WATER RESOURCES RESEARCH, VOL. ???, XXXX, DOI:10.1029/,

A Blueprint for Process-Based Modeling of Uncertain Hydrological Systems

Alberto Montanari[1] and Demetris Koutsoyiannis[2]

A. Montanari, Department DICAM, University of Bologna, via del Risorgimento 2, Bologna, I-40136, Italy. (alberto.montanari@unibo.it)

D. Koutsoyiannis, Department of Water Resources and Environmental Engineering, National Technical University of Athens, Heroon Polytechneiou 5, Zographou, GR-157 80, Greece.

[1]Department DICAM, University of Bologna, Bologna, Italy.

[2]Department of Water Resources and Environmental Engineering National Technical University of Athens, Zographou, Greece.

Abstract. We present a probability based theoretical scheme for building process-based models of uncertain hydrological systems, thereby unifying hydrological modeling and uncertainty assessment. Uncertainty for the model output is assessed by estimating the related probability distribution via simulation, thus shifting from one to many applications of the selected hydrological model. Each simulation is performed after stochastically perturbing input data, parameters and model output, this latter by adding random outcomes from the population of the model error, whose probability distribution is conditioned on input data and model parameters. Within this view randomness, and therefore uncertainty, is treated as an inherent property of hydrological systems. We discuss the related assumptions as well as the open research questions. The theoretical framework is illustrated by presenting real-world and synthetic applications. The relevant contribution of this study is related to proposing a statistically consistent simulation framework for uncertainty estimation which does not require model likelihood computation and simplification of the model structure. The results show that uncertainty is satisfactorily estimated although the impact of the assumptions could be significant in conditions of data scarcity.

1. Introduction

Process-based modeling has been a major focus for hydrologists for four decades already. In fact, more than forty years passed since *Freeze and Harlan* [1969] proposed their "physically-based digitally simulated hydrologic response model", which set the basis for detailed process-based simulation in hydrology. The terms "physically-based" and "process-based" models are often used interchangeably, in contrast to purely empirical models. Other times, "process-based" is regarded to include a family of models broader than "physically-based". In fact, through the years it has become clear that there are no purely "physically-based" models for large hydrological systems. All models include assumptions and simplifications that depart from pure deductive physics and thus the term "process-based" is more accurate and general.

A comprehensive review of the research activity related to physically-based models was presented by *Beven* [2002]. Perhaps one of the most known process-based models in hydrology is the spatially-distributed Système Hydrologique Européen (SHE, see *Abbot et al.* [1986]), which has been the subject of many contributions. Another relevant contribution was given by *Reggiani et al.* [1998, 1999] who introduced the concept of "representative elementary watershed". Many other approaches were recently proposed which refer to process-based models in general. In fact, in the past ten years, process-based modeling, in contrast to empirical modeling, has been one of the targets of the well known "Prediction in Ungauged Basins" (PUB; see *Kundzewicz* [2007]) initiative of the International Association of Hydrological Sciences (IAHS).

Process-based models are almost always formulated in deterministic form, by setting up a set of mathematical equations. However, during the last four decades it became increasingly clear that deterministic models in hydrology are never accurate and imply uncertainty whose estimation is important for real world decision making (see, for instance, *Grayson et al.* [1992]; *Beven* [1989, 2001]). Some authors expressed their belief that uncertainty in hydrology is epistemic and therefore can be in principle eliminated through a more accurate representation of the related processes [*Sivapalan et al.*, 2003]. However, recent research suggested that uncertainty is unavoidable in hydrology, originating from natural variability and related to inherent unpredictability in deterministic terms, which is typically referred to as randomness (see, for instance, *Montanari et al.* [2009]; *Koutsoyiannis et al.* [2009]). The latter concept implies that to produce a fully deterministic model that would eliminate uncertainty is impossible and that modeling schemes need to explicitly recognise its role [*Beven*, 2002].

