Handling uncertainty in extreme or unrepeatable hydrological processes—the need for an alternative paradigm

Jim Hall\textsuperscript{1} and Malcolm Anderson\textsuperscript{2*}

\textsuperscript{1} Department of Civil Engineering, University of Bristol, Queen’s Building, University Walk, Bristol BS8 1TR, UK
\textsuperscript{2} School of Geographical Sciences, University of Bristol, University Road, Bristol BS8 1SS, UK

*Correspondence to: Malcolm Anderson, School of Geographical Sciences, University of Bristol, University Road, Bristol BS8 1SS, UK. E-mail: m.g.anderson@bristol.ac.uk

The conventional approach to assessing uncertainty in a hydrological model involves comparing model predictions with a test dataset of measurements. Typically, both the dataset and the model predictions will be represented as a time series of precise measurements, even though it is acknowledged that field measurements have associated inaccuracies and, more significantly, a model cannot be expected to make an exact prediction of a hydrological phenomenon. By comparing predicted and measured time series it is possible to extract, often multiple, measures of the distance (in length or time) between model prediction and measurement, which is thought of as residual uncertainty. This conventional characterization also applies to the GLUE methodology (Beven and Binley, 1992), in which multiple model runs with multiple parameter sets, after some preselection on 'behavioural' grounds, are conditioned according to the distance between predicted and measured response.

However, we believe that there are a number of classes of problems in hydrological processes that require different approaches towards uncertainty estimation. Figure 1 illustrates the dimensions on which hydrological problems can be classified, following a general classification for modelling problems proposed by Blockley (1980). The base (and rather special) case is the situation in which there are precise simultaneous measurements of the phenomenon of interest and the model prediction of that phenomenon. However, often the most interesting hydrological problems are where there are no simultaneous measurements and predictions. This may, for example, be because the phenomenon of interest is extremely rare (such as an extreme flood) or unrepeatable (such as a hydrologically activated landslide). The situation is analogous to the problem that occupies engineers responsible for high reliability systems: how is the probability of a nuclear reactor meltdown or the collapse of a long-span suspension bridge to be estimated when there may be no instances of the event of interest, and even if there are, the population is hopelessly small in statistical terms? Even if relevant measurements are available they may not be recorded in precise terms. Measurements may only be bounded, as would be the case with a discrete measurement device (any digital device), or may be described in linguistic terms, as is the case in a linguistic soil classification. Finally, the available measurements may only be partially relevant to the
J. HALL AND M. ANDERSON

Figure 1. Dimensions of hydrological problem classification

Decreasing repeatability of system states

- System state may be repeated many times, e.g. ‘ordinary’ winter hydrograph levels
- System state may be repeated very few times, e.g. extreme out-of-bank flows
- System state cannot be repeated, e.g. hydrologically activated landslide

Decreasing precision with which the system state can be identified

- System state is precisely measurable, e.g. water level measurement with an analogue device
- System state can only be bounded, e.g. digital measurement
- System state is only measurable in vague terms, e.g. soil is ‘sandy’

Decreasing relevance of available measurements to the phenomenon of interest

- Measured system state is the state variable to be predicted, e.g. hindcast time series
- Measured system state is only partially relevant to prediction, e.g. river levels recorded at a time of different land use
- Measured system state is not relevant to prediction, e.g. climate change

phenomenon of interest. This is the case in problems of regional generalization and prediction of system behaviour in states that are not represented in the measured record. A similar problem confronts climate modellers who wish to predict climate change under conditions (for example of greenhouse gas concentrations) that have not been observed in the past.

We argue that most attention in uncertainty analysis has hitherto focused on the base case problem, using, for example, methods of model optimization or conditioning. These current approaches presuppose the existence of a reasonably long, precise and relevant time series of measurements. Uncertainty in the broader range of situations introduced above is arguably more important than in the situations where there are test datasets. Under these more general circumstances it is necessary to assemble evidence about the causal factors for the event of interest (in the hydrological case: rainfall intensity and duration, slope saturation, etc.) and the relationship between causal factors and system response. We are therefore beyond the realm of statistical or even conceptual models that can be calibrated with data and are reliant on process-based understanding of the system of interest. Uncertainties in both our knowledge of the causal phenomena and the system response to those phenomena become of utmost interest. Unfortunately, evidence about these uncertainties is scarce, so we need to be as inclusive as possible in assembling relevant evidence. Evidence may include published or measured data relating to parameter distributions or bounds, knowledge about the accuracy of measurement methods and analogous cases. The information will appear in diverse formats, including statistical data, but also perhaps merely bounds on the possible values of a variable, or expert judgements about model dependability.

