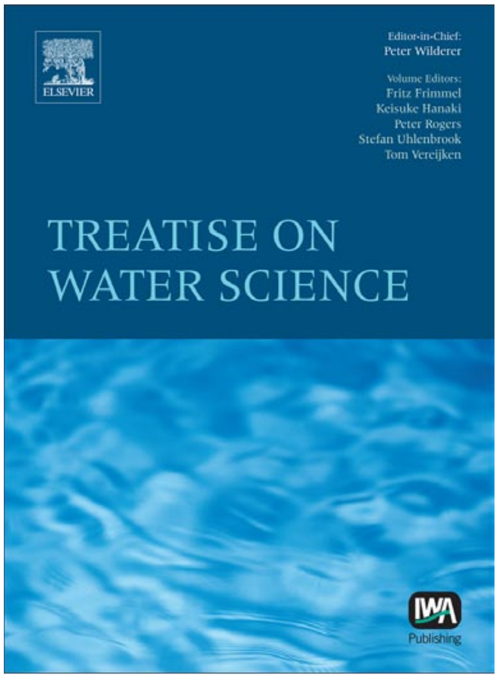

Montanari A (2011) Uncertainty of Hydrological Predictions. In: Peter Wilderer (ed.) Treatise on Water Science, vol. 2, pp. 459–478 Oxford: Academic Press.

2.17 Uncertainty of Hydrological Predictions

A Montanari, University of Bologna, Bologna, Italy

2.17.1	**Introduction**	459
2.17.2	**Definitions and Terminology**	461
2.17.2.1	Probability	461
2.17.2.2	Randomness	461
2.17.2.3	Random Variable	461
2.17.2.4	Stochastic Process	461
2.17.2.5	Stationarity	461
2.17.2.6	Ergodicity	462
2.17.2.7	Uncertainty	462
2.17.2.8	Global Uncertainty and Individual Uncertainties	462
2.17.2.9	Uncertainty Assessment	462
2.17.2.10	Probabilistic Estimate/Estimation/Assessment of Uncertainty (Probabilistic Uncertainty)	462
2.17.2.11	Nonprobabilistic Estimate/Estimation/Assessment of Uncertainty (Nonprobabilistic Uncertainty)	462
2.17.2.12	Confidence Band	463
2.17.2.13	Equifinality	463
2.17.2.14	Behavioral Model	463
2.17.3	**Classification of Uncertainty and Reasons for the Presence of Uncertainty in Hydrology**	463
2.17.3.1	Inherent Randomness	464
2.17.3.2	Model Structural Uncertainty	464
2.17.3.3	Model Parameter Uncertainty	464
2.17.3.4	Data Uncertainty	465
2.17.3.5	Operation Uncertainty	465
2.17.4	**Uncertainty Assessment**	465
2.17.5	**Classification of Approaches to Uncertainty Assessment**	465
2.17.5.1	Research Questions about Uncertainty in Hydrology	465
2.17.5.2	An Attempt of Classification	466
2.17.6	**Assessment of the Global Uncertainty of the Model Output**	467
2.17.6.1	Analytical Methods	467
2.17.6.2	The Generalized Likelihood Uncertainty Estimation	467
2.17.6.3	The Bayesian Forecasting System	468
2.17.6.4	Techniques Based on the Statistical Analysis of the Model Error	469
2.17.6.5	Bayesian Model Averaging	470
2.17.6.6	Machine Learning Techniques	471
2.17.7	**Assessment of Data Uncertainty**	472
2.17.7.1	Precipitation Uncertainty	472
2.17.7.2	River Discharge Uncertainty	473
2.17.8	**Assessment of Parameter Uncertainty**	474
2.17.8.1	The MOSCEM-UA Method	474
2.17.8.2	The AMALGAM Method	474
2.17.9	**Assessment of Model Structural Uncertainty**	474
2.17.10	**Uncertainty Assessment as a Learning Process**	475
2.17.11	**Conclusions**	475
Acknowledgments		476
References		476

2.17.1 Introduction

Hydrological modeling is receiving increasing attention from researchers and practitioners. The increasing availability of mathematical tools and computing power together with an improved understanding of the dynamics of hydrological processes has favored the continuous development of new modeling approaches in the past few decades (see Chapter 2.16 Hydrological Modeling). Hydrological modeling is an attractive option today for solving many practical problems of environmental engineering, flood protection, water resource management, and applied hydrology in general.

Setting up a hydrological model in order to solve a practical problem requires the application of proper procedures of model identification, parameter calibration, hypothesis testing, model testing (also called model validation), and uncertainty assessment. The above procedures are often strictly related and are the subject of an increasing research activity by hydrologists. In particular, uncertainty estimation is very much related with parameter calibration and model validation. It consists of a verification of the hydrological model appropriateness and performances finalized to providing a quantitative assessment of its reliability.

As a matter of fact, uncertainty estimation in hydrological surface and subsurface modeling is today one of the most important subfields of hydrology, according to the numerous contributions in recent scientific literature. Uncertainty reduction is also one of the main goals of the Prediction in Ungauged Basins (PUB) initiative promoted by the International Association of Hydrological Sciences.

While quantitative uncertainty assessment in hydrology is often considered a relatively new topic, it is worth noting that hydrologists were aware of uncertainty and used to deal with it because the first hydrological studies and applications were carried out. In particular, empirical techniques were used to compensate for the lack of information about model reliability. For instance, hydrologists are well used to adopt safety factors or allowance for freeboard, which are usually set basing on consensus, expert opinion, and empirical evidence. These safety factors were the first and very useful tools to take into account inherent uncertainty and imperfect knowledge of hydrological processes in hydrological design. However, expert knowledge is by itself subjective and referred to specific contexts and situations. The call for a generalized and systematic approach to uncertainty estimation in hydrology is the motivation for the renewed interest in the past few years.

One of the reasons why uncertainty assessment in hydrology was not much investigated on theoretical basis until the recent past is that hydrological modeling itself is a relatively young discipline. In fact, the first hydrological models were the rational formula proposed by Kuichling (1889) (although the principles of the method were introduced by Mulvaney (1851)), and the unit hydrograph model proposed by Sherman (1932). Most of the hydrological models that are used today were proposed after the 1960s. The interest in new techniques for uncertainty assessment was stimulated by Spear and Hornberger (1980), who introduced the generalized sensitivity analysis methodology, also known as regional sensitivity analysis. Their work inspired the development of the generalized likelihood uncertainty estimation (GLUE; Beven and Binley, 1992; see Section 2.17.6.2), which works under the hypothesis that different sets of model parameters/structures may be equally likely as simulators of the real system. In the 1990s the emerging need for reliable techniques for uncertainty estimation, for the multitude of modeling situations and approaches that are experienced in hydrology, stimulated the development of many methods (for a long, though still incomplete, list one can refer to Liu and Gupta (2007) and Matott *et al.* (2009)).

Another reason limiting the use of uncertainty assessment methods is that the transfer of the know-how about uncertainty in hydrology from scientists to end-users was and still is, difficult, notwithstanding the relevant research activity mentioned above. Pappenberger and Beven (2006) provided an extensive analysis of this issue. A relevant problem today is that uncertainty assessment in hydrology suffers from the lack of a coherent terminology and a systematic approach. The result of this situation is that it is extremely difficult (if not impossible) to obtain a coherent picture of the available methods. This lack of clarity is an example of linguistic uncertainty (Regan *et al.*, 2003). Therefore, much is still to be done to reach a coherent treatment of the topic.

Quantitative uncertainty assessment in conditions of data scarcity is a very difficult task, if not impossible, in some cases. Usually, uncertainty estimation in applied scientific modeling is dealt with by comparing the model output with observed data, by borrowing concepts from statistics. According to this procedure, model reliability is quantified in a probabilistic framework. However, statistical testing becomes not as reliable in situations of data scarcity and therefore the use of statistical concepts for uncertainty assessment in hydrology sometimes may not be appropriate. This is one of the reasons why hydrologists are looking for different procedures that can be complementary or alternative to statistics. Moreover, uncertainty in hydrology might arise from limited knowledge (epistemic uncertainty, see Section 2.17.3) or from natural variability. In the former case, we deal with uncertainties that might not be aleatory in nature. They can be treated with statistical methods (e.g., the BATEA method, see Section 2.17.7.1), but many authors question the validity of statistics in this case and prefer nonstatistical approaches. These procedures are generally conceived in order to allow incorporation of expert knowledge in a theoretically based framework. They are characterized by a certain degree of subjectivity, which needs to be reduced as much as possible in order to allow their application in situations where knowledge is lacking.

Therefore, different philosophies and approaches for quantifying the reliability of hydrological models were recently proposed. As a result, an active debate recently began about the relative advantages of each of them. Such debate in many cases assumed a philosophical behavior, because the philosophy underlying each method is one of the main subjects of the discussion. On the one hand, such a debate stimulated additional developments and insights in itself; on the other, it is still not clear which approach is most appropriate given the needs of the user. For this reason, the hydrologic scientific community still calls for more pragmatism in uncertainty estimation.

On the one hand, hydrology is a science where uncertainty is very significant. Progress in monitoring techniques, process understanding, and modeling will certainly reduce uncertainty in the future but will never eliminate it. On the other hand, hydrologists are in charge of providing design variables that play a fundamental role in water engineering, civil protection, and water resource management. Therefore, it is clear that the efficient real world use of an uncertain design variable should necessarily be based on uncertainty assessment.

This chapter aims at presenting a comprehensive introduction to the subject of uncertainty assessment in hydrology. After presenting a brief glossary and a discussion about the reasons for the presence of uncertainty in hydrology, a review of the most-used approaches to uncertainty assessment is presented.

2.17.2 Definitions and Terminology

There is currently a linguistic uncertainty affecting the topic of uncertainty assessment in hydrology (Regan *et al.*, 2003; Beven, 2009), meaning that an agreed terminology is still lacking. Some basic definitions are provided in the following.

2.17.2.1 Probability

Probability can be defined in different ways. In fact, probability is currently interpreted according to two broad and distinguished views.

The classical frequentist view of probability defines the probability of an event occurring in a particular trial as the frequency with which it occurs in a long sequence of similar trials. In a Bayesian or subjectivist view, the probability of an event is dependent upon the state of information available and this information can include expert opinion. Probability theory forms the basis of classical statistics, which has estimators based on a likelihood function that represents how likely an observed data sample is for a given model and parameter set.

2.17.2.2 Randomness

Randomness is a term that is used within science with different meanings. In statistics, and hydrology as well, a random process is such that its outcome cannot be predicted deterministically. Randomness does not imply lack of knowledge about the process dynamics or impossibility to set up a deterministic model for it. However, if a deterministic model can be set up for a process, randomness implies that such a model cannot perfectly predict the process outcome.

For instance, in the case of a roulette wheel, if the geometric and dynamic behaviors of the system are perfectly known, then the number on which the ball will stop would be a certainty. However, one is fully aware that even a small imperfection in the description of the geometry of the system and/or its initial conditions makes the outcome of the experiment unpredictable. A probabilistic description can thus be more useful than a deterministic one for analyzing the pattern of outcomes of repeated rolls of a roulette wheel. Physicists face the same situation in kinetic theory of gases, where the system, while deterministic in principle, is very complex so that only a statistical description of its properties is feasible.

Another example is the experiment of dropping balls into a spiked sieve. Here, the geometry of the system is perfectly known as well as the initial and boundary conditions. However, once a ball is dropped in the sieve, it is impossible to predict deterministically its trajectory, because no one can predict which way the ball will follow after hitting a spike. However, the distribution of the balls at the bottom of the sieve is well known to be Gaussian. In this case, the full comprehension of the geometry and dynamics of the system does not allow one to set up a deterministic description, while a stochastic description can provide a satisfactory model. Actually, one cannot exactly predict the number of balls in each bar, but the probabilistic prediction will have a small uncertainty.

An important discovery of the twentieth-century physics was the random character of all physical processes that occur at subatomic scales and are governed by the laws of quantum mechanics. This means that probability theory is required to describe nature. This type of interpretation was questioned by many scientists, as the famous quote by Albert Einstein, from a letter to Max Born, clearly testifies: "I am convinced that He does not play dice."

A similar controversy currently occurs in hydrology (for an interesting discussion, see Koutsoyiannis *et al.* (2009)). The trend toward the so-called physically based models induced in the last few decades the inspiration to pursue a completely deterministic description of hydrological systems, through a better understanding of the internal dynamics of hydrological processes. However, such deterministic description is so complicated that only a probabilistic treatment is possible. This does not mean that knowledge is unuseful. On the contrary, it allows one to set up a plausible probabilistic description of the random outcome.

2.17.2.3 Random Variable

A random variable maps all possible outcomes from a random event into the real numbers. As such, it is affected by uncertainty and cannot be deterministically predicted. Random variables can assume discrete and continuous values.

2.17.2.4 Stochastic Process

A stochastic process can be defined as a collection of random variables. For instance, if we assume that the river flow at time t is a random variable, then the time series of river flow observations during an assigned observation period is a realization of a stochastic process. While a deterministic process gives only one possible value of its output under assigned initial and boundary conditions (as it is the case, e.g., for the solution of an ordinary differential equation), the output of a stochastic process is affected by some uncertainty that is described by the corresponding probability distributions. This means that there are many possible paths for the evolution of the process, with some of them being more likely and others less. A stochastic process can assume discrete or continuous values. Although the random variables of a stochastic process may be independent, in most commonly considered situations in hydrology, they exhibit statistical correlations. A stochastic process can include a deterministic representation but always includes a random component which makes its output uncertain.

2.17.2.5 Stationarity

A stochastic process is strictly stationary when the joint probability distribution of an arbitrary number of its random variables does not change when shifted in time or space. As a result, parameters such as the statistics of the process also do not change over time or position. Stationarity is a property of the mathematical representation of the system, or an ensemble of outcomes from a repeatable experiment, and therefore does not constitute an actual property of the natural process itself. This latter follows just one trajectory and therefore its outcome is unique, because nature and life do not enable repeatability. Stationarity is a property that is used in statistics in order to make inference about the physical process and therefore does not imply any assumption on the natural process itself.

