Towards a new paradigm in hydrology

KEITH BEVEN
Department of Environmental Science, University of Lancaster, Lancaster, LAI 4YQ, UK

ABSTRACT Hydrological scientists are faced with the problem (common to many of the field sciences) of complexity at small scales leading to relative simplicity (the hydrograph) at large scales. Little or no success has been gained in relating the former to the latter. Hydrologists will be increasingly forced to think in terms of spatial complexity and spatial pattern but the available tools of analysis are not adequate to accommodate such information. Consequently a theoretical crisis in hydrology is imminent. The development of such a crisis is analysed in terms of the philosophical frameworks of theory change in science. A basis for a new paradigm is laid in a perceptual model of catchment response, leading to an initial conceptual framework incorporating spatial integration and predictive uncertainty.

Vers un nouveau paradigme en hydrologie

RESUME Les hydrologues doivent faire face au problème, commun à bien des sciences sur le terrain, de la complexité à petites échelles menant à une simplicité relative (la courbe de débit) aux grandes échelles. On a rencontré peu ou pas de succès dans les tentatives visant à relier l'une à l'autre. Les hydrologues seront de plus en plus fréquemment obligés de penser en termes de complexité spatiale et de modèle spatial mais les outils nécessaires à l'analyse ne permettent pas d'adapter une telle information. En conséquence, une crise en hydrologie est imminente. Le développement d'une telle crise est analysé du point de vue du cadre philosophique de l'évolution des théories en science. On établit une base pour un nouveau paradigme grâce à un modèle perceptuel de réaction de captation menant à un cadre conceptuel initial incorporant l'intégration spatiale et l'incertitude de la prévision.

A SHORT PATHOLOGY OF HYDROLOGICAL SCIENCE

There is an increasing body of evidence that hydrological science is in some disarray. Exhibit A: studies of the isotopic content of the waters of storm hydrographs (see for example, Sklash & Farvolden, 1979; Herrmann & Stichler, 1980; Kennedy et al., 1986) have provided convincing evidence that traditional methods of "baseflow separation", still presented in recent textbooks and used in
research papers (e.g. Shaw, 1983; Loague & Freeze, 1985) are a nonsense. Exhibit B: remote sensing images and studies of the spatial variability of soil properties (e.g. Russo & Bresler, 1982) have shown that the common assumption of homogeneity that underlies much of the hydrological theory used in predicting catchment responses cannot be justified on physical grounds, albeit theoretically convenient. Exhibit C: experimental studies of water flows in situ and in undisturbed soil cores have shown that water flows in the field may not be well described by Darcy's law (see review by Beven & Germann, 1982) although all our most sophisticated "physically-based" models are based on this law.

Hydrologists have long recognized that, in common with other field scientists, they face a reality that demonstrates great complexity at small space scales, a complexity that ultimately must defy detailed description and complete experimental study at the scales of interest (say the catchment). Yet at these larger scales, hydrological systems exhibit relatively simple responses (for example hydrographs) that individually can be described quite easily. It is indicative of an impending theoretical crisis in hydrological science that we have made little progress in relating the former to the latter.

THEORIES AND MODELS IN HYDROLOGY

In this discussion the term theory will be used to describe a mathematical description of reality that, in so far as it has been tested, has proven to be acceptably consistent with our understanding of reality (although it must be recognized that our standards of acceptability are often not very high). Where such a mathematical description can be used for quantitative prediction, it will be called a model. It does not follow that all models are theories. Many models in hydrology attempt to reproduce hydrological behaviour without more than a superficial consistency with our knowledge of the physical principles of flow.

Most hydrologists will accept that at a fundamental level flow of liquid water may be described by the Navier-Stokes equations of fluid dynamics. These equations represent a theoretical description at the microscale but are useless at the scales of hydrological interest. Certain theories in hydrology (such as Darcy's law) can be shown to be consistent with the Navier-Stokes equations under certain limiting assumptions. These laws and principles have been shown to be good descriptors of experiments in the laboratory. They also form the basis of physically-based theory in hydrology on which a number of hydrological models are based following the general theoretical framework outlined by Freeze & Harlan (1969).

