

Modelling everything everywhere: a new approach to decision-making for water management under uncertainty

KEITH J. BEVEN^{*,†,‡} AND RUTH E. ALCOCK^{*}

^{*}Lancaster Environment Centre, Lancaster University, Lancaster, U.K.

[†]Geocentrum, Uppsala University, Uppsala, Sweden

[‡]Centre for Analysis of Time Series, London School of Economics, London, U.K.

SUMMARY

1. There are increasing demands to predict ecohydrological responses to future changes in catchments but such predictions will be inevitably uncertain because of natural variability and different sources of knowledge (epistemic) uncertainty.
2. Policy setting and decision-making should therefore reflect these inherent uncertainties in both model predictions and potential consequences.
3. This is the focus of a U.K. Natural Environment Research Council knowledge exchange project called the Catchment Change Network (CCN). The aim is to bring academics and practitioners together to define Guidelines for Good Practice in incorporating risk and uncertainty into assessments of the impacts of change.
4. Here, we assess the development of such Guidelines in the context of having catchment models of everywhere.

Keywords: catchments, ecohydrology, Guidelines, models, uncertainty

Introduction

Ecohydrological processes in catchments are complex. They are forced by inputs that are not always well known that induce a response in a system of which characteristics are difficult to measure, changing over time, and difficult to estimate (particularly in the subsurface). The hydrological processes driving transport processes are both non-stationary and not that well understood. The biogeochemical processes affecting water quality in surface waters are not that well understood. The impact of those processes on biodiversity and ecological systems is also not well understood. Faced with such a range of uncertainties in knowledge and the natural randomness of environmental forcing, there is a real question as to whether the predictions made by models of catchment processes, and the way in which they might change in the future, might be useful in informing management decisions about future investment to effect improvements in water quality and ecological status.

Indeed, it has been argued (for example in the post-normal science perspective of Funtowicz & Ravetz, 1990) that sensible management for sustainability in an uncer-

tain world should not be based on prediction but on consensus about action. Faced with the very real uncertainties about catchment responses, management should be treated as a social and political process. The scientific evidence simply cannot be sufficiently convincing in the context of so much uncertainty. It is therefore a better strategy to try to get consensus about robust and adaptive management strategies without resort to model predictions.

We think that this argument will generally fail on two counts. First, it will be very difficult to get a consensus (or even compromise) between the many stakeholders in the catchment ecohydrological system; second, sensible investment strategies (including robust, adaptive and precautionary strategies) require some sense of the effects of a given management input. In general, the more the investment, the more the effect will be, but how much input will be required to have the required effect and over what time scale? These questions are current given the financial constraints in developing the river basin management plans required to achieve good chemical and ecological status in the implementation of the Water Framework Directive (WFD) across the European Union.

Models might then still have a role to play in policy setting by providing a means to predict the impact resulting from a management or investment decision. They do not necessarily have to be complex models of all the various processes involved but they do have to allow for the uncertainty in the modelling process. In fact, as will become more apparent later, the very process of discussing the assumptions of a predictive tool, and the assumptions about the uncertainties involved, can in itself become a useful component of the social and political process that is catchment management. What is required is a framework for making this happen. Here, we will suggest that an appropriate framework is provided by the creation of models of everywhere and the development of Guidelines for Good Practice to shape that process.

Models of everywhere and everything

We now have the computer power to be able to model the hydrology and water quality of the whole of the United Kingdom. The Grid to Grid (G2G) model is being applied nationally in the United Kingdom for flood risk assessment (Bell *et al.*, 2007). The PSYCHIC model is being used to identify the risk of high phosphorus loads from agricultural land in a number of U.K. catchments for WFD planning (Davison *et al.*, 2008). The hydrology of Denmark is being modelled (Henriksen *et al.*, 2008). Spatial data sets are becoming more commonly available (including projections of future meteorological variables at the 5-km scale in U.K. Climate Predictions 2009 (UKCP09), see Kilsby *et al.*, 2007). Models of everywhere are becoming more and more computationally feasible.

On the other hand, it is a hydrological modelling aphorism that every catchment is unique, making regionalisation or the prediction of the responses of ungauged catchments difficult (for an extended discussion, see Beven, 2000). 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?

In a recent paper looking ahead to the availability of models of everywhere (Beven, 2007), it was argued that it is indeed useful, and even necessary, to think in terms of models of everywhere. 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 hydrology 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. 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.