It is interesting to mention that the opposite of stationarity is nonstationarity, which implies that the above statistics change accordingly to deterministic functions of time, where deterministic means that the above-mentioned functions should be known independently of the data and should apply to any time, past, present, and future (Papolulis, 1991). Conversely, if the above functions are random (i.e., realizations of stationary stochastic processes), then the process is stationary.

The concept of stationarity is a way to find invariant properties in complex natural systems. In view of what was anticipated above, it is important to note that stationarity does not imply that the statistics of a realization of a process are constant in time. Actually, such statistics are affected by sampling variability and therefore they certainly change after a time shift. The crucial issue is to detect if such a change exists in the process and can be expressed through a deterministic function of time.

Recently, the scientific literature presented contributions stating that stationarity is dead because of hydrological change and climate change. Actually, stationarity is an assumtpion and therefore can hardly be dead.

2.17.2.6 Ergodicity

A stochastic process is said to be ergodic if its statistical properties can be deduced from a single, sufficiently long sample (realization) of the process.

2.17.2.7 Uncertainty

Uncertainty can be defined as an attribute of information (Zadeh, 2005; Montanari, 2007). In the context of hydrology, uncertainty is generally meant to be a quantitative indication of reliability for a given hydrological quantity, either observed or inferred by using models. The indication of reliability can be provided by estimating the error affecting the quantity or the expected range of variability (due to uncertainty) for the quantity itself. Uncertainty can be broadly grouped into two major categories, namely, aleatory and epistemic uncertainty (see Section 2.17.3), and can be inferred by using probabilistic or nonprobabilistic methods.

2.17.2.8 Global Uncertainty and Individual Uncertainties

Global uncertainty can be defined as the discrepancy between the model output and the true value of the corresponding variable. Different uncertainties can compensate each other in the formation of the global uncertainty; for instance, parameter errors can compensate, at least in part, for data errors and model structural errors. These different uncertainties are termed individual uncertainties and are specifically referred to with a terminology which recall their causal origin, such as parameter uncertainty and model structural uncertainty (see Section 2.17.3.2 for an extended description). The terms above are not formally defined and therefore some linguistic uncertainty is present. For instance, the terms parameter uncertainty, input uncertainty, and model structural uncertainty should be used to indicate the uncertainty affecting the model parameters, input, and structure, respectively. Hereafter this is the meaning that will be used in this chapter. However, these terms are sometimes used to indicate the part of uncertainty in the model output that is caused by imperfect parameters, input, and model structure, respectively.

While global uncertainty is relatively easy to estimate *a posteriori*, for instance, by computing the difference between the model output and the corresponding observed variable (under the assumption that this latter is correct), the identification of the contribution of individual uncertainties above is impossible, unless assumptions are introduced or independent observations are available (see Section 2.17.4). This means that it is usually difficult, if not impossible, to assess whether the model performance is affected by, say, a parameter error rather than a model structural error.

2.17.2.9 Uncertainty Assessment

In what follows, we refer to uncertainty assessment to mean a quantitative evaluation of uncertainty affecting a hydrological variable, parameter, or model. Uncertainty estimation and uncertainty quantification will be considered synonymous with uncertainty assessment, which is different from uncertainty analysis and uncertainty modeling. The former is a preliminary step of uncertainty assessment aimed at identifying the reasons for the presence of uncertainty and the nature of uncertainty itself, while the latter term refers to the tools that are used for uncertainty assessment.

2.17.2.10 Probabilistic Estimate/Estimation/Assessment of Uncertainty (Probabilistic Uncertainty)

We will use the term probabilistic estimate of uncertainty to mean that uncertainty estimation for a given hydrological quantity has been carried out consistently with formal probability theory. In the probabilistic approach, uncertainties are characterized by the probabilities associated with events. Therefore, if one refers to the output of a hydrological model, the related probability distribution should actually provide an estimate of the frequency with which the true values fall within a given range.

2.17.2.11 Nonprobabilistic Estimate/Estimation/Assessment of Uncertainty (Nonprobabilistic Uncertainty)

Nonprobabilistic methods to uncertainty estimation in hydrology are frequently applied. Nonprobabilistic 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 (Jacquin and Shamseldin, 2007). 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 endeavoring to attach probabilities to the likelihood of it happening, particularly in situations of very scarce information.

In a more general context, we will refer to nonprobabilistic uncertainty when the estimation is carried out by using other approaches than formal probabilistic ones. This category includes probabilistic methods where some of the underlying assumptions are relaxed.

2.17.2.12 Confidence Band

A range around an estimated quantity that encompasses the true value with a probability $1-\alpha$, where α is the significance level and $1-\alpha$ is the confidence level. It is worth pointing out that the terminology is sometimes ambiguous. Some authors use the term confidence band or confidence interval when referring to the distribution of estimates that cannot be observed (e.g., a model parameter), while the term prediction interval is used when referring to the distribution of future values. Moreover, some authors indicate with the term tolerance interval a range in the observations that encompasses a $1-\alpha$ proportion of the population of the related random variable. For more details, the reader is reffered to Hahn and Meeker (1991).

Figure 1 shows an example of confidence bands computed with the meta-Gaussian approach (Montanari and Brath, 2004; see Section 2.17.6.2) for river flow simulations referred to the Samoggia River at Calcara (Italy). It is interesting to note that the shape of the confidence bands themselves provides indications about the goodness of the fit provided by the model. Moreover, the skew in the prediction distribution results indicates that a systematic error is likely to be present.

2.17.2.13 Equifinality

Equifinality implies that in a system interacting with its environment a given end state can be reached by more than one potential mean. The term is due to von Bertalanffy (1968), the founder of general systems theory. The idea of equifinality suggests that similar results may be achieved with different initial conditions, different model parameters, and different model structures. In hydrology the concept of equifinality was introduced by Beven (1993) as an unavoidable effect of the presence of uncertainty. For an extended discussion, see Beven (2006a). Equifinality leads to the idea of multimodeling solutions in hydrology (see Section 2.17.6.5).

2.17.2.14 Behavioral Model

Within the context of equifinality, a behavioral model is one that provides an acceptable simulation of observed natural processes. In a multimodel approach, the collection of behavioral models provides a means for assessing the uncertainty of their output (see Section 2.17.6.5).

2.17.3 Classification of Uncertainty and Reasons for the Presence of Uncertainty in Hydrology

There have been many attempts presented by the literature to classify uncertainty in hydrology. The proposed solutions were not always in agreement because, given the uncertain nature of hydrological processes, it is sometimes impossible to unambiguously decipher the reason for the presence of errors. It is generally agreed that uncertainties can be grouped into two major categories: (1) natural variability (also called structural uncertainty, aleatory, external, objective, inherent, random, irreducible, or stochastic uncertainty) and (2) knowledge uncertainty (also called epistemic, functional, internal, reducible, or subjective uncertainty (**Table 1** in NRC, 2000; Koutsoyiannis *et al.*, 2009; Hall and Solomatine, 2008). These two categories have different ramifications. In fact, the global uncertainty of a given model or variable may be characterized in three ways: purely structural, partly epistemic and partly structural, and purely epistemic (Cullen and Frey, 1999). When evaluating model performances and when possible, the different types of uncertainty should be separated (Cullen and Frey, 1999; Hoffman and Hammonds, 1994; Nauta, 2000; Sonich-Mullin, 2001). However, this is not always possible and therefore epistemic uncertainty and natural variability are often dealt with in an integrated fashion.

Other classifications were proposed. According to the causes for the presence of uncertainty in hydrology (which nevertheless are not always identifiable), one may identify the following categories: (1) inherent randomness (the geometry

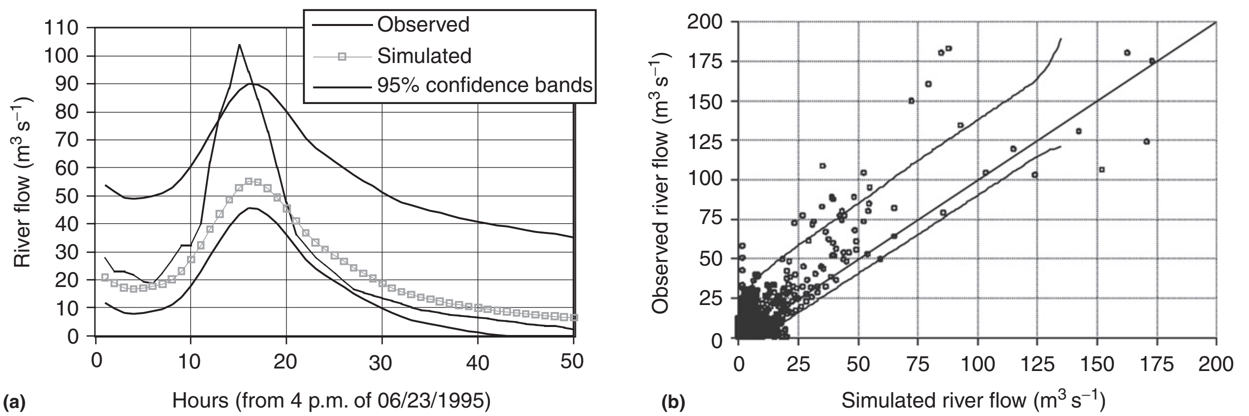

Figure 1 (a) Example of confidence bands computed with the meta-Gaussian approach (Montanari and Brath (2004); see Section 2.17.6.2) for a flood event occurred in the Samoggia River at Calcara (Italy) in 1995. (b) Example of confidence bands computed with the meta-Gaussian approach (Montanari and Brath (2004); see Section 2.17.6.2) drawn on a scatterplot of observed versus simulated hourly river flows for the Samoggia River at Calcara (Italy) during the years 1995–97.

Table 1 Uncertainty assessment methods in hydrology, along with their classification (see Section 2.17.5) and purpose (see Sections 45.6–45.10)

Assessment method	*Classification*	*Type of uncertainty estimated*
AMALGAM	Nonprobabilistic, parameter estimation	Parameter
BATEA	Probabilistic, parameter estimation, uncertainty assessment, sensitivity analysis	Precipitation induced
BFS	Probabilistic, Bayesian	Global
BMA	Probabilistic, multimodel	Global
DYNIA	Nonprobabilistic, identifiability analysis	Parameter
GLUE	Nonprobabilistic (when an informal likelihood is used), parameter estimation, uncertainty assessment, sensitivity analysis	Global, parameter, data, structural
IBUNE	Probabilistic, parameter estimation, uncertainty assessment, sensitivity analysis	Global, precipitation induced, model structure induced
Machine learning	Nonprobabilistic	Usually global, in principle all
Meta-Gaussian	Probabilistic, data analysis	Global
MOSCEM-UA	Nonprobabilistic, parameter estimation, sensitivity analysis	Parameter
SCE-UA	Probabilistic, parameter estimation	Parameter

Classification is ambiguous in some cases; it distinguishes between probabilistic and nonprobabilistic methods, as well as among the seven categories introduced by Matott *et al.* (2009) (see Section 2.17.5.2).

AMALGAM, a multialgorithm genetically adaptive method for multiobjective optimization; BATEA, Bayesian total error analysis; BFS, Bayesian forecasting system; BMA, Bayesian multimodel analysis; DYNIA, dynamic identifiability analysis; GLUE, generalized likelihood uncertainty estimation; IBUNE, integrated Baysian uncertainty estimator; MOSCEM-UA, multiobjective shuffled complex evolution University of Arizona; SCE-UA, shuffled complex evolution university of Arizona. References for the methods are in the text.

of the control volumes, the weather, etc.); (2) model structural uncertainty that reflects the inability of a model to represent precisely the true behavior of the system; (3) model parameter uncertainty; and (4) data uncertainty. When using models to make engineering or management decisions about hydrologic systems we also have to deal with (5) operation uncertainties (associated with construction and maintenance; Loucks and Van Beek, 2005). The above sources of uncertainty are briefly discussed in the following.

It is generally agreed that uncertainty in hydrology cannot be eliminated, no matter if it is epistemic in nature or induced by inherent randomness. For instance, rainfall inputs to a catchment might be highly structured, with different structures in different events that lead to nonrandom errors in estimates of areal rainfall. This type of error can be reduced with new measurement techniques but cannot be fully removed.

2.17.3.1 Inherent Randomness

Inherent randomness is one of the main reasons for the presence of uncertainty and is a intrinsic behavior of hydrological processes. For instance, a deterministic description of subsurface flow paths is impossible. Different soils and rocks, irregular macropores, faults and cracks with their heterogeneous patterns in both space and time, combined with two phase flows, varying wetting fronts, form a extremely complex system, for which a deductive description is impossible (Koutsoyiannis *et al.*, 2009). Inherent randomness emerges also from meteorology, variability of the surface flow paths, and so on. It is not only related to a coarse description of the system (that also induces uncertainty, which is nevertheless epistemic at least in part; in fact, it could be reduced by an increased capability to monitor the processes at finer spatial and temporal scales) but rather related to the effective impossibility to describe deterministically the inherent variability of the process. Inherent randomness has been long discussed by the hydrologic community in the recent past (Koutsoyiannis *et al.*, 2009; Koutsoyiannis, 2009). It has been argued that dynamical systems theory has well shown that uncertainty can emerge even from an insignificant perturbation of the initial conditions of a pure, simple, and fully known deterministic (chaotic) dynamics.

2.17.3.2 Model Structural Uncertainty

In the ideal situation in which perfect input data are available and model parameters are perfect, model structural uncertainty is defined as the uncertainty in the model output induced by the inhability of the hydrological model to perfectly reproduce the dynamics of hydrological systems. This means that the model output would still be uncertain even in the ideal situation in which no other uncertainties are present. Model structural uncertainty can be induced by imperfect model structure or lack of computational power. If the reason for the presence of uncertainty is an incorrect selection of the model or the computational tools, then model structural uncertainty is epistemic; on the other hand, the effective impossibility to describe the system with a mathematical model induces the presence of irreducible uncertainty. Given that model structural uncertainty is epistemic at least in part, the search for improved modeling tools has been the main focus of the hydrologic scientific community in the last few decades.