Laboratory scale theories have proved to be useful for a certain range of macroscale problems of hydrological interest (groundwater flows, pipe flows, flows in large regular channels). It has proved very difficult to demonstrate their utility in the very basic hydrological problem of predicting the response of catchments to storm rainfall which involves great temporal and spatial heterogeneity of flow velocities.

It is this central problem that this paper addresses. It will be argued that the extension of laboratory scale theory to the
catchment scale is unjustified and that a radical change in theoretical structure (a new paradigm) will be required before any major advance can be made in this area.

THEORIES AND MODELS OF CATCHMENT RESPONSE

A cursory perusal of the literature reveals a plethora of available models of catchment response. Perhaps the earliest such models were the variants of the "rational" method; the latest are the theory-based, distributed models such as SHE, FESHM and the IHDM (Beven, 1985). These recent models represent a vast advance over the rational method: many more processes, initial and boundary conditions can be simulated; spatial variability of hillslope, soil and vegetation characteristics can be taken into account.

In one very important sense, however, recent theoretical descriptions of catchment response have advanced little beyond the rational method. The rational method requires that a parameter be estimated in order to make a prediction. An initial estimate may be obtained by using physical reasoning on the basis of experience of model use. This reasoning may, where appropriate, be aided by quantitative analyses such as correlations between the parameters and catchment, antecedent and rainfall characteristics. But, if the data are available, this estimate may always be refined by adjusting that parameter to obtain a good match between observed and predicted responses. We refer to this process as model calibration.

Our most sophisticated theories and models also require calibration. Many more parameters are involved and we may be unsure of the way in which the effects of changes in parameters values may interact in the operation of the model. However, we will usually base our initial estimates of the parameters values on some hydrological reasoning and then employ some selective adjustments to improve the match of predictions and observations during model calibration.

Both our simplest and our most sophisticated models work best when the parameter values are back-calculated in this way, and it is common to present figures and data demonstrating model performance during calibration. It is less common to carry out a validation of the model by making further predictions of different observations. This may be because model performance may be considerably poorer during such validation periods (for recent examples with different types of models see Rogers et al., 1985; Loague & Freeze, 1985; Hornberger et al., 1985).

We could consider such a validation exercise as a form of hypothesis testing. We are testing a combination of a particular theoretical structure (the model) and a particular set of auxiliary conditions which will include the initial and boundary conditions given or specified and the set of calibrated parameter values. We might consider that the performance of many models during validation (or even during calibration) is such that we should reject the hypothesis that we have an adequate hydrological model. It is arguable that, in this way, we should reject all hydrological models.

However, hydrologists have tended not to reject their theories or
modelling practice in this way. They have instead tended to explain away these poor results by appealing to possible errors in observations of inputs and outputs, by pointing to the difficulties of specifying boundary conditions, or by indicating the problems of obtaining the "correct" combinations of parameter values. We can in this way use circular reasoning to justify the continued use of theories and models that will provide "acceptable" predictions, providing that the standards of acceptability are not too high. We always question the auxiliary conditions but tend not to question the underlying theoretical structure.

This has led, even recently, to some remarkable consequences. The use of baseflow separation and unit hydrograph techniques persists in current engineering hydrology practice, despite being inconsistent with what is known about the hydrological response of catchments to storm rainfall. Even so-called physically-based models have been applied in a way that is inconsistent with hydrological knowledge. Loague & Freeze (1985), for example, use an infiltration excess-overland flow routing model to predict the response of one of the Hubbard Brook catchments where measured infiltration rates are far greater than normal rainfall rates. The performance of the model was poor, but how far can that inadequate performance be assigned to the effects of observation errors, initial conditions, parameter values or model structure? Would the performance have been better with a different model structure based on subsurface flows rather than surface flows? An alternative model would also require some calibration of parameter values and specification of initial and boundary conditions. What tests do we have of model performance that would allow us to distinguish between these effects, and evaluate the adequacy of hydrological theory on which the models are based?

The poverty of macroscale hydrological methodology arises from the fact that continual back-calculation (calibration) of parameter values with non-error free auxiliary conditions precludes such tests and the proper evaluation of theoretical structures. There has been no adequate hypothesis testing in hydrology outside of the laboratory; and certainly not in the study of catchment responses. Consequently, we have inherited and continue to develop a plethora of incompatible and inadequate models of catchment response without the tools to choose between them in any rigorous way. They all work (to some extent); equally, they are all falsified.

