**Hydrology & Earth System Sciences**

# Towards integrated environmental models of everywhere: uncertainty, data and modelling as a learning process

## Keith Beven

Environmental Science/Lancaster Environment Centre, Lancaster University, Lancaster, LA1 4YQ, UK

Email: K.Beven@lancaster.ac.uk

## Abstract

Developing integrated environmental models of everywhere such as are demanded by the requirements of, for example, implementing the Water Framework Directive in Europe, is constrained by the limitations of current understanding and data availability. The possibility of such models raises questions about system design requirements to allow modelling as a learning and data assimilation process in the representation of places, which might well be treated as active objects in such a system. Uncertainty in model predictions not only poses issues about the value of different types of data in characterising places and constraining predictive uncertainty but also about how best to present the pedigree of such uncertain predictions to users and decision-makers.

Keywords: hydrological models, hydroecological models, characterising places, prediction uncertainty

## Introduction

One of the recurring themes of research at the Institute of Hydrology at Wallingford under the Directorship of Jim McCulloch was the development of models of river flows and water quality. IH played a role in developing calibration methods for conceptual rainfall–runoff models (Blackie and Eeles, 1985); in the development of both the original Système Hydrologique Européen (SHE) distributed model (Beven *et al.*, 1980; Abbott *et al.*, 1986; Bathurst *et al.* 1986) and also the Institute of Hydrology Distributed Model (IHDM; Morris, 1980; Beven and O'Connell, 1982; Beven *et al.*, 1987; Rogers *et al.*, 1985; Calver, 1988; Calver and Wood, 1989; Binley *et al.*, 1991); the Probability Distributed Model (PDM; Moore and Clarke, 1981, 1983; Moore, 1985); the TATE model (Calver, 1996); the QUASAR water quality model (Lewis *et al.*, 1997) and others.

Since then the need for accurate predictions of river flow and water quality has increased as a result of increasing environmental legislation. In the European Union, the recent Water Framework Directive (WFD) will have a major impact on the way in which water resources are managed for the foreseeable future (see, for example, http://www.defra.gov.uk/environment/water/wfd ). Implementation of the WFD requires that all designated water bodies be moved towards 'good ecological status' with the aim of achieving 'sustainable use' by 2015 (although 'heavily modified' water bodies can be derogated until 2027). While it is not yet clear what all the phrases used in the WFD will actually mean in practice, managing the process of achieving the requirements of the WFD will demand estimates of the benefits to be gained from new investments to improve water quality towards good ecological status. If these estimates are to be in any way quantified, they will require the implementation of models. However, the science underlying water quality models is by no means adequate; the extension to modelling the ecological impacts is even less so. There are also expected to be more than six thousand surface water bodies of 27 types in England and Wales (plus lakes, estuaries and groundwater bodies) designated under the WFD, all with their own catchment areas with different characteristics of soil, geology, hydrology, topography, land use, land management and infrastructure. How can such predictions, therefore, be made?

New computer technologies seem likely to change the way environmental models are constructed and used. New hardware and software solutions will be based on distributed

high performance parallel computers, linked by fast network connections that, to the user, should appear as a single machine (the concept of the GRID). The user should not have to worry about where the data necessary for a project are stored, nor where any computational tasks are run. The possibility of using high speed computer networking to link together distributed database and computational engines means that it will be possible to couple together models of different environmental systems across disciplinary and national administrative boundaries. This is, in fact, happening already on a limited basis, as demonstrated in the regional water resources models under construction in Denmark (Henriksen *et al.*, 2003), in the national environmental management models being used in the Netherlands (van der Giessen, 2005) and in the Europe-wide Flood Advisory Service being implemented by the European Union Joint Research Centre in Ispra, Italy (http://efas.jrc.it/).