2.17.3.3 Model Parameter Uncertainty

Model parameter uncertainty is the result of the lack of a sufficiently extended database of good quality, or the inefficiency of the optimization algorithm and/or the related objective function, which induce parameter estimates to be significantly uncertain even if a perfect model and a perfect knowledge of the system were available. This is a relevant problem in all the hydrological applications and motivates the intense efforts that were dedicated to parameter estimation and related uncertainty assessment (e.g., Ibbitt and O'Donnell, 1974; Alley, 1984; Kleisenn *et al*., 1990; Gan and Burges, 1990; Duan *et al*., 1992; Brath *et al*., 2004; Vrugt *et al*., 2003a; 2003b; Vrugt and Robinson, 2007). Parameter estimation can be coupled with sensitivity analysis and model diagnostic to identify the most sensitive parameters in periods of model failures, thus gaining insights into the reasons for model inadequacy (Sieber and Uhlenbrook, 2005). To this end, Wagener *et al*. (2003) proposed the dynamic identifiability analysis (DYNIA).

Usually, an objective function is used to calibrate the model parameters to observed data. Independently of the objective function and the tools employed to optimize the model parameters, most hydrological models suffer from the existence of multiple optima of the objective function itself and the presence of high interaction or correlation among the parameters. These problems make parameter calibration uncertain even when a relatively large database is at disposal (Kuczera and Mroczkowski, 1998). Parameter uncertainty also arises when the parameters are not calibrated but rather estimated on the basis of field surveys or expert knowledge, for instance, while defining land-cover parameters. Parameter uncertainty can be epistemic at least in part.

2.17.3.4 Data Uncertainty

Data uncertainty is an emerging problem that is gaining renewed attention by hydrologists in the recent past (see, e.g., Di Baldassarre and Montanari, 2009; Dottori *et al*., 2009; Koussis, 2009; Petersen-Øverleir and Reitan, 2009). In fact, even modern technologies cannot avoid the presence of a significant approximation in observations of, say, rainfall, river flows (for both low and high flows), and so forth. Data uncertainty emerges from limitation of the monitoring techniques (instrumentation error, rating curve approximations, etc.) or variability of the spatial and temporal distribution of the observed hydrological variables (spatial variability of rainfall, time variability of streamflow, etc.). It follows that hydrological models are optimized against imperfect data and therefore an error is induced in hydrological simulations. Data uncertainty has both epistemic and aleatory components and therefore it is particularly important how observation errors are treated. Some authors claim that treating data error with purely statistical approaches may induce overconditioning in hydrological modeling (Beven, 2006a).

2.17.3.5 Operation Uncertainty

Operation uncertainty arises when hydrological models are used in the real world. In fact, it is well known that in real-time applications uncertainties of different nature are present that do not affect off-line exercises. Often the data and the initial and boundary conditions cannot be preliminary checked, the computational time might become a relevant constraint, the end-users operate under stress and therefore the human error becomes more likely, there is a weak ability to identify decision criteria, communication becomes difficult, and so forth. Operation uncertainty is difficult to assess, is rarely considered by researchers, and represents an emerging awareness among hydrological modelers and end-users. As a matter of fact, the identification of the most suitable model should be carried out in view of operation uncertainty as well. Data assimilation can be used to constrain uncertainty during model application.

2.17.4 Uncertainty Assessment

It is well known that uncertainty assessment in hydrology is a topical issue. Already in 1905, W.E. Cooke, who was issuing daily weather forecasts in Australia, stated: "It seems to me that the condition of confidence or otherwise form a very important part of the prediction, and ought to find expression." Uncertainty assessment in hydrology involves the analysis of multiple sources of error, the main ones being outlined in Section 2.17.3. The contribution of these latter to the formation of the global uncertainty cannot be quantified independently, unless (1) one is willing to introduce subjective assumptions about the nature of the individual error components or (2) independent observations are available for estimating each source of error. As an example for the latter solution, the reader is referred to Winsemius *et al*. (2006, 2008) where gravity and evaporation measurements are used to constrain the water balance and the land surface parameters, respectively, for a rainfall–runoff model.

However, in some hydrological applications it is not necessary to separate different sources of error. For this reason in many cases, uncertainty is assessed in an aggregated solution, therefore quantifying global uncertainty.

2.17.5 Classification of Approaches to Uncertainty Assessment

This section aims to propose a classification of uncertainty assessment methods in hydrology. Classifying the methods is useful to clarify their behavior and operational purpose. However, it should be premised that such a classification might be subjective, because some methods lend themselves to different interpretations of their nature and scope.

2.17.5.1 Research Questions about Uncertainty in Hydrology

The uncertain nature of hydrology has pushed hydrologists to raise many questions related to uncertainty assessment. The most urgent ones are those related to quantifying the reliability of the output variables of hydrological models (forecasts, simulations, etc.). Hydrological simulation is often used in real-time prediction systems for natural hazards or for assessing long-term effects of climate change or for assessing the reliability of proposed water resource management strategies. In these cases, quantifying the uncertainty of the hydrological model response is extremely important from a societal point of view.

Uncertainty assessment in hydrology includes additional research issues. Among them, there is the call for assessing the uncertainty of observed data, model parameters, and model structure. These issues are also significant for gaining further insight into the dynamics of hydrological processes. Indeed, to identify the most appropriate model is a means to provide support to hydrological theory. Therefore, uncertainty assessment became strongly related to parameter estimation, multiobjective optimization, model identification, model building, model diagnostics, model averaging, data collection, and information theory in general. All topics in this list have gained the attention of researchers in recent years and are often allocated under the one umbrella of uncertainty assessment in hydrology. Indeed, it would be helpful for end-users to formally identify such subtopics and the related research questions.

2.17.5.2 An Attempt of Classification

The traditional way of dealing with uncertainty in science is through statistics and probability (see, e.g., Montanari *et al.*, 2009) but, as mentioned above, nonprobabilistic approaches to uncertainty analysis are also popular in hydrology.

In some cases, it is not easy to classify an approach as either probabilistic or not. In fact, there are some methods that are based on probability theory, but in real-word applications simplifying assumptions are often introduced which finally lead to a nonprobabilistic estimation of the likelihood of a given scenario. Such assumptions are introduced in order to overcome operational problems, for instance, due to lack of enough data to support a statistical application.

The decision to use probabilistic or nonprobabilistic methods is currently the most controversial issue in hydrologic uncertainty analysis. This debate has raised the very relevant question about the capability of probabilistic and nonprobabilistic methods to correctly infer the frequency properties of hydrological simulations and predictions (see, e.g., Beven, 2006a; Montanari, 2005, 2007; Mantovan and Todini, 2006; Beven *et al.*, 2007, 2008). Criticism about probabilistic methods is focused on the concern that for many data sets it is not clear if the assumptions of classical statistics (e.g., stationarity) can be justified. The main reason for criticism of nonprobabilistic methods is that they are subjective and not necessarily coherent from a statistical point of view (see, e.g., the criticism of Mantovan and Todini (2006) with respect to GLUE). Moreover, on known problems for which the data do support the necessary probabilistic assumptions, probabilistic and nonprobabilistic methods provide different answers (e.g., Stedinger *et al.*, 2008).

The suitability of probabilistic versus nonprobabilistic methods and the difference in their response are dictated by the knowledge that the user has about the structure of the error model. Using a correctly based inference should lead to similar results in uncertainty assessment. Conversely, some authors claim that with unknown error structure it is dangerous to rely on statistical methods based on simple assumptions about the nature of the errors themselves.

There is an increasing consensus about the opportunity to use probabilistic approaches, as a way to efficiently summarize the information content of the data, when sufficient information is available to support statistical hypotheses with appropriate statistical tests (Montanari *et al.*, 2009). Conversely, data scarcity calls for expert knowledge to support uncertainty assessment. Above all, data scarcity calls for the integration of different types of information, within a framework that is unavoidably subjective, given that the information itself is often soft.

Besides the above, additional classifications were recently proposed for uncertainty assessment methods. For instance, Matott *et al.* (2009) identified seven categories of models: (1) Data analysis methods, including analytical and statistical procedures for evaluating the accuracy of data. These include also parametrization of probability distributions. (2) Identifiability analysis, aiming at detecting data inadequacy and suggesting model improvements. (3) Parameter estimation methods, quantifying uncertain model parameters. (4) Uncertainty analysis techniques, meaning methods to propagate sources of uncertainty through the model to generate probability distributions for the model output. These methods include approximation and sampling methods. (5) Sensitivity analyses, investigating to what extent different sources of variation in the input of a mathematical model affect the variation of the output. Sensitivity analysis aims at identifying what source of uncertainty weights more on the model output (see, e.g., Van Griensven *et al.* (2006); Götzinger and Bárdossy, 2008). Sensitivity analysis and uncertainty estimation are well distinguished. Their results can be comparable, because a probability distribution of model outputs corresponding to different inputs can be similar to the analogous distribution derived through the analysis of probabilistic uncertainty. This similarity of results has originated a confusion of terms in some applications. (6) Multimodel analysis, consisting of generating multiple possible outputs accordingly to different models, parameters, and boundary conditions. (7) Bayesian methods, which were previously defined (this category could be joined with category 4 above). The seven categories above are not strictly separated, meaning that a method can belong to more than one of them.

Another classification for uncertainty assessment methods for the model output was recently proposed by Shrestha and Solomatine (2008) who consider the following categories: (1) analytical methods, using derived distribution methods to compute the probability distribution function of the model output; (2) approximation methods, providing only the moments of the distribution of the uncertain output variable; (3) simulation and sampling–based methods, estimating the full distribution of the model output via simulation; (4) Bayesian methods, which combine Bayes' theorem and various simulation approaches to either estimate or update the probability distribution function of the parameters of the model and consequently estimate the uncertainty of the model output; (5) methods based on the analysis of the model errors, such as the meta-Gaussian approach described in Section 2.17.6.4; and (6) fuzzy-theory-based methods, providing a nonprobabilistic approach for modeling the kind of uncertainty associated with vagueness and imprecision.

Whatever approach is chosen to uncertainty assessment, the end-user should be made fully aware of the assumptions and drawbacks of the method that is being used. The presence of subjectivity should be clearly stated and the limitations of the underlying hypotheses, both in the probabilistic and nonprobabilistic approaches, clearly described and discussed. An appropriate terminology should also be used to make the meaning of the provided confidence bands clear. Whenever a subjective method is adopted, the user should be made aware that the uncertainty bands reflect user belief instead of providing a frequentist assessment of the probability of the true value to fall between them. Appropriate use of the methods being proposed by the scientific community, depending on the user needs and data availability, would allow us to successfully reach a better communication between scientists and end-users. It is as important to communicate uncertainty as communicate the assumptions on which an assessment has been based.

2.17.6 Assessment of the Global Uncertainty of the Model Output

Assessment of the global uncertainty for the model output is by far the application that is most frequently presented by the hydrological literature, as a means for quantifying model reliability and providing end-users with operational indications. Several methods are available to this end, ranging from statistically based to subjective approaches.

2.17.6.1 Analytical Methods

The most direct method to assess the uncertainty of a system output is to derive its statistics from a 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; and, 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). Examples of applications in hydrology are given by Melching (1992) and Vrugt and Bouten (2002). Point estimate methods are an interesting option too, in view of their computational efficiency (Tsai and Franceschini, 2005).

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 recently proposed by Huard and Mailhot (2006) in a hydrological context.

2.17.6.2 The Generalized Likelihood Uncertainty Estimation

GLUE was introduced by Beven and Binley (1992), who were inspired by the generalized sensitivity analysis methodology proposed by Spear and Hornberger (1980).

GLUE rejects the concept of an optimum model and parameter set and assumes that, prior to input of data into a model, all models and parameter sets have an equal likelihood of being acceptable. The acceptance of the existence of multiple likely models has been called equifinality (Beven, 1993) to suggest that this should be accepted as a generic problem in hydrological modeling rather than simply reflecting the problem of identifying the true model in the face of uncertainty.

GLUE is performed by first selecting different modeling options (different hydrological models and different parameters). In order to reduce the computational requirements of the procedure, it might be necessary to limit the dimension of the sample space of the parameters and models. Then, a high number N of simulation is generated by sampling the model and parameter spaces accordingly to a prior probability distribution. In the absence of prior knowledge, uniform sampling can be used. By increasing N one increases the probability of trying all of the most relevant solutions. The different models are then run for each of the parameter sets and the model output is then compared to a record of observed data (e.g., for observed hydrographs or annual maximum peak flows, see Cameron *et al.* (1999); another interesting example is given by Blazkova and Beven (2009)). The performance of each trial is assessed via likelihood measures, either formal or informal. This includes rejecting some parameter sets as nonbehavioral. For instance, the Nash and Sutcliffe (1970) efficiency can be used as informal likelihood measure of the simulation of a continuous hydrograph. All parameter sets that lead to obtaining an efficiency above a subjective threshold are retained. Finally, likelihood weighted uncertainty bounds are calculated depending on the likelihood (Freer *et al.*, 1996). For instance, the calculated likelihoods can be rescaled to produce a cumulative sum of 1.0, thereby obtaining informal weights. A cumulative distribution function of simulated discharges is then constructed using the rescaled weights. Linear interpolation is used to extract the discharge estimates corresponding to cumulative probabilities of $\alpha/2$, 0.5, and $1.0 - \alpha/2$. This allows $100(1 - \alpha)\%$ uncertainty bounds to be derived, in addition to a median simulation.