If this analysis is accepted, then it must be concluded that this branch of hydrology faces a theoretical crisis, engendered partly by new lines of evidence about the operation of hydrological systems such as that obtained from isotopes and remote sensing, and partly by the difficulty of hypothesis testing in the current modus operandi. In the next section we shall consider theories of scientific change with a view to illuminating the problems that must be faced in the future.

PARADIGM LOST! CURRENT THEORIES OF CHANGE IN SCIENCE

The classical account of the scientific method suggests that theoretical development proceeds by inductive generalization from a
Towards a new paradigm in hydrology

body of observations into a formal structure capable of deductive prediction of further events. Comparison of observations and predictions is expected to lead to further refinements of the theory (or model) that should result in improved predictive success. This account is now generally viewed as an inadequate account of the progress of science, but there is a continuing debate as to a more satisfactory theory of scientific change. We will deal briefly with the reasons for the failure of the classical account, and then present some alternatives including Popper's falsificationist stance; Kuhn's theory of paradigm change; and Feyerabend's anarchist viewpoint (for a review see Chalmers, 1982). These theories of change will be assessed in terms of their relevance for the problems faced by hydrology.

It has long been recognized that the principle of induction does not provide a firm basis for theory development. There are two major failings of induction. The first is that induction is not logically consistent. It is not the case that if the premises of an inductive argument are true, then the conclusions are necessarily true. There is no way that it can be guaranteed that a future observation will not contradict the argument. The second problem of induction arises in the underlying assumption that the observation statements on which an inductive generalization must be based are independent of any theoretical construct. Induction is held to derive theory from observations. The assumption is, however, untenable; observations are necessarily theory-dependent, and a change of theory will often result in a different interpretation of the same observations (for many examples, see Hanson, 1958). In hydrology, we may point to the concept of stormflow, which within a theory based on Hortonian infiltration excess concepts might be considered as a result of overland flow; whereas in a theory based on the ideas of Hewlett (1974) is interpreted as resulting from subsurface responses and the extension of the channel net.

The problems of induction led to the development of a falsificationist view of theory development (e.g. Popper, 1968). The falsificationist recognizes the theory dependence of observation statements, but assigns a critical role to such statements in the assessment of a theoretical structure. No theory can ever be verified by a set of observations however large, but a single crucial observation can be sufficient to falsify a theory. In the case of our stormflow example, the alternative theories might easily be evaluated in a given situation (for example by the observation that overland flow occurred only rarely in Robert Horton's laboratory in Voorheesville, New York); but in hydrology this is rarely done. The hydrologist, as dilettante scientist, has usually been more interested in the production of numbers than in the testing of competing theories (Klemes, 1986).

The falsification approach to science also has problems, however, since only rarely do the accuracy of the observations allow an unequivocal falsification to take place. Reference back to the earlier discussion of the difficulties of separating out the effects of errors in parameter values, initial and boundary conditions illustrates how difficult it might be to achieve a sufficient degree of falsification to result in rejection of a theory. Indeed, it is extremely rare that rejection will take place unless there is a
competing theory available to be favoured; without such an alternative it is more likely that the theory will be modified by the addition of *ad hoc* hypotheses that protect the deficiencies revealed by one (or many) conflicting observations. In the falsificationist view, such *ad hoc* hypotheses are acceptable, if they themselves may be tested, but this is rarely possibly in hydrology.

An example is provided by the study of Stephenson & Freeze (1974) who applied a physically-based distributed Darcian subsurface flow model to simulate runoff from a hillslope segment under snowmelt conditions. They show only the results of calibration simulations which are not very accurate, but comment that

"a model is not validated until it has successfully simulated a different event from that used for the calibration ... . This calibration-validation sequence pre-supposes perfect knowledge of the boundary parameters .... In our case we have only qualitative information on the spatial dependence of these parameters, and this spatial dependence thus becomes one of the flexible parameters in the calibration procedure. When this uncertainty in the input parameters also exists during the validation event, the resulting flexibility almost ensures that a satisfactory validation will be obtained" (p.289).

In other words, adjustment of the boundary conditions during a validation period is an *ad hoc* procedure that cannot be independently tested and results in a theory that cannot be falsified. Klemes (1986) provides a similar illustration in the field of flood frequency analysis. As a result of the way that parameter calibration permeates hydrology, such examples abound.