In fact, this learning process about place is a way of doing science in a complex system. Models can be treated as hypotheses about how the catchment system functions (Beven, 2002, 2009, 2010). Those 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. Some of these might be poor models that have survived by chance (a false-positive or Type I error in hypothesis testing), while we might also reject good models because of poor data (a false-negative or Type II error). Ideally, we wish to minimise both Type I and Type II errors, but the nature of the epistemic errors in the modelling process means that this is very difficult to achieve securely (Beven, 2010). Such hypothesis testing is well developed within a statistical framework when we can consider that the sources of uncertainty involved are fundamentally random in nature (they are *aleatory* in nature). This is not, however, the case in environmental modelling (even if it is often assumed to be the case for statistical convenience) because so many of the sources of uncertainty result from a lack of knowledge. Such errors are often referred to as *epistemic* errors. They might be reduced in future by more detailed study, or better measurement techniques, or a breakthrough in the understanding of controlling processes, but they might not be properly represented by a simple statistical model or likelihood function. In particular, epistemic errors can lead to model residuals that have non-stationary characteristics that are not easily handled within a statistical framework. 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.

Hypothesis testing, epistemic and aleatory errors

Hypothesis testing is generally treated as a problem in statistics. Statistical theory including Bayesian methods, however, depends on being able to assume that all sources of error are aleatory, or can be treated *as if* they were aleatory. Aleatory errors can be considered to be randomly drawn from some underlying distribution. If this is the case, then assumptions about the error model lead to a well-defined likelihood function and a probabilistic interpretation of uncertain predictions of future behaviour. This is, in the terminology of Beven (2006), the ideal case. The problem is that ecohydrological

models are not ideal in this sense because they involve multiple sources of epistemic errors. Epistemic uncertainties result from lack of knowledge about the system under study, which might involve lack of knowledge about inputs, processes or observations with which a model is being compared. Epistemic errors in environmental modelling are not easily treated as if they were aleatory (see the example in Table 1). This is not just a difficulty in applying formal statistical methods as suggested by O'Hagan & Oakley (2004) or more recently in the reification approach of Goldstein & Rougier (2009); the list of epistemic errors is long and this is a generic problem.

Table 1 Sources of uncertainty with random (aleatory) and knowledge (epistemic errors) in the case of flood risk mapping

Source of uncertainty	Aleatory errors	Epistemic nature
Design flood magnitude	What is the range of sampling variability around underlying distribution of flood magnitudes?	Are floods generated by different types of events? What frequency distribution should be used for each type of event? Are frequencies stationary? Will frequencies be stationary into the future?
Conveyance estimates	What is the random sampling variability around estimates of conveyance at different flood levels?	Is channel geometry stationary over time? Do conveyance estimates properly represent changes in momentum losses and scour at high discharges? Are there seasonal changes in vegetation in channel and on floodplain? Is flood plain infrastructure, walls, hedges, culverts etc. taken into account?
Rating curve interpolation and extrapolation	What is standard error of estimating the magnitude of discharge from measured levels?	Is channel geometry stationary over time? What is estimation error in extrapolating rating curve beyond the range of measured discharges? Does extrapolation properly represent changes in momentum losses and scour at high discharges?
Flood plain topography	What is the standard error of survey errors for flood plain topography?	Are there epistemic uncertainties in correction algorithms in preparing digital terrain map?
Model structure		How far do results depend on choice of model structure, dimensions, discretisation and numerical approximations?
Flood plain infrastructure	What is the random error in specifying the positions of elements, including elevations of flood defences?	How should storage characteristics of buildings, tall vegetation, walls and hedges in geometry be treated? Are there missing features in the terrain map (e.g. walls, culverts)?
Observations used in model calibration/conditioning	What is the standard error of estimating a flood level given post-event survey of wrack marks or gauging station observations?	Is there some potential for the misinterpretation of wrack marks surveyed after past events? Are there any systematic survey errors?
Future catchment change	–	What process representations for effects of land management should be used? What future scenarios of future change should be used? Are some scenarios more likely than others?
Future climate change	What is the variability in outcomes owing to random weather generator realisations?	What process representations in weather generators should be used? What future scenarios of future change should be used? Are some scenarios more likely?
Fragility of defences	What are the probabilities of failure under different boundary conditions?	What are the expectations about failure modes and parameters?
Consequences/Vulnerability	What is the standard error of estimation for losses in different loss classes?	What knowledge about uncertainty in loss classes and vulnerability is available?