If either (or both) the likelihood measure or the procedure for computing the rescaled weights is informal, the probabilities computed with GLUE do not possess the classical frequentist meaning. Therefore, strictly speaking, it is inappropriate to refer to them with the term probability and many authors classify GLUE as a nonprobabilistic approach. Conversely, if formal statistical procedures are used, GLUE assumes the behavior of a probabilistic methodology. For extended discussions, the reader is referred to Beven *et al.* (2008) and Stedinger *et al.* (2008).

GLUE could be applied in principle even in the absence of observed historical data, in those real-world applications in which the likelihood measure is estimated on the basis of expert knowledge. GLUE is highly computationally demanding, especially if the number of significant model parameters is high. This problem may prevent the application of GLUE when dealing with complex models.

Beven (2006a) formally introduced a different procedure for the identification of behavioral models, by following previous practical experiments by Pappenberger and Beven (2004) and Page *et al.* (2007). A recent interesting application is presented by Liu *et al.* (2009). In this approach, limits of acceptability are preliminarily identified for the model output or selected performance measures. All the models that meet the limits of acceptability are retained so that an envelope of behavioral model simulations can be identified. Finally, a likelihood weighted cumulative density function for the model output can be computed as previously in GLUE so that simulation quantiles can be estimated (see also Blazkova and Beven, 2009).

There are many other variants of GLUE; for example, Tolson and Shoemaker (2008) and Mugunthan and Shoemaker (2006) combined optimization methods with a nonprobabilistic GLUE-like approach to increase computational efficiency of nonprobabilistic uncertainty analysis.

The hydrological literature presented many applications to GLUE to numerous hydrological problems, including rainfall–runoff modeling (Cameron *et al.*, 1999), groundwater modeling (Christensen, 2003), inundation modeling (Aronica *et al.*, 1998, 2002), and urban water-quality modeling (Freni *et al.*, 2008, 2009).

2.17.6.3 The Bayesian Forecasting System

The Bayesian Forecasting System (BFS) was proposed by Krzysztofowicz (1999, 2001a, 2002), Krzysztofowicz and Kelly (2000), and Krzysztofowicz and Herr (2001). The purpose is to produce a probabilistic river stage forecast (PRSF) based on a probabilistic quantitative precipitation forecasting (PQPF) as an input to a hydrological model that is in charge of simulating the response of a river basin to precipitation. It can be adapted to produce a probabilistic river discharge forecast.

The BFS assumes that the dominant source of uncertainty derives from the imperfect knowledge of the future precipitation, so that it can be assumed that all other sources of uncertainty play a minor role. The system can work with any hydrological model and aims at estimating the global uncertainty of the forecast, which is considered to be caused by: (1) precipitation uncertainty, which is dominant and quantified by the probability distribution of the future rainfall specified by the PQPF and (2) hydrologic uncertainty, which is the aggregate of all uncertainties arising from sources other than precipitation uncertainty. In particular, it aggregates the model uncertainty and parameter uncertainty.

The BFS has three structural components: the precipitation uncertainty processor (PUP; Kelly and Krzysztofowicz, 2000), the hydrologic uncertainty processor (HUP; Krzysztofowicz and Kelly, 2000), and the integrator (INT; Krzysztofowicz, 2001b). Figure 2 reports a sketch of the BFS structure adapted from Krzysztofowicz (2002). The PUP has the purpose of mapping precipitation uncertainty to output uncertainty under the hypothesis that there is no hydrologic uncertainty. This involves running the hydrological model for a set of specified quantiles of the probability distribution of the future rainfall. The HUP quantifies hydrologic uncertainty under the hypothesis that there is no precipitation uncertainty. Finally, the INT integrates the two uncertainties in order to produce a PRSF. For extended details on the PUP and INT, the interested reader is invited to refer to Kelly and Krzysztofowicz (2000) and Krzysztofowicz (2001b, 2002). Next, we provide a brief description of the HUP for the purpose of illustrating the meta-Gaussian approach adopted by BFS.

Let h_n denote the true river stage on day n, counting from day $n=0$ when the forecast is issued. At the forecast time the actual river stage on day n is unknown and thus uncertain.

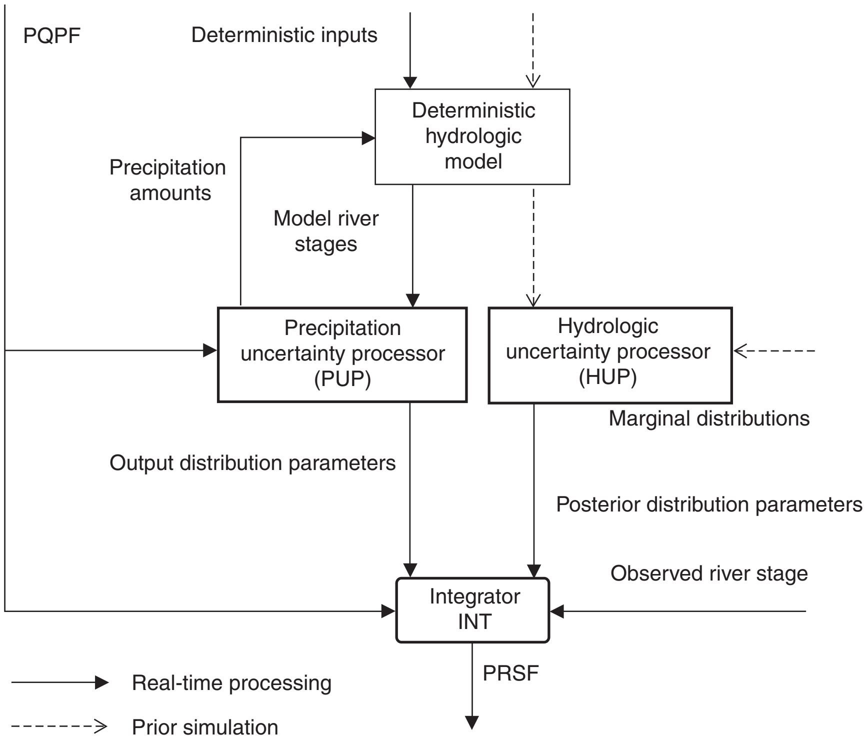

Figure 2 Sketch of the Bayesian forecasting system. PQPF and PRSF are probabilistic quantitative precipitation forecasting and probabilistic river stage forecasting, respectively. Adapted from Krzysztofowicz R (2002) Bayesian system for probabilistic river stage forecasting. *Journal of Hydrology* 268: 16–40.

Therefore, it is treated as a random variable which we refer to with the symbol H_n. Let s_n be an estimate of H_n from the hydrologic model based on all the input variables and the true precipitation amount. Estimate s_n is treated as a realization of the random variable S_n. One would observe $h_n = s_n$ if there were no hydrologic uncertainty (let us remark that HUP is developed under the assumptions that there is no precipitation uncertainty). The presence of hydrologic uncertainty gives rise to a probability distribution of the actual river stage H_n, conditional on a realization of the model river stage $S_n = s_n$. Therefore, we can treat H_n as a random variable whose probability distribution is conditioned on the corresponding realization of the model river stage $S_n = s_n$. The purpose of the HUP is to provide an estimate for such probability distribution. This is achieved by applying a Bayesian technique.

First of all, H_n and S_n are transformed to the Gaussian variables W_n and X_n respectively, by applying the standard normal quantile transform (NQT; see Kelly and Krzysztofowicz, 1997) and assuming that all conditional and joint densities are Gaussian. Then, it is assumed that the actual river stage process is well represented by a Markov stochastic process of order 1 and strictly stationary. This allows one to derive the prior probability distribution $g_n(w_n|w_0)$ of w_n conditional on $W_0 = w_0$. The prior density is derived under the assumption that the following normal linear equation applies in the transformed domain:

$$W_n = rW_{n-1} + \Theta_n \tag{1}$$

where r is a parameter and Θ_n is a random variable stochastically independent of W_{n-1} and normally distributed with mean zero and variance $\sigma^2(\Theta_n)$. Therefore, the probability distribution of w_n is Gaussian with mean equal to rw_{n-1} and variance $\sigma^2(\Theta_n)$. The recursive application of the above derivation allows one to estimate $g_n(w_n|w_0)$.

Subsequently, a probability distribution of the normalized model river stage x_n conditioned on w_n and w_0 is built, and is denoted as $f_n(x_n|w_n, w_0)$. This is derived under the hypothesis that the stochastic dependence between the transformed variables is governed by the following normal linear equation:

$$X_n = a_n W_n + d_n W_0 + b_n + \Phi_n \tag{2}$$

in which a_n, b_n, and d_n are parameters and Φ_n is a variate stochastically independent of (W_n, W_0) and normally distributed with mean zero and variance $\sigma^2(\Phi_n)$. It follows that the probability distribution $f_n(x_n|w_n, w_0)$ is Gaussian with mean and variance (see Krzysztofowicz and Kelly, 2000):

$$E(X_n|W_n = w_n, W_0 = w_0) = a_n w_n + d_n w_0 + b_n \tag{3}$$

$$\mathrm{Var}(X_n|W_n = w_n, W_0 = w_0) = {\sigma_n}^2(\Phi) \tag{4}$$

The coefficients r, a_n, b_n, and d_n are derived by running the hydrological model for an extended record of the model input with a perfect forecast of precipitation amount, thus obtaining joint realizations of the model-actual river stage process that are transformed to the Gaussian probability distribution through the NQT. These joint realizations are used to estimate the parameters of the above regressions (1) and (2). This design of the analysis assures that there is no precipitation uncertainty, but only hydrologic uncertainty.

Once $g_n(w_n|w_0)$ and $f_n(x_n|w_n, w_0)$ are known, the total probability law allows one to derive the distribution $k_n(x_n|w_0)$ of the transformed model river stage conditioned on w_0, while the Bayes theorem allows one to derive the posterior density of w_n conditioned on x_n and w_0, namely:

$$r(w_n|x_n, w_0) = \frac{f_n(x_n|w_n, w_0)g_n(w_n, w_0)}{k_n(x_n|w_0)} \tag{5}$$

where

$$k(x_n|w_0) = \int_{-\infty}^{\infty} f_n(x_n|w_n, w_0)g_n(w_n|w_0)\mathrm{d}w_n \tag{6}$$

Finally, the inverse of the NQT allows one to derive the posterior density of h_n conditioned on s_n and h_0. Such distribution allows one to quantify the hydrologic uncertainty. More details are provided by Krzysztofowicz and Kelly (2000).

Despite a theoretical development that may appear complicated, the BFS has the advantage of being easy to apply and allowing rapid implementation in real time. However, it was conceived for estimating the uncertainty of forecasted variables only.

2.17.6.4 Techniques Based on the Statistical Analysis of the Model Error

Several methods for uncertainty assessment were proposed based on the statistical analysis of the model error. Accordingly, the model error is treated as a stochastic process for which realizations are obtained by performing off-line simulations which are matched with the corresponding observations. Of course, observed data are themselves uncertain and therefore the model reliability analysis could not be correct in absolute terms (in the ideal situation of a perfect model, if we compared its response with uncertain output observations, that we assumed to be correct, we would 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 gives an important information for the sake of inferring reality based on the model output (Refsgaard *et al.*, 2006).

A technique for global uncertainty assessment based on the analysis of the model error is the meta-Gaussian approach proposed by Montanari and Brath (2004) for the case of hydrological simulations and extended by Montanari and Grossi (2008) for hydrological forecasting. Next, the latter methodology is presented, therefore making reference to real-time flood forecasting systems. The meta-Gaussian approach is probabilistic.

In order to estimate the uncertainty of a hydrological forecast, it is assumed that the forecast error is a stationary and ergodic stochastic process, denoted with the symbol $E(t)$. Its statistical properties are inferred by analyzing a past realization $e_{\mathrm{obs}}(t) = Q_{\mathrm{obs}}(t) - Q_{\mathrm{pred}}(t)$ that it is assumed to be available, where $Q_{\mathrm{obs}}(t)$ and $Q_{\mathrm{pred}}(t)$ are true and forecasted river flows, respectively. The use of a meta-Gaussian model is then proposed to derive the time-varying probability distribution of the forecast error. In particular, the probability distribution of $E(t)$ is inferred on the basis of its dependence on M selected explanatory random variables. The statistical inference is performed in the Gaussian domain, by preliminarily transforming $E(t)$ and the explanatory variables to the Gaussian probability distribution. The above transformation is operated through the NQT.

The probabilistic model for $E(t)$ is built as follows. First of all, it is assumed that positive and negative errors come from two different statistical populations $E^{(+)}(t)$ and $E^{(-)}(t)$. Therefore, the probability model for $E(t)$ is given by a mixture of two probability distributions, one for $E^{(+)}(t)$ and one for $E^{(-)}(t)$. The mixture is composed such that the area of the probability distribution of $E^{(+)}(t)$ is equal to the percentage, $P^{(+)}$, of positive errors over the total sample size of the available past realization $e_{\text{obs}}(t)$ of the forecast error.

The two realizations $e^{(+)}{}_{\text{obs}}(t)$ and $e^{(-)}{}_{\text{obs}}(t)$ are transformed through the NQT, therefore obtaining the normalized realizations $Ne^{(+)}{}_{\text{obs}}(t)$ and $Ne^{(-)}{}_{\text{obs}}(t)$. Then, M explanatory variables, $X^{(i)}(t)$ with $i=1,\ldots, M$ (which should be readily available at the forecast time), are selected in order to explain the variability in time of the marginal statistics of $E^{(+)}(t)$ and $E^{(-)}(t)$. The values of such explanatory variables for the realizations $e^{(+)}{}_{\text{obs}}(t)$ and $e^{(-)}{}_{\text{obs}}(t)$ above are estimated and then transformed by using the NQT, therefore obtaining the normalized explanatory variables $Nx^{(i)}{}_{\text{obs}}(t)$ with $i=1,\ldots, M$.

