. *Journal of Hydrology*, 38:1–32, 1978.

Morton F. I. Operational estimates of areal evapotranspiration and their significance to the science and practice of hydrology. *Journal of Hydrology*, 66:1–76, 1983a.

Morton F. I. Operational estimates of lake evaporation. *Journal of Hydrology*, 66:77–100, 1983b.

Motovilov Y. G., Gottschalk L., Engeland K. and Rodhe A. Validation of a distributed hydrological model against spatial observations. *Agricultural and Forest Meteorology*, 98-99:257–277, 1999.

Mou L., Tian F., Hu H. and Sivapalan M. Extension of the representative elementary watershed approach for cold regions: Constitutive relationships and an application. *Hydrology and Earth System Sciences*, 12(2):565–585, 2008.

Moyeed R. A. and Clarke R. T. The use of Bayesian methods for fitting rating curves, with case studies. *Advances in Water Resources*, 28:807–818, 2005.

Mroczkowski M., Raper G. P. and Kuczera G. The quest for more powerful validation of conceptual catchment models. *Water Resources Research*, 33:2325–2335, 1997.

Muleta M. K. and Nicklow J. W. Sensitivity and uncertainty analysis coupled with automatic calibration for a distributed watershed model. *Journal of Hydrology*, 306:127–145, 2005.

Muttiah R. S., Srinivasan R. and Allen P. M. Prediction of two-year peak stream discharges using neural networks. Several regions tried–not all better than regression. *Journal of the American Water Resources Association*, 33:625–630, 1997.

Myrabø S. Temporal and spatial scale of response area and groundwater variation in till. *Hydrological Processes*, 11:1861–1880, 1997.

Naden P. S. Spatial variability in flood estimation for large catchments: The exploitation of channel network structure. *Journal of Hydrology*, 37:53–71, 1992.

Naden P. S. A routing model for continental scale hydrology. In Wilkinson W. B., editor, *Macroscale Modelling of the Hydrosphere*, IAHS Publication No. 214:67–79, Wallingford, UK, 1993.

Naden P. S., Broadhurst P., Tauveron N. and Walker A. River routing at the continental scale: Use of globally-available data and an a priori method of parameter estimation. *Hydrology and Earth System Sciences*, 3:109–124, 1999.

Nash J. E. Systematic determination of unit hydrograph parameters. *Journal of Geophysical Research*, 64:111–115, 1959.

Nash J. E. A unit hydrograph study with particular reference to British catchments. *Proc. Inst. Civ. Eng.*, 17:249–282, 1960.

Nash J. E. and Shamseldin A. Y. The geomorphological unit hydrograph: a critique. *Hydrology and Earth System Sciences*, 2:1–8, 1998.

Nash J. E. and Sutcliffe J. V. River flow forecasting through conceptual models: 1. A discussion of principles. *Journal of Hydrology*, 10:282–290, 1970.

Natale L. and Todini E. A constrained parameter estimation technique for linear models in hydrology. In Ciriani T. A., Malone U. and Wallis J. R., editors, *Mathematical Models for Surface Water Hydrology*. John Wiley & Sons, Chichester, UK, 1977.

Navar J. and Bryan R. B. Fitting the analytical model of rainfall interception of Gash to individual shrubs of semi-arid vegetation in northeastern Mexico. *Agricultural and Forest Meteorology*, 68:133–143, 1994.

Nayak P. C., Sudheer K. P. and Ramasastri K. S. Fuzzy computing based rainfall–runoff model for real time flood forecasting. *Hydrological Processes*, 19:955–968, 2005.

Neal C., Robson A. J., Shand P., Edmunds W. M., Dixon A. J., Buckley D. K., Hill S., Harrow M., Neal M., Wilkinson J. and Reynolds B. The occurrence of groundwater in the lower Palaeozoic rocks of upland central Wales. *Hydrology and Earth System Sciences*, 1:3–18, 1997.

Neal J. C., Fewtrell T. J., Bates P. D. and Wright N. G. A comparison of three parallelisation methods for 2D flood inundation models. *Environmental Modelling & Software*, 25:398–411, 2010.

Nearing M. A., Foster G. R., Lane L. J. and Finkner S. C. A process-based soil erosion model for USDA-water erosion prediction project technology. *Trans. Am. Soc. Agr. Engrs*, 32(5):1587–1593, 1989.

Nelder J. A. and Mead R. A simplex method for function minimization. *Computer Journal*, 7:308–313, 1965.

Nelson E. J., Jones N. L. and Berrett R. J. Adaptive tessellation method for creating TINs from GIS data. *ASCE Journal of Hydrologic Engineering*, 4:2, 1999.

NERC. *The Flood Studies Report (5 Volumes)*. Natural Environment Research Council, Wallingford, UK, 1975.

Newson M. D. The geomorphological effectiveness of floods: A contribution stimulated by two recent events in mid-Wales. *Earth Surface Processes and Landforms*, 5:1–16, 1980.

Nichols W. E. and Cuenca R. H. Evaluation of the evaporative fraction for parameterization of the surface energy balance. *Water Resources Research*, 29:3681–3690, 1993.

Niemi A. J. Residence time distribution of variable flow processes. *International Journal of Applied Radiation and Isotopes*, 28:855–860, 1977.

Nijssen B., Lettenmaier D. P., Liang X., Wetzel S. W. and Wood E. F. Streamflow simulation for continental-scale river basins. *Water Resources Research*, 33:711–724, 1997.

Norbiato D. and Borga M. Analysis of hysteretic behaviour of a hillslope-storage kinematic wave model for subsurface flow. *Advances in Water Resources*, 31:118–131, 2008.

Obled C., Bontrona G. and Garçon R. Quantitative precipitation forecasts: A statistical adaptation of model outputs through an analogue sorting approach. *Atmospheric Research*, 63:303–324, 2002.

O'Connell P. E. A historical perspective. In Bowles D. S. and O'Connell P. E., editors, *Recent Advances in the Modeling of Hydrologic Systems*, pages 3–30. Kluwer, Dordrecht, 1991.

O'Donnell T. Methods of computation in hydrograph analysis and synthesis. In *Recent Trends in Hydrograph Synthesis*. Committee for Hydrological Research, Central Organisation for Applied Research in the Netherlands, The Hague, 1966.

O'Kane P. Hysteresis in hydrology. *Acta Geophysica Polonica*, 53:373–383, 2005.

Olden J. D. and Poff N. L. Redundancy and the choice of hydrologic indices for characterizing streamflow regimes. *River Research and Applications*, 19:101–121, 2003.

Oleson K. W., Niu G.-Y., Yang Z.-L., Lawrence D. M., Thornton P. E., Lawrence P. J., Stöckli R., Dickinson R. E., Bonan G. B., Levis S., Dai A. and Qian T. Improvements to the Community Land Model and their impact on the hydrological cycle. *Journal of Geophysical Research*, 113:G01021, DOI:10.1029/2007JG000563, 2008.

Olivera F. and Maidment D. R. Geographical information systems (GIS)-based spatially distributed model for runoff routing. *Water Resources Research*, 35:1155–1164, 1999.

O'Loughlin E. M. Saturation regions in catchments and their relation to soil and topographic properties. *Journal of Hydrology*, 53:229–246, 1981.

O'Loughlin, E. M. Prediction of surface saturation zones in natural catchments by topographic analysis. *Water Resources Research*, 22:794–804, 1986.

Olsson J. A. and Andersson L. Possibilities and problems with the use of models as a communication tool in water resources management. *Water Resources Management*, 21:97–110, 2007.

Olsson J. A. and Berg K. Local stakeholders' acceptance of model-generated data used as a communication tool in water management: The Rönneå study. *Ambio*, 34:507–512, 2005.

Onof C., Faulkner D. and Wheater H. S. Design rainfall modelling in the Thames catchment. *Journal of Hydrology*, 41:715–733, 1996.

Oreskes N., Schrader-Frechette K. and Belitz K. Verification, validation and confirmation of numerical models in the earth sciences. *Science*, 263:641–646, 1994.

Oudin L., Andréassian V., Perrin C., Michel C. and Le Moine N. Spatial proximity, physical similarity, regression and ungaged catchments: A comparison of regionalization approaches based on 913 french catchments. *Water Resources Research*, 44:W03413, DOI:10.1029/2007WR006240, 2008.

Oudin L., Kay A., Andréassian V. and Perrin C. Are seemingly physically similar catchments truly hydrologically similar? *Water Resources Research*, 46(11):W11558, 2010.

Özelkan E. C. and Duckstein L. Fuzzy conceptual rainfall–runoff models. *Journal of Hydrology*, 253:41–68, 2001.

Ozyurt N. N. and Bayari C. S. LUMPED unsteady: A Visual Basic code of unsteady-state lumped-parameter models for mean residence time analyses of groundwater systems. *Computers and Geosciences*, 31:329, 2005a.

Ozyurt N. N. and Bayari C. S. Steady- and unsteady-state lumped parameter modelling of tritium and chlorofluorocarbons transport: Hypothetical analyses and application to an alpine karst aquifer. *Hydrological Processes*, 19:3269–3284, 2005b.

Pachepsky Y. and Timlin D. Water transport in soils as in fractal media. *Journal of Hydrology*, 204:98–107, 1998.

Page T., Beven K. J. and Freer J. Modelling the chloride signal at the Plynlimon catchments, Wales using a modified dynamic TOPMODEL. *Hydrological Processes*, 21:292–307, 2007.

Pahl-Wostl C. Towards sustainability in the water sector: The importance of human actors and processes of social learning. *Aquatic Sciences*, 64:394–411, 2002.

Pahl-Wostl C. Transitions towards adaptive management of water facing climate and global change. *Water Resources Management*, 21:49–62, 2007.

Paik K. and Kumar P. Optimality approaches to describe characteristic fluvial patterns on landscapes. *Philosophical Transactions of the Royal Society London B*, 365:1387–1395, 2010.

Palacios-Velez O. L. and Cuevas-Renaud B. Automated river course, ridge and basin delineation from digital elevation data. *Journal of Hydrology*, 86:299–314, 1986.

Palacios-Velez O. L., Gandoy-Bernasconi W. and Cuevas-Renaud B. Geometric analysis of surface runoff and the computation order of unit elements in distributed hydrological models. *Journal of Hydrology*, 211:266–274, 1998.

Paniconi C., Aldama A. A. and Wood E. F. Numerical evaluation of iterative and noniterative methods for the solution of the nonlinear Richards equation. *Water Resources Research*, 27:1147–1163, 1991.

Paniconi C. and Wood E. F. A detailed model for simulation of catchment scale subsurface hydrologic processes. *Water Resources Research*, 29:1601–1620, 1993.

Panziera L. and Germann U. The relationship between airflow and orographic precipitation on the southern side of the Alps as revealed by weather radar. *Quarterly Journal of the Royal Meteorological Society*, 136:222–238, 2010.

Pappenberger F., Beven K. J., Frodsham K., Romanovicz R. and Matgen P. Grasping the unavoidable subjectivity in calibration of flood inundation models: A vulnerability weighted approach. *Journal of Hydrology*, 333:275–287, 2007a.

Pappenberger F., Beven K. J., Hunter N., Gouweleeuw B., Bates P., de Roo A. and Thielen J. Cascading model uncertainty from medium range weather forecasts (10 days) through a rainfall–runoff model to flood inundation predictions within the European Flood Forecasting System (EFFS). *Hydrology and Earth System Sciences*, 9:381–393, 2005b.