These new possibilities raise some really interesting questions about how this type of interdisciplinary prediction, such as that required for implementing the WFD, might best be achieved. In the past, comprehensive modelling systems have been constructed as large complex computer programmes. These programmes were intended to be general but were expensive to develop, difficult to maintain and to apply because of their data demands and needs for parameter estimation or calibration. With new computing technology, it will be possible to continue in the same vein but with more coupled processes and finer spatial and temporal resolutions for the predictions. It is not clear, however, whether this will result in a real improvement in model accuracy and utility because the problems inherent in the current generation of distributed environmental models may not go away despite improvements in space and time resolutions of the component models.

Those problems include the possibility of Type III modelling errors, i.e. the neglect of processes because a lack of understanding of how the system works is an ultimate constraint on how well the system can be predicted. Not much will be learned, however, about the structural limitations of such models as long as existing 'calibrated' models are adopted in making predictions rather than treating models simply as hypotheses about how the system works and which might be rejected (Beven, 2002a,b, 2006)

There may be another approach. One of the features of having the possibility of large catchment, regional or national scale models is that *everywhere* is represented, so there will be, *de facto*, environmental models of everywhere. Once all places are represented, *data* may assume a greater importance than model structures as a means to refine the representation of each place within a learning framework.

The result may be a new way of looking at environmental modelling that transcends the traditional goal of incorporating all understanding of the complexity of coupled environmental systems into a single mathematical framework with a multitude of parameters that cannot, easily, be identified for any particular place (Beven, 2000, 2002b; Young, 2003; Young *et al.,* 2004).

Consider, for flood-prediction purposes, the possibility of modelling the subtle (and interdisciplinary) coupling between atmospheric forcing, catchment response, river runoff and coastal interaction with tidally-dominated sea levels; capturing these subtleties will require the dynamical coupling of many processes and components from different institutes and different computing systems. Components would be a representation of the coastal seas, the regional atmosphere and the terrestrial surface and subsurface hydrology that would interact through different boundary conditions. Built on the fluxes within those models, air and water pollutant transport models and biogeochemical models could, additionally, be implemented locally within the regional scale domain. Each component should be able to assimilate data transmitted from field sites and to assess the uncertainty in the predictions. Such an integrated system should operate both in real time, assimilating data and boundary conditions from larger scale models and displaying the 'current state of the environment', as well as providing the potential to update model predictions into the future under different scenarios.

The components would share 4-D/5-D visualisation tools with appropriate interactive user interfaces. Users would be able to access the current data, visualise predictions for particular locations and play what-if scenario games over different time scales. The structure of the system would facilitate and even stimulate improvements to the representation of different components and the constraint of predictive uncertainty by field data collection. The potential capabilities of the new computing technology being developed for GRID-scale computing underlie all these components, though much could be achieved using the Web technology of today. Examples of steps towards this type of integrated system (albeit essentially raster-based) include the US Inter-Agency Object Modelling System (OMS, Leavesley *et al.*, 2002) and the UK Coastal Observatory System (http://coastobs.pol.ac.uk/).

## Modelling as a learning process

Even if the very best graphical user interface were available, however, all environmental modellers recognise that their predictive capability is limited and that something is learned about the limitations of the modelling process in every

application to a new site (though they virtually never say so in published papers or reports to clients – there are clearly strong incentives to be positive, even if this results in making predictions with models that have not actually provided very good simulations of calibration data). In the past, the learning process has tended to be treated as learning about the parameter values required in different situations, or, more rarely, about the adequacy of particular model structures (Beven, 2001). However, once everywhere is represented, this emphasis will change to a process of learning about the idiosyncrasies of particular places, albeit that initially this will most probably be treated as a problem of learning about parameter values appropriate to a place.