In the Gaussian domain, it is assumed that the forecast error can be expressed as a linear combination of the selected explanatory variables. Let us focus on the positive error. The linear combination can be expressed through the following relationship:

$$Ne^{(+)}(t_j) = C_1^{(+)}Nx^{(1)}(t_j) + C_2^{(+)}Nx^{(2)}(t_j) + \cdots + C_M^{(+)}Nx^{(M)}(t_j) + \varepsilon^{(+)}(t_j) \quad (7)$$

where $\varepsilon^{(+)}(t_j)$ is an outcome of a homoscedastic and Gaussian random variable and t_j is an assigned time step. An analogous relationship holds for $Ne^{(-)}(t)$. It is assumed that positive and negative errors are conditioned by the same explanatory variables, but the fit of the linear regression (7) leads to a different set of coefficient values. Such coefficients are estimated by inserting in (7) the past realizations of transformed forecast error, $Ne^{(+)}_{\text{obs}}(t)$, and explanatory variables, $Nx^{(i)}_{\text{obs}}(t)$, and then by identifying the coefficient values that lead to the best fit (for instance by minimizing the sum of the squares of $\varepsilon^{(+)}(t_j)$). The goodness of the fit provided by (7) can be verified by drawing a normal probability plot and a residual plot for $\varepsilon^{(+)}(t_j)$ as in Montanari and Brath (2004).

Once the linear regression (7) has been calibrated, for positive and negative errors, the probability distribution of the transformed positive forecast error can be easily derived for real-time and real-world applications. Such distribution is Gaussian and is expressed by the following relationship:

$$Ne^{(+)}(t_j) \sim G[\mu[Ne^{(+)}(t_j)], \sigma[Ne^{(+)}(t_j)]] \quad (8)$$

where '$\sim$' means equality in probability distribution and G indicates the Gaussian distribution whose parameters are given by

$$\mu[Ne^{(+)}(t_j)] = C_1^{(+)}Nx^{(1)}(t_j) + C_2^{(+)}Nx^{(2)}(t_j) + \cdots + C_M^{(+)}Nx^{(M)}(t_j) \quad (9)$$

$$\sigma[Ne^{(+)}(t_j)] = \sigma[\varepsilon^{(+)}(t_j)] \quad (10)$$

Analogous relationships (from (8) to (10)) hold for the negative error. Therefore, the confidence bands for the transformed forecast and an assigned significance level can be straightforwardly derived. In detail, the upper confidence band of the transformed forecast at the α significance level is given by the $1-\alpha/(2\cdot P^{(+)})$ quantile of the Gaussian distribution given by (8), (9), and (10). Given that $P^{(+)}$ can be arbitrarily close to 0, in the technical computation one may obtain values greater than 1 of $\alpha/(2\cdot P^{(+)})$. This means that the probability of getting a positive forecast error is small enough to make equal to 0 the width of the upper confidence band at the α significance level.

For instance, if $P^{(+)}=0.5$ and $\alpha=10\%$, the transformed upper confidence band is given by the well-known relationship:

$$Ne^{(+)}_{90‰}(t_j) = \mu[Ne^{(+)}(t_j)] + 1.96\sigma[Ne^{(+)}(t_j)] \quad (11)$$

Finally, by applying back the NQT one obtains the confidence bands for the assigned significance level in the untransformed domain.

The reason why positive and negative errors are treated separately is that a good fit is frequently not achieved through the linear regression (7) when the errors are pooled together. In fact, in this case, it appears that the NQT is not effective in making the errors homoscedastic and therefore the assumption of linearity does not hold. The reason for this result is that the NQT is not efficient in assuring homoscedasticity if the mean of the model error is not significantly changing across the range of the error itself, as it often happens when dealing with hydrological models. By treating positive and negative errors separately, the problem disappears and the assumptions of the linear regression are met. Finally, it is important to note that the only assumption made about the sign of the future forecast error is that it has a probability equal to $P^{(+)}$ to be positive. Therefore, no inference is made on the sign of the forecast error on the basis of the explanatory variables.

Figure 3 shows the confidence bands computed with the meta-Gaussian approach for the forecast with 1-h lead time of two flood events occurred on the Toce River at Candoglia, in Italy.

2.17.6.5 Bayesian Model Averaging

Bayesian model averaging (BMA, Hoeting *et al.*, 1999) is a statistical way of postprocessing model output ensembles to derive predictive probability density functions for hydrological variables. It represents the predictive probability distribution as a weighted average of the individual predictive probabilities of each model, where the weights are posterior probabilities of the models themselves and reflect the models' relative contributions to predictive skill over a training period. The combination of multiple models is an important component of

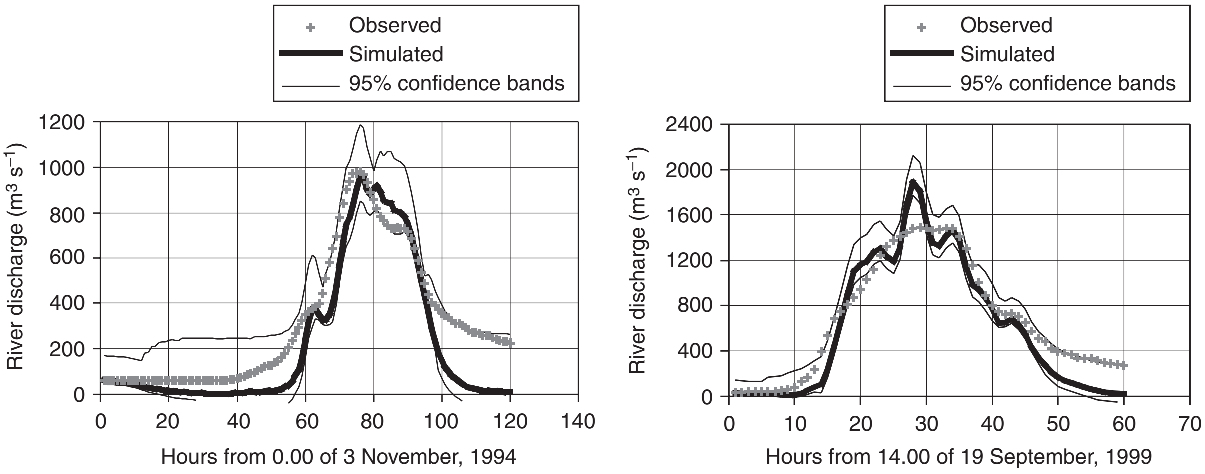

Figure 3 95% confidence bands computed with the meta-Gaussian approach for the forecast with 1-h lead time of two flood events occurred on the Toce River at Candoglia, on 3 November 1994 (left) and 19 September 1999 (right). From Montanari A and Grossi G (2008) Estimating the uncertainty of hydrological forecasts: A statistical approach. *Water Resources Research* 44: W00B08 (doi:10.1029/2008WR006897).

model validation (Burnham and Anderson, 2002). Multi-modeling solutions are often applied in real time forecasting (ensemble forecasting). See, for instance, the activity carried out in the framework of the HEPEX Project (Franz *et al.*, 2005; Zappa *et al.*, 2008).

BMA is applied in a Bayesian framework. Let $\mathbf{M} = \{M_i; i = 1,2,\ldots, N\}$ be a set of N hydrological models for obtaining the vector $\hat{\mathbf{z}}$ of hydrological variables. Given a set of data, $\mathbf{D}$, the posterior probability $\Pr(\hat{z}|\mathrm{D})$ of $\hat{z}$ is obtained through the BMA according to the law of total probability:

$$\Pr(\hat{z}|\mathrm{D}) = E_{\mathrm{M}}[\Pr(\hat{z}|M_i,\mathrm{D})] = \sum_{i=1}^{N} \Pr(\hat{z}|M_i,\mathrm{D})Pr(M_i|\mathrm{D}) \quad (12)$$

where $\Pr(\hat{z},\mathrm{D})$ is the posterior probability of $\hat{z}$ for the given data set $\mathbf{D}$, $\Pr(\hat{z}|M_i,\mathrm{D})$ is the posterior probability of $\hat{z}$ for given data set $\mathbf{D}$ and model M_i, $\Pr(M_i,\mathrm{D})$ is the posterior model probability for model M_i, and E_{M} is the expectation operator over simulation models. Essentially, Equation (12) says that the probability distribution given by the model ensemble for the output variable is a weighted mixture of the individual distributions given by each model, where the weights are the posterior model probabilities. Therefore, Equation (12) presupposes that individual probability distributions for the output from each model, conditioned on the model itself and the available data set, are available. According to Bayes' rule, the posterior model weight is

$$\Pr(M_i,\mathrm{D}) = \frac{\Pr(\mathrm{D}|M_i)\Pr(M_i)}{\sum_{i=1}^{N}\Pr(\mathrm{D}|M_i)\Pr(M_i)} \quad (13)$$

where $\Pr(\mathrm{D}|M_i)$ is the marginal model likelihood function for model M_i, $\Pr(M_i)$ is the prior model probability for model M_i, and $\sum_p \Pr(M_i) = 1$. A uniform distribution can be assumed for the priors if better information is not available. Equation (13) implies the total model weight $\sum_p Pr(M_i|\mathrm{D}) = 1$. The marginal model likelihood function $Pr(\mathrm{D}|M_i)$ plays an important role in the determination of the degree of importance for each model, given the same data set. For noninformative model priors, higher posterior model weights reflect better agreement between results and observed data.

According to Equation (12), the law of total expectation allows one to obtain the means of the predicted $\hat{z}$ over the models for given data D:

$$E(\hat{z}|\mathrm{D}) = E_M[E[\hat{z}|M^{(p)},D] = \sum_p E[\hat{z}|M^{(p)},\mathrm{D}]\Pr(M^{(p)}|\mathrm{D}) \quad (14)$$

where E is the expectation operation over $\hat{z}$. Analogous relationships allow one to obtain the covariance matrix of the predicted $\hat{z}$, therefore allowing a quantification of uncertainty. For more details, and an application that refers to the prediction of groundwater heads, the reader is referred to Li and Tsai (2009).

There are plenty of applications of BMA in hydrology (see, for instance, Ajami *et al.* (2006), Duan *et al.* (2007), Zhang *et al.* (2009), Reggiani *et al.* (2009), and Li and Tsai (2009)). Figure 4 shows confidence bands computed with BMA for simulations of river flows obtained with the soil and water assessment tool (SWAT) model in the Yellow River Headwater Basin (from Zhang *et al.*, 2009)

BMA tends to be computationally demanding and relies heavily on prior information about models. Neuman (2003) proposed a maximum likelihood version (MLBMA) of BMA to render it computationally feasible and to allow dealing with cases where reliable prior information is lacking (Ye *et al.*, 2004). BMA is also used within the Integrated Bayesian Uncertainty Estimator (IBUNE) proposed by Ajami *et al.* (2007).

2.17.6.6 Machine Learning Techniques

In the recent past, there has been an increased interest about machine learning technique for global uncertainty assessment (see, for instance, Shrestha *et al.*, 2009; Solomatine and Shrestha, 2009; Hall and Solomatine, 2008). These methods are frequently used as a mean to approximate complex models for uncertainty assessment, therefore obtaining a less computationally intensive approach.

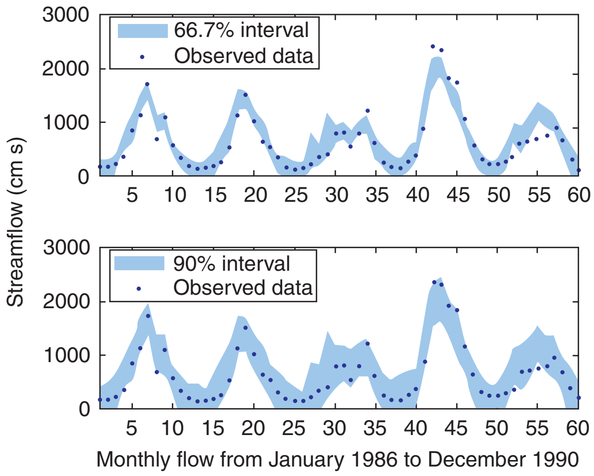

Figure 4 Confidence bands computed with BMA for simulations of river flows obtained with the SWAT model in the Yellow River Headwater Basin. From Zhang X, Srinivasan R, and Bosch D (2009) Calibration and uncertainty analysis of the SWAT model using genetic algorithms and Bayesian model averaging. *Journal of Hydrology* 374: 307–317.

Machine learning is concerned with the design and development of algorithms that allow computers to learn based on data, such as from sensor data or databases. A major focus of machine learning research is to automatically learn to recognize complex patterns and make intelligent decisions based on data. Machine learning techniques include, among others, approaches that have been widely used in hydrology, such as neural networks, nearest-neighbor methods, and statistical methods.

2.17.7 Assessment of Data Uncertainty

Among the sources of uncertainty in hydrology, data uncertainty is often believed to play a marginal role. Accordingly, only a few attempts have been made to quantify the effects of uncertainty in observations on hydrological modeling (see, for instance, Clarke, 1999).

Different types of observations are currently used in hydrological modeling. The most common applications usually refer to precipitation as input and river flows as output data, although very often solar radiation, temperature, wind speed, soil moisture, groundwater levels, geomorphological features, land use, and others are also employed. Some of the above variables are affected by a limited uncertainty with respect to the others. In particular, uncertainty in precipitation and river flow is often considered to be dominant, because of the spatial variability of rainfall and snowfall on the one hand, and the errors in the determination of the rating curve on the other.

The presence of uncertainty in input and output data induces two types of problems to hydrologists: the first is related to its estimation (to what extent the observed data are uncertain?), whereas the second is connected to accounting for such uncertainty in hydrological modeling.

2.17.7.1 Precipitation Uncertainty

Hydrologists are well aware that a multitude of problems and research issues are related to precipitation uncertainty, which are connected to precipitation monitoring and prediction (Chua and Bras, 1982; Gottschalk and Jutman, 1982).