So what are these epistemic errors in the case of ecohydrological models in assessing the impacts of catchment change? They include the following:

1. Non-stationarity in the errors of estimates of inputs to the catchment system.
2. Unknown temporal variability in the system characteristics as represented by the model parameters.
3. Unknown temporal and spatial variability in controlling processes.
4. Indecision about how some processes should be represented mathematically.
5. Non-stationarity in the errors associated with observations with which model predictions are compared.
6. Lack of commensurability between observed and predicted variables because of scale issues or simple difference in meaning.

The recognition of such errors is not new. They are analogous to what Knight, 1921 called the *real uncertainties*, not readily amenable to a statistical analysis (or in his case, what an insurance broker would be prepared to take odds on). Epistemic uncertainties are an inherent part of modelling environmental systems and result in what Beven (2006) called non-ideal cases. Experience suggests that treating epistemic errors as if they were aleatory will generally lead to over-confidence and bias in how well a model represents a system (Beven, Smith & Freer, 2008). In some cases, lack of knowledge of the boundary conditions for the catchment system (rainfalls, discharges, nutrient inputs, species migration etc.) might mean that information for some periods or events might not be informative about whether a model is a good representation or not. Such effects can also persist for some time (such as when a poor rainfall estimate for an event affects how well a hydrological model can predict discharges for that event and subsequent events). This is not an uncommon situation in hydrological data. This then suggests that some alternative approach is required to test models as hypotheses in a way that reflects more properly the sources and nature of non-statistical error.

A limits-of-acceptability approach to testing models as hypotheses

However, although epistemic errors are endemic to environmental models, there is no formal theory for dealing with them. Effectively, there can be no formal theory for dealing with epistemic error, because we do not have adequate knowledge of the nature of the errors. If we had adequate knowledge, we would have a better idea of how to deal with them and they would no longer be

epistemic (but still would not necessarily be simply aleatory). This is the core dilemma in modelling catchment systems for decision-making. Will ignoring epistemic errors lead to too many Type I errors of accepting poor models based on the available data, so that future predictions will be compromised and might lead to poor decisions? In such a situation, would it not be better to formulate decision-making in a way that does not depend on model predictions?

We do not know the answer to these questions because we have not traditionally considered them in this way. There has been no framework for doing so, and no way of deciding when model predictions might be informative and when they might not. Both Frequentist and Bayesian statistical methods depend on formulation of a model of the prediction errors as if they were, at base, aleatory. Part of such a model might be a structured transformation (such as the removal of a constant bias or other model discrepancy function, e.g. Kennedy & O'Hagan, 2001), but both forms of statistical analysis assume that the model is correct and that every prediction error will be informative in the model conditioning process. This may not be the case: in these complex systems with poorly defined inputs, epistemic errors may mean that some prediction errors might be disinformative (see Beven *et al.*, 2008). A new approach is required.

The first steps are just being taken to provide such a framework and working methodology. The first stage is to evaluate model performance against past data in a way that reflects the sources of uncertainty in the modelling process when the uncertainties are not always probabilistic. Beven (2006) suggested an approach based on specifying the limits of acceptability around some observational data within which we would wish model predictions to lie (see Liu *et al.*, 2009; for a case study in rainfall-runoff modelling; Dean *et al.*, 2009 for a case study in water quality modelling; and Blazkova & Beven, 2009, for a case study in flood frequency estimation). Any models that are acceptable in this sense would be used in prediction; those that do not would be rejected. One nice feature of this approach is that the limits of acceptability can be applied to every available observation (or only those of greatest interest), so that models are not evaluated purely in terms of some global performance or likelihood measure which can obscure relatively poor performance about important features of the behaviour. However, in setting such limits of acceptability, it is important that we do not expect a model to perform better than the limitations of the forcing data and the observations with which it is being compared, including potential epistemic errors, so as to minimise Type II errors of

rejecting models that might be useful in future prediction. We will never be sure, of course, that in doing so we are not also making Type I errors, but we will only be able to make such an assessment as new observations become available.