Treating modelling more explicitly as a learning process will allow a new approach to this problem based on a methodology that will match scale-dependent model objects, data bases and spatial objects in applications within the areas of interest. One of the most exciting benefits of the possibility of representing everywhere in environmental modelling is the potential to implement models available from different institutions as a process of *learning about specific places.* It will be possible, in fact, to have models of all places of interest. However, as argued by Beven (2000, 2001, 2002a,b), as a result of scale, non-linearity and incommensurability issues, the representation of place will be inherently uncertain so that this learning process should be implemented *within a framework of uncertainty estimation.*

## PLACES AS ACTIVE OBJECTS

Modern computer programming systems are generally 'object-oriented', in that components of the system are treated as a structured hierarchy of objects. There will be variable and data objects of different types embedded within algorithm objects which make up a programme. In a classical distributed model implementation in hydrology, for example, coding in an object-oriented manner would treat each grid square or spatial element as an object, with embedded variables and interacting with other spatially contiguous objects. Here, in the context of models of everywhere, a system is envisaged in which sites of interest for a particular prediction can be implemented as *active spatial objects*. Defining the *spatial context* (which might be the reach of a river, an estuary, a catchment, a farm) and the *problem context* (which might be predicting discharge, water quality, velocity distributions, erosion and deposition), then solution of the prediction problem requires three types of information: process representations (model structures), site characterisation (parameters) and boundary conditions (either as observations or as the result of other model predictions). Treating a place as an active object within a distributed computing network would then allow this information to be sought across the network to achieve a specified purpose, recognising that this might result in multiple model structures, parameter possibilities and a lack of definitive information about boundary conditions, i.e. the need to address uncertainty in the predictions.

Such a system would, however, also allow the power of parallel computing resources to be applied to estimate the uncertainty associated with the prediction as constrained by **site-specific observations,** including those accessed over a network in real time. Initially, model results, based perhaps on only GIS databases and limited local information, may be relatively uncertain but experience in monitoring and auditing of predictions will gradually improve the representation of sites and boundary conditions. It is this learning process that will be critical in the development of a new generation of environmental models that are geared towards the management of specific places, rather than general process representations.

## LEARNING ABOUT MODEL STRUCTURES

That is not to say that models of places will not require process representations, nor that current process representations are always acceptably accurate. The approach does not resolve the problem of making errors in the choice of models (whether Type I, Type II or Type III errors) but, by setting the modelling problem in the context of a learning framework for specific places, it does allow a gradual refinement of how places are represented, including (at least in principle) allowing the rejection of models as they are shown to be incompatible with new data. There will always be a real research question about how detailed a process representation is necessary to be useful in predicting the dominant modes of response of a system, given the uncertainties inherent in representing the processes in places that are all unique and the process of learning about that place. This appropriate complexity issue has become obscured by the desire to build more and more scientific understanding into models, including physical, chemical and biological components (but see Young, 2003; Young *et al.*, 2004). This desire is perfectly understandable; it is a way of demonstrating understanding of the science of the environment, but it results in models that have lots of parameter values that cannot easily be measured or estimated in applications to real places. There is always a certain underlying principle in science, that as more understanding is added and empiricisms are eliminated, then the application of scientific principles should become simpler and more robust. This does not seem to have been the experience in

the practical application of environmental models where, in very many cases, model predictions are so heavily dependent on poorly specified boundary conditions of different types. This is one reason why environmental models often exhibit **equifinality** of model structures and sets of parameter values in fitting available observations (Beven, 1993, 2002a,b; Beven and Freer, 2001; Feyen *et al.*, 2001; Aronica *et al.*, 2002; Christiaens and Feyen, 2002; McIntyre *et al.,* 2002; Morse *et al.*, 2003; Uhlenbrook and Sieber, 2005). Such model representations might be useful in producing some prior estimates of what might happen, but it is the learning process about specific places in improving those representations (and perhaps rejecting many potential model representations) that will be crucial. This will particularly be the case where the observations are shown to be incompatible with the model predictions, suggesting that the model can be rejected as a hypothesis of how the system is working. In some cases, this might arise as a result of checking model  performance on more than one performance measure (Freer *et al.,* 2002; Parkin *et al.,* 1996); in other cases, it might be the collection of new types of data that reveals inconsistencies (e.g. the geochemical modelling of the Birkenes catchment discussed by Hooper *et al.*, 2001 and Page *et al.*, 2005: and the inference of macropore flows in plot-scale tracer experiments, e.g. Koide and Wheater, 1992). Two important points have to be stressed about model inconsistencies. The first is that more may be learned from model rejection than acceptance; rejection of a hypothesis, when properly justified, is an important stage in model development and improvement. The second is that model rejection might be purely a result of inadequate boundary condition data (or observations with which the model is being compared). Very often modellers accept and use input data files as if they were without error but even a 'perfect' model (if this were feasible) would not necessarily perform adequately with poor input data. Moreover, to avoid Type II errors, the possibility of input errors in model evaluations must always be considered (Beven, 2005, 2006).