Pappenberger F., Beven K. J., Ratto M. and Matgen P. Multi-method global sensitivity analysis of flood inundation models. *Advances in Water Resources*, 31(1):1–14, 2008.

Pappenberger F., Frodsham K., Beven K. J., Romanovicz R. and Matgen P. Fuzzy set approach to calibrating distributed flood inundation models using remote sensing observations. *Hydrology and Earth System Sciences*, 11(2):739–752, 2007b.

Pappenberger F., Iorgulescu I. and Beven K. J. Sensitivity analysis based on regional splits (SARS) and regression trees. *Environmental Modelling & Software*, 21:976–990, 2006a.

Pappenberger F., Matgen P., Beven K. J., Henry J.-B., Pfister L. and de Fraipont P. Influence of uncertain boundary conditions and model structure on flood inundation predictions. *Advances in Water Resources*, 29:1430–1449, DOI:10.1016/j.advwatres.2005.11.012, 2006b.

Pappenberger F., Scipal K. and Buizza R. Hydrological aspects of meteorological verification. *Atmospheric Science Letters*, 9:43–52, 2008b.

Parada L. M. and Liang X. A novel approach to infer streamflow signals for ungauged basins. *Advances in Water Resources*, 33:372–386, 2010.

Parajka J., Merz R. and Blöschl G. A comparison of regionalisation methods for catchment model parameters. *Hydrology and Earth System Sciences*, 9:157–171, 2005.

Parkin G., O'Donnell G., Ewen J., Bathurst J. C., O'Connell P. E. and Lavabre J. Validation of catchment models for predicting land-use and climate change impacts: 2. Case study for a Mediterranean catchment. *Journal of Hydrology*, 175:595–613, 1996.

Parlange J.-Y. and Haverkamp R. Infiltration and ponding time. In Morel-Seytoux H. J., editor, *Unsaturated Flow in Hydrologic Modeling*, pages 105–126. Kluwer Academic, Dordrecht, 1989.

Parlange J.-Y., Hogarth W. L., Barry D. A., Parlange M. B., Haverkamp R., Ross P. J., Steenhuis T. S., DiCarlo D. A. and Katul G. Analytical approximation to the solutions of Richards' equation with applications to infiltration, ponding, and time compression approximation. *Advances in Water Resources*, 23:189–194, 1999.

Parry M. L., Canziani O. F., Palutikof J. P., van der Linden P. J. and Hanson C. E., editors. *Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change: Impacts, Adaptation and Vulnerability*. Cambridge University Press, Cambridge, UK, 2007.

Parsons A. J., Wainright J., Abrahams A. D. and Simanton J. R. Distributed dynamic modelling of interrill overland flow. *Hydrological Processes*, 11:1833–1859, 1997.

Pellenq J., Kalma J., Boulet G., Saulnier G.-M., Wooldridge S., Kerr Y. and Chehbouni A. A disaggregation scheme for soil moisture based on topography and soil depth. *Journal of Hydrology*, 276:112–127, 2003.

Pereira L. S., Perrier A., Allen R. G. and Alves I. Evapotranspiration: Concepts and future trends. *ASCE Journal of Hydrologic Engineering*, 4:45, 1999.

Perrin C., Michel C. and Andréassian V. Does a large number of parameters enhance model performance? Comparative assessment of common catchment model structures on 429 catchments. *Journal of Hydrology*, 242:275–301, 2001.

Persson K. and Destouni G. Propagation of water pollution uncertainty and risk from the subsurface to the surface water system of a catchment. *Journal of Hydrology*, 377:434–444, 2009.

Peters N. E., Freer J. E. and Beven K. J. Modelling hydrologic responses in a small forested catchment (Panola Mountain, Georgia, USA): A comparison of the original and a new dynamic TOPMODEL. *Hydrological Processes*, 17(2):345–362, 2003.

Peterson J. R. and Hamlett J. M. Hydrologic calibration of the SWAT model in a watershed containing fragipan soils. *Journal of the American Water Resources Association*, 34:531–544, 1998.

Philip J. R. The theory of infiltration: 1. The infiltration equation and its solution. *Soil Science*, 83:345–357, 1957.

Philip J. R. Hillslope infiltration: Divergent and convergent slopes. *Water Resources Research*, 27:1035–1040, 1991.

Pielke Sr R., Beven K. J., Brasseur G., Calvert J., Chahine M., Dickenson R., Entekhabi D., Foufoula-Georgiou E., Gupta H., Gupta V., Krajewski W., Krider E. P., Lau W. K. M., McDonnell J. J., Rossow W., Schaake J.,

Smith J., Sorooshian S. and Wood E. F. Climate change: The need to consider human forcings other than greenhouse gases. *EoS, Transactions AGU*, 90:413, 2009.

Pinder G. F. and Gray W. G. *Finite Element Simulation in Surface and Subsurface Hydrology*. Academic Press, New York, 1977.

Pinder G. F. and Jones J. F. Determination of the ground-water component of peak discharge from the chemistry of total runoff. *Water Resources Research*, 5(2):438–445, 1969.

Pingel N. and Watkins D. Multiple flood source expected annual damage computations. *ASCE Journal of Water Resources Planning and Management*, 136:319, DOI:10.1061/(ASCE)WR.1943–5452.0000031, 2010.

Piñol J., Beven K. J. and Freer J. Modelling the hydrological response of Mediterranean catchments, Prades, Catalonia: The use of distributed models as aids to hypothesis formulation. *Hydrological Processes*, 11:1287–1306, 1997.

Plate E. J., Ihrringer J. and Lutz W. Operational models for flood calculations. *Journal of Hydrology*, 100:489–506, 1988.

Poeter E. P. and Hill M. C. Inverse methods: A necessary next step in ground-water flow modeling. *Ground Water*, 35:250–260, 1997.

Ponce V. M. *Engineering Hydrology: Principles and Practices*. Prentice-Hall, Englewood Cliffs, NJ, 1989.

Ponce V. M. The kinematic wave controversy. *ASCE Journal of the Hydraulics Division*, 117:511–525, 1991.

Ponce V. M. Notes of my conversation with Vic Mockus. http://mockus.sdsu.edu, 1996.

Post D. A. and Jakeman A. J. Relationships between physical attributes and hydrologic response characteristics in small Australian mountain ash catchments. *Hydrological Processes*, 10:877–892, 1996.

Post D. A. and Jakeman A. J. Predicting the daily streamflow of ungauged catchments in S. E. Australia by regionalising the parameters of a lumped conceptual rainfall–runoff model. *Ecol. Model.*, 123:91–104, 1999.

Post D. A., Jones J. A. and Grant G. E. An improved methodology for predicting the daily hydrologic response of ungauged catchments. *Environmental Modelling & Software*, 13:395–403, 1998.

Pradhan N. R., Ogden F. L., Tachikawa Y. and Takara K. Scaling of slope, upslope area and soil water deficit: Implications for transferability and regionalization in topographic index modeling. *Water Resources Research*, 44:1–12, 2008.

Pradhan N. R., Tachikawa Y. and Takara K. A downscaling method of topographic index distribution for matching the scales of model application and parameter identification. *Hydrological Processes*, 20(6):1385–1405, 2006.

Press W. H., Flannery B. P., Teukolsky S. A. and Vetterling W. T. *Numerical Recipes, the Art of Scientific Computing*. Cambridge University Press, Cambridge, 2nd edition, 1992.

Prevost M., Barry R., Stein J. and Plamondon A. P. Snowmelt runoff modelling in a balsam fir forest with a variable source area simulator. *Water Resources Research*, 26:1067–1077, 1990.

Price R. K. Volume-conservative nonlinear flood routing. *ASCE Journal of Hydraulic Engineering*, 135:838–845, 2009.

Di Prinzio M., Castellarin A. and Toth E. Data-driven catchment classification: Application to the PUB problem. *Hydrology and Earth System Sciences Discussion*, 8:391–427, 2011.

Prudhomme C., Wilby R. L., Crooks S., Kay A. L. and Reynard N. S. Scenario-neutral approach to climate change impact studies: Application to flood risk. *Journal of Hydrology*, 390:198–209, 2010.

Puente C. E. A new approach to hydrologic modeling: Derived distributions revisited. *Journal of Hydrology*, 187:65–80, 1997.

Qu Y. and Duffy C. J. A semi discrete finite volume formulation for multiprocess watershed simulation. *Water Resources Research*, 43:W08419, DOI: 10.1029/2006WR005752, 2007.

Quinn P. F., Beven K. J., Chevallier P. and Planchon O. The prediction of hillslope flow paths for distributed hydrological modelling using digital terrain models. *Hydrological Processes*, 5:59–79, 1991.

Quinn P. F., Beven K. J. and Lamb R. The ln(a/tan) index: How to calculate it and how to use it in the TOPMODEL framework. *Hydrological Processes*, 9:161–182, 1995.

Quinn P. F., Ostendorf B., Beven K. J. and Tenhunen J. D. Spatial and temporal predictions of soil moisture patterns and evaporative losses using TOPMODEL and the GAS-Flux model for an Alaskan catchment. *Hydrology and Earth System Sciences*, 2:41–54, 1998.

Ragab R. and Bromley J. IHMS: Integrated hydrological modelling system. Part 1. Hydrological processes and general structure. *Hydrological Processes*, 24:2663–2680, 2010.

Ragab R., Bromley J., Dörflinger G. and Katsikides S. IHMS: Integrated hydrological modelling system. Part 2. Application of linked unsaturated, DiCaSM and saturated zone, MODFLOW models on Kouris and Akrotiri catchments in Cyprus. *Hydrological Processes*, 24:2681–2692, 2010.

Ran Q., Heppner C. H., VanderKwaak J. E. and Loague K. Further testing of the integrated hydrology model (InHM): Multiple-species sediment transport. *Hydrological Processes*, 21:1522–1531, 2007.

Rango A. The snowmelt runoff model (SRM). In Singh V. P., editor, *Catchment Models of Watershed Hydrology*, pages 477–520. Water Resource Publications, Highlands Ranch, CO, 1995.

Rango A. and Martinec J. Revisiting the degree-day method for snowmelt computations. *Water Resources Bulletin*, 31:657–669, 1995.

Rao A. R. and Srinivas V. Regionalization of watersheds by hybrid cluster analysis. *Journal of Hydrology*, 318:37–56, 2006.

Rawls W. J. and Brakensiek D. L. Estimating soil water retention from soil properties. *ASCE Journal of Irrigation and Drainage Engineering*, 108:166–171, 1982.

Rawls W. J. and Brakensiek D. L. Estimation of soil water retention and hydraulic properties. In Morel-Seytoux H. J., editor, *Unsaturated Flow in Hydrologic Modeling: Theory and Practice*, pages 275–300. Kluwer Academic, Dordrecht, 1989.

Rawls W. J., Brakensiek D. L. and Miller N. Green–Ampt infiltration parameters from soil data. *ASCE Journal of the Hydraulics Division*, 109:62–70, 1983.

Reed D. W., Johnson P. and Firth J. M. A non-linear rainfall–runoff model, providing for variable lag time. *Journal of Hydrology*, 25:295–305, 1975.

Reeves A. D., Henderson D. E. and Beven K. J. Flow separation in undisturbed soil using multiple anionic tracers: (1) Analytical methods and unsteady rainfall and return flow experiments. *Hydrological Processes*, 10:1435–1450, 1996.

Refsgaard J. C. Parameterisation, calibration and validation of distributed hydrological models. *Journal of Hydrology*, 198:69–97, 1997.