Setting limits of acceptability

Setting the limits of acceptability to reflect the potential for epistemic uncertainties is clearly critical in this framework. This means making decisions about the nature of different sources of uncertainty while lacking sufficient information to do so. That is the essence of epistemic uncertainties. Setting limits of acceptability is therefore difficult, but not one that should be ignored. In analogy with statistical methodologies for uncertainty estimation, there will be information about the nature of the relevant uncertainties in the model residuals available from evaluations against calibration data. In statistical estimation, these residuals are used to formulate a model of the errors that is assumed to hold into the future for making predictions. It is not generally necessary to try to disaggregate different sources of error but, as noted earlier, there is a danger that future uncertainties will be underestimated by treating the model errors as if they are simply aleatory.

This recognition does not, however, provide a means as to what to do instead. Some assumptions about sources of uncertainty will be necessary in setting appropriate limits of acceptability in model evaluation, but the way in which epistemic (non-stationary) errors in model inputs get propagated in some nonlinear way through a non-error free model and then compared with non-error free observations means that the disaggregation of the contribution of different sources of error is a poorly posed problem and may be impossible (see, e.g. Beven, 2005). So how should such assumptions be decided and later used in model prediction when, even given some knowledge of model residuals, there can be no unique to the characterisation of sources of uncertainty?

Guidelines for good practice and stakeholder involvement

One way forward in this situation, which presents some advantages, 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 do so, experience that might be encapsulated in sets of rules or Guidelines for Good Practice. Such Guidelines might set out the decisions needed in considering sources of uncertainty to be considered 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 *translatory discourse* advocated by Faulkner *et al.* (2007).

What does being robust mean in the face of epistemic uncertainties?

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 be necessarily 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 uncertainties. 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 will 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.

There is quite a lot in common in this view with the Info-Gap methodology of Ben-Haim (2006; Hine & Hall, 2010). Info-Gap was designed to handle decision-making under high uncertainty, looking at the trade-off of robustness and opportuneness functions as a system departs from some baseline condition or simulation (see also Beven, 2009; for a summary but note also the suggestion of Sniedovich (2010 and references therein) that Info-Gap is a specific case of Wald's max-min decision theory). However, 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; ...) which would 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 type of Guidelines for Good Practice argued for here represents a formal extension of this approach.

Does it matter to robustness 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. This might be 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.

Heuristics for change: the Catchment Change Network

If the Guidelines for Good Practice methodology is to be useful in robust decision-making, then these need to be developed for different types of application. Although each decision-support situation is unique in terms of context, elements and location etc., guiding principles as heuristics, are a valuable means to define and summarise a collective consensus body of expert knowledge. They represent efficient frameworks to guide decision-makers by helping them simplify choices.

Developing Guidelines is one of the aims of the Catchment Change Network (CCN), a U.K. Natural Environment Research Council (NERC) Knowledge Transfer project being led by Lancaster University. The Network – made up of three discrete but interlinked Focus Areas covering flood risk, water quality and water scarcity – will exchange knowledge across a wide range of project partners about how best to handle uncertainties in integrated catchment management. A key aim is to integrate modern methods of uncertainty estimation to improve decision-making for adaptive management across catchments. Workshop activities in each of the Network Focus Areas have recently explored the form, scope and content of such Guidance with debate centring on sources of uncertainty, the range and composition of audiences for the guidance produced and the communication and transparency of the underlying assumptions made.

Progressively updated Guides to Good Practice will be produced for each of the three Focus Areas with content defined and developed via workshop activities and interactive web-based involvement across a range of stakeholders. The web site <http://www.catchment-change.net> acts as both an information hub and knowledge exchange portal to communicate and interact across our project partners both in the United Kingdom and in Europe. The intention is that these guides will ultimately become embedded across a wide range of catchment management professionals with the intention they will help practitioners and decision-makers in problem framing by focussing on key variables while clarifying the strength of available evidence. These will be living documents that with broad user input will be able to both refine the heuristics and add new ones as the concept of 'good practice' continues to evolve.