And yet we recognize some progress in hydrology. Recently, theories have been published to account for the dynamics of water-soil-vegetation interactions involved in the average annual water balance (e.g. Sharma & Luxmoore, 1979); the physical basis of flood frequency characteristics (e.g. Diaz-Granados *et al.*, 1984; Beven, 1986) and so on. The range of problems tackled becomes wider, the theoretical constructs become more complex, progress is made. We are involved in what Kuhn (1970) calls a period of normal science in the development of a paradigm. Kuhn stresses the sociological nature of science, within which a paradigm represents the theoretical concepts, assumptions and working techniques accepted by a particular scientific community.

Normal science consists of the gradual development of the paradigm to accommodate a wider range of problems and observations. During periods of normal science, apparent falsifications of the paradigm will be encountered. Some may be accommodated, but eventually increasing conflicts between theory and observation may lead to a period of crisis when a paradigm is rejected in favour of an alternative. Kuhn suggests that periods of crisis may be characterized by feelings of professional insecurity due to the failure of existing rules, a proliferation of competing theories (and by increasing interest in philosophical analysis!). Eventually, a major part of the scientific community involved will come to accept a particular set of assumptions, theories and techniques constituting a new paradigm that resolves the crisis (at least in part) and offers scope for further development. The realization of
that development occurs during a further period of normal science. Kuhn argues that the all-pervasive nature of the assumptions of a particular paradigm requires that periods of change be revolutionary: terms that have meaning in one paradigm may not appear in a competing paradigm; observations that are crucial in one may be insignificant in another; arguments of the proponents of one cannot be framed in the language of another. Paradigms are incommensurable.

Feyerabend (1975) also stresses the incommensurability of different theories and the theory-dependence of observation statements, but rejects all the methodological accounts of the progress of science as inadequate. His widely quoted maxim is "anything goes", which may be interpreted to mean that new ideas should be encouraged: those that are perceived to have value will be developed, those that are not will be abandoned. "The invention of alternatives constitutes an essential part of the empirical method" (p.41). Inevitably, of course, every scientist will be faced with a problem of theory choice. All such decisions are subjective and may be governed by factors such as training, the availability of research grants and job prospects as well as comparison with observations. Feyerabend points out some of the dangers in the latter criterion and highlights the role of ad hoc procedures in producing a correspondence with reality. "Suspicion arises that the alleged success is due to the fact that the theory, when extended beyond its starting point, was turned into rigid ideology" (p.43). In the case of more radical choice between incommensurable theories, Feyerabend argues that the only criterion of importance is individual aesthetic judgement (p.285) and that "common scientific wisdom is not very common, and it certainly is not very wise" (p.203).

There are several other recent views of the nature of scientific change (such as the methodology of research programmes of Lakatos, 1978; the objectivism of Chalmers, 1982; and the constructivist views of Knorr-Cetina, 1981), but the brief summary of the views of Kuhn and Feyerabend is sufficient for our analysis here. Both stress the theory dependence of observations, the role of ad hoc procedures in the development and defence of theories, the importance of the sociological setting of theory development, and the occurrence of incommensurable theories in the history of different sciences. They differ in their description of procedures of theory choice and the role of the scientific community in theory development. For Kuhn, a period of normal science is a time of steady theoretical development; for Feyerabend it is a time of small-minded constraint resulting from a general lack of imaginative flair within a science. I leave the reader to decide which description best fits recent progress in hydrology.

TOWARDS A NEW PARADIGM IN HYDROLOGY

I would like to hope that hydrology is coming to the end of a period of normal science, (arguably its first). There have been no revolutionary changes of paradigm in the history of the subject (although a number of competing research programmes in the sense of Lakatos can be distinguished in its development). Hydrologists share a common set of beliefs and assumptions about hydrological
systems that have developed over time but which have never been seriously questioned. We know that the assumptions underlying our macroscale theories are inconsistent with reality, but we have ways of protecting them through the process of calibration that has enabled us to avoid questioning them in a serious way. If hydrology is not to stagnate as a science, then it is time to consider seriously the limitations of our theoretical heritage. It is not sufficient that we can prove a correspondence between predictions and observations in terms of numbers, if the theory is incompatible with our perceptual knowledge of the operation of hydrological systems.