Refsgaard J. C. and Henriksen H. J. Modelling guidelines – terminology and guiding principles. *Advances in Water Resources*, 27:71–82, 2004.

Refsgaard J. C. and Knudsen J. Operational validation and intercomparison of different types of hydrological models. *Water Resources Research*, 32:2189–2202, 1996.

Refsgaard J. C., Seth S. M., Bathurst J. C., Erlich M., Storm B., Jogenson G. H. and Chandra S. Application of the SHE to catchments in India: Part 1. General results. *Journal of Hydrology*, 140:1–23, 1992.

Refsgaard J. C. and Storm B. MIKE SHE. In Singh V. P., editor, *Computer Models of Watershed Hydrology*, pages 809–846. Water Resources Publications, Highlands Park, CO, 1995.

Refsgaard J. C., Storm B. and Abbott M. B. Comment on 'A discussion of distributed hydrological modelling' by Beven K. In *Distributed Hydrological Modelling*, volume Abbott M. B. and Refsgaard J. C., pages 279–287. Kluwer Academic, Dordrecht, 1996.

Refsgaard J. C., Storm B. and Clausen T. Système Hydrologique Européen (SHE): Review and perspectives after 30 years development in distributed physically based hydrological modelling. *Hydrology Research*, 45:355–377, 2010.

Refsgaard J. C., van der Sluijs J. P., Brown J. and van der Keur P. A framework for dealing with uncertainty due to model structure error. *Advances in Water Resources*, 29:1586–1597, 2006.

Reggiani P., Hassanizadeh S. M., Sivapalan M. and Gray W. G. A unifying framework for watershed thermodynamics: Constitutive relationships. *Advances in Water Resources*, 23:15–39, 1999.

Reggiani P. and Rientjes T. H. M. Flux parameterization in the representative elementary watershed approach: Application to a natural basin. *Water Resources Research*, 41:W04013, DOI:10.1029/2004WR003693, 2005.

Reggiani P. and Schellekens J. Modelling of hydrological responses: The representative elementary watershed approach as an alternative blueprint for watershed modelling. *Hydrological Processes*, 17(18):3785–3789, 2003.

Reggiani P., Sivapalan M. and Hassanizadeh S. M. A unifying framework for watershed thermodynamics: Balance equations for mass, momentum, energy and entropy and the second law of thermodynamics. *Advances in Water Resources*, 22(4):367–398, 1998.

Reggiani P., Sivapalan M. and Hassanizadeh S. M. Conservation equations governing hillslope responses: Exploring the physical basis of water balance. *Water Resources Research*, 36(7):1845–1863, 2000.

Reggiani P., Sivapalan M., Hassanizadeh S. M. and Gray W. G. Coupled equations for mass and momentum balance in a stream network: Theoretical derivation and computational experiments. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 457(2005):157, 2001.

Reggiani P. and Weerts A. H. A Bayesian approach to decision-making under uncertainty: An application to real-time forecasting in the River Rhine. *Journal of Hydrology*, 356:56–69, 2008.

Reichl J. P. C., Western A. W., McIntyre N. R. and Chiew F. H. S. Optimization of a similarity measure for estimating ungauged streamflow. *Water Resources Research*, 45:W10423, DOI:10.1029/2008WR007248, 2009.

Renard B., Kavetski D. and Kuczera G. Comment on "An integrated hydrologic Bayesian multimodel combination framework: Confronting input, parameter, and model structural uncertainty in hydrologic prediction" by Ajami Q. Y., Duan N. K. and Sorooshian S. *Water Resources Research*, 45:W03603, DOI:10.1029/2007WR006538, 2009.

Renard B., Kavetski D., Kuczera G., Thyer M. and Franks S. W. Understanding predictive uncertainty in hydrologic modeling: The challenge of identifying input and structural errors. *Water Resources Research*, 46:W05521, DOI:10.1029/2009WR008328, 2010.

Reusser D. E., Blume T., Schaefli B. and Zehe E. Analysing the temporal dynamics of model performance for hydrological models. *Hydrology and Earth System Sciences*, 13:999–1018, 2009.

Reynolds C. S., Carling P. A. and Beven K. J. Flow in river channels: New insights into hydraulic retention. *Arch. Hydrobiol.*, 121:171–179, 1991.

Richards B. D. *Flood Estimation and Control*. Chapman and Hall, London, 1944.

Richards L. A. Capillary conduction of liquids through porous mediums. *Physics*, 1:318–333, 1931.

Rinaldo A., Beven K. J., Bertuzzo E., Nicotina L., Davies J., Fiori A., Russo D. and Botter G. Catchment travel time distributions and water flow in soils. *Water Resources Research*, 47, W07537, doi:10.1029/2011WR010478. 2011.

Rinaldo A. and Rodriguez-Iturbe I. Geomorphological theory of the hydrological response. *Hydrological Processes*, 10(6):803–829, 1996.

Robinson J. S. and Sivapalan M. Catchment-scale runoff generation model by aggregation and similarity analyses. *Hydrological Processes*, 9:555–574, 1995.

Robinson J. S., Sivapalan M. and Snell J. D. On the relative roles of hillslope processes, channel routing, and network geomorphology in the hydrologic response of natural catchments. *Water Resources Research*, 31:3089–3101, 1995.

Robson A. J. Evidence for trends in UK flooding. *Philosophical Transactions of the Royal Society London*, 360:1327–1343, 2002.

Robson A. J., Beven K. J. and Neal C. Towards identifying sources of subsurface flow: A comparison of components identified by a physically based runoff model and those determined by chemical mixing techniques. *Hydrological Processes*, 6:199–214, 1992.

Rodda J. C. and Smith S. W. The significance of the systematic error in rainfall measurement for assessing wet deposition. *Atmos. Environ.*, 20:1059–1064, 1986.

Rode M. and Suhr U. Uncertainties in selected river water quality data. *Hydrology and Earth System Sciences*, 11:863–874, 2007.

Rodell M., Chen J., Kato H., Famiglietti J. S., Nigro J. and Wilson C. R. Estimating groundwater storage changes in the Mississippi River basin (USA) using GRACE. *Hydrogeology Journal*, 15:159–166, 2007.

Rodhe A., Nyberg L. and Bishop K. Transit times for water in a small till catchment from a step shift in the oxygen 18 content of the water input. *Water Resources Research*, 32:3497–3511, 1996.

Rodriguez-Iturbe I. The geomorphological unit hydrograph. In Beven K. J. and Kirkby M. J., editors, *Channel Network Hydrology*, pages 43–68. John Wiley & Sons, Chichester, UK, 1993.

Rodriguez-Iturbe I. and Porporato A. *Ecohydrology of water-controlled ecosystems: Soil moisture and plant dynamics*. Cambridge University Press, Cambridge, UK, 2004.

Rodriguez-Iturbe I. and Rinaldo A. *Fractal river basins: Chance and self-organization*. Cambridge University Press, Cambridge, UK, 1997.

Rodriguez-Iturbe I. and Valdes J. The geomorphic structure of hydrologic response. *Water Resources Research*, 15:1409–1420, 1979.

Rodriguez-Paz M. X. and Bonet J. Mesh-free numerical simulation of debris flows avalanches. *Numerical Methods for Partial Differential Equations*, 20:140–163, 2004.

Rodriguez-Paz M. X. and Bonet J. A corrected smooth particle hydrodynamics formulation of the shallow-water equations. *Computers and Structures*, 83:1396–1410, 2005.

Rojas-Cerna C., Michel C., Perrin C. and Andréassian V. Ungauged catchments: How to make the best of a few streamflow measurements. In *Large Sample Basin Experiments for Hydrological Model Parameterization: Results of the Model Parameter Experiment (MOPEX)*, IAHS Publication No. 307:230–236. IAHS Press, 2006.

Romano N. and Santini A. Effectiveness of using pedo-transfer functions to quantify the spatial variability of soil water retention characteristics. *Journal of Hydrology*, 202:137–157, 1997.

Romanowicz R. A Matlab implementation of TOPMODEL. *Hydrological Processes*, 11(9):1115–1129, 1997.

Romanowicz R. and Beven K. J. Estimation of flood inundation probabilities as conditioned on event inundation maps. *Water Resources Research*, 39:1073–1085, 2003.

Romanowicz R. J., Young P. C. and Beven K. J. Data assimilation and adaptive forecasting of water levels in the River Severn catchment, United Kingdom. *Water Resources Research*, 42:W06407, 2006.

Romanowicz R. J., Young P. C., Beven K. J. and Pappenberger F. A data based mechanistic approach to nonlinear flood routing and adaptive flood level forecasting. *Advances in Water Resources*, 31:1048–1056, 2008.

Römkens M. J. M., Prasad S. N. and Whisler F. D. Surface sealing and infiltration. In Anderson M. G. and Burt T. P. editors, *Process Studies in Hillslope Hydrology*, pages 127–172. John Wiley & Sons, Chichester, UK, 1990.

Rooney G. G. and Claxton B. M. Comparison of the Met Office's surface exchange scheme, MOSES, against field observations. *Quarterly Journal of the Royal Meteorological Society*, 132:425–446, 2006.

Rose K. A., Smith E. P., Gardner R. H., Brenkert A. L. and Bartell S. M. Parameter sensitivities, Monte Carlo filtering and model forecasting under uncertainty. *Journal of Forecasting*, 10:117–134, 1991.

Rosenbrock H. H. An automatic method of finding the greatest or least value of a function. *Computer Journal*, 3:175–184, 1960.

Ross C. N. The calculation of flood discharge by the use of time contour plan isochrones. *Transactions of the Institution of Engineers, Australia*, 2:85–92, 1921.

Ross J. L., Ozbek M. M. and Pinder G. F. Aleatoric and epistemic uncertainty in groundwater flow and transport simulation. *Water Resources Research*, 45:W00B15, DOI:10.1029/2007WR006799, 2009.

Rosso R. Nash model relation to Horton order ratios. *Water Resources Research*, 20:914–920, 1984.

Roubtsova V. and Kahawita R. The SPH technique applied to free surface flows. *Computers & Fluids*, 35:1359–1371, 2006.

Rutter A. J., Kershaw K. A., Robins P. C. and Morton A. J. A predictive model of rainfall interception in forests: 1. Derivation of the model from observations in a plantation of Corsican pine. *Agricultural Meteorology*, 9:367–384, 1971.

Rutter A. J., Morton D. J. and Robins P. C. A predictive model of rainfall interception by forests. II. Generalisations of the model and comparisons with observations in some coniferous and hardwood stands. *Journal of Applied Ecology*, 12:367–380, 1975.

Salamon P. and Feyen L. Assessing parameter, precipitation, and predictive uncertainty in a distributed hydrological model using sequential data assimilation with the particle filter. *Journal of Hydrology*, 376:428–442, 2009.

Saltelli A., Ratto M., Tarantola S. and Campolongo F. Sensitivity analysis practices: Strategies for model-based inference. *Reliability Engineering and System Safety*, 91:1109–1125, 2006.

Saltelli A., Tarantola S., Campolongo F. and Ratto M. *Sensitivity Analysis in Practice: A Guide to Assessing Scientific Models*. John Wiley & Sons, Chichester, UK, 2004.

Samaniego L., Bárdossy A. and Kumar R. Streamflow prediction in ungauged catchments using copula-based dissimilarity measures. *Water Resources Research*, 46:W02506, DOI:10.1029/2008WR007695, 2010.

Santra P., Das B. S. and Chakravarty D. Delineation of hydrologically similar units in a watershed based on fuzzy classification of soil hydraulic properties. *Hydrological Processes*, 25:64–79, 2011.