Systematic guidelines may prove very helpful and consistent for decision-making in the face of uncertainty, particularly in terms of agreeing and communicating the assumptions of any risk and uncertainty analysis that

feeds into a decision-making process. Relying on pre-defined aids to good practice also needs to recognise that heuristics are fallible and have limitations. A range of biases may be buried within them, contingent on perceived past experience. In particular, confirmation biases may result from choices that reinforce past preconceptions. That is why it is important that the Guidelines for Good Practice should be living documents, evolving over time as experience of applying them increases. One way of ensuring this is to structure the Guidelines in terms of a set of decisions about options that have to be agreed upon between analysts, stakeholders and users. Such a decision structure allows such evolution over time while making the assumptions of any analysis to be defined explicitly. The overall framework for one set of Guidelines for the preparation of flood risk maps reflecting the uncertainties outlined in Table 1 is shown in Fig. 1. Such flood risk

maps, and associated uncertainties, might have an impact on the ecohydrological management of floodplain habitats. Lower levels in the decision framework are defined by decision trees for specific sources of uncertainty. In addition, sections on how to condition the uncertainty estimates on observational data, how to visualise the outputs of the analysis and how to take action to manage and reduce the uncertainties are included in the framework. The conditioning process includes the concepts of hypothesis testing that might take the form of a limits-of-acceptability evaluation of models as outlined earlier.

Back to models of everywhere

So how is this relevant to models of everywhere? We noted earlier that models of everywhere allow a learning process to develop in the representation of place (here,