## Modelling as prophecy

Initiatives such as the Water Framework Directive are increasing demands for predictions about the **responses of specific locations to change** in a way that integrates hydrological and ecological considerations in management. However, such a demand does not mean that there is an adequate scientific basis for making the necessary predictions. In many cases, such predictions will be associated with significant uncertainties and, in certain cases, might actually prove to be wrong, even when using the best

science and model components currently available. In short, predictions to users should be presented with great care lest they be regarded as akin to prophecy (Beven, 1993). At the very least, an integrated environmental modelling system of everywhere would need to be powerful enough to be used for assessing **uncertainties in model predictions and the consequential risks of potential outcomes**. It is important to recognise that, in doing so, even when taking account of the important uncertainties in boundary conditions and multiple process representations, not all possibilities will be covered due to imperfect knowledge. Any uncertainty assessment will, therefore, always be **conditional** on the possibilities considered and the assumptions made. This, then, raises the question of how best to present the **pedigree** of model components, data bases and observations to the user where that might have an important effect on decisions made on the basis of the resulting predictions.

### COMMUNICATING PREDICTIVE PEDIGREE TO THE USER

It is easy to imagine a system based on models as active objects in which modelling methods, data bases, observations and uncertainty estimation are integrated seamlessly and automatically for the user interested in the results of predictions (even if this is still a long way off). If much of the modelling process in assembling model components and data in a particular application is to be hidden from the user and if the sheer range and quantity of data involved in applying a model continues to increase, how is the user to assess whether the resulting predictions (including estimates of uncertainty associated with those predictions) have any value?

The word pedigree is here used in the sense of Funtovicz and Ravetz who were the instigators of the NUSAP methodology (see http://www.nusap.net). NUSAP provides a way of specifying, at least qualitatively, a degree of belief in a predictive method based on past experience of its use and the data on which it is based (where this is relevant).  A version of the approach is now being applied operationally in RIVM in the Netherlands (van der Sluijs *et al.*, 2004).

The NUSAP methodology is one way of reflecting the various issues involved in a user making some assessment of the utility of model predictions, given the cumulative past experience of model use. Different approaches to making the required predictions can be compared within such a framework, allowing the user to make some statements about relative degrees of belief in each set of predictions.  It is quite possible, however, that if the range of uncertainty associated with each set of predictions is assessed and compared with the available observations (for

example using the extended GLUE methodology proposed by Beven, 2006), then the issue of equifinality of multiple model structures and parameter sets consistent with what is known about the system response will arise. The user might still want to allow different degrees of belief in each model structure or particular parameter sets (whether assessed by NUSAP or other method) but what is really needed for decision making, or prioritisation of investments for improvements in water supply or quality, is some estimate of the risk associated with the possible model outcomes (and possibly an assessment of the possibility of all the available models being wrong).

One expression of the pedigree of a set of model predictions is, then, the range of possible outcomes that is predicted in a way that is consistent with all available observations, together with an assessment of the possibility of making a Type II error in prediction (i.e. accepting a set of models as representing the system response but which prove to give unacceptable predictions). Such an assessment might present the <u>scientific</u> uncertainty in predictions to a user in a way that can feed directly into a risk-based decision-making framework.