Saulnier G.-M., Beven K. J. and Obled C. Digital elevation analysis for distributed hydrological modelling: Reducing scale dependence in effective hydraulic conductivity values. *Water Resources Research*, 33:2097–2101, 1997b.

Saulnier G.-M., Beven K. J. and Obled C. Including spatially variable soil depths in TOPMODEL. *Journal of Hydrology*, 202:158–172, 1998.

Saulnier G.-M. and Datin R. Analytical solution to a bias in the TOPMODEL framework balance. *Hydrological Processes*, 18:1195–1218, 2004.

Saulnier G.-M., Obled C. and Beven K. J. Analytical compensation between DTM grid resolution and effective values of saturated hydraulic conductivity within the TOPMODEL framework. *Hydrological Processes*, 11:1331–1346, 1997a.

Sayama T. and McDonnell J. J. A new time–space accounting scheme to predict stream water residence time and hydrograph source components at the watershed scale. *Water Resources Research*, 45:W07401, doi:10.1029/2008WR007549, 2009.

Schaake J. C., Hamill T. H., Buizza R. and Clark M. HEPEX – the hydrological ensemble prediction experiment. *Bull. Amer. Meteor. Soc.*, 88:DOI:10.1175/BAMS–88–10–1541, 2007.

Schaap M. G. and Bouten W. Modelling water retention curves of sandy soils using neural networks. *Water Resources Research*, 32:3033–3040, 1996.

Schaap M. G. and Leij F. J. Database related accuracy and uncertainty of pedotransfer functions. *Soil Science*, 163:765–779, 1998.

Schaap M. G., Leij F. J. and van Genuchten M. T. Neural network analysis for hierarchical prediction of soil water retention and saturated hydraulic conductivity. *Soil Science Society of America Journal*, 62:847–855, 1998.

Schaefli B. and Gupta H. V. Do Nash values have value? *Hydrological Processes*, 21:2075–2080, 2007.

Schaefli B., Harman C. J., Sivapalan M. and Schymanski S. J. Hydrologic predictions in a changing environment: Behavioral modeling. *Hydrology and Earth System Sciences*, 15:635–646, 2011.

Schaefli B. and Zehe E. Hydrological model performance and parameter estimation in the wavelet-domain. *Hydrology and Earth System Sciences*, 13:1921–1936, 2009.

Schapp M. G., Leij F. J. and van Genuchten M. T. Rosetta: A computer program for estimating soil hydraulic parameters with hierarchical pedotransfer functions. *Journal of Hydrology*, 251:163–176, 2001.

Scherrer S. and Naef F. A decision scheme to indicate dominant hydrological flow processes on temperate grassland. *Hydrological Processes*, 17:391–401, 2003.

Schoups G., van de Giesen N. C. and Savenije H. H. G. Model complexity control for hydrologic prediction. *Water Resources Research*, 44:W00B03, DOI:10.1029/2008WR006836, 2008.

Schoups G. and Vrugt J. A. A formal likelihood function for parameter and predictive inference of hydrologic models with correlated, heteroscedastic, and non-Gaussian errors. *Water Resources Research*, 46:W10531, DOI:10.1029/2009WR008933, 2010.

Schoups G., Vrugt J. A., Fenicia F. and van de Giesen N. C. Corruption of accuracy and efficiency of Markov Chain Monte Carlo simulation by inaccurate numerical implementation of conceptual hydrologic models. *Water Resources Research*, 46:W10530, DOI: 10.1029/2009WR008648, 2010.

Schreider S. Y., Jakeman A. J., Pittock A. B. and Whetton P. H. Estimation of possible climate change impacts on water availability, extreme flow events and soil moisture in the Goulburn and Ovens Basins, Victoria. *Climatic Change*, 34:513–546, 1996.

Schreider S. Y., Whetton P. H., Jakeman A. J., and Pittock A. B. Runoff modelling for snow-affected catchments in the Australian alpine region, eastern Victoria. *Journal of Hydrology*, 200:1–23, 1997.

Schulz K. and Beven K. Data-supported robust parameterisations in land surface–atmosphere flux predictions: Towards a top-down approach. *Hydrological Processes*, 17:2259–2277, 2003.

Schulz K., Beven K. and Huwe B. Equifinality and the problem of robust calibration in nitrogen budget simulations. *Soil Science Society of America Journal*, 63:1934–1941, 1999.

Schulz K. and Huwe B. Water flow modelling in the unsaturated zone with imprecise parameters using a fuzzy approach. *Journal of Hydrology*, 201:211–229, 1997.

Schumann A. H. and Funke R. GIS-based components for rainfall–runoff models. In *HydroGIS96: Application of Geographic Information Systems in Hydrology and Water Resources*, IAHS Publication No. 235:477–484, Wallingford, UK, 1996.

Schumann G. J.-P., Di Baldassarre G., Alsdorf D. E. and Bates P. D. Near real-time flood wave approximation on large rivers from space: Application to the River Po, northern Italy. *Water Resources Research*, 46:W05601, DOI:10.1029/2008WR007672, 2010.

Schumm S. A. Evolution of drainage systems and slopes in badlands at Perth Amboy, NJ. *Geological Society of America Bulletin*, 67(5):597, 1956.

Schüttemeyer D., Schillings C., Moene A. F. and de Bruin H. A. R. Satellite-based actual evapotranspiration over drying semiarid terrain in west Africa. *Journal of Applied Meteorology and Climatology*, 46:97–111, 2007.

Schymanski S. J., Roderick M. L., Sivapalan M., Hutley L. B. and Beringer J. A canopy scale test of the optimal water use hypothesis. *Plant, Cell & Environment*, 31:97–111, 2007.

Schymanski S. J., Sivapalan M., Roderick M. L., Beringer J. and Hutley L. B. An optimality-based model of the coupled soil moisture and root dynamics. *Hydrology and Earth System Sciences*, 12:913–932, 2008.

Schymanski S. J., Sivapalan M., Roderick M. L., Hutley L. B. and Beringer J. An optimality based model of the dynamic feedbacks between natural vegetation and the water balance. *Water Resources Research*, 45:W01412, DOI:10.1029/2008wr006841, 2009.

Scoging H. M. and Thornes J. B. Infiltration characteristics in a semiarid environment. *IAHS Publication No. 128*, 128:159–168, 1982.

Sefton C. E. M. and Howarth S. M. Relationships between dynamic response characteristics and physical descriptors of catchments in England and Wales. *Journal of Hydrology*, 211:1–16, 1998.

Seibert J. Regionalisation of parameters for a conceptual rainfall–runoff model. *Agricultural and Forest Meteorology*, 98-99:279–293, 1999.

Seibert J. and Beven K. J. Gauging the ungauged basin: How many discharge measurements are needed? *Hydrology and Earth System Sciences*, 13:883–892, 2009.

Seibert J., Bishop K. H. and Nyberg L. A test of TOPMODEL's ability to predict spatially distributed groundwater levels. *Hydrological Processes*, 11:1131–1144, 1997.

Seibert J. and McDonnell J. J. On the dialog between experimentalist and modeler in catchment hydrology: Use of soft data for multicriteria model calibration. *Water Resources Research*, 38:W01241, DOI:10.1029/2001WR000978, 2002.

Sellers P. J. Canopy reflectance, photosynthesis and transpiration. *Int. J. Remote. Sens.*, 8:1335–1372, 1985.

Sen M. and Stoffa P. L. *Global Optimisation Methods in Geophysical Inversion*. Elsevier, Amsterdam, 1995.

Senbeta D. A., Shamseldin A. Y. and O'Connor K. M. Modification of the probability-distributed interacting storage capacity model. *Journal of Hydrology*, 224:149–168, 1999.

Seuffert G., Gross P., Simmer C. and Wood E. F. The influence of hydrologic modeling on the predicted local weather: Two-way coupling of a mesoscale weather prediction model and a land surface hydrologic model. *Journal of Hydrometeorology*, 3:505–523, 2002.

Shamir E., Imam B., Morin E., Gupta H. V. and Sorooshian S. The role of hydrograph indices in parameter estimation of rainfall–runoff models. *Hydrological Processes*, 19:2187–2207, 2004.

Sharma M. L., Gander G. A. and Hunt C. G. Spatial variability of infiltration in as watershed. *Journal of Hydrology*, 45:101–122, 1980.

Shaw E. M., Beven K. J., Chappell N. A. and Lamb R. *Hydrology in Practice*. Spon Press, Taylor and Francis, 4th edition, 2010.

Sherman L. K. Streamflow from rainfall by unit-graph method. *Engineering News Record*, 108:501–505, 1932.

Shorter J. A. and Rabitz H. A. Risk analysis by the guided Monte Carlo technique. *Journal of Statistical Computation and Simulation*, 57:321–226, 1997.

Shouse P. J. and Mohanty B. P. Scaling of near-saturated hydraulic conductivity measured using disc infiltrometers. *Water Resources Research*, 34:1195–1205, 1998.

Shuttleworth W. J., Gash J. H. C., Lloyd C. R., McNeill D. D., Moore C. J. and Wallace J. S. An integrated micrometeorological system for evaporation measurement. *Agricultural and Forest Meteorology*, 43:295–317, 1988.

Sidle R. C. Field observations and process understanding in hydrology: Essential components in scaling. *Hydrological Processes*, 20:1439–1445, 2006.

Silburn D. M. and Connolly R. D. Distributed parameter hydrology model (Answers) applied to a range of catchment scales using rainfall simulator data: I. Infiltration modelling and parameter measurement. *Journal of Hydrology*, 172:87–104, 1995.

Simic E. and Destouni G. Water and solute residence times in a catchment: Stochastic–mechanistic model interpretation of ^{18}O transport. *Water Resources Research*, 35:2109–2119, 1999.

Simmons C. S., Nielsen D. R. and Biggar J. W. Scaling field measured soil water properties (2 parts). *Hilgardia*, 47:77–173, 1979.

Šimůnek J., Kodesova R., Gribb M. M. and van Genuchten M. T. Estimating hysteresis in the soil water retention function from core permeameter experiments. *Water Resources Research*, 35:1329–1346, 1999.

Šimůnek J., Sejna M. and van Genuchten M. T. The HYDRUS-2D software package for simulating water flow and solute transport in two dimensional variably saturated media, version 1. Technical Report IGWMC-TP553, International Ground Water Modelling Center, Golden, CO, 1996.

Šimůnek J. and van Genuchten M. T. Modeling nonequilibrium flow and transport processes using HYDRUS. *Vadose Zone Journal*, 7:782–797, 2008.

Sinclair S. and Pegram G. G. S. A comparison of ASCAT and modelled soil moisture over South Africa, using TOPKAPI in land surface mode. *Hydrology and Earth System Sciences*, 14:613–626, 2010.

Singh V. P. *Computer Models of Watershed Hydrology*. Water Resource Publications, Highlands Ranch, CO, 1995.

Singh V. P. *Kinematic Wave Modelling in Water Resources*. John Wiley & Sons, Chichester, UK, 1996.

Singh V. P., Bengtsson L. and Westerstrom G. Kinematic wave modelling of saturated basal flow in a snowpack. *Hydrological Processes*, 11:177–187, 1997.

Singh V. P. and Frevert D., editors. *Watershed Models*. Taylor and Francis Group, CRC Press, 2005.

Singh V. P. and Frevert D. K., editors. *Mathematical models of small watershed hydrology and applications*. Water Resource Publications, Littleton, CO, 2002a.