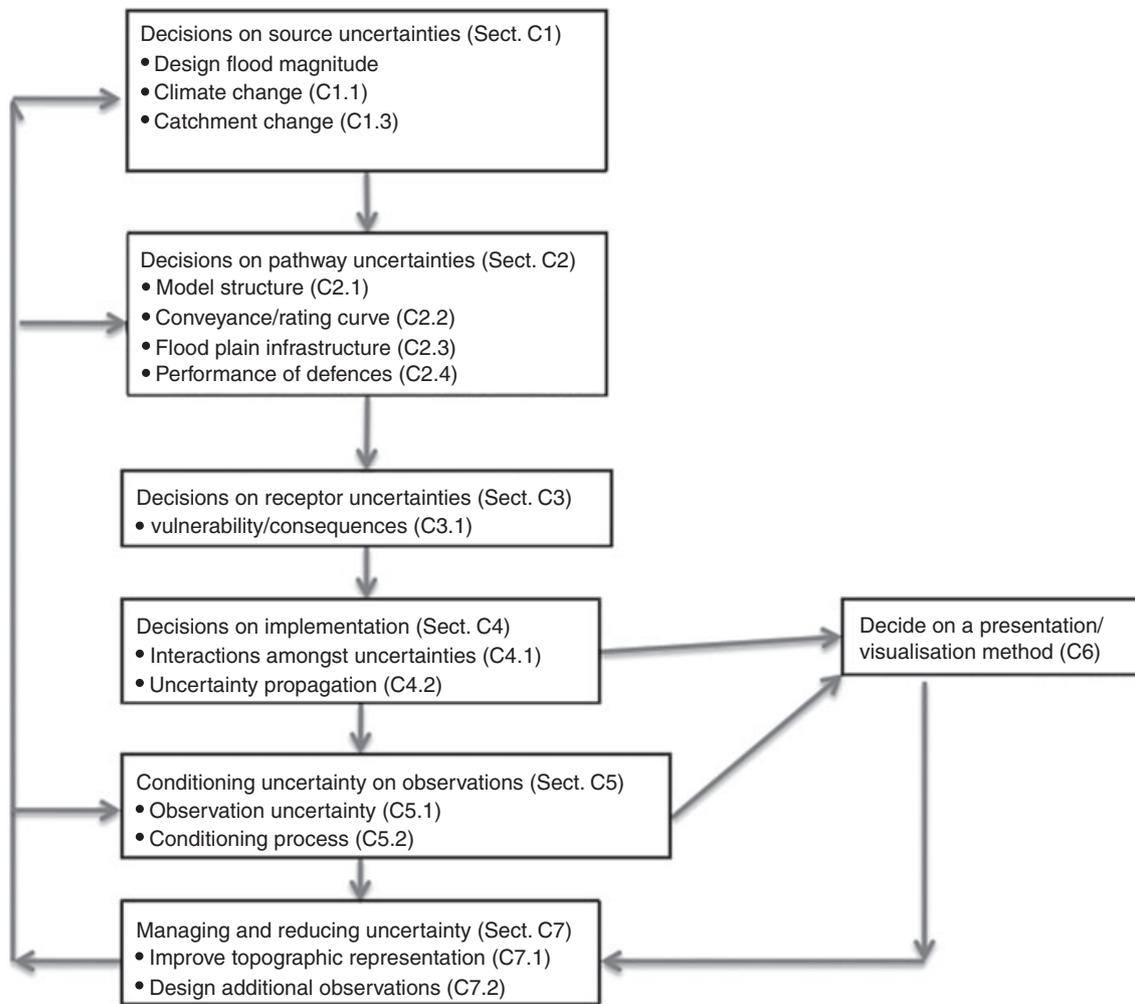


Fig. 1 High-level decision structure for Guidelines for Good Practice in uncertain flood risk mapping (after Beven, Leedal & McCarthy, 2010); each of the bulleted items represents a decision tree about assumptions at a lower level as indicated by the Section labels. The return arrows in the figure represent the pathways for review of the assumptions that might result from additional information being made available, e.g. from adding additional observations.

places within catchments in integrated water management). This learning process should be powerful exactly because everywhere is represented. This means that local results should be available to local stakeholders in visual form, such as on-line maps, animations and graphs of observations in real time. Local stakeholders will have (often qualitative) knowledge of the system response that will be valuable in the evaluation of local model predictions. Thus, if there are local deficiencies, they will be closely scrutinised and local stakeholders will be only too pleased to point out such deficiencies to the modeller (or agency providing the predictions). The nature of those deficiencies will then be a guide to how to make local improvements (which is not necessarily the same as saying making the model locally more complex; simpler approaches may be sensible for some purposes).

This will also provide a strong incentive for the modeller to anticipate the issues that might be raised by local stakeholders beforehand. One way of doing so is to involve stakeholders from the very beginning of an implementation, including in the decisions framed in the type of Guidelines for Good Practice discussed earlier. Agreeing on such decisions and making the assumptions inherent in the decisions explicitly allows and provides a useful framework for such interactions (see for example, Ryedale Flood Research Group, 2008).

Models of Everywhere can also provide the basis for local adaptive management strategies. Once implemented, what-if strategies can be played out to explore decision options (see, for example, Olsson & Berg, 2005; Olsson & Andersson, 2007). In doing so, it is worth noting that models can be useful guides in adaptive management, particularly in estimating the time scale of a response to a management decision. Adaptive management requires a signal to be observed in response to an action. Yet we know that for some of the problems intrinsic to integrated catchment management, the time scales of the response might be long (perhaps decades) and uncertain (if only because the nature of the response might depend on the particular sequence of wet and dry years to come). Evaluating the uncertainty in the response might then be a useful guide to stakeholders as to what outcomes might be expected, what action might be enough to show benefit and what might be the most robust or least regret strategy given the uncertainty.

In hydrology and water quality, there is a long tradition of making predictions with inadequate data, without recognising their limitations and uncertainties. This has resulted in expectations of reduced performance in prediction as being normal. This is not surprising, given the epistemic nature of the different sources of uncertainty

in the modelling process. What is unsatisfactory is that so little has been done about the problem until very recently. The heuristic has been to ignore the problems because they are perceived as being too difficult; there is still no theory of how to deal with epistemic uncertainties since, by definition, they are poorly known. Yet they may have an effect on what decision might be made, particularly if we are interested in decisions that are robust to uncertainty. Once more 'models of everywhere' are implemented and the need for Guidelines for Good Practice involving local stakeholders in decisions about assumptions is accepted, then perhaps this will change.