That does not mean to say that the scientific estimate of risk of possible outcomes is entirely objective. This would require risk to be evaluated in terms of a true probability and this is possible only under unrealistically restrictive assumptions (Beven, 2006). In the general case, the type of predictions required for environmental decision-making remain as an exercise in prophecy with many subjective elements. This includes the choice of model structures to consider, the treatment of error in the model inputs and boundary conditions, the choice of feasible parameter values and their interactions, and the choice of evaluation data and performance measures. Even where such choices are made on the basis of some consensus, there is certainly a possibility of a Type II error in prediction (as is the case, for example, of the IPCC consensus evaluations of the predictions of global climate change and, indeed, of the very large number of Monte Carlo predictions of climate change carried out under the climateprediction.net project (http://www.climateprediction.net) using cruder models even when only models consistent with historical data are retained (Stainforth *et al.*, 2005).

What <u>can</u> be presented to the user is the range of possible outcomes predicted by those models that are consistent with the observations available, even if the choice of feasible models is conditioned by the subjective decisions outlined above, together with a list of the assumptions made in setting up the models (perhaps with some NUSAP type weighting scheme). This allows the user to assess the uncertainty in the predictions that results from the specified assumptions

in a form that can be used in decision making. Where the model simulations have been subject to an evaluation against observations, the user will realise that the range of predictions comes from models that have been assessed as being acceptable in some defined way. Where this is not possible (such as in predictions relating to ungauged catchments, or where there is a requirement to predict the impact of 'what if' scenarios for future changed conditions) then the list of assumptions will need to be scrutinised more carefully; any NUSAP-type evaluations of assumptions provided by modelling experts will take on a greater significance.

The possibility of predicting the full range of possible outcomes will be limited both by understanding of how the system works and by computer time. This will particularly be the case where the acceptability or otherwise of a model might depend on the representation of input error, and also where coupled model components mean that uncertainty in one component might cascade as an input error into another component. Integrated models of everywhere will require cascades of multiple components for many components, all of which will involve some uncertainties. There will, therefore, be situations in which the assessment of possible outcomes will be necessarily limited, even using massively parallel computing resources (as in climateprediction.net). It does seem, however, as if computing power will continue to increase (and costs to decrease) making the type of multiple model and parameter set evaluation outlined in this section feasible for an increasing range of coupled models.

## Future proofing modelling systems: adaptive modelling, adaptive management

An essential element of this strategy will be the need, as far as possible, to 'future proof' the model and database systems used, avoiding, for example, a strict raster- based approach or a commitment to one particular modelling framework. The key must be flexibility and adaptive modelling. Raster databases will continue to be driven by remote sensing imaging inputs to the modelling process and, in some cases, by convenient numerical solution schemes for partial differential equations. However, it is often inappropriate to force an environmental problem into a raster straight-jacket. Treating places as flexible active objects might be one way around this future proofing problem. Defining the spatial domain of a prediction problem would allow that place, as an active object, to search for appropriate methods and data for resolving that problem, and also for appropriate methods and data for providing the boundary conditions for the

problem (which might then involve other modelling or data extrapolation techniques).

There are some interesting implications of such an approach. One is that the variety of modelling methods available to solve a prediction problem might be compared more readily, leading to a better understanding of issues of appropriate model complexity for different modelling problems. This will especially be the case if, as part of the learning process, simulations are saved to be compared with later observations of the real outcome. Rarely has this 'post-audit' analysis been used in environmental modelling, although it has been instructive in groundwater modelling (Konikow and Bredehoeft, 1992; Anderson and Woessner, 1992) and is routine in atmospheric modelling in the evaluation of forecast skill (although the evaluation of global climate model predictions still requires an element of compromise at the regional level (Shackley *et al.*, 1997).