Indeed, recent contributions were proposed where deterministic hydrological modeling was efficiently coupled with uncertainty assessment. The most relevant example is the Generalised Likelihood Uncertainty Estimation (GLUE) [*Beven and Binley*, 1992], where multiple modeling schemes are retained provided they are behavioral in the face of uncertainty. GLUE was long discussed and sometimes criticized for using informal approaches for statistical inference and in particular for computing the model likelihood (see, for instance, *Stedinger et al.* [2008] and *Mantovan and Todini* [2006]), but was used in several practical applications. Relevant recent contributions to GLUE were given by *Liu et al.* [2009], who proposed the limits of acceptability approach to retain behavioral simulations (see also *Winsemius et al.* [2009]), and *Krueger et al.* [2010], who explicitly considered the contribution of data, parameter and model uncertainty via ensemble simulation. However, GLUE still suffers from subjectivity, mainly related to the identification of behavioral models and probability estimation for their output [*Stedinger et al.*, 2008], which in GLUE is obtained through (possibly informal) likelihood estimation for any candidate model.

A second relevant approach to uncertainty assessment was proposed by *Krzysztofowicz* [2002] who introduced the Bayesian Forecasting System (BFS), which aims to estimate the probability distribution for a future river stage or river flow. The method is based on the preliminary identification of a prior distribution for the unknown future variable, which is obtained by approximating the river flow process with a linear model, therefore expressing future observations depending on past ones. Then, the posterior distribution is obtained by including the information provided by the hydrological model prediction. The probability distributions are estimated in the Gaussian domain by normalizing predictors and predictand through the normal quantile transform [*Krzysztofowicz*, 2002]. The method assumes that the dominant source of uncertainty is related to rainfall forecasting, thus focusing on a very specific type of application. The above assumption implies that hydrological uncertainty is estimated by introducing restrictive approximations. In particular, parameter uncertainty for the hydrological model is neglected.

Estimation of hydrological uncertainty is also the target of the BATEA method [*Kavetski et al.*, 2006], where a novel approach is introduced to account for all sources of data uncertainty. In particular, rainfall uncertainty is accounted for by introducing a rainfall multiplier. The probability distribution for model parameters is estimated through the Bayes theorem and therefore a formulation for the model likelihood needs to be identified.

For example, *Kavetski et al.* [2006] adopted a likelihood function depending on the sum of squared residuals.

In recent times, there has been a renewed interest for multi-model approaches, which estimate unknown hydrological variables by averaging outputs from several models. This possibility is also offered by GLUE [*Krueger et al.*, 2010]. Another relevant example is the Maximum Likelihood Bayesian Model Averaging (MLBMA, see *Neuman* [2003]; *Ye et al.* [2008]). Multi-model techniques may require likelihood estimation to derive the probability that each model is correct.

The above considerations show that model likelihood estimation is a key step for many uncertainty assessment methods. It is well known that likelihood computation for hydrological models is a very challenging task, due to the complex structure of the model error which makes its statistical description complicated. Interesting contributions were recently proposed by *Schoups and Vrugt* [2010] and *Pianosi and Raso* [2012] who proposed innovative likelihood formulations. However, they are still based on assumptions that may be restrictive in some practical applications, like the use of a model bias correction factor in *Schoups and Vrugt* [2010] and the hypothesis of independence for the model error in *Pianosi and Raso* [2012]. Moreover, there is a drawback related to the use of the likelihood to estimate the reliability of a model in hydrology: in fact, the likelihood is usually estimated in calibration, as it is done in statistics, but is used to assess uncertainty of out-of-sample predictions. Therefore it is implicitly assumed that model performances in calibration are analogous to those experienced in the evaluation period. Namely, one assumes that the model errors during calibration are statistically representatives of those that will be experienced in applications. Actually, this assumption is valid only under the condition that the hydrological model is stationary and not overparametrised, but fails in all instances in which the actual model reliability is expected to deteriorate with respect to calibration (as it frequently happens in hydrology). This limitation, in the context of GLUE, is recognised by *Beven* [2006], p. 27. To overcome it, likelihood should be computed by running the model, after optimizing its parameters, during an evaluation period. Namely, one should refer to data that were not used in calibration and similar conditions with respect to those that are expected in applications.