Singh V. P. and Frevert D. K., editors. *Mathematical Models of Large Watershed Hydrology*. Water Resource Publications, Littleton, CO, 2002b.

Singh V. P. and Yu C.-Y. Evaluation and generalisation of 13 mass-transfer equations for determining free water evaporation. *Hydrological Processes*, 11:311–323, 1997.

Sisson S. A., Fan Y. and Tanaka M. Sequential Monte Carlo without likelihoods. *Proceedings of the National Academy of Sciences*, 104:1760–1765, 2007.

Sivapalan M. Prediction in ungauged basins: A grand challenge for theoretical hydrology. *Hydrological Processes*, 17:3163–3170, DOI:10.1002/hyp.5155, 2003. 7795, 2003.

Sivapalan M. *Encyclopaedia of Hydrological Sciences*, volume 1, pages 193–220. John Wiley & Sons, Chichester, UK, 2005.

Sivapalan M., Blöschl G., Zhang L. and Vertessy R. Downward approach to hydrological prediction. *Hydrological Processes*, 17:2101–2111, 2003.

Sivapalan M., Wood E. F. and Beven K. J. On hydrological similarity: 3. A dimensionless flood frequency distribution. *Water Resources Research*, 26:43–58, 1990.

Sklash M. G. Environmental isotope studies of storm and snowmelt generation. In Anderson M. G. and Burt T. P., editors, *Process Studies in Hillslope Hydrology*, pages 401–435. John Wiley & Sons, Chichester, UK, 1990.

Sklash M. G., Beven K. J., Gilman K. and Darling W. G. Isotope studies of pipeflow at Plynlimon, Wales, UK. *Hydrological Processes*, 10:921–944, 1996.

Sklash M. G. and Farvolden R. N. The role of groundwater in storm runoff. *Journal of Hydrology*, 43:43–65, 1979.

Smakhtin V. U. Some early Russian studies of subsurface storm-flow processes. *Hydrological Processes*, 16:2613–2620, 2002.

De Smedt F., Brevis W. and Debels P. Analytical solution for solute transport resulting from instantaneous injection in streams with transient storage. *Journal of Hydrology*, 315:25–39, 2005.

Smith A. A. A generalized approach to kinematic flood routing. *Journal of Hydrology*, 45(1-2):71–89, 1980.

Smith J. A., Baeck M. L., Steiner M. and Miller A. J. Catastrophic rainfall from an upslope thunderstorm in the central Appalachians: The rapidan storm of June 27th, 1995. *Water Resources Research*, 32(10):3099–3113, 1996.

Smith L. A. Disentangling Uncertainty and Error: on the predictability of nonlinear systems, in A. Mees, editor, *Nonlinear Dynamics & Statistics*, pages 31–64. Springer, Berlin, 2000.

Smith P. J., Beven K. J. and Tawn J. A. Detection of structural inadequacy in process-based hydrological models: A particle-filtering approach. *Water Resources Research*, 44(1):W01410, DOI:10.1029/2006WR005205, 2008b.

Smith P. J., Tawn J. and Beven K. J. Informal likelihood measures in model assessment: Theoretic development and investigation. *Advances in Water Resources*, 31:1087–1100, 2008a.

Smith R. E. Approximate soil water movement by kinematic characteristics. *Soil Science Society of America Journal*, 47:3–8, 1983.

Smith R. E., Corradini C. and Melone F. A conceptual model for infiltration and redistribution in crusted soils. *Water Resources Research*, 35:1385–1393, 1999.

Smith R. E. and Goodrich D. C. Model for rainfall excess patterns on randomly heterogeneous areas. *ASCE Journal of Hydrologic Engineering*, 5:355–362, 2000.

Smith R. E., Goodrich D. C., Woolhiser D. A. and Unkrich C. L. KINEROS – a KINematic runoff and EROSion model. In Singh V. P., editor, *Computer Models of Watershed Hydrology*, pages 697–732. Water Resource Publications, Highlands Ranch, CO, 1995.

Smith R. E. and Parlange J.-Y. A parameter efficient infiltration model. *Water Resources Research*, 14:533–538, 1978.

Smith R. E. and Woolhiser D. A. Overland flow on an infiltrating surface. *Water Resources Research*, 7(4):899, 1971.

Smith R. N. B., Blyth E. M., Finch J. W., Goodchild S., Hall R. L. and Madry S. Soil state and surface hydrology diagnosis based on MOSES in the Met Office Nimrod nowcasting system. *Meteorological Applications*, 13:89–109, 2006.

Smith T. J. and Marshall L. A. Bayesian methods in hydrologic modeling: A study of recent advancements in Markov Chain Monte Carlo techniques. *Water Resources Research*, 44:W00B05, DOI:10.1029/2007WR006705, 2008.

Snell J. D. and Sivapalan M. Application of the meta-channel concept: Construction of the meta-channel hydraulic geometry for a natural channel. *Hydrological Processes*, 9:485–505, 1995.

Solomatine D. P. and Dulal K. N. Model trees as an alternative to neural networks in rainfall–runoff modelling. *Hydrological Sciences Journal*, 48(3):399–411, 2003.

Sorooshian S., Duan Q. and Gupta V. K. Calibration of the SMA-NWSRFS conceptual rainfall–runoff model using global optimisation. *Water Resources Research*, 29:1185–1194, 1992.

Sorooshian S. and Gupta V. K. Model calibration. In Singh V. P., editor, *Computer Models of Watershed Hydrology*, pages 23–68. Water Resource Publications, Highlands Ranch, CO, 1995.

Sorooshian S., Gupta V. K. and Fulton J. L. Evaluation of maximum likelihood parameter estimation techniques for conceptual rainfall–runoff models: Influence of calibration data variability and length on model credibility. *Water Resources Research*, 19:251–259, 1983.

Soulsby C., Malcolm R., Ferrier R. C., Helliwell R. C. and Jenkins A. Isotope hydrology of the Allt a'Mharcaidh catchment, Cairngorms, Scotland: Implications for hydrological pathways and residence times. *Hydrological Processes*, 14:747–762, 2000.

Soulsby C., Petry J., Brewer M. J., Dunn S. M., Otta B. and Malcolm I. A. Identifying and assessing uncertainty in hydrological pathways: A novel approach to end member mixing in a Scottish agricultural catchment. *Journal of Hydrology*, 274:109–128, 2003.

Spear R. C., Grieb T. M. and Shang N. Parameter uncertainty and interaction in complex environmental models. *Water Resources Research*, 30:3159–3170, 1994.

Srinivasan R., Zhang X. and Arnold J. SWAT ungauged: Hydrological budget and crop yield predictions in the upper Mississippi river basin. *Transactions of the ASABE*, 53:1533–1546, 2010.

Stadler D., Wunderli H., Auckenthaler A. and Flühler H. Measurement of frost-induced snowmelt runoff in a forest soil. *Hydrological Processes*, 10:1293–1304, 1996.

Stedinger J. R., Vogel R. M., Lee S. U. and Batchelder R. Appraisal of the generalized likelihood uncertainty estimation (GLUE) method. *Water Resources Research*, 44:W00B06, DOI:10.1029/2008WR006822, 2008.

Steel M. E., Black A. R., Werrity A. and Littlewood I. G. Reassessment of flood risk for Scottish rivers using synthetic runoff data. In *Hydrological Extremes: Understanding, Predicting, Mitigating*, Int. Assoc. Sci. Hydrol. Publ. No. 255:209–215, Wallingford, UK, 1999.

Steenhuis T. S., Winchell M., Rossing J., Zollweg J. A. and Walter M. F. SCS runoff equation revisited for variable source runoff areas. *ASCE Journal of Irrigation and Drainage Engineering*, 121:234–238, 1995.

Stephenson G. R. and Freeze R. A. Mathematical simulation of subsurface flow contributions to snowmelt runoff, Reynolds Creek, Idaho. *Water Resources Research*, 10(2):284–298, 1974.

Stewart M. K., Morgenstern U. and McDonnell J. J. Truncation of stream residence time: How the use of stable isotopes has skewed our concept of streamwater age and origin. *Hydrological Processes*, 24(12):1646–1659, 2010.

Stoker J. J. *Water Waves*. Interscience, New York, 1957.

Stone J. J., Nichols M. H., Goodrich D. C. and Buono J. Long-term runoff database, Walnut Gulch experimental watershed, Arizona, United States. *Water Resources Research*, 44:W05S05, DOI:10.1029/2006WR005733, 2008.

Struthers I. and Sivapalan M. A conceptual investigation of process controls upon flood frequency: Role of thresholds. *Hydrology and Earth System Sciences*, 11:1405–1416, 2007.

Su Z. The surface energy balance system (SEBS) for estimation of turbulent heat fluxes. *Hydrology and Earth System Sciences*, 6:85–99, 2002.

Sun W. C., Ishidaira H. and Bastola S. Towards improving river discharge estimation in ungauged basins: Calibration of rainfall–runoff models based on satellite observations of river flow width at basin outlet. *Hydrology and Earth System Sciences*, 14:2011–2022, 2010.

Surkan A. J. Synthetic hydrographs: Effects of network geometry. *Water Resources Research*, 5:115–128, 1969.

Szilagyi L. and Jozsa J. New findings about the complementary relationship based evaporation estimation methods. *Journal of Hydrology*, 354:171–186, 2008.

Szolgay J., Hlavcŏvá K., Kohnová S. and Danihlík R. Regional estimation of parameters of a monthly water balance model. *Journal of Hydrology and Hydromechanics*, 51:256– 273, 2003.

Tabrizi M. H. N., Said S. E., Badr A. W., Mashor Y. and Billings S. A. Nonlinear modelling and prediction of a river flow system. *Journal of the American Water Resources Association*, 34:1333–1339, 1998.

Taha A., Gresillon J. M. and Clothier B. E. Modelling the link between hillslope water movement and stream flow: Application to a small Mediterranean forest watershed. *Journal of Hydrology*, 203:11–20, 1997.

Takeuchi K., Hapuarachchi P., Zhou M., Ishidaira H. and Magome J. A BTOP model to extend TOPMODEL for distributed hydrological simulation of large basins. *Hydrological Processes*, 22:3236–3251, 2008.

Tarantola A. *Inverse Problem Theory*. Elsevier, New York, 1987.

Tarantola A. Popper, Bayes and the inverse problem. *Nature Physics*, 2:492–494, 2006.

Tarboton D. G. A new method for the determination of flow directions and upslope areas in grid digital elevation models. *Water Resources Research*, 33:309–319, 1997.

Tardieu F. and Davies W. J. Integration of hydraulic and chemical signalling in the control of stomatal conductance and water status of droughted plants. *Plant, Cell and Environment*, 16:341–349, 1993.

Tartakovsky A. M., Meakin P., Scheibe T. D. and Wood B. D. A smoothed particle hydrodynamics model for reactive transport and mineral precipitation in porous and fractured porous media. *Water Resources Research*, 43:W05437, DOI:10.1029/2005WR004770, 2007.

Tayfur G. and Kavvas M. L. Areally-averaged overland flow equations at hillslope scale. *Journal of Hydrology*, 43:361–378, 1998.

Taylor C. J., Pedregal D. J., Young P. C. and Tych W. Environmental time series analysis and forecasting with the CAPTAIN toolbox. *Environmental Modelling & Software*, 22:797–814, 2007.

Teuling A. J., Lehner I., Kirchner J. W. and Seneviratne S. I. Catchments as simple dynamical systems: Experience from a Swiss prealpine catchment. *Water Resources Research*, 46:W10502, 2010.