The underlying learning framework is best suited to systems that are not changing. In that way, new data should allow a refinement of the acceptable model representations and a reduction in the predictive uncertainty in modelling particular places. However, most prediction problems involve learning about a system as it is changing (even though such non-stationarities are often neglected in traditional model calibration exercises). Predictions of the effects of current or future change under different scenarios will be even more uncertain than the simulation of current conditions but there has been very limited work on estimating the uncertainties of potential outcomes in future scenario simulations, and still less on the conditioning of those predictions in monitoring changes under changing conditions. Data assimilation, in this context of changing conditions, then becomes a tool for following drift in system response (within the limitations of data uncertainties and variability) with the possibility of changing management strategies as the prediction of future responses also changes over time.

There is one very important corollary to modelling as a process of learning about places in this way. It is that such learning cannot proceed without data assimilation, and data assimilation requires the continued collection of data. Thus, learning about places will imply both the continued monitoring of the systems of interest (particularly in detecting non-stationarities in the future that might indicate model inadequacies) and more directed, cost-effective, local measurement campaigns to learn more about places of particular significance (flood risk sites, lake and river sites of particular ecological significance, storm water overflow sites, etc).

## Conclusions

There is a theme running through this discussion of environmental models of everywhere: a focus on data and on making the data that do exist (including past model predictions) available. Data will be required to characterise places, to drive model predictions, to evaluate the results of model predictions and constrain predictive uncertainty, to reject some models previously considered feasible, and to monitor changes in system response. The role of models has always been — albeit sometimes rather implicitly — to extrapolate data in both time and space. This role will now become more explicit in extrapolating from those sites where data are available to the more numerous sites without data and where the characteristics are poorly known. There will still be an argument for using models based on understanding to do that extrapolation (particularly for predicting the impact of changes into the future) but, given the demonstrated limitations and uncertainties of current models based on process understanding, there will also be the opportunity to reconsider the extrapolation problem for particular places as a more general learning and data assimilation process.

In essence, it would appear that learning about **places,** and taking account of the inherent uncertainty in doing so, will become more important than using particular model structures. The significance of having models of everywhere is that all places are represented, and will remain represented. Thus, flexibility and assimilating data in the learning process will be very important in trying to characterise places with all their idiosyncrasies. One of the most important things to study in this modelling methodology will be the **value** of different types of data in terms of both continued monitoring and directed field campaigns as well as in learning about places and constraining predictive uncertainties.

## Acknowledgement

# References

Abbott, M.B., Bathurst, J.C., Cunge, J.A., O'Connell, P.E. and Rasmussen, J., 1986. An introduction to the European Hydrological System – Système Hydrologique Européen SHE. 1. History and philosophy of a physically-based, distributed modelling system. *J. Hydrol.,* **87**, 45–59.

Anderson, M.P. and Woessner, W.W., 1992. The role of the post-audit in model validation. *Adv. Water Resour.*, **15**, 167–174.

Aronica, G., Bates, P.D. and Horritt, M.S., 2002. Assessing the uncertainty in distributed model predictions using observed binary pattern information within GLUE. *Hydrol. Process.,* **16**, 2001–2016.

Bathurst, J.C., 1986. Physically-based distributed modelling of an upland catchment using the Système Hydrologique Européen. *J. Hydrol.,* **87**, 79–102.

Beven, K.J., 1993. Prophecy, reality and uncertainty in distributed hydrological modelling. *Adv.Water Resour.*, **16**, 41–51.

Beven, K.J., 2000. Uniqueness of place and the representation of hydrological processes. *Hydrol. Earth Syst. Sci.*, **4**, 203–213.

Beven, K.J., 2001. How far can we go in distributed hydrological modelling? *Hydrol. Earth Syst. Sci.*, **5**, 1–12.

Beven, K.J., 2002a. Towards a coherent philosophy of environmental modelling. *Proc. Roy. Soc. Lond. A*, **458**, 2465–2484.

Beven, K.J., 2002b. Towards an alternative blueprint for a physically-based digitally simulated hydrologic response modelling system. *Hydrol. Process.,* **16**, 189–206.

