INVITED COMMENTARY

HP TODAY

HYDROLOGICAL PROCESSES
Hydrol. Process. **21**, 841–845 (2007)
Published online 30 January 2007 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/hyp.6623

What do we mean by 'uncertainty'? The need for a consistent wording about uncertainty assessment in hydrology

Alberto Montanari*

Faculty of Engineering—University of Bologna, Via del Risorgimento 2, I-40136 Bologna

*Correspondence to:
Alberto Montanari, Faculty of Engineering—University of Bologna, Via del Risorgimento 2, I-40136 Bologna, Italy.
E-mail: alberto.montanari@unibo.it

Received 11 October 2006
Accepted 2 November 2006

Uncertainty estimation in hydrological modelling is receiving increasing attention by researchers and practitioners. However, the transfer of relative know-how from scientists to end-users is still difficult, notwithstanding the relevant research activity developed in the last 10 years. This problem was rightly pointed out in a recent commentary on HPToday by Beven (2006), who even wondered whether the results of uncertainty analyses will undermine the confidence of stakeholders in the science of hydrology.

One of the main reasons that so far prevented the hydrologic community from efficiently communicating the knowledge about uncertainty estimation is certainly the impracticality of a systematic testing of the many methods proposed recently. As Beven (2006) pointed out, to perform extensive validation in hydrology is not easy, and impossible in some cases. However, I think this is not sufficient to justify the unquestionable trouble that applied hydrologists have to deal with when they try to identify from the current literature the best uncertainty estimation method for their needs.

The topic of uncertainty assessment in hydrology suffers today from the lack of a coherent terminology and a systematic approach, which would allow us to clearly classify the practical problems to be solved and, consequently, the methods that can be used. Even the term 'uncertainty' itself is sometimes used without any reference to the precise scientific definition of its meaning. The result of this situation is that it is extremely difficult (if not impossible) to assess the prerogatives and limitations of individual methods that are currently used by researchers to quantify the reliability of hydrological simulations, design variables and forecasts.

The aim of this comment is to provide some personal opinions and ideas in order to better identify (a) the questions concerning uncertainty estimation currently raised in applied hydrology and (b) the peculiarities of the most widely-known uncertainty estimation methods. The purpose is to contribute to the discussion about the strategy that could be used in order to convey the results of uncertainty estimation to stakeholders without undermining their confidence in hydrological studies.

The Meaning of 'Uncertainty' for Hydrologists

Uncertainty is a common everyday experience in human life. A unique definition of uncertainty is hard to find in the literature. In a broad sense, we may say that uncertainty can be considered an attribute of information (Zadeh, 2005). This definition can be well applied to hydrology, where uncertainty was traditionally dealt with by using the probability theory. This used to be a common practice even in other scientific fields. Today, many authors are convinced that uncertainty can be efficiently dealt with in a much broader perspective, in which statistical information is one (albeit an important one) of the many possible forms of information (Zadeh, 2005; Langley, 2000). This is a very interesting view (discussed in the conclusion of this commentary). However, I would like to stress that non-statistical approaches (also known as possibilistic methods, explained later) are also referred to as

techniques based on 'imprecise probability' (Hall and Anderson, 2002). In other words, they are very useful when the available information does not allow one to describe the uncertainty in terms of probability, which, in my opinion, remains the most descriptive approach to uncertainty assessment.

The practical meaning of 'uncertainty' for hydrologists can be better understood by focusing on the related questions hydrologists themselves are most interested in. As a matter of fact, practitioners are often involved in assessing the uncertainty of two types of variables that are estimated in applications, namely, (a) uncertainty in the design variable or design process (like peak flow or flood hydrograph) and (b) uncertainty in the forecast (typically rainfall or river flow forecast). In both cases we deal with the uncertainty of a model output. Hydrologists are often interested in the reliability of model parameters and observed measures as well; however, this latter interest is often motivated by the need to obtain necessary information in order to estimate the uncertainty of a model output. It is important to note that uncertainty of the model parameters and observed measures is the subject of a relevant and ongoing research activity that focuses on techniques that are generally different with respect to those that are used to estimate the reliability of a model output (although some approaches may allow a joint assessment, an example is the simultaneous optimisation and data assimilation (SODA) method (Vrugt *et al*., 2005)). I personally believe it would be advisable to keep these scientific fields well distinguished in order to convey a clearer view to the end users.

In the remainder of this commentary we will focus our attention on the uncertainty assessment of the output of a generic hydrological model. This is the problem end users are most interested in.