Precipitation monitoring is carried out through direct measurements (raingauges and snowgauges) or remote sensing (satellite, radar, and microsensors). The uncertainty of gauge measurements is typically limited and therefore the estimation error of the precipitation field is mainly induced by spatial variability. When remote-sensed data are used, spatial variability is generally better estimated but the uncertainty in point measurements is relevant. There is a large body of literature about uncertainty assessment for precipitation, starting from the pioneeristic work of Thiessen (1911).

Uncertainty of gauge measurements of precipitation has been the subject of numerous case studies (see, for instance, Morrissey *et al.*, 1995; Brath *et al.*, 2004). These studies proved that the estimation error of mean areal precipitation significantly depends on the climatic conditions, the spatial structure of the precipitation itself, the morphology of the catchment, and the gauging network.

The task of quantifying remotely sensed precipitation uncertainty has proved to be difficult. A fundamental problem is the lack of a term of comparison (Habib and Krajewski, 2002). Numerous studies compared remotely sensed and gauged data and showed significant disagreement. For instance, Austin (1987) found that for individual storms, radar and raingauge measurements can differ of a factor of 2 or more. In a more recent investigation, Brandes *et al.* (1999) found that radar-to-gauge ratios of storm totals were in the range of about 0.7–1.9. These differences become even more significant when satellite versus raingauge comparisons are carried out.

In general, estimating precipitation uncertainty is a difficult task and no general rule exists. Integrating different monitoring techniques is certainly a potentially valuable solution.

Turning to the purpose of accounting for precipitation uncertainty in hydrological modeling, different methods were proposed by the literature. The BFS described in Section 2.17.6.3 is a relevant example. GLUE can be applied as well, by introducing an input error model and then generating many different realizations of the input data themselves, with which one can derive a likelihood weighted model output.

Kavetski *et al.* (2002, 2006a, 2006b) introduced the Bayesian total error analysis (BATEA), that is, a method for explicitly accounting for sampling and measurement uncertainty in both input and output data. In view of the inability to build a formal and sufficiently representative input error model in many real-world applications, BATEA is based on the use of vague error models, with the awareness that such an approach can cause a degeneration of the reliability of the inference equations.

The basic working hypothesis of BATEA consists of assuming that the input uncertainty is multiplicative Gaussian and independent of each storm, even though its general framework allows alternative uncertainty models. The multiplier approach assumes that the storm depth is the only quantity in error, whereas the rainfall pattern is correct up to a multiplicative constant m_i, such that $d_i' = m_i d_i$, where d_i' and d_i are the observed and true precipitation depths for the ith storm. Accordingly, the parameter vector of the hydrologic model is extended to include the parameters of the uncertainty models. Parameter inference is then carried out within a Bayesian framework, which requires indentifying a prior distribution of the model parameters that is subsuquently updated by using Bayes' theorem in view of the available observations.

The above treatment of input uncertainty implies that the dimensionality of the parameter vector is increased with latent variables, whose number depends on the sample size of the observed data (and the type of error model assumed). Moreover, if both the input and the output data are observed with large uncertainty, the utility of any parameter estimation methodology becomes questionable. Finally, one might be concerned that rainfall multipliers can possibly interact with other sources of error and therefore separation of errors in BATEA is conditional on other sources of uncertainty. For instance, classic underprediction by a hydrological model after a long dry period can be compensated by increasing the rainfall multiplier. A similar treatment for precipitation uncertainty is used in IBUNE (Ajami *et al.*, 2007).

2.17.7.2 River Discharge Uncertainty

Pelletier (1987) reviewed 140 publications dealing with uncertainty in the determination of the river discharge, thereby providing an extensive summary. He referred to the case where the river discharge at a given cross section is measured by using the velocity–area method, that is,

$$Q_{\mathrm{obs}}(t) = A(t) \cdot v(t) \tag{15}$$

where t is the sampling time, $Q_{\mathrm{obs}}(t)$ the measured river discharge, $A(t)$ the cross-sectional area of the river, and $v(t)$ the velocity of the river flow averaged over the cross section. Errors in $Q_{\mathrm{obs}}(t)$ are originated by uncertainties in both $A(t)$ and $v(t)$, which in turn are originated by uncertainty in the current meter, variability of the river flow velocity over the cross section, and uncertainty in the estimation of the cross-section geometry. Pelletier (1987) highlighted that the overall uncertainty in a single determination of river discharge, at the 95% confidence level, can vary in the range (8–20%), mainly depending on the exposure time of the current meter, the number of sampling points where the velocity is measured, and the value of $v(t)$. Another interesting contribution was provided by the European ISO EN Rule 748 (1997) that quantified the expected errors in the determination of the river discharge with the velocity–area method. The conclusions were similar to those of Pelletier (1987).

In some cases, including the usual practice in many countries of Europe, river discharge values are estimated by using the rating curve method, which is very easy to apply. According to the rating curve method, observations of river stage are converted into river discharge by means of a rating curve, which is preliminarily estimated by using observations collected using the velocity–area method. Hence, an additional error is induced by imperfect estimation of the rating curve.

Di Baldassarre and Montanari (2009) proposed a model for estimating the error affecting river flow observations derived by the rating curve method. The model aims at taking into account the main sources of uncertainty within a simplified approach. The most important assumptions underlying the model are as follows. (1) The uncertainty induced by imperfect observation of the river stage is negligible. This is consistent with the fact that these errors are usually very limited (around 1–2 cm; e.g., Schmidt, 2002; Pappenberger *et al.*, 2006) and therefore of the same order of magnitude as standard topographic errors. (2) The geometry of the river is assumed to be invariant, which means that the rating curve changes in time only because of seasonal variation of roughness (see below). This assumption has been made because the uncertainty induced by possible variations of the river geometry is heavily dependent on the considered case study and no general rule can be suggested. However, it is worth noting that, using this assumption, the study neglects one of the most relevant sources of uncertainty that may affect river discharge observations where relevant sediment transport and erosion processes are present. (3) Uncertainty in the rating curve derives from the following causes: errors in the river discharge measurements that are used to calibrate the rating curve itself; interpolation and extrapolation error of the rating curve; unsteady flow conditions; and seasonal changes of roughness.

The uncertainty affecting the river discharge measurements was estimated by Di Baldassarre and Montanari (2009) according to the guidelines reported by the European ISO EN Rule 748 (1997), which lead to an estimate of about 5–6% when the measurements are collected in ideal conditions. This outcome matches the indications reported in Leonard *et al.* (2000) and Schmidt (2002).

The remaining sources of uncertainty were evaluated by Di Baldassarre and Montanari (2009) by developing a numerical simulation study for a 330-km reach of the Po River (Italy). The study focused on 17 cross sections and found that the estimation of river discharge using the rating curve method is affected by an increasing error for increasing river discharge values. At the 95% confidence level, the error ranges from 6.2% to 42.8% of the observation, with an average value of 25.6%. Furthermore, the uncertainty induced by the extrapolation of the rating curve is dominating the other errors in high flow conditions. In fact, previous contributions in hydrology (e.g., Rantz *et al.*, 1982) do not recommend extrapolating rating curves beyond a certain range. Nevertheless, several hydrological applications are unavoidably based on flood flow observations (e.g., calibration and validation of rainfall–runoff models, flood frequency analysis, and boundary conditions of flood inundation models) and therefore one needs to extrapolate the rating curve beyond the measurement range (Pappenberger *et al.*, 2006).

The above analysis proved that river flow uncertainty can indeed be very significant and therefore should be accounted for in practical applications. An interesting opportunity is offered by the application of GLUE according to the limits of acceptability concept (Blazkova and Beven, 2009). Once the uncertainty in the river flows is estimated, it is possible to fix limits of acceptability for the observed river flows, and the models that do not respect them can be rejected as nonbehavioral. The collection of the behavioral outputs allows the user to obtain an envelope of likely model simulations. The above approach is nonprobabilistic.

2.17.8 Assessment of Parameter Uncertainty

Calibration in hydrology is increasingly done automatically, while manual calibration (through trial and error procedures) is used only when dealing with complex models requiring high computational costs. Parameter calibration techniques lead to either a single solution or multiple solutions (i.e., parameter sets). The approaches leading to a single solution are basically optimization problems, while techniques leading to multiple likely solutions can serve as tools for uncertainty assessment. Many search algorithms have been successfully devised and applied to automatically find the optimal parameter set for hydrological models, which can be subdivided into local, global, and hybrid search techniques (Duan *et al.*, 1992; Mugunthan and Shoemaker, 2006; Tolson and Shoemaker, 2008; Thyer *et al.*, 2009; Tonkin and Doherty, 2009).

Approaches for multiple-solution parameter estimation can be broadly divided into two categories: importance sampling and Markov chain Monte Carlo (MCMC) sampling (Kuczera and Parent, 1998). With this approach full parameter distributions rather than simple point estimates can be obtained. Methods based on importance sampling aim to identify a set of behavioral model parameter configurations according to a selected objective function. Then, parameter distributions are estimated using a weighted combination of the behavioral parameter sets. GLUE is perhaps the most-used method based on importance sampling.

MCMC parameter estimation incorporates importance sampling into a procedure for evaluating conditional probability distributions. Prior parameter distributions are selected (for instance, by assigning a uniform distribution or a distribution derived through expert knowledge) and the sampler evolves them into posterior distributions that are estimated by using the observed data.

Thus, multiple-solution approaches can be used to assess parameter uncertainty. A relevant example within this respect is the SCEM-UA algorithm by Vrugt *et al.* (2003a). Once the uncertainty in the parameters is known, simulation approaches can be applied to estimate the related uncertainty induced in the model output. An example is given by Thorsen *et al.* (2001) who assessed the uncertainty in simulations of nitrate leaching induced by using model parameters obtained from databases at the European level.

End-users frequently experience the case where multiple or competing objectives need to be optimized. According to this need, numerous multiobjective optimization algorithms have been devised, with numerous developments in the recent past (see, for instance, Zhang *et al.*, 2008). Relevant examples are the MOSCEM-UA and AMALGAM methods (Vrugt *et al.*, 2003b; Vrugt and Robinson, 2007). These two methods are briefly described in the following.

2.17.8.1 The MOSCEM-UA Method

Multiobjective calibration problems can be dealt with by defining more than one optimization criteria (objective functions) that correspond to different performance measures of the selected model. Then, a multicriteria optimization method can be used to identify the set of nondominated, efficient, or Pareto optimal solutions (Gupta *et al.*, 1998). The Pareto solutions represent tradeoffs among the different performance measures that are often conflicting. As such, moving from one solution to another results in the improvement of one objective and deterioration in one or more others.

A simple way to deal with multiobjective calibration is to weigh the different criteria into a single objective function and to run a large number of independent single-criteria optimization runs using different values for the weights (Madsen, 2000). This method is simple to implement, but has the drawback that a complete single-objective optimization is to be solved to obtain each discrete Pareto solution. Moreover, maintaining the independence of the various criteria will allow the user to analyze the tradeoffs among the different criteria, therefore enabling an improved understanding of the limitations of the model structure.

MOSCEM-UA (Vrugt *et al.*, 2003b) is an effective and efficient MCMC sampler, which is capable of generating a fairly uniform approximation of the Pareto frontier within a single optimization run. The algorithm is closely related to the SCEM-UA algorithm (Vrugt *et al.*, 2003a). In addition, MOSCEM-UA uses a newly developed, improved concept of Pareto dominance, thereby also containing the single-criteria solutions at the extremes of the Pareto solution set. For more details, the interested reader is invited to refer to Vrugt *et al.* (2003b).

The ensemble of the models lying on the Pareto frontier allows the user to identify an envelope of model outputs corresponding to the nondominated solutions.

2.17.8.2 The AMALGAM Method

AMALGAM (Vrugt and Robinson, 2007) is a follow-up of MOSCEM-UA and is specifically designed to take full advantage of the power of distributed computer networks. AMALGAM runs multiple different search strategies simultaneously for population evolution and adaptively updates the weights of these individual methods based on their reproductive success. This ensures a fast, reliable, and computationally efficient solution to multiobjective optimization problems.

2.17.9 Assessment of Model Structural Uncertainty

Model structural uncertainty is induced by inadequateness of the hydrological model to represent the hydrological system. This situation is also characterized by reduced model identifiability, because the imperfectness of the modeling solutions makes many of them potentially suboptimal, regardless of the different values of the selected performance measure. In the presence of model, structural uncertainty a performance measure becomes less effective and therefore the highest of its value does not necessarily identify the best model. For instance, a performance measure that lays emphasis on floods may be biased toward a (imperfect) model that could not be as reliable in reproducing the low flows. This is the reason why multiobjective calibration is frequently applied in hydrology.

A statistical and rigorous evaluation of model structural uncertainty is not possible in practical hydrological applications, at least because it should necessarily be performed with perfect data. The literature proposed approximate techniques for estimating the uncertainty in the model output induced by model structural uncertainty. The most popular of them are based on multimodel analysis. In fact, the variability of the response provided by different models, if other uncertainty sources are negligible, provides indications on the uncertainty induced by a wrong model structure.

Multimodel analysis is based on the use of many different plausible models that may consider, for instance, alternative processes and alternative simplified approximations. An example of model combination is the BMA presented in Section 2.17.6.5.

Another quantitative approach for performing multimodel application was presented by Burnham and Anderson (2002) and Ye *et al.* (2008). It is implemented by assigning performance scores and importance weights to each candidate model with which the ensemble of model outputs can be constructed basing on the importance of each model.

Multimodel applications can be performed also by applying GLUE, with which different models can be considered and evaluated according to a single likelihood measure or one or more likelihood measures. An example of application of GLUE with different modeling solutions is provided by Rojas *et al.* (2009).

From a practical point of view, the above techniques are often applied for assessing the global uncertainty in the model output instead of model structural uncertainty only, because it is impossible to carry out such techniques in the absence of data and parameter uncertainty. As such, the combination of different models with uncertain parameters and uncertain data bases does not allow one to separate the above sources of uncertainty, unless one makes heavy assumptions (see, e.g., the IBUNE method; Ajami *et al.* (2007); see also Clark *et al.* (2008)).