An approach to hydrological uncertainty assessment which does not require likelihood estimation was presented by *Götzinger and Bárdossy* [2008] who assumed that the model error is given by the sum of the random components due to input uncertainty and process description uncertainty. To estimate this latter contribution, they assumed that the standard deviation of the random contribution of a certain process (model structural uncertainty) to the total uncertainty is proportional to the sensitivity of the output to the related parameter group. The above assumptions may not be satisfied in practical applications (see, for instance, *Beven* [2006]).

Likelihood computation might also be avoided by using data assimilation methods, for which a comprehensive review, from a system-perspective, was presented by *Liu and Gupta* [2007]. In fact, the Bayesian uncertainty assessment method developed by *Bulygina and Gupta* [2009, 2010, 2011] assumes that the hydrological system evolves in time according to a first-order Markov state-space process and the relationships among the relevant variables (inputs, states and outputs) is represented through the direct estimation of their joint probability density function. This latter takes uncertainty into account and is conditioned on both the observed data and the available conceptual understanding of system physics, therefore obtaining a flexible and statistically consistent approach. *Bulygina and Gupta* [2010] note that additional research is needed to make the method applicable to complex system models. Moreover, if no or weakly informative prior is used, any prediction is mainly based on the conditions observed during the considered observation period only, and therefore particular care should be used when extrapolating to out-of-sample conditions.

The purpose of this paper is to introduce a novel methodological scheme for estimating the probability distribution of the output from a process-based (deterministic) hydrological model. The distinguishing feature of the approach herein proposed is that likelihood computation can be avoided, without imposing any restriction to model complexity, therefore complementing the features of the techniques reviewed above. Conversely, the most significant limitation is that the probability distributions of input data, model parameters and model error are needed as input information. It is well known that their definition is still a challenging task in practical applications. The underlying theory is derived in a probabilistic framework, in which Bayesian concepts can be introduced to take into account prior information. Statistical consistency of the scheme is ensured by introducing assumptions whose reliability is discussed below. The scheme itself is based on converting a deterministic hydrological model into a stochastic one, therefore incorporating randomness in hydrological modeling as a fundamental component. In fact, in our framework uncertainty is recognized as an inherent property of the water cycle, taking into account randomness of atmospheric processes, drop paths, soil properties, turbulence in fluid mechanics and many others.

In the next Section of the paper we provide more details on the rationale for stochastic process-based modeling. The third section of the paper is dedicated to the theory underlying the new blueprint that we are proposing. The fourth section describes the practical application. The fifth section reviews the underlying assumptions and their limitations, while in the sixth section two examples of application are presented. In the seventh section we discuss the value of uncertainty estimation as a learning process. Finally we discuss open research questions and draw some conclusions.