Therrien R., McLaren R., Sudicky E. A. and Panday S. HydroGeoSphere: a three-dimensional numerical model describing fully-integrated subsurface and surface flow and solute transport. Technical report, Groundwater Simulation Group, University of Waterloo, Waterloo, Ontario, Canada, 2006.

Thielen J., Bartholmes J., Ramos M.-H. and de Roo A. The European Flood Alert System EFAS – Part 1: Concept and development. *Hydrology and Earth System Sciences*, 13:125–140, 2009.

Thiemann M., Trosset M., Gupta H. and Sorooshian S. Bayesian recursive parameter estimation for hydrologic models. *Water Resources Research*, 37(10):2521–2535, 2001.

Thirumalaiah K. and Deo M. C. River stage forecasting using artificial neural networks. *ASCE Journal of Hydrologic Engineering*, 3:26, 1998.

Thorndahl S., Beven K. J., Jensen J. B. and Schaarup-Jensen K. Event based uncertainty assessment in urban drainage modelling, applying the GLUE methodology. *Journal of Hydrology*, 357:421–437, 2008.

Thyer M., Kuczera G. and Bates B. C. Probabilistic optimisation for conceptual rainfall runoff models: A comparison of the shuffled complex evolution and simulated annealing algorithms. *Water Resources Research*, 35:767–774, 1999.

Thyer M., Leonard M., Kavetski D., Need S. and Renard B. The open source RFortran library for accessing R from fortran, with applications in environmental modelling. *Environmental Modelling & Software*, 26:219–234, 2011.

Thyer M., Renard B., Kavetski D., Kuczera G., Franks S. and Srikanthan S. Critical evaluation of parameter consistency and predictive uncertainty in hydrological modelling: A case study using Bayesian total error analysis. *Water Resources Research*, 45:W00B14, DOI:10.1029/2008WR006825, 2009.

Tian F., Hu H., Lei Z. and Sivapalan M. Extension of the representative elementary watershed approach for cold regions via explicit treatment of energy related processes. *Hydrology and Earth System Sciences*, 10(5):619–644, 2006.

Tietje O. and Tapkenhinrichs M. Evaluation of pedo-transfer functions. *Soil Science Society of America Journal*, 57:1088–1095, 1993.

Tillotson P. M. and Nielsen D. R. Scale factors in soil science. *Soil Science Society of America Journal*, 48:953–959, 1984.

Todini E. New trends in modelling soil processes from hillslope to GCM scales. In Oliver H. R. and Oliver S. A., editors, *The Role of Water and the Hydrological Cycle in Global Change*, pages 317–347. NATO ASI Series Vol. I 31, 1995.

Todini E. The ARNO rainfall–runoff model. *Journal of Hydrology*, 175:339–382, 1996.

Todini E. Role and treatment of uncertainty in real-time flood forecasting. *Hydrological Processes*, 18:2743–2746, 2004.

Todini E. A mass conservative and water storage consistent variable parameter Muskingum–Cunge approach. *Hydrology and Earth System Sciences*, 11:1645–1659, 2007.

Todini E. and Mantovan P. Comment on "On undermining the science?" by Keith Beven. *Hydrological Processes*, 21(12):1633–1638, 2007.

Tokar A. S. and Johnson P. A. Rainfall–runoff modeling using artificial neural networks. *ASCE Journal of Hydrologic Engineering*, 4:232, 1999.

Toni T., Welch D., Strelkowa N., Ipsen A. and Stumpf M. P. H. Approximate Bayesian computation scheme for parameter inference and model selection in dynamical systems. *J. Roy. Soc. Interface*, 6:187–202, 2009.

Tonkin M. and Doherty J. Calibration-constrained Monte Carlo analysis of highly parameterized models using subspace techniques. *Water Resources Research*, 45:W00B10, DOI:10.1029/2007WR006678, 2009.

Sempere Torres D., Rodriguez J.-Y. and Obled C. Using the DPFT approach to improve flash flood forecasting models. *Natural Hazards*, 5:17–41, 1992.

Trigg M. A., Wilson M. D., Bates P. D., Horritt M. S., Alsdorf D. E., Forsberg B. R. and Vega Corredor M. C. Amazon flood wave hydraulics. *Journal of Hydrology*, 374:92–105, DOI:10.1016/j.jhydrol.2009.06.004, 2009.

Troch P., van Loon E. and Hilberts A. Analytical solutions to a hillslope storage kinematic wave equations for subsurface flow. *Advances in Water Resources*, 25:637–649, 2002.

Turner H. M. and Burdoin A. S. The flood hydrograph. *J. Boston Soc. Civ. Eng.*, 28:232–256, 1941.

Tych W. and Young P. C. A MATLAB software framework for dynamic model emulation. *Environmental Modelling & Software*, doi:10.1016/j.envsoft.2011.08.008, 2011.

Tyler S. W. and Wheatcraft S. W. Fractal scaling of soil particle size distributions: Analysis and limitations. *Soil Science Society of America Journal*, 56:362–369, 1992.

Uhlenbrook S. and Leibundgut C. Process-oriented catchment modelling and multiple-response validation. *Hydrological Processes*, 16:423–440, 2002.

USDA-SCS. *National Engineering Handbook, Supplement A, Section 4, Chapter 10: Hydrology*. US Department of Agriculture, Washington DC, 1985.

USDA-SCS. *Urban Hydrology for Small Watersheds, Technical Release 55*. US Department of Agriculture, Washington DC, 1986.

USDA-SCS. *STATSGO: State Soils Geographic Data Base*. Soil Conservation Service, Publication No. 1492, Washington DC, 1992.

Vaché K. B. and McDonnell J. J. A process-based rejectionist framework for evaluating catchment runoff model structure. *Water Resources Research*, 42:W02409, DOI:10.1029/2005WR004247, 2006.

van de Beek C. Z., Leijnse H., Stricker J. N. M., Uijlenhoet R. and Russchenberg H. W. J. Performance of high-resolution X-band radar for rainfall measurement in The Netherlands. *Hydrology and Earth System Sciences*, 14:205–221, 2010.

van der Knijff J. M., Younis J. and de Roo A. P. J. LISFLOOD: A GIS-based distributed model for river basin scale water balance and flood simulation. *International Journal of Geographical Information Science*, 24:189–212, 2008.

van Genuchten M. T. A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. *Soil Science Society of America Journal*, 44:892–898, 1980.

van Genuchten M. T., Leij F. J. and Lund L. J. editors. *Indirect Methods for Estimating the Hydraulic Properties of Unsaturated Soils*, USDA, Salinity Lab., Riverside, CA., 1989.

van Genuchten M. T., Schaap M. G., Mohanty B. P., Simunek J. and Leij F. J. Modelling flow and transport processes at the local scale. In Feyen J. and Wiyo K., editors, *Modelling Transport Processes in Soils*, pages 23–45, Wageningen, The Netherlands, 1999.

van Griensven A., Meixner T., Grunwald S., Bishop T., Diluzio M. and Srinivasan R. A global sensitivity analysis tool for the parameters of multi-variable catchment models. *Journal of Hydrology*, 324:10–23, 2006.

Tromp van Meerveld H. J. and McDonnell J. J. Threshold relations in subsurface stormflow: 2. The fill and spill hypothesis. *Water Resources Research*, 42:W02411, DOI: 10.1029/2004WR003800, 2006.

van Straten G. and Keesman K. Uncertainty propagation and speculation in projective forecasts of environmental change: A lake eutrophication example. *Journal of Forecasting*, 10:163–190, 1991.

VanderKwaak J. E. and Loague K. Hydrologic-response simulations for the R-5 catchment with a comprehensive physics-based model. *Water Resources Research*, 37:999–1013, 2001.

Vapnik V. *The Nature of Statistical Learning Theory*. Springer, New York, 1995.

Varado N., Braud I., Galle S., Le Lay M., Séguis L., Kamagate B. and Depraetere C. Multi-criteria assessment of the representative elementary watershed approach on the Donga catchment (Benin) using a downward approach of model complexity. *Hydrology and Earth System Sciences*, 10(3):427–442, 2006.

Vázquez R., Beven K. J. and Feyen J. GLUE based assessment on the overall predictions of a MIKE-SHE application. *Water Resources Management*, 23(7):1325–1349, 2009.

Vázquez R. F., Feyen L., Feyen J. and Refsgaard J. C. Effect of grid size on effective parameters and model performance of the MIKE-SHE code. *Hydrological Processes*, 16:355–372, 2002.

Venetis C. The IUH of the Muskingum channel reach. *Journal of Hydrology*, 7:444–447, 1969.

Venturini V., Islam S. and Rodriguez L. Estimation of evaporative fraction and evapotranspiration from MODIS products using a complementary based model. *Remote Sensing of Environment*, 112:132–141, 2008.

Vereeken H., Maes J. and Feyen J. Estimating unsaturated hydraulic conductivity from easily measured soil properties. *Soil Sci.*, 149:1–11, 1990.

Vereeken H., Maes J., Feyen J. and Darius P. Estimating the soil moisture retention characteristics from texture, bulk density and carbon content. *Soil Science*, 148:389–403, 1989.

Verhoest N. E. C., Vandenberghe S., Cabus P., Onof C., Meca-Figueras T. and Jameleddine S. Are stochastic point rainfall models able to preserve extreme flood statistics? *Hydrological Processes*, 24:3439–3445, 2010.

Vernieuwe H., Georgieva O., De Baets B., Pauwels V. R. N., Verhoest N. E. C. and De Troch F. P. Comparison of data-driven Takagi–Sugeno models of rainfall-discharge dynamics. *Journal of Hydrology*, 302:173–186, 2005.

Vertessy R. A. and Elsenbeer H. Distributed modelling of storm flow generation in an Amazonian rain forest catchment: Effects of model parameterisation. *Water Resources Research*, 35:2173–2187, 1999.

Vertessy R. A., Hatton T. J., O'Shaughnessy P. J. and Jayasuriya M. D. A. Predicting water yield from a mountain ash forest using a terrain analysis based catchment model. *Journal of Hydrology*, 150:665–700, 1993.

Daluz Vieira J. H. Conditions governing the use of approximations for the St. Venant equations for shallow surface water flow. *Journal of Hydrology*, 60:43–58, 1983.

Vieux B. E. *Distributed Hydrologic Modeling Using GIS*. Springer-Verlag, Berlin, 2004.

Vincendon B., Ducrocq V., Bouilloudd L., Saulnier G.-M., Chancibaulte K., Habets F. and Noilhan J. Benefit of coupling the ISBA land surface model with a TOPMODEL hydrological model version dedicated to Mediterranean flash floods. *Journal of Hydrology*, 394:256–266, 2010.

Viney N. R., Bormann H. and Breuer L. *et al.* Assessing the impact of land use change on hydrology by ensemble modelling (LUCHEM) II: Ensemble combinations and predictions. *Advances in Water Resources*, 32:147–158, 2009.

Vinogradov U. B., Semenova O. M. and Vinogradova T. A. An approach to the scaling problem in hydrological modelling: The deterministic modelling hydrological system. *Hydrological Processes*, 25(7):1055–1073, 2011.

Viviroli D., Zappa M., Gurtz J. and Weingartner R. An introduction to the hydrological modelling system PREVAH and its pre- and post-processing tools. *Environmental Modelling & Software*, 24:1209–1222, 2009.

Vogel, R. M. Regional calibration of watershed models, in V. P. Signh and D. F. Frevert, editors, *Watershed Models*, pages 549–567. CRC Press, Boca Raton, FL, 2005.