2.17.10 Uncertainty Assessment as a Learning Process

Uncertainty assessment is an effective mean to quantitatively assess model reliability and therefore perform model diagnostic and evaluation. These latter, in turn, provide indications about the model ability to simulate hydrology at a given place and therefore about the correctness of our understanding of the hydrological processes at that place. Thus, uncertainty assessment plays a fundamental role in the learning process.

In the past, the learning process was mainly linked to parameter estimation for a given model. The optimal parameter values, actually, provide information about the conditions of the system. Treating modeling more explicitly as a learning process allows one to follow a new approach to this problem based on a methodology that will link models, databases, and parameters with the areas of interest, thereby providing information on the dominant hydrological processes (see Beven (2007); applications are presented in Montanari *et al.* (2006), Fenicia *et al.* (2008), and Schoups *et al.* (2008)). This is part of the downward modeling approach that recently gained increased attention within the context of PUB.

One of the most exciting future perspectives is the possibility to implement many different models as a process of learning about specific places (Beven, 2007). The representation will be uncertain so that this learning process should be implemented within a framework of uncertainty estimation. Indeed, uncertainty estimation, providing quantitative information about model reliability, if coupled with a multimodel approach, could provide indications about the dominant hydrological processes and their dynamics. Within this framework it is also necessary to set up a mechanism for model rejection (e.g., the model providing the simulations presented in Figure 1 could be rejected because it is too biased). There is the potential problem that model rejection is not by default embedded in uncertainty assessment methods. In particular, it is not embedded in statistical approaches, which in many cases do not assess the motivation for the presence of uncertainty. Model rejection is often based on expert knowledge, which is subjective but indeed necessary in the context of a learning process (see, for instance, Merz and Blöschl (2008a, 2008b)). This implies that the use of statistical methods for uncertainty assessment in a learning process should be based on including in the statistical representation the available information about the underlying physical process (for an extended discussion, see Koutsoyiannis (2009)).

2.17.11 Conclusions

Uncertainty assessment in hydrology is a relevant practical problem and still a research challenge. The limited extension of hydrological databases and the complexity of hydrological processes, whose dynamics and domains are to a great extent nonobservable, make the interpretation of the results of hydrological modeling studies not easy. The intense research activity recently done on uncertainty resulted in the development of many new techniques for uncertainty assessment, which differ in behavior and scope. It is essential to formally define a terminology and make clear the prerogatives of each method in order to make clear to end-users the meaning of uncertainty in hydrology and convey them a useful information.

In order to provide a contribution to this end, we provide in Table 1 a brief summary of the most-used uncertainty assessment methods, including those presented here, by also providing an attempt of classification and by specifying their purpose.

Uncertainty assessment in hydrology will represent a research challenge for a long time to come. Uncertainty is an inherent property of hydrological processes which in principle will not prevent gaining a much better understanding of how water flows downstream. Uncertainty in hydrology should not be viewed as a limitation to be eliminated but rather as a intrinsic feature that needs to be properly and objectively quantified, whenever possible, with scientific method, that is, through the collection of data by means of observation and experimentation, and the formulation and testing of hypotheses.

Communicating uncertainty to end-users should not undermine their confidence in models (Beven, 2006b; Pappenberger and Beven, 2006; Faulkner *et al.*, 2007), but rather increase it through an improved perception of the underlying natural processes and an increased awareness of model reliability. Uncertainty does not mean lack of knowledge or lack of modeling capability but that the predicted value of a hydrological variable is uncertain. A proper estimation of uncertainty is the way forward to a reliable hydrological design and therefore a proper management of the environment and water resources.

Acknowledgments

The author is grateful to Demetris Koutsoyiannis, Jasper Vrugt, Keith Beven, Simone Castiglioni, an anonymous referee and the Editor Stefan Uhlenbrook for providing very useful comments on the text. The support of the Italian Government, through the National Research Project "Uncertainty estimation for precipitation and river discharge data. Effects on water resources planning and flood risk management" is also acknowledged.

References

Ajami NK, Duan Q, Gao X, and Sorooshian S (2006) Multi-model combination techniques for hydrological forecasting: Application to distributed model intercomparison project results. *Journal of Hydrometeorology* 8: 755–768.

Ajami NK, Duan Q, and Sorooshian S (2007) An integrated hydrologic Bayesian multimodel combination framework: Confronting input, parameter, and model structural uncertainty in hydrologic prediction. *Water Resources Research* 43: W01403 (doi:10.1029/2005WR004745).

Alley WM (1984) On the treatment of evapotranspiration, soil moisture accounting and aquifer recharge in monthly water balance models. *Water Resources Research* 20: 1137–1149.

Aronica G, Bates PD, and Horritt MS (2002) Assessing the uncertainty in distributed model predictions using observed binary pattern information within GLUE. *Hydrological Processes* 16: 2001–2016.

Aronica G, Hankin B, and Beven KJ (1998) Uncertainty and equifinality in calibrating distributed roughness coefficients in a flood propagation model with limited data. *Advances in Water Resources* 22: 349–365.

Austin PM (1987) Relation between measured radar reflectivity and surface rainfall. *Monthly Weather Review* 115: 1053–1070.

Beven KJ (1993) Prophesy, reality and uncertainty in distributed hydrological modeling. *Advances in Water Resources* 16: 41–51.

Beven KJ (2006a) A manifesto for the equifinality thesis. *Journal of Hydrology* 320: 18–36.

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

Beven KJ (2007) Towards integrated environmental models of everywhere: Uncertainty, data and modeling as a learning process. *Hydrology and Earth System Sciences* 11: 460–467.

Beven KJ (2009) *Environmental Modelling: An Uncertain Future?: An Introduction to Techniques for Uncertainty Estimation in Environmental Prediction*. London: Taylor and Francis.

Beven KJ and Binley AM (1992) The future of distributed models: Model calibration and uncertainty prediction. *Hydrological Processes* 6: 279–298.

Beven KJ, Smith PJ, and Freer JE (2007) Comment on "Hydrological forecasting uncertainty assessment: Incoherence of the GLUE methodology" by Pietro Mantovan and Ezio Todini. *Journal of Hydrology* 338: 315–318.

Beven KJ, Smith PJ, and Freer JE (2008) So just why would a modeller choose to be incoherent? *Journal of Hydrology* 354: 15–32.

Blazkova S and Beven KJ (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).

Brandes EA, Vivekanandan J, and Wilson JW (1999) A comparison of radar reflectivity estimates of rainfall from collocated radars. *Journal of Atmospheric Oceanic Technology* 16: 1264–1272.

Brath A, Montanari A, and Toth E (2004) Analysis of the effects of different scenarios of historical data availability on the calibration of a spatially-distributed hydrological model. *Journal of Hydrology* 291: 272–288.

Burnham KP and Anderson DR (2002) *Model Selection and Multimodel Inference*, 2nd edn. New York: Springer.

Cameron DS, Beven KJ, Tawn J, Blazkova S, and Naden P (1999) Flood frequency estimation by continuous simulation for a gauged upland catchment (with uncertainty). *Journal of Hydrology* 219: 169–187.

Christensen S (2003) A synthetic groundwater modelling study of the accuracy of GLUE uncertainty intervals. *Nordic Hydrology* 35: 45–59.

Chua S-H and Bras RL (1982) Optimal estimations of mean areal precipitation in regions of orographic influence. *Journal of Hydrology* 57: 23–48.

Clark MP, Slater AG, Rupp DE, *et al.* (2008) Framework for understanding structural errors (FUSE): A modular framework to diagnose differences between hydrological models. *Water Resources Research* 44: W00B02 (doi:10.1029/2007WR006735).

Clarke RT (1999) Uncertainty in the estimation of mean annual flood due to rating curve indefinition. *Journal of Hydrology* 222: 185–190.

Cullen AC and Frey HC (1999) *Probabilistic Techinques in Exposure Assessment: A Handbook for Dealing with Variability and Uncertainty in Model and Inputs*. New York: Plenum.

Di Baldassarre G and Montanari A (2009) Uncertainty in river discharge observations: A quantitative analysis. *Hydrology and Earth System Sciences* 13: 913–921.

Dottori F, Martina MLV, and Todini E (2009) A dynamic rating curve approach to indirect discharge measurement. *Hydrology and Earth System Sciences* 13: 847–863.

Duan Q, Ajami NK, Gao X, and Sorooshian S (2007) Multi-model ensemble hydrologic prediction using Bayesian model averaging. *Advances in Water Resources* 30: 1371–1386.

Duan Q, Sorooshian S, and Gupta VK (1992) Effective and efficient gobal optimization for conceptual rainfall-runoff models. *Water Resources Research* 28: 1015–1031.

European ISO EN Rule 748 (1997) Measurement of liquid flow in open channels – velocity-area methods, Reference number ISO 748:1997 (E), International Standard.

Faulkner H, Parker D, Green C, and Beven KJ (2007) Developing a translational discourse to communicate uncertainty in flood risk between science and the practitioner. *AMBIO* 36: 692–704.

Fenicia F, Savenije HHG, Matgen P, and Pfister L (2008) Understanding catchment behavior through stepwise model concept improvement. *Water Resources Research* 44: W01402 (doi:10.1029/2006WR005563).

Franz K, Ajami N, Schaake J, and Buizza R (2005) Hydrologic ensemble prediction experiment focuses on reliable forecasts. *Eos* 86: 239.

Freer J, Beven KJ, and Ambroise B (1996) Bayesian uncertainty in runoff prediction and the value of data: An application of the GLUE approach. *Water Resources Research* 32: 2163–2173.

Freni G, Mannina G, and Viviani G (2008) Uncertainty in urban stormwater quality modelling: The effect of acceptability threshold in the GLUE methodology. *Water Research* 42: 2061–2072.

Freni G, Mannina G, and Viviani G (2009) Uncertainty assessment of an integrated urban drainage model. *Journal of Hydrology* 373: 392–404.

Gan TY and Burges SJ (1990) An assessment of a conceptual rainfall-runoff model's ability to represent the dynamics of small hypothetical catchments 2. Hydrologic responses for normal and extreme rainfall. *Water Resources Research* 26: 1605–1619.

Gottschalk L and Jutman T (1982) Calculation of areal means of meteorologic variables for watersheds. *7th Nordic Hydrological Conference. NHP-Report No. 2*, pp. 720–736.

Götzinger J and Bárdossy A (2008) Generic error model for calibration and uncertainty estimation of hydrological models. *Water Resources Research* 44: W00B07 (doi:10.1029/2007WR006691).

Gupta HV, Sorooshian S, and Yapo PO (1998) Toward improved calibration of hydrologic models: Multiple and noncommensurable measures of information. *Water Resources Research* 34: 751–763.

Habib E and Krajewski WF (2002) Uncertainty analysis of the TRMM ground-validation radar-rainfall products: Application to the TEFLUN-B field campaign. *Journal of Applied Meteorology* 41: 558–572.

Hahn GJ and Meeker WQ (1991) *Statistical Intervals; A Guide for Practitioners*. New York: Wiley.

Hall J and Solomatine DP (2008) A framework for uncertainty analysis in flood risk management decisions. *International Journal of River Basin Management* 6: 85–98.

Hoeting JA, Madigan D, Raftery AE, and Volinsky CT (1999) Bayesian model averaging: A tutorial. *Statistical Science* 14: 382–401 (doi:10.1214/ss/1009212519).

Hoffman FO and Hammonds JS (1994) Propagation of uncertainty in risk assessments: The need to distinguish between uncertainty due to lack of knowledge and uncertainty due to variability. *Risk Analysis* 14: 707–712.

Huard D and 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).

Ibbitt RP and O'Donnell T (1974) Designing conceptual catchment models for automatic fitting methods. In: *Mathematical Models in Hydrology Symposium*, IAHS-AISH, Publication No. 2, pp. 461–475.

Jacquin AP and Shamseldin AY (2007) Development of a possibilistic method for the evaluation of predictive uncertainty in rainfall-runoff modeling. *Water Resources Research* 43: W04425. doi:10.1029/2006WR005072.

Kavetski D, Franks SW and Kuczera G (2002) Confronting input uncertainty in environmental modelling. In: *Calibration of Watershed Models*. Duan Q, Sorooshian S, Gupta H, Rosseau H, and Turcotte R (eds.). Water Science and Application, vol. 6, pp. 49–68. Washington, DC: AGU.

Kavetski D, Kuczera G, and Franks SW (2006a) Bayesian analysis of input uncertainty in hydrological modeling: 1. Theory. *Water Resources Research* 42: W03407 (doi:10.1029/2005WR004368).

Kavetski D, Kuczera G, and Franks SW (2006b) Bayesian analysis of input uncertainty in hydrological modeling: 2. Application. *Water Resources Research* 42: W03408 (doi:10.1029/2005WR004376).

Kelly KS and Krzysztofowicz R (1997) A bivariate meta-Gaussian density for use in hydrology. *Stochastic Hydrology and Hydraulics* 11: 17–31.

Kelly KS and Krzysztofowicz R (2000) Precipitation uncertainty processor for probabilistic river stage forecasting. *Water Resources Research* 36: 2643–2653.

Kleisenn FM, Beck MB, and Weather HS (1990) The identifiability of conceptual hydrochemical models. *Water Resources Research* 26: 2979–2992.

Koussis A (2009) Comment on "A dynamic rating curve approach to indirect discharge measurement" by Dottori *et al.* (2009). *Hydrology and Earth System Sciences Discussions* 6: 7429–7437.

Koutsoyiannis D (2009) A random walk on water. *Hydrology and Earth System Sciences Discussions* 6: 6611–6658.

Koutsoyiannis D, Makropoulos C, Langousis A, *et al.* (2009) HESS opinions: "Climate, hydrology, energy, water: Recognizing uncertainty and seeking sustainability". *Hydrology and Earth System Sciences* 13: 247–257.

Krzysztofowicz R (1999) Bayesian theory of probabilistic forecasting via deterministic hydrologic model. *Water Resources Research* 35: 2739–2750.