We all have perceptions of how hydrological systems respond to rainfall (or snowmelt) inputs. Those perceptions are necessarily personal, and have been conditioned by our teachers, textbooks, field and modelling experience. They are rarely recorded in print, perhaps because they often conflict with hydrological theory. My own perceptual model is one of complex spatial and temporal variability of input rates, flow paths and nonlinear dynamic responses resulting from the effects of spatial variability in rainfalls, vegetation canopy, soil structure, and topography. My model embodies the idea that preferential pathways are important in flow processes at all scales, from the micro-scale in soil physics, through the hillslope scale in overland and subsurface flows, to the catchment scale in expanding networks of small channels.

Inherent in this perceptual model is an element of unknowability of the system resulting from heterogeneity. We cannot expect to identify the small scale characteristics of a catchment system by either field measurement or from the bulk characteristics of its integrated response represented by the stream hydrograph. "Point" measurements of responses (tensiometers, piezometers, raingauges) must also have limited utility in system identification because of their limited spatial scale and (usually) number of available measurement points. Unknowability implies uncertainty and yet hydrologists persist in using deterministic models for predicting catchment responses.

This model conflicts in many important ways even with the distributed theories embodied in the latest "physically-based" models of catchment response; the question is then how do we proceed. This paper has developed the argument that a theoretical crisis is imminent; can we then begin to discover the shape of a new paradigm?

At the present time I believe that there are two possible lines of progress that will prove fruitful. The first will be to instil the idea within the hydrological community that hydrological predictions are fundamentally uncertain. This will require the development of techniques for making such predictions; the means for assessing uncertainties in parameter values and input values; and a far greater appreciation of methods of stochastic prediction by both hydrological practitioners and students. The first steps in this direction have already been taken (e.g. Gardner et al., 1980; Goren & Burgess, 1981; Rogers et al., 1985) and theories are, of course, already well developed for linear systems. Such a proposal may not therefore be considered revolutionary.

Consider, however, what it would mean if all hydrological predictions were associated with a realistic estimate of uncertainty.
Towards a new paradigm in hydrology

In many cases, it is probable that our current predictions would have little value; the uncertainty bounds would be just too large. The value of observations and calibration procedures would then be recognized in constraining the levels of uncertainty rather than in obtaining a deterministic "best fit". Parameter values and input data (and model structures) would be viewed in terms of sources of uncertainty rather than as "physical" values or theories. Levels of uncertainty could be used directly in decision-making processes. The nature of teaching of hydrology would change, with the possibility of explicit consideration of processes that are not properly understood as well as those that are. The whole enterprise of prediction would be a far more intellectually honest process than our current delusion ridden methodologies.

Within such a framework, it is possible that our current theories will not be the most appropriate for development in a stochastic framework. Hydrology in the future will require a macroscale theory that deals explicitly with the problems posed by spatial integration of heterogeneous nonlinear interacting processes (including the effects of preferential flow pathways alluded to above) to provide a rigorous basis for both "lumped" and "physically-based" predictions. Such a theory will be inherently stochastic and will deal with the value of observations and qualitative knowledge in reducing predictive uncertainty; the interactions between parameterizations and uncertainty; and the changes in hydrological response to be expected as spatial scale increases. Such a theoretical framework should initiate new lines of thought, and innovative methods of measurement, analysis and hypothesis testing to be developed during a future period of "normal" science in hydrology.

I feel that this second line of progress would indeed herald a true revolution in hydrology, a new paradigm on which to base the future of hydrological science. It is, as yet, a goal that seems far off, too complex a problem to be tackled by the current methods of "normal" hydrology. It may well be that the start of the revolution will be marked by a conceptual breakthrough that will appear to simplify the development required, as has happened in other sciences. I suspect, however, that there may be a different type of revolution impending, related to the technological revolution in desktop computing power that is imminent. We can use that power to predict (approximate) uncertainties for our nonlinear systems. We may be able to, indeed it may be necessary, to use brute force methods to investigate the type of theoretical problems that have been outlined above to resolve the current theoretical crisis. The results may not be pretty, but they may be realistic in reflecting both our understanding and our lack of knowledge of hydrological systems.

REFERENCES