Vrugt J. A., Gupta H. V., Bastidas L. A., Bouten W. and Sorooshian S. Effective and efficient algorithm for multiobjective optimization of hydrologic models. *Water Resources Research*, 39:W01214, DOI:10.1029/2002WR001746, 2003.

Vrugt J. A., ter Braak C. J. F., Diks C. G. H., Roberts D. A., Hyman J. M. and Higdon D. Accelerating Markov Chain Monte Carlo simulation by differential evolution with self-adaptive randomized subspace sampling. *International Journal of Nonlinear Sciences and Numerical Simulation*, 10:273, 2009.

Vrugt J. A., ter Braak C. J. F., Gupta H. V. and Robertson D. Equifinality of formal (DREAM) and informal (GLUE) Bayesian approaches in hydrologic modeling? *Stochastic Environmental Research and Risk Assessment*, 23(7):1011–1026, 2008b.

Vrugt J. A., ter Braak C. J. F., Clark M. P., Hyman J. M. and Robinson B. A. Treatment of input uncertainty in hydrologic modeling: Doing hydrology backward with Markov Chain Monte Carlo simulation. *Water Resources Research*, 44:W00B09, DOI:10.1029/2007WR006720, 2008a.

Wagener T., McIntyre N., Lees M. J., Wheater H. S. and Gupta H. V. Towards reduced uncertainty in conceptual rainfall–runoff modelling: Dynamic identifiability analysis. *Hydrological Processes*, 17:455–476, 2003.

Wagener T. and Montanari A. A paradigm shift in predictions in ungauged basins (PUB). *Water Resources Research*, 2011.

Wagener T., Sivapalan M., Troch P. and Woods R. Catchment classification and hydrologic similarity. *Geography Compass*, 1:901–931, 2007.

Wagener T., Sivapalan M., Troch P. A., McGlynn B. L., Harman C. J., Gupta H. V., Kumar P., Rao P. S. C., Basu N. B. and Wilson J. S. The future of hydrology: An evolving science for a changing world. *Water Resources Research*, 46:W05301, DOI:10.1029/2009WR008906, 2010.

Wagener T., Wheater H. and Gupta H. V. *Rainfall–runoff Modelling in Gauged and Ungauged Catchments*. Imperial College Press, London, 2004.

Wagener T. and Wheater H. S. Parameter estimation and regionalization for continuous rainfall–runoff models including uncertainty. *Journal of Hydrology*, 320:132–154, 2006.

Wagner B., Tarnawski V., Wessolek G. and Plagge R. Suitability of models for the estimation of soil hydraulic parameters. *Geoderma*, 86:229–239, 1998.

Wallace J. S. Calculating evaporation: Resistance to factors. *Agricultural and Forest Meteorology*, 73:353–366, 1995.

Wallis S. G., Young P. C. and Beven K. J. Experimental investigation of the aggregated dead zone model for longitudinal solute transport in stream channels. *Proc. Inst. Civ. Eng., Part 2*, 87:1–22, 1989.

Wang A., Li K. Y. and Lettenmaier D. P. Integration of the variable infiltration capacity model soil hydrology scheme into the community land model. *Journal of Geophysical Research*, 113:D09111, DOI:10.1029/2007JD009246, 2008.

Wang Q. J. The genetic algorithm and its application to calibrating conceptual rainfall–runoff models. *Water Resources Research*, 27:2467–2471, 1991.

Wang Z., Feyen J., van Genuchten M. T. and Nielsen D. R. Air entrapment effects on infiltration rates and flow instability. *Water Resources Research*, 34:213–222, 1998.

Warrick A. W. and Hussen A. A. Scaling of Richards equation for infiltration and drainage. *Soil Science Society of America Journal*, 57:15–18, 1993.

Watson F. G. R., Vertessy R. A. and Grayson R. B. Large-scale modelling of forest hydrological processes and their long-term effect on water yield. *Hydrological Processes*, 13:689–700, 1999.

Weerts A. H., Winsemius H. C. and Verkade J. S. Estimation of predictive hydrological uncertainty using quantile regression: Examples from the National Flood Forecasting System (England and Wales). *Hydrology and Earth System Sciences*, 15:255–265, 2011.

Weiler M. and McDonnell J. J. Virtual experiments: A new approach for improving process conceptualization in hillslope hydrology. *Journal of Hydrology*, 285:3–18, 2004.

Weiler M. and Naef F. Simulating surface and subsurface initiation of macropore flow. *Journal of Hydrology*, 273:139–154, 2003.

Weinberg A. Trans-science. *Minerva*, 10:209–222, 1972.

Weltz M. A., Ritchie J. C. and Fox H. D. Comparison of laser and field measurements of vegetation height and canopy cover. *Water Resources Research*, 30:1311–1319, 1994.

Werner M. G. F., van Dijk M. and Schellekens J. DELFT-FEWS: An open shell flood forecasting system. In Liong S. Y., Phoon K. K. and Babovic C. V. editors, *Proceedings of the 6th International Conference on Hydroinformatics*, pages 1205–1212, World Scientific Publishing Company, Singapore, 2004.

Westerberg I. K., Guerrero J.-L., Seibert J., Beven K. J. and Halldin S. Stage-discharge uncertainty derived with a non-stationary rating curve in the Choluteca River, Honduras. *Hydrological Processes*, 25:603–613, DOI: 10.1002/hyp.7848., 2010a.

Westerberg I. K., Guerrero J.-L., Younger P. M., Beven K. J., Seibert J., Halldin S., Freer J. and Xu C.-Y. Calibration of hydrological models using flow-duration curves. *Hydrology and Earth System Sciences Discussion*, 7:9467–9522, 2010b.

Western A. W., Grayson R. B., Blöschl G., Willgoose G. and McMahon T. A. Observed spatial organisation of soil moisture and its relation to terrain indices. *Water Resources Research*, 35:797–810, 1999.

Westhoff M. C., Savenije H. H. G., Luxemburg W. M. J., Stelling G. S., van de Giesen N. C., Selker J. S., Pfister L. and Uhlenbrook S. A distributed stream temperature model using high resolution temperature observations. *Hydrology and Earth System Sciences*, 11:1469–1480, 2007.

Weyman D. R. Throughflow on hillslopes and its relation to the stream hydrograph. *Hydrological Sciences Bulletin*, 15:25–33, 1970.

Wheater H. S., Jakeman A. J. and Beven K. J. Progress and directions in rainfall–runoff modelling. In Jakeman A. J., Beck M. B., and McAleer M. J., editors, *Modelling Change in Environmental Systems*, pages 101–132. John Wiley & Sons, Chichester, UK, 1993.

White E. D., Easton Z. M., Fuka D. R., Collick A. S., Adgo E., McCartney M., Awulachew S. B., Selassie Y. and Steenhuis T. S. Development and application of a physically based landscape water balance in the SWAT model. *Hydrological Processes*, 25:915–925, 2011.

Whitehead P. G., Young P. C. and Hornberger G. M. A systems model of streamflow and water quality in the Bedford-Ouse River: 1. Streamflow modelling. *Water Resources Research*, 13:1155–1169, 1979.

Whittaker G., Confesor Jr R., Di Luzio M. and Arnold J. G. Detection of overparameterisation and overfitting in an automatic calibration of SWAT. *Transactions of the ASABE*, 53:1487–1499, 2010.

Wigmosta M. S. and Lettenmaier D. P. A comparison of simplified methods for routing topographically driven subsurface flow. *Water Resources Research*, 35(1):255–264, 1999.

Wigmosta M. S., Vail L. W. and Lettenmaier D. P. A distributed hydrology–vegetation model for complex terrain. *Water Resources Research*, 30(6):1665–1679, 1994.

Wilby R. L. When and where might climate change be detectable in UK river flows? *Geophysical Research Letters*, 33:L19407, 2006.

Wilby R. L., Beven K. J. and Reynolds N. Climate change and fluvial flood risk in the UK: More of the same? *Hydrological Processes*, 22(14):2511–2523, 2008.

Willgoose G. and Kuczera G. Estimation of subgrid scale kinematic wave parameters for hillslopes. *Hydrological Processes*, 9(3-4):469–482, 1995.

Williams J. R. The EPIC model. In Singh V. P., editor, *Computer Models of Watershed Hydrology*, pages 909–1000, Water Resource Publications, Highlands Ranch, CO, 1995.

Wilson M., Bates P. D., Alsdorf D., Forsberg B., Horritt M., Melack J., Frappart F. and Famiglietti J. Modeling large-scale inundation of Amazonian seasonally flooded wetlands. *Geophysical Research Letters*, 34:L15404, DOI:. 10.1029/2007GL030156, 2007.

Wimsatt W. *Re-engineering Philosophy for Limited Beings: Piecewise approximations to reality*. Harvard University Press, Cambridge, MA, 2007.

Winsemius H. C., Schaefli B., Montanari A. and Savenije H. H. G. On the calibration of hydrological models in ungauged basins: A framework for integrating hard and soft hydrological information. *Water Resources Research*, 45:W12422, DOI:10.1029/2009WR007706, 2009.

WMO. Guide for hydrometeorological practices. Technical report, World Meteorological Organisation, Geneva, 1964.

WMO. Intercomparison of conceptual models used in operational hydrological forecasting. Technical Report Operational Hydrology Report 7, WMO429, World Meteorological Organisation, Geneva, 1975.

WMO. Intercomparison of models of snowmelt runoff. Technical Report Operation Hydrology Report No. 23, World Meteorological Organisation, Geneva, 1986.

Wolock D. M. Simulating the variable source area concept of streamflow generation with the watershed model TOPMODEL. Water-Resources Investigations Report 93-4124. Technical report, US Geological Survey, Lawrence, Kansas, 1993. 33 pp.

Wolock D. M. and Hornberger G. M. Hydrological effects of changes in levels of atmospheric carbon dioxide. *Journal of Forecasting*, 10:105–116, 1991.

Wong T. H. F. and Laurenson E. M. Wave speed–discharge relations in natural channels. *Water Resources Research*, 19:701–706, 1983.

Wong T. S. W. How nature uses overland flow regime to maximize equilibrium detention storage and flood attenuation. *Hydrological Processes*, 22:5004–5012, 2008.

Wood E. F., Lettenmaier D. P., Xu Lianga, Nijssen B. and Wetzel S. W. Hydrological modelling of continental-scale basins. *Annual Review of Earth and Planetary Sciences*, 25:279–300, DOI: 10.1146/annurev.earth.25.1.279, 1997.

Wood E. F., Lettenmaier D. P. and Zatarian V. G. A land-surface hydrology parameterisation with sub-grid variability for general circulation models. *Journal of Geophysical Research*, 97(D3):2717–2728, 1992.

Wood E. F., Lin D.-S., Mancini M., Thongs D., Troch P. A., Jackson T. J., Famiglietti J. S. and Engman E. T. Intercomparisons between passive and active microwave remote sensing and hydrological modeling for soil moisture. *Advances in Space Research*, 13:167–176, 1993.

Wood E. F., Sivapalan M. and Beven K. J. Similarity and scale in catchment storm response. *Reviews of Geophysics*, 28:1–18, 1990.

Wood S. J., Jones D. A. and Moore R. J. Accuracy of rainfall measurement for scales of hydrological interest. *Hydrology and Earth System Sciences*, 4:531–543, 2000.

Woolhiser D. A. and Goodrich D. C. Effect of storm rainfall intensity patterns on surface runoff. *Journal of Hydrology*, 102(1–4):335–354, 1988.