Krzysztofowicz R (2001a) The case for probabilistic forecasting in hydrology. *Journal of Hydrology* 249: 2–9.

Krzysztofowicz R (2001b) Integrator of uncertainties for probabilistic river stage forecasting: Precipitation-dependent model. *Journal of Hydrology* 249: 69–85.

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

Krzysztofowicz R and Herr HD (2001) Hydrologic uncertainty processor for probabilistic river stage forecasting: Precipitation-dependent model. *Journal of Hydrology* 249: 46–68.

Krzysztofowicz R and Kelly KS (2000) Hydrologic uncertainty processor for probabilistic river stage forecasting. *Water Resources Research* 36: 3265–3277.

Kuczera G and Mroczkowski M (1998) Assessment of hydrologic parameter uncertainty and the worth of multiresponse data. *Water Resources Research* 34: 1481–1489.

Kuczera G and Parent E (1998) Monte Carlo assessment of parameter uncertainty in conceptual catchment models: The Metropolis algorithm. *Journal of Hydrology* 211: 69–85.

Kuichling E (1889) The relation between rainfall and the discharge of sewers in populous districts. *Transactions ASCE* 20: 1–60.

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

Leonard J, Mietton M, Najib H, and Gourbesville P (2000) Rating curve modelling with Manning's equation to manage instability and improve extrapolation. *Hydrological Sciences Journal* 45: 739–750.

Li X and Tsai FT-C (2009) Bayesian model averaging for groundwater head prediction and uncertainty analysis using multimodel and multimethod. *Water Resources Research* 45: W09403 (doi:10.1029/2008WR007488).

Liu Y, Freer J, Beven KJ, and Matgen P (2009) Towards a limits of acceptability approach to the calibration of hydrological models: Extending observation error. *Journal of Hydrology* 367: 93–103.

Liu Y and Gupta HV (2007) Uncertainty in hydrologic modeling: Toward an integrated data assimilation framework. *Water Resources Research* 43: W07401 (doi:10.1029/2006WR005756).

Loucks DP and Van Beek E (2005) Water resources systems planning and management an introduction to methods, models and applications. In: *Studies and Reports in Hydrology*, United Nations Educational, Scientific and Cultural Organization (ISBN 92-3-103998-9). Paris, France: UNESCO.

Madsen H (2000) Automatic calibration of a conceptual rainfall-runoff model using multiple objectives. *Journal of Hydrology* 235: 276–288.

Mantovan P and Todini E (2006) Hydrological forecasting uncertainty assessment: Incoherence of the GLUE methodology. *Journal of Hydrology* 330: 368–381.

Matott LS, Babendreier JE, and Purucker ST (2009) Evaluating uncertainty in integrated environmental models: A review of concepts and tools. *Water Resources Research* 45: W06421 (doi:10.1029/2008WR007301).

Melching CS (1992) An improved first-order reliability approach for assessing uncertainties in hydrologic modeling. *Journal of Hydrology* 132: 157–177.

Merz R and Blöschl G (2008a) Flood frequency hydrology: 1. Temporal, spatial, and causal expansion of information. *Water Resources Research* 44: W08432, doi:10.1029/2007WR006744.

Merz R and Blöschl G (2008b) Flood frequency hydrology: 2. Combining data evidence. *Water Resources Research* 44: W08433, doi:10.1029/2007WR006745.

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 (2007) What do we mean by 'uncertainty'? The need for a consistent wording about uncertainty assessment in hydrology. *Hydrological Processes* 21: 841–845.

Montanari A and Brath A (2004) A stocastic approach for assessing the uncertainty of rainfall-runoff simulations. *Water Resources Research* 40: W01106 (doi:10.1029/2003WR002540).

Montanari A and Grossi G (2008) Estimating the uncertainty of hydrological forecasts: A statistical approach. *Water Resources Research* 44: W00B08 (doi:10.1029/2008WR006897).

Montanari A, Shoemaker CA, and Van de Giesen NC (2009) Uncertainty assessment in surface and subsurface hydrology: An overview. *Water Resources Research* 45: W00B00 (doi:10.1029/2009WR008471).

Montanari L, Sivapalan M, and Montanari A (2006) Investigation of dominant hydrological processes in a tropical catchment in a monsoonal climate via the downward approach. *Hydrology and Earth System Sciences* 10: 769–782.

Morrissey ML, Maliekal JA, Greene JS, and Wang J (1995) The uncertainty of simple spatial averages using rain gauge networks. *Water Resources Research* 31: 2011–2017.

Mugunthan P and Shoemaker CA (2006) Assessing the impacts of parameter uncertainty for computationally expensive groundwater models. *Water Resources Research* 42: W10428 (doi:10.1029/2005WR004640).

Mulvaney TJ (1851) On the use of self-registering rain and flood gauges in making observations of the relations of rainfall and of flood discharges in a given catchment. *Proceeding of the Institute of Civil Engineers of Ireland* 4: 18–31.

Nash JE and Sutcliffe JV (1970) River flow forecasting through conceptual models. 1: A discussion of principles. *Journal of Hydrology* 10: 282–290.

Nauta MJ (2000) Separation of uncertainty and variability in quantitative microbial risk assessment models. *International Journal of Food Microbiology* 57: 9–18.

Neuman SP (2003) Maximum likelihood Bayesian averaging of alternative conceptual-mathematical models. *Stochastic Environmental Resources Risk Assessment* 17: 291–305.

NRC (National Research Council) (2000) *Risk Analysis and Uncertainty in Flood Damage Reduction Studies*. Washington, DC: National Academy Press.

Page T, Beven KJ, Freer J, and Neal C (2007) Modelling the chloride signal at Plynlimon, Wales, using a modified dynamic TOPMODEL incorporating conservative chemical mixing (with uncertainty). *Hydrological Processes* 21: 292–307.

Papolulis A (1991) *Probability, Random Variables, and Stochastic Processes*, 3rd edn. New York: McGraw-Hill.

Pappenberger F and Beven KJ (2004) Functional classification and evaluation of hydrographs based on Multicomponent Mapping (M^x). *International Journal of River Basin Management* 2: 89–100.

Pappenberger F and Beven KJ (2006) Ignorance is bliss: Or seven reasons not to use uncertainty analysis. *Water Resources Research* 42: W05302 (doi:10.1029/2005WR004820).

Pappenberger F, Matgen P, Beven KJ, Henry JB, Pfister L, and de Fraipont P (2006) Influence of uncertain boundary conditions and model structure on flood inundation predictions. *Advances in Water Resources* 29: 1430–1449.

Pelletier MP (1987) Uncertainties in the determination of river discharge: A literature review. *Canadian Journal of Civil Engineering* 15: 834–850.

Petersen-Øverleir A and Reitan T (2009) Accounting for rating curve imprecision in flood frequency analysis using likelihood-based methods. *Journal of Hydrology* 366: 89–100.

Rantz SE, *et al.* (1982) *Measurement and Computation of Streamflow*, US Geological Survey, Water Supply Paper 2175. http://water.usgs.gov/pubs/wsp/wsp2175/ (accessed March 2010).

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

Regan HM, Akcakaya HR, Ferson S, Root KV, Carroll S, and Ginzburg LR (2003) Treatments of uncertainty and variability in ecological risk assessment of single-species populations. *Human Ecology Risk Assessment* 9: 889–906.

Reggiani P, Renner M, Weerts AH, and van Gelder PAHJM (2009) Uncertainty assessment via Bayesian revision of ensemble streamflow predictions in the operational river Rhine forecasting system. *Water Resources Research* 45: W02428 (doi:10.1029/2007WR006758).

Rojas R, Batelaan O, Feyen L, and Dassargues A (2009) Assessment of conceptual model uncertainty for the regional aquifer Pampa del Tamarugal – North Chile. *Hydrology and Earth System Sciences Discussion* 6: 5881–5935.

Schmidt AR (2002) Analysis of stage-discharge relations for open channel flow and their associated uncertainties. PhD Thesis, University of Illinois.

Schoups G, Van de Giesen NC, and Savenije HHG (2008) Model complexity control for hydrologic prediction. *Water Resources Research* 44: W00B03 (doi:10.1029/2008WR006836).

Sherman LK (1932) Streamflow from rainfall by the unit-graph method. *Engineering News Record* 108: 501–505.

Shrestha DL, Kayastha N, and Solomatine DP (2009) A novel approach to parameter uncertainty analysis of hydrological models using neural networks. *Hydrology and Earth System Sciences* 13: 1235–1248.

Shrestha DL and Solomatine DP (2008) Data-driven approaches for estimating uncertainty in rainfall-runoff modelling. *International Journal of River Basin Management* 6: 109–122.

Sieber A and Uhlenbrook S (2005) Sensitivity analyses of a distributed catchment model to verify the model structure. *Journal of Hydrology* 310: 216–235.

Solomatine DP and Shrestha DL (2009) A novel method to estimate model uncertainty using machine learning techniques. *Water Resources Research* 45: W00B11 (doi:10.1029/2008WR006839).

Sonich-Mullin C (2001) Harmonizing the incorporation of uncertainty and variability into risk assessment: An international approach. *Human and Ecological Risk Assessment* 7: 7–13.

Spear RC and Hornberger GM (1980) Eutrophication in Peel Inlet: II. Identification of critical uncertainties via generalized sensitivity analysis. *Water Research* 14: 43–49.

Stedinger JR, Vogel RM, Lee SU, and Batchelder R (2008) Appraisal of the generalized likelihood uncertainty estimation (GLUE) method. *Water Resources Research* 44: W00B06 (doi:10.1029/2008WR006822).

Thiessen AH (1911) Precipitation averages for large areas. *Monthly Weather Report* 39: 1082–1084.

Thorsen M, Refsgaard JC, Hansen S, Pebesma E, Jensen JB, and Kleeschulte S (2001) Assessment of uncertainty in simulation of nitrate leaching to aquifers at catchment scale. *Journal of Hydrology* 242: 210–227.

Thyer M, Renard B, Kavetski D, Kuczera G, Franks SW, and Srikanthan S (2009) Critical evaluation of parameter consistency and predictive uncertainty in hydrological modeling: A case study using Bayesian total error analysis. *Water Resources Research* 45: W00B14 (doi:10.1029/2008WR006825).

Tolson BA and Shoemaker CA (2008) Efficient predicition uncertainty approximation in the calibration of environmental simulation models. *Water Resources Research* 44: W04411 (doi:10.1029/2007WR005869).

Tonkin M and Doherty J (2009) Calibration-constrained Monte Carlo analysis of highly parameterized models using subspace techniques. *Water Resources Research* 45: W00B10 (doi:10.1029/2007WR006678).

Tsai CW and Franceschini S (2005) Evaluation of probabilistic point estimate methods in uncertainty analysis for environmental engineering applications. Journal of Environmental Engineering 131: 387–395.

Van Griensven A, Meixner T, Grunwald S, Bishop T, Di luzio M, and Srinivasan R (2006) A global sensitivity analysis tool for the parameters of multi-variable catchment models. *Journal of Hydrology* 324: 10–23.

Von Bertalanffy L (1968) *General System Theory.* New York: George Braziller.

Vrugt JA and Bouten W (2002) Validity of first-order approximations to describe parameter uncertainty in soil hydrologic models. *Soil Science Society of America Journal* 66: 1740–1752.

Vrugt JA, Gupta HV, Bastidas LA, Bouten W, and Sorooshian S (2003b) Effective and efficient algorithm for multiobjective optimization of hydrologic models. *Water Resources Research* 39: 1214 (doi:10.1029/2002WR001746).

Vrugt JA, Gupta HV, Bouten W, and Sorooshian S (2003a) *A Shuffled Complex Evolution Metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters, Water Resources Research* 39: 1201 (doi:10.1029/2002WR001642).

Vrugt JA and Robinson BA (2007) Improved evolutionary optimization from genetically adaptive multimethod search. *Proceedings of the National Academy of Science of the United States of America* 104: 708–711.

Wagener T, McIntyre N, Lees MJ, Wheater HS, and Gupta HV (2003) Towards reduced uncertainty in conceptual rainfall-runoff modelling: Dynamic identifiability analysis. *Hydrological Processes* 17: 455–476.

Winsemius HC, Savenije HHG, and Bastiaanssen WGM (2008) Constraining model parameters on remotely sensed evaporation: Justification for distribution in ungauged basins? *Hydrology and Earth System Sciences* 12: 1403–1413.

Winsemius HC, Savenije HHG, Gerrits AMJ, Zapreeva EA, and Klees R (2006) Comparison of two model approaches in the Zambezi river basin with regard to model reliability and identifiability. *Hydrology and Earth System Sciences* 10: 339–352.

Ye M, Meyer PD, and Neumann SP (2008) On model selection criteria in multimodel analysis. *Water Resources Research* 44: W03428 (doi:10.1029/2008WR006803).

Ye M, Neuman SP, and Meyer PD (2004) Maximum likelihood Bayesian averaging of spatial variability models in unsaturated fractured tuff. *Water Resources Research* 40: W05113 (doi:10.1029/2003WR002557).

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

Zappa M, Rotach MV, Arpagaus M, *et al.* (2008) MAP D5-PHASE: Real-time demonstration of hydrological ensemble prediction systems. *Atmospheric Science Letters* 9: 80–87.

Zhang X, Srinivasan R, and Bosch D (2009) Calibration and uncertainty analysis of the SWAT model using genetic algorithms and bayesian model averaging. *Journal of Hydrology* 374: 307–317.

Zhang Z, Wagener T, Reed P, and Bhushan R (2008) Reducing uncertainty in predictions in ungauged basins by combining hydrologic indices regionalization and multiobjective optimization. *Water Resources Research* 44: W00B04 (doi:10.1029/2008WR006833).

Relevant Websites

http://education.mit.edu
Gaussian Distribution.

http://www.itia.ntua.gr
Presentation: Hurst-Kolmogorov dynamics and uncertainity.

http://www.agu.org
Special issue on uncertainty.