References

- Bell V.A., Kay A.L., Jones R.G. & Moore R.J. (2007) Use of a grid-based hydrological model and regional climate model outputs to assess changing flood risk. *International Journal of Climatology*, **27**, 1657–1671. 10.1002/joc.1539 <http://dx.doi.org/10.1002/joc.1539>
- Ben-Haim Y. (2006) *Info-Gap Decision Theory*, 2nd edn. Academic Press, Amsterdam.
- Beven K.J. (2000) Uniqueness of place and process representations in hydrological modelling. *Hydrology and Earth System Sciences*, **4**, 203–213.
- Beven K.J. (2002) Towards a coherent philosophy for environmental modelling. *Proceedings of the Royal Society of London*, **458**, 2465–2484.
- Beven K.J. (2005) On the concept of model structural error. *Water Science and Technology*, **52**, 165–175.
- Beven K.J. (2006) A manifesto for the equifinality thesis. *Journal of Hydrology*, **320**, 18–36.
- Beven K.J. (2007) Working towards integrated environmental models of everywhere: uncertainty, data, and modelling as a learning process. *Hydrology and Earth System Sciences*, **11**, 460–467.
- Beven K.J. (2009) *Environmental Modelling: An Uncertain Future?* Routledge, London.
- Beven K.J. (2010) Preferential flows and travel time distributions: defining adequate hypothesis tests for hydrological process models. *Hydrological Processes*, **24**, 1537–1547.
- Beven K.J. (2011) I believe in climate change but how precautionary do we need to be in planning for the future? *Hydrological Processes*, in press. DOI: 10.1002/hyp.7939.
- Beven K.J., Smith P.J. & Freer J. (2008) So just why would a modeller choose to be incoherent?. *Journal of Hydrology*, **354**, 15–32.
- Beven K.J., Leedal D.T. & McCarthy S. (2010) Guidelines for Assessing Uncertainty in Flood Risk Mapping, *Flood Risk Management Research Consortium User Report*, available at <http://www.floodrisk.net>
- Blazkova S. & Beven K.J. (2009) 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.
- Davison P.S., Withers P.J.A., Lord E.I., Betson M.J. & Strömqvist J. (2008) PSYCHIC – A process-based model of phosphorus and sediment mobilisation and delivery within agricultural catchments. Part 1: model description and parameterisation. *Journal of Hydrology*, **350**, 290–302.
- Dean S., Freer J.E., Beven K.J., Wade A.J. & Butterfield D. (2009) Uncertainty assessment of a process-based Integrated Catchment Model of Phosphorus (INCA-P). *Stochastic Environmental Research and Risk Assessment*, **23**, 991–1010.
- Faulkner H., Parker D., Green C. & Beven K.J. (2007) Developing a translational discourse to communicate uncertainty in flood risk between science and the practitioner. *Ambio*, **16**, 692–703.
- Funtowicz S.O. & Ravetz J.R. (1990) *Uncertainty and Quality in Science for Policy*. Kluwer Academic, Dordrecht.
- Goldstein M. & Rougier J. (2009) Reified Bayesian modelling and inference for physical systems. *Journal of Statistical Planning and Inference*, **139**, 1221–1239.
- Henriksen H.J., Trolborg L., Hojberg A.L. & Refsgaard J.C. (2008) 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.
- Hine D. & Hall J. (2010) Information gap analysis of flood model uncertainties and regional frequency analysis. *Water Resources Research*, **46**: doi: 10.1029/2008WR007620
- Kennedy M.C. & O'Hagan A. (2001) Bayesian calibration of mathematical models. *Journal of the Royal Statistical Society. Series B, Statistical methodology*, **63**, 425–450.
- Kilsby C.G., Jones P.D., Burton A., Ford A.C., Fowler H.J., Harpham C. *et al.* (2007) A daily weather generator for use in climate change studies. *Environmental Modelling & Software*, **22**, 1705–1719.
- Knight F.H. (1921) *Risk, Uncertainty and Profit*. Houghton-Mifflin Co. (reprinted University of Chicago Press, Chicago, 1971)
- Liu Y., Freer J.E., Beven K.J. & Matgen P. (2009) Towards a limits of acceptability approach to the calibration of hydrological models: extending observation error. *Journal of Hydrology*, **367**, 93–103.
- O'Hagan A. & Oakley A.E. (2004) Probability is perfect but we can't elicit it perfectly. *Reliability Engineering & System Safety*, **85**, 239–248.
- Olsson J.A. & Andersson L. (2007) Possibilities and problems with the use of models as a communication tool in water resource management. *Water Resources Management*, **21**, 97–110.
- Olsson J.A. & Berg K. (2005) Local stakeholders' acceptance of model-generated data used as a communication tool to water management: the Rönneå study. *Ambio*, **34**, 507–512.
- Ryedale Flood Research Group (2008) *Making Space for People*. Available at: http://knowledge-controversies.ouce.ox.ac.uk/ryedaleexhibition/Making_Space_for_People.pdf.
- Sniedovich M. (2010) A bird's view of info-gap decision theory. *Journal of Risk Finance*, **11**, 268–283.

(Manuscript accepted 8 February 2011)