An Attempt to Classify Uncertainty Assessment Methods for the Model Output

The most direct method to assess the uncertainty of a system output is to derive its statistics from knowledge of the statistical properties of the system itself and the input data (Langley, 2000). However, this approach may be limited by two main problems. First, the derivation of the statistics of the output can imply significant mathematical and numerical difficulties; second, the statistical properties of the system and the input may not be known in detail.

The first difficulty has stimulated the development of a first type of uncertainty assessment technique, namely, the approximate analytical methods. An example is the asymptotic reliability analysis, like the first-order reliability method (FORM) and second-order reliability method (SORM).

The second problem mentioned above may be even more difficult to deal with. For instance, the definition of the statistics of the system is a delicate step of the uncertainty assessment method that was recently proposed by Huard and Mailhot (2006) in a hydrological context.

Different methods have been devised to deal with the two problems mentioned above. A solution was found by studying the uncertainty of the model output by directly analysing the statistics of the model errors. In this way, another type of uncertainty assessment methods was developed. Accordingly, a model run is used to obtain a Monte Carlo simulation of the error itself. This approach was considered in many hydrological studies. It was recently used by Krzysztofowicz (2002) in the framework of the Bayesian Forecasting System (BFS). A similar philosophy was also applied by Montanari and Brath (2004) in a meta-Gaussian approach. Of course, the observed data are themselves uncertain, and therefore the model reliability analysis will not be correct in absolute terms (in the ideal situation of a perfect model, if we compare its response with the uncertain output observations that we assume to be correct, we may wrongly conclude that the model is uncertain). However, in any case, from a practical point of view, the difference between the model response and what we measure in the field provides important information for the sake of inferring reality based on the model output (Refsgaard *et al*., 2006).

A limitation of this type of technique could be the suitability of the stationarity and ergodicity assumptions that are operated on the error model in order to be able to derive its statistics from a simulation (which in hydrology is often short). However, it should be noted that the adoption of stationarity should not be always seen as a weakness; rather, it should be seen as a point that offers us a solution when we do not have any concrete information that could be translated into a different description. Therefore, in the specific case of the above techniques, the assumption of stationarity for the error model allows us to extend to unknown events the information we gained from the observation of the natural process, under a clearly understandable hypothesis. In other words, I believe it is much better to profit from this type of statistical information, on the basis of the stationarity assumption, than to neglect the information by questioning the assumption itself.

Another possibility to assess uncertainty is to derive the statistics of the system output numerically, by performing another type of Monte Carlo simulation. Once the statistics of the input data and system are known (or are assumed to be known), one may generate multiple model runs by randomly sampling the input data space and the system space, thereby obtaining a collection of outputs that can allow one to derive the statistics of the output itself. Here the system space is intended as the collection of all the behavioural modelling solutions that are obtained by varying the model structure and parameters (Refsgaard *et al*., 2006). This is the basic philosophy of

Hydrol. Process. **21**, 841–845 (2007)
DOI: 10.1002/hyp

the Generalized Likelihood Uncertainty Estimation (GLUE; Beven and Binley, 1992), which is probably the most used uncertainty assessment method in hydrology. In GLUE, the output of each simulation is weighted through a formal likelihood measure in order to give more credibility to the simulations produced by more reliable modelling solutions. A problem with this type of approach is the selection of an appropriate sampling procedure from the input space and, especially, the system space. In fact, though performing simulations with varying parameters is relatively simple, it is not yet clear how one can efficiently sample the space of the model structure. Moreover, a second problem is that the rescaled likelihood measure that is employed within GLUE (the literature has proposed many solutions (Beven and Freer, 2001)) is not a consistent estimate of the probability density of the model output. Therefore, the obtained confidence limits do not necessarily encompass the true observation with a given frequency (Montanari, 2005). In fact, the output of GLUE depends on the likelihood measure that is used and therefore is subjective. Unless this problem is solved, GLUE should not be considered as a probabilistic method, but instead should be considered as a weighted sensitivity analysis. Therefore the confidence limits provided by GLUE could be better named as sensitivity envelopes.

Being based on the user's judgement, it can be considered that GLUE is not dissimilar to another class of uncertainty assessment approaches that is increasingly gaining attention in hydrology recently, namely, the non-probabilistic methods.

Non-probabilistic methods are various generalizations of probability theory that have emerged since the 1950s, including random set theory, evidence theory, fuzzy set theory and possibility theory. In particular, fuzzy set theory and possibility theory have received considerable attention from hydrologists because 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.