Wörman A. Comparison of models for transient storage of solutes in small streams. *Water Resources Research*, 36:455–468, 2000.

Wösten J. H. M. The HYPRES database of hydraulic properties of European soils. In Feyen J. and Wiyo K., editors, *Modelling Transport Processes in Soils*, pages 675–681, Wageningen, The Netherlands, 1999.

Wright I. R., Gash J. H. C., Da Rocha H. R., Shuttleworth W. J., Nobre C. A., Maitelli G. T., Zamparoni C. and Carvalho P. R. A. Dry season micrometeorology of central Amazonian ranchland. *Quarterly Journal of the Royal Meteorological Society*, 118(508):1083–1099, 1992.

Wu L., Seo D.-J., Demargne J., Brown J. D., Cong S. and John Schaake. Generation of ensemble precipitation forecast from single-valued quantitative precipitation forecast for hydrologic ensemble prediction. *Journal of Hydrology*, 399:281–298, 2011.

Xevi E., Christiaens K., Espino A., Sewnandan W., Mallants D., Sorensen H. and Feyen J. Calibration, validation and sensitivity analysis of the MIKE SHE model using the Neuenkirchen catchment as case study. *Water Resources Management*, 11:219–242, 1997.

Xie L., Wei G., We. Deng and Zhao X. Daily δ^{18}O and δD of precipitations from 2007 to 2009 in Guangzhou, South China: Implications for changes of moisture sources. *Journal of Hydrology*, 400:477–489, 2011.

Xie Z., Yuan F., Duan Q., Zheng J., Liang M. and Chen F. Regional parameter estimation of the VIC land surface model: Methodology and application to river basins in China. *Journal of Hydrometeorology*, 8:447–468, 2007.

Xinmei H., Lyons T. J., Smith R. C. G. and Hacker J. M. Estimation of land surface parameters using satellite data. *Hydrological Processes*, 9:631–643, 1995.

Xiong L., Shamseldin A. and O'Connor K. A non linear combination of the forecasts of rainfall–runoff models by the first order Takagi–Sugeno fuzzy system. *Journal of Hydrology*, 245:196–217, 2001.

Xu C.-Y. Estimation of parameters of a conceptual water balance model for ungauged catchments. *Water Resources Management*, 13:353–368, 1999.

Xu C.-Y. and Singh V. P. Cross comparison of empirical equations for calculating potential evapotranspiration with data from Switzerland. *Water Resources Management*, 16:197–219, 2002.

Yadav M., Wagener T. and Gupta H. Regionalization of constraints on expected watershed response behavior for improved predictions in ungauged basins. *Advances in Water Resources*, 30:1756–1774, 2007.

Yapo P. O., Gupta H. V. and Sorooshian S. Calibration of conceptual rainfall–runoff models: Sensitivity to calibration data. *Journal of Hydrology*, 181:23–48, 1996.

Yapo P. O., Gupta H. V. and Sorooshian S. Multi-objective global optimization for hydrologic models. *Journal of Hydrology*, 204:83–97, 1998.

Yatheendradas S., Wagener T., Gupta H. V., Unkrich C., Goodrich D., Schaffner M. and Stewart A. Understanding uncertainty in distributed flash flood forecasting for semiarid regions. *Water Resources Research*, 44:W05S19, DOI:10.1029/2007WR005940, 2008.

Young P. C. Recursive approaches to time series analysis. *Bulletin of the Institute of Mathematics and its Applications*, 10:209–224, 1975.

Young P. C. The validity and credibility of models for badly defined systems. In Beck M. B. and G. van Straten, editors, *Uncertainty and Forecasting of Water Quality*, pages 69–98, Springer-Verlag, New York, 1983.

Young P. C. *Recursive Estimation and Time Series Analysis*. Springer-Verlag, Berlin, 1984.

Young P. C. Parallel processes in hydrology and water-quality: A unified time-series approach. *Journal of the Institution of Water & Environmental Management*, 6(5):598–612, 1992.

Young P. C. Time variable and state dependent parameter modelling of nonstationary and nonlinear time series. In Subba T. Rao, editor, *Developments in Time Series*, pages 374–413. Chapman and Hall, London, 1993.

Young P. C. Data-based mechanistic modelling of environmental, ecological, economic and engineering systems. *Environmental Modelling & Software*, 13:105–122, 1998.

Young P. C. Data-based mechanistic modelling and validation of rainfall-flow models. In Anderson M. G. and Bates P. D., editors, *Validation of Hydrological Models*. John Wiley & Sons, Chichester, UK, 2000.

Young P. C. *Model Validation: Perspectives in Hydrological Science*, pages 117–161. John Wiley, Chichester, UK, 2001.

Young P. C. Advances in real-time flood forecasting. *Philosophical Transactions of the Royal Society London*, A360:1433–1450, 2002.

Young P. C. Top-down and data-based mechanistic modelling of rainfall–flow dynamics at the catchment scale. *Hydrological Processes*, 17:2195–2217, 2003.

Young P. C. *Recursive Estimation and Time-Series Analysis*. Springer-Verlag, Berlin, 2011a.

Young P. C. Real-time updating in flood forecasting and warning. In Pender G. and Faulkner H., editors, *Flood Risk Science and Management* Wiley-Blackwell, Chichester, UK, 2011b.

Young P. C. and Beven K. J. Computation of the instantaneous unit-hydrograph and identifiable component flows with application to two small upland catchments: Comment. *Journal of Hydrology*, 129(1-4):389–396, 1991.

Young P. C. and Beven K. J. Data-based mechanistic modelling and the rainfall-flow nonlinearity. *Environmetrics*, 5(3):335–363, 1994.

Young P. C., Castelletti A. and Pianosi F. *Topics on System Analysis and Integrated Water Resource Management*, pages 27–48. Elsevier, Amsterdam, 2007.

Young P. C., Jakeman A. J. and Post D. A. Recent advances in the data-based modelling and analysis of hydrological systems. *Water Science and Technology*, 36:99–116, 1997.

Young P. C., Leedal D. and Beven K. J. Reduced order emulation of distributed hydraulic models. In *Proceedings 15th IFAC Symposium on System Identification SYSID09*, St Malo, France, 2009.

Young P. C. and Ratto M. A unified approach to environmental systems modeling. *Stochastic Environmental Research and Risk Assessment*, 23:1037–1057, 2009.

Young P. C. and Wallis S. G. Solute transport and dispersion in channels. In Beven K. J. and Kirkby M. J., editors, *Channel Network Hydrology*, pages 129–174. John Wiley & Sons, Chichester, UK, 1993.

Young R. A., Onstad C. A. and Bosch D. D. AGNPS: An agricultural nonpoint source model. In Singh V. P., editor, *Computer Models of Watershed Hydrology*, pages 1001–1020, Water Resource Publications, Highlands Ranch, CO, 1995.

Younger P. M., Wang C. G., Clark P. A., Gadian A. M. and Beven K. J. The useability of 250m resolution data from the UK Meteorological Office Unified Model as input data for a hydrological model. *Meteor. Applic.*, 15:207–217, 2008.

Younger P. M., Freer J. and Beven K. J. Detecting the effects of spatial variability of rainfall on hydrological modelling within an uncertainty analysis framework. *Hydrological Processes*, 23:1988–2003, 2009.

Yu B. Theoretical justification of SCS method for runoff estimation. *ASCE Journal of Irrigation and Drainage Engineering*, 124:306–309, 1998.

Yu P. S., Chen S.-T. and Chang I.-F. Support vector regression for real-time flood stage forecasting. *Journal of Hydrology*, 328:704–716, 2006.

Yu P. S. and Yang T. C. Using synthetic flow duration curves for rainfall–runoff model calibration at ungauged sites. *Hydrological Processes*, 14:173–181, 2000.

Yu P. S., Yang T. C. and Wang Y. C. Uncertainty analysis of regional flow duration curves. *ASCE Journal of Water Resources Planning and Management*, 128:424–430, 2002.

Zehe E., Becker R., Bárdossy A. and Plate E. Uncertainty of simulated catchment sale runoff response in the presence of threshold processes: Role of initial soil moisture and precipitation. *Journal of Hydrology*, 315:183–202, 2005.

Zehe E., Blume T. and Blöschl G. The principle of "maximum energy dissipation": A novel thermodynamic perspective on rapid water flow in connected soil structures. *Philosophical Transactions of the Royal Society London B*, 365:1377–1386, 2010.

Zehe E. and Flühler H. Slope scale variation of flow patterns in soil profiles. *Journal of Hydrology*, 247:116–132, 2001.

Zehe E., Lee H. and Sivapalan M. Dynamical process upscaling for deriving catchment scale state variables and constitutive relations for meso-scale process models. *Hydrology and Earth System Sciences*, 10:981–996, 2006.

Zhang D., Beven K. J. and Mermoud A. A comparison of nonlinear least square and GLUE for model calibration and uncertainty estimation for pesticide transport in soils. *Advances in Water Resources*, 29:1924–1933, 2006b.

Zhang G. P. and Savenije H. H. G. Rainfall–runoff modelling in a catchment with a complex groundwater flow system: Application of the representative elementary watershed (REW) approach. *Hydrology and Earth System Sciences*, 9(3):243–261, 2005.

Zhang W. and Montgomery D. R. Digital elevation model grid size, landscape representation and hydrologic simulations. *Water Resources Research*, 30:1019–1028, 1994.

Zhang Y. and Chiew F. H. S. Relative merits of different methods for runoff predictions in ungauged catchments. *Water Resources Research*, 45:W07412, DOI:10.1029/2008WR007504, 2009.

Zhang Y., Liu S. and Ding Y. Observed degree-day factors and their spatial variation on glaciers in western China. *Annals of Glaciology*, 43:301–306, 2006a.

Zhang Z., Wagener T., Reed P. and Bushan R. Ensemble streamflow predictions in ungauged basins combining hydrologic indices, regionalization and multiobjective optimization. *Water Resources Research*, 44:W00B04, DOI:10.1029/2008WR006833, 2008.

Zhao R. J. The Xinanjiang model applied in China. *Journal of Hydrology*, 135:371–381, 1992.

Zhao R. J. and Liu X-R. The Xinanjiang model. In Singh V. P. editor, *Computer Models of Watershed Hydrology*, pages 215–232. Water Resource Publications, Highlands Ranch, CO, 1995.

Zhao R. J., Zhuang Y-L., Fang L-R., Liu X-R. and Zhang Q-S. The Xinanjiang model. In *Hydrological Forecasting, IAHS Publication No. 129*, pages 351–356, Int. Assoc. Sci. Hydrol. Press, Wallingford, UK, 1980.

Zhu Y. and Fox P. J. Simulation of pore-scale dispersion in periodic porous media using smoothed particle hydrodynamics. *Journal of Computational Physics*, 182:622–645, 2002.

Zoch R. T. On the relation between rainfall and streamflow. *Monthly Weather Review*, 62(9):315–322, 1934.

Zoppou C. and O'Neill I. C. Criteria for the choice of flood routing methods in natural channels. In *Hydrology and Water Resources Symposium*, pages 75–81. Institute of Engineers, Australia, 1982.

Zuber A. On the interpretation of tracer data in variable flow systems. *Journal of Hydrology*, 86:45–57, 1986.

Zuidema P. K. *Hydraulik der Abflussbildung Wahrend Starniederschlagen*. PhD thesis, Versuchanstalt fur Wasserbau, Hydrologie und Glaziologie, ETH, Zurich, 1985.

Index

Rainfall–Runoff Modelling: The Primer, Second Edition. Keith Beven.