Even in situations where there are model test datasets, there is merit in also endeavouring to include in the uncertainty estimate prior evidence about variability in parameter values, perhaps from analogous sites, and expert reasoning about
model performance. The test dataset cannot reasonably be expected to capture all of the characteristics of a given model.

In practical terms, two problems arise: (i) making use of information in a range of formats from a variety of sources; (ii) projecting uncertain data through numerical models and, if appropriate, conditioning model results to account further for uncertainty. The solution to the first problem can be found in the various generalizations of probability theory that have emerged since the 1950s, including Choquet’s theory of capacities, random set theory, evidence theory, fuzzy set theory, possibility theory and Walley’s theory of imprecise probabilities. All of these approaches are based on various weakenings or reworkings of Kolmogorov’s axioms of probability in order to develop mathematizations that are more appropriate to the range of non-probabilistic information formats in which evidence about uncertainty will appear. A critical aspect is the use of expert knowledge. Much human reasoning about hydrological systems is possibilistic rather than strictly probabilistic. We reason about whether a given scenario could happen, without necessarily endeavouring to attach probabilities to the likelihood of it happening, particularly in situations of very scarce information. For example, until recently the Intergovernmental Panel on Climate Change resisted attaching probabilities to climate change scenarios, preferring only to use possible bounds. This type of reasoning is a natural response to uncertainty, but is much looser than probabilistic reasoning. The adoption of both possibilistic and probabilistic approaches results, in practice, in a combination of probabilistic analysis and interval analysis. Statistical data are handled in probabilistic terms, whilst imprecise measurements are handled as intervals, and fuzzy sets, which are nested families of intervals, are used to represent expert judgements.

The conventional numerical approach to projecting uncertain data through numerical models is Monte Carlo simulation, which, when combined with interval analysis, can be used to handle a range of generalized approaches to uncertainty handling. For very large computer models Monte Carlo integration is prohibitively expensive, but methods for statistical analysis of computer codes can be efficient (Craig et al., 2001; Kennedy and O’Hagan, 2001). In situations of scarce data, however, it will be impossible to construct full joint probability distributions in high-dimensional parameter space, and assumptions of independence between parameters can be difficult to justify. Nonetheless, expert reasoning can be used to constrain joint probability distributions so that combinations of variables are at least physically realizable. This leads to the notion of identifying permissible regions of parameter space and then either assuming local independence in those regions or proceeding with families of probability distributions constrained only by the fuzzy boundaries on the permissible region (Wolkenhauer, 2001). The approach has been adopted in hydrological modelling of hillslope processes in which soil classification (according to the USDA system) has been used to identify the region in model parameter space of interest and then published data from the literature have been used to develop parameter constraints that are appropriate for the region.

The modelling paradigm for handling uncertainty in predictions of extreme or unrepeatable hydrological processes, therefore, involves assembling evidence, in whatever form it may appear, about the relevant influencing factors and then projecting it though a causal model of the phenomenon of interest. There is nothing novel about this approach—it is widespread in disciplines from reliability engineering to climate modelling. It is, however, based on a rather different conception and methodology than is commonplace in hydrology, and has very important implications. The causal model used to conduct the extension greatly influences the dependability of the result. We can never be sure about the truth of the model. We can only make judgements, on the basis of evidence, about its partial dependability. Those judgements of dependability are not straightforward. Moreover, because the prediction is of a very rare or unrepeatable event, it is impossible to falsify directly. It can only be tested through reasoning about the various elements in the process of constructing the prediction. If the input variables can be shown to be dependable and the model can be shown to be dependable, then one would expect (provided no
mistake has been made in the integration) that the prediction is dependable. These are procedural arguments and demonstrate the importance of the process of data collection and analysis when making probabilistic inferences. Methods for evaluating the process of model construction and use and recording this information as meta-commentaries on model predictions are not well established. Yet, if we are to gain useful insights into the uncertainty in estimates of the likelihood of extreme or unrepeatable hydrological events, then these issues merit greater attention than they have received in the past.

References