On the basis of the above considerations, four types of techniques for assessing the uncertainty of the output of a hydrological model can be identified:

- approximate analytical methods;
- techniques based on the statistical analysis of model errors;
- approximate numerical methods/sensitivity analyses;
- non-probabilistic methods.

The above list can probably be updated, especially in view of the considerable research activity that is continuously being undertaken. Moreover, it is important to note that some techniques may fall across the different categories listed above. However, it is important to agree about a consistent classification in order to better convey the behaviour of each method.

How Should We Select the Most Appropriate Uncertainty Assessment Method in Hydrology?

To answer this question is certainly not easy, as the topic is continuously undergoing development. However, I believe that we should try to propose an identification procedure for the uncertainty assessment methods. What follows is only a collection of thoughts based on my personal opinion, with the aim of stimulating a discussion. An interesting review about the subject is also presented by Refsgaard *et al.* (2006).

First, I would say that the experiments with synthetic data might be a useful tool to test and show to end users the performances of the different methods. Beven (2006) questions the value of synthetic experiments in the face of real applications, "... when it is impossible to have secure knowledge about the different sources of uncertainty". However, I think that an uncertainty assessment method that is good in real world applications should be equally good when applied to a hypothetical numerical experiment. Therefore, I would say that synthetic tests could be an extremely useful tool in order to identify the main behaviour and optimal ambits of application of each technique (Montanari, 2005).

To start the discussion about what method to choose, let us first focus on the typical question an end user asks about uncertainty: 'Please give me an indication of a confidence (or a prediction) interval for my design (or forecasted) measure (river flow, rainfall, and so forth)'. We should remember that the confidence interval (CI) is defined as a range between two random variables with an associated probability p (the confidence level) such that if the CI is recalculated for many design measures according to the same method, a proportion p of the CIs would contain the true value. This is the indication we should provide when we talk about CIs in our studies.

The first indication to be provided in order to identify the most suitable uncertainty assessment method is that it must be able to take into account any type of useful information. This is a fundamental premise. For instance, in real time forecasting, a very important piece of prior information is provided by the last forecast error; the errors of hydrological models are significantly correlated (in many rainfall-runoff model applications the lag-one autocorrelation coefficient of the model errors can be as high as 0·8). This prior information is not taken into account in the classical formulation of GLUE and in the meta-Gaussian approach by Montanari and Brath (2004). Therefore, these methods should not be used in forecasting, unless some modifications are made. Correlation in the forecast errors is instead accounted for

Hydrol. Process. **21**, 841–845 (2007)
DOI: 10.1002/hyp

in the SODA method (Vrugt *et al.*, 2005) and in the BFS by Krzysztofowicz (2002), where it constitutes the prior information that is subsequently updated, through the Bayes' theorem, in light of the rainfall-runoff model reliability analysis.

When an extended series of observed variables is available, an important piece of information is indeed provided by the statistical analysis of the model errors. Therefore, in such conditions, these techniques should be considered. Of course, one has to verify the appropriateness of their underlying hypotheses. Testing the assumption of stationarity may be particularly worrisome when the database is limited. Probabilistic methods for uncertainty estimation have limited usefulness in the case of ungauged basins and should be used with care for estimating the uncertainty of exceptional events for which historical information is always scarce.

In the case of ungauged or scarcely gauged basins, non-probabilistic approaches, including sensitivity analyses, probably convey the most promising perspectives. In this case, GLUE can allow one to explore the variability of the model output by sampling the space of the possible modelling solutions. However, it should be clear for the end user that, in the case of ungauged basins, the uncertainty is always relevant and that the indication we derive by applying non-probabilistic methods cannot be expressed by means of a probability. We must admit that when only a limited amount of information is available the expression of uncertainty in terms of a probability is not possible.

To summarize the above considerations, the identification of the uncertainty assessment methods should take into account the following main issues:

- the type of model whose output uncertainty is to be inferred (simulation, forecasting);
- the type of information available (observed data, information about model uncertainty);
- the prerogatives of uncertainty assessment methods.

Are the Uncertainties in Hydrology Being Overestimated?

The title of this paragraph is quoted from Beven (2006), who stated that 'The answer to this has to be that we do not actually know'. My opinion about this issue is more optimistic. I would say that a realistic answer would carefully evaluate the amount of data available. When the information is scarce, I agree that the answer would be that 'we do not actually know'. In such a case we can only provide a sensitivity envelope that may have nothing to do with a confidence interval defined on probabilistic grounds (albeit that it might be very useful for practical applications).