Beven, K.J., 2004. Towards environmental models of everywhere: advances in modelling and data assimilation. In: *Hydrology: science and practice for the 21st century, Proc.BHS International Hydrology Symposium*, Imperial College London, UK. I: 244–250.

Beven, K.J., 2005. On the concept of model structural error. *Water Sci. Technol.*, **52**. 165–175.

Beven, K.J., 2006. A manifesto for the equifinality thesis. *J. Hydrol.,* **320**, 18–36.

Beven, K.J. and O'Connell, P.E., 1982. On the role of physically-based distributed models in hydrology. *Institute of Hydrology Report No. 81*, Wallingford, UK.

Beven, K.J. and Freer, J., 2001 Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems. *J. Hydrol.,* 249, 11–29.

Beven, K.J., Warren, R. and Zaoui, J., 1980. SHE: towards a methodology for physically-based distributed forecasting in hydrology. *IAHS Publ. No***. *129*, Wallingford, UK. 133–137.

Beven, K.J., Calver, A. and Morris, E.M., 1987. The Institute of Hydrology Distributed Model, *Institute of Hydrology Report No***. *98***, Wallingford, UK.

Binley, A.M., Beven, K.J., Calver, A. and Watts, L., 1991, Changing responses in hydrology: assessing the uncertainty in physically-based predictions. *Water Resour. Res.*, **27**, 1253–1262.

Blackie, J.R. and Eeles, C.W.O., 1985. Lumped catchment models. In: *Hydrological Forecasting*, M.G. Anderson and T.P. Burt (Eds.), Wiley, Chichester, UK. 311–346.

Calver, A., 1988. Calibration, sensitivity and validation of a physically-based rainfall-runoff model. *J. Hydrol.*, **103**, 103–115.

Calver, A., 1996. Development and experience of the 'TATE' rainfall-runoff model. *Proc. Inst. Civil Eng.- Water, Mar.*, **118**, 168–176.

Calver, A. and Wood, W.L., 1989. On the discretization and cost-effectiveness of a finite element solution for hillslope subsurface flow. *J. Hydrol.*, **110**, 165–179.

Christiaens, K and Feyen, J., 2002. Constraining soil hydraulic parameter and output uncertainty of the distributed hydrological MIKE SHE model using the GLUE framework. *Hydrol. Process.*, **16**, 372–391.

Feyen, J., Beven, K.J., De Smedt, F. and Freer, J., 2001. Stochastic capture zones delineated within the Generalised Likelihood Uncertainty Estimation methodology: conditioning on head observations. *Water Resour. Res.*, **37**, 625–638.

Freer, J.E., Beven, K.J. and Peters, N.E., 2002. Multivariate seasonal period model rejection within the Generalised Likelihood Uncertainty Estimation procedure. In: *Calibration of Watershed Models,* Q.Duan, H.Gupta, S.Sorooshian, A.N.Rousseau and R.Turcotte (Eds.), AGU Books, Washington, USA. 69–87.

Funtowicz, S.O. and Ravetz, J.R., 1993. Science for the post-normal age. *Futures*, **25**, 735–755.

Henriksen, H.J., Troldborg, L., Nyegaard, P., Sonnenborg, T.O., Refsgaard, J.C. and Madsen, B., 2003. Methodology for constructionm, calibration and validation of a national hydrological model for Denmark. *J. Hydrol.*, **280**, 52–71.

Hooper, R.P., Stone, A., Christophersen, N., De Grosbois, E. and Seip, H.M., 1988. Assessing the Birkenes model of stream acidification using a multisignal calibration methodology. *Water Resour. Res.*, **24**, 1308–1316.

Kirchner, J.W., Feng, X. and Neal, C., 2000. Fractal stream chemistry and its implications for contaminant transport in catchments. *Nature*, **403**, 524–527.

Koide, S. and Wheater, H.S., 1992. Subsurface simulation of a small plot at Loch Chon, Scotland. *Hydrol. Process.,* **6**, 299–326.