However, when sufficient information is available, I believe we can efficiently judge whether uncertainty is correctly estimated. There are many examples of applications in which uncertainty is operationally assessed, with interesting and extensive results (see, for instance, the operational testing of the BFS operated by the US National Weather Service (Krzysztofowicz, 2002)).

I believe it is important to let the end users know that the hydrologic community has been developing many approaches for uncertainty estimation (see http://www.es.lancs.ac.uk/hfdg/uncertainty_workshop/uncert_intro.htm) and that uncertainty in hydrology is consistently addressed, and in some cases we can also efficiently address the practitioner's requests. What is required is to develop established and accepted guidelines for the identification of the appropriate uncertainty assessment method, depending on the purpose of the application and availability of data. We also need a clear view of the requirements that must be satisfied in order to compute probabilistic CIs or, alternatively, sensitivity envelopes. What hydrologists need is not a method that is best, but a theory and a terminology that are comprehensive and clear.

Perspectives for the Future

A view about future perspectives can be derived by recalling the fundamental premise mentioned above: an uncertainty assessment method must be able to take into account any type of useful information. This assertion opens the door to a relevant avenue of research: the development of a generalized theory that may allow us to profit from different types of information, either probabilistic or possibilistic (Langley, 2000).

A recent interesting contribution to this issue was given by Zadeh (2005), who showed how different types of constraints on uncertainty estimation (such constraints may be probabilistic, possibilistic, veristic, fuzzy-graph, bimodal and many others) can be combined within a Generalized Theory of Uncertainty (GTU). The possibility of combining the strengths of different methods for uncertainty estimation in hydrology, including also the user's knowledge and belief (Hall and Anderson, 2002), would undoubtedly allow us to move forward by a significant step.

Acknowledgements

I would like to thank Demetris Koutsoyiannis, Francesco Laio and Attilio Castellarin for their very constructive suggestions. Jim Buttle, Editor of HPToday, is also acknowledged for inviting me to write this commentary and providing helpful advice. This work was partially supported by the Italian Government through the research project 'Advanced techniques for estimating the magnitude and forecasting extreme hydrological events, with uncertainty analysis'.

DOI: 10.1002/hyp

References

Beven KJ. 2006. On undermining the science? *Hydrological Processes (HPToday)* 20: 3141–3146.

Beven KJ, Binley A. 1992. The future of distributed models: model calibration and uncertainty prediction. *Hydrological Processes* 6: 279–298.

Beven KJ, Freer J. 2001. Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. *Journal of Hydrology* 249: 11–29.

Hall J, Anderson M. 2002. Handling uncertainty in extreme or unrepeatable hydrological processes—the need for an alternative paradigm. *Hydrological Processes (HPToday)* 16: 1867–1870.

Huard D, Mailhot A. 2006. A Bayesian perspective on input uncertainty in model calibration: Application to hydrological model "abc". *Water Resources Research* 42: W07416, DOI:10·1029/2005WR004661.

Krzysztofowicz R. 2002. Bayesian system for probabilistic river stage forecasting. *Journal of Hydrology* 268: 16–40.

Langley RS. 2000. Unified approach to probabilistic and possibilistic analysis of uncertain systems. *Journal of Engineering Mechanics* 126: 1163–1172.

Montanari A. 2005. Large sample behaviors of the generalized likelihood uncertainty estimation (GLUE) in assessing the uncertainty of rainfall-runoff simulations. *Water Resources Research* 41: W08406, DOI:10·1029/2004WR003826.

Montanari A, Brath A. 2004. A stochastic approach for assessing the uncertainty of rainfall-runoff simulations. *Water Resources Research* 40: W01106, DOI:10·1029/2003WR002540.

Refsgaard JC, van der Sluijs JP, Brown J, van der Keur P. 2006. A framework for dealing with uncertainty due to model structure error. *Advances in Water Resources* 29: 1586–1597.

Vrugt JA, Diks CGH, Gupta HV, Bouten W, Verstraten JM. 2005. Improved treatment of uncertainty in hydrologic modeling: Combining the strengths of global optimization and data assimilation. *Water Resources Research* 41: W01017, DOI:10·1029/2004WR003059.

Zadeh LA. 2005. Toward a generalized theory of uncertainty (GTU)—an outline. *Information Sciences* 172: 1–40.

Hydrol. Process. **21**, 841–845 (2007)
DOI: 10.1002/hyp