Konikow, L.F. and Bredehoeft, J.S.D., 1992. Groundwater models cannot be validated. *Adv. Water Resour.*, **15**, 75–83.

Leavesley, G., Markstrom, S.L., Restrepo, P.J. and Viger, R.J., 2002. A modular approach to addressing model design, scale, and parameter estimation issues in distributed hydrological modelling. *Hydrol. Process.*, **16**, 173–187.

Lewis, D.R., Williams, R.J. and Whitehead, P.G., 1997. Quality simulation along rivers (QUASAR). An application to the Yorkshire Ouse. *Sci. Total Envir.*, **194-195**, 399–418.

McIntyre, N., Wheater, H. and Lees, M., 2002. Estimation and propagation of parametric uncvertainty in environmental models. *J Hydroinf.,* **4**, 177–198.

Moore, R.J., 1985. The probability-distributed principle and runoff production at point and basin scales. *Hydrolog. Sci. J.*, **30**, 273–297.

Moore, R.J. and Clarke, R.T., 1981. A distribution function approach to rainfall-runoff modelling. *Water Resour. Res.*, **17**, 1367–1382.

Moore, R.J. and Clarke, R.T., 1983. A distribution function approach to modelling basin sediment yield. *J. Hydrol.*, **65**, 239–257.

Morse, B.S., Pohll, G., Huntington, J. and Rodriguez-Castillo, R., 2003. Stochastic capture zone analysis of an arsenic-contaminated well using the Generalised Likelihood Uncertainty Estimator (GLUE) methodology. *Water Resour. Res.*, **39**(6), DOI 10.1029/2002WROO1470.

Page, T., Beven, K.J. and Freer, J.,2005. Modelling the chloride signal at the Plynlimon catchments, Wales using a modified dynamic TOPMODEL. *Hydrol.Process.*, (in press).

Parkin, G., O'Donnell, G., Ewen, J., Bathurst, J., O'Connell, P.E. and Lavabre, J., 1996. Validation of catchment models for predicting land-use and climate change impacts, 2. Case study for a Mediterranean catchment. *J. Hydrol.,* **175,** 595–613.

Rogers, C.C.M., Beven, K.J., Morris, E.M. and Anderson, M.G., 1985. Sensitivity analysis calibration and predictive uncertainty of the Institute of Hydrology Distributed Model. *J. Hydrol.,* **81**, 179–191.

Shackley, S., Young, P.C., Parkinson, S. and Wynne, B., 1997. Uncertainty, complexity and concepts of good science in climate change modelling: are GCMs the best tools? *Climatic Change*, **38**, 159–235.

Stainforth, D.A. and 15 co-authors, 2005. Uncertainty in predictions of the climate response to rising levels of greenhouse gases. *Nature*, **433**, 403–406.

Uhlenbrook, S. and Sieber, A., 2005. On the value of experimental data to reduce the prediction uncertainty of a process-oriented catchment model*. Environ. Model. Software,* **20,** 19–32.

Van der Giessen, A., (Ed.) 2005. *Naar en gezamenlijk nationaaal hydrologisch modelinstrumnetarium.* Report 500026002/2005, MNP, Bilhoven, The Netherlands.

Van der Sluijs, J., Janssen, P.H.M., Peteresen, A.C., Kloprogge, P., Risby, J.S., Tuinstra, W. and Ravetz, J.R., 2004. *Tool catalogue of RIVM/MNP Guidance for Uncertainty Assessment and Communication.* Copernicus Institute, Utrecht University, The Netherlands.

Young, P.C., 2003. Top-down and data –based mechanistic modelling of rainfall-flow dynamics at the catchment scale. *Hydrol. Process*., **17**, 2195–2217.

Young, P.C., Chotai, A. and Beven, K.J., 2004. Data-based mechanistic modelling and the simplification of environmental systems. In: *Environmental Modelling: Finding Simplicity in Complexity*, J. Wainwright and M. Mulligan (Eds.), Wiley, Chichester, UK. 371–388.