2. The rationale for stochastic process-based modeling

In a deterministic model the outcomes are precisely determined through known relationships among states and events, without any room for random variation. A given input (including initial and boundary conditions) will always produce the same output and therefore uncertainty is not taken into account in a formal manner. Uncertainty assessment, when needed, is often carried out indirectly, e.g., by post processing the results. Such separation of targets has been favoured by the illusion that uncertainty can be eliminated by refining deterministic modeling (see, for instance, *Sivapalan et al.* [2003]). Such refining has commonly been envisaged through a "reductionist" approach, in which all heterogeneous details of a catchment would be modeled explicitly and the modeling of details would provide the behavior of the entire system (for an extended discussion see *Beven* [2002]). However, some researchers have pointed out that this is a hopeless task [*Savenije*, 2009]. Indeed, physical processes governing the water cycle involve inherent randomness; for instance, meteorological processes are governed by the laws of thermodynamics, which are, essentially, statistical physical laws. Moreover, some degree of approximation is unavoidable in process-based modeling in hydrology [*Beven*, 1989]. In fact, physical laws give simple and meaningful descriptions of problems in simple systems, but their application in hydrological systems demands simplification, lumping and statistical parameterization *Beven* [1989], and sometimes even replacing by conceptual or statistical laws (e.g. the Manning formula). Therefore, uncertainty in hydrology is not just related to temporary knowledge limitations (epistemic uncertainty) but it rather reflects inherent randomness and therefore it is unavoidable (see *Koutsoyiannis et al.* [2009]). Thus, the traditional deterministic form of process-based modeling in hydrology is a relevant limitation per se which should be overcome by incorporating uncertainty modeling in a fully integrated approach.

In fact, we believe that recognizing uncertainty as an essential attribute of the water cycle, which needs to be respected in process-based models, is not just a nuisance. In our view, uncertainty estimation is not the remedy against limited representativity of deterministic schemes (which some may believe to be a transient weakness of current models that would be cleared up in the future), but rather a way to fully take into account and reproduce in a process-based framework the dynamics of hydrological systems. There are many possible alternatives to deal with uncertainty thereby overcoming the limitations of purely deterministic approaches, including subjective methods like fuzzy logic, possibility theory, and others [*Beven*, 2009; *Montanari*, 2007]. We believe that one of the most comprehensive, elegant and complete ways of dealing with uncertainty is provided by the theory of probability. In fact, probabilistic descriptions allow predictability (supported by deterministic laws) and unpredictability (given by randomness) to coexist in a unified theoretical framework, therefore giving us the means to efficiently exploit and improve the available physical understanding of uncertain systems [*Koutsoyiannis et al.*, 2009].

The theory of stochastic processes also allows the incorporation into our descriptions of (possibly human induced) changes affecting hydrological processes [*Koutsoyiannis*, 2011], by modifying their physical representation and/or their statistical properties (see, for instance, *Merz and Blöschl* [2008a, b]). Finally, subjectivity and expert knowledge can be taken into account in prior distribution functions through Bayesian theory [*Box and Tiao*, 1973]. Therefore, in our opinion, a theoretical setting needs to be established where probability-based modeling of uncertainty is an essential piece of possibly complex deterministic models. Such a setting should be flexible enough to allow the deterministic model to increase in complexity therefore reducing epistemic uncertainty as much as possible in the future, while retaining the essential role of inherent randomness.

In view of the above considerations and in agreement with *Beven* [2002], we believe that a new blueprint should be established, which should be built on a key concept that is actually well known: it is stochastic process-based modeling, which needs to be brought to a new light in hydrology. Here the term "stochastic" is used to collectively represent probability, statistics and stochastic processes. We formalize the theoretical framework for the application of this type of approach here below.

3. Formulating a Process-Based Model Within a Stochastic Framework

In this section we show how a deterministic model can be converted into an essential part of a wider stochastic approach through an analytical transformation. Such a conversion is necessary to understand how the probability distribution of the model output can be estimated by simulation, without necessarily requiring likelihood computation. From an analytical point of view, while the deterministic formulation of the model transforms the values of the inputs into an output value, the stochastic version of the model acts on probability densities, rather than single values, of the inputs, producing a probability density of the output. That is, the deterministic model acts on values of variables while the stochastic model acts on probability densities thereof. From a numerical point of view, a deviation (error term) from a single-valued relationship is introduced and the density of the output is calculated by repeated applications of the single-valued (deterministic) version of the model, where the model output is stochastically perturbed to account for uncertainties in a statistically consistent framework. The scheme presented here focuses on the conversion of a single deterministic model into a stochastic model. However, a multi-model extension, which uses more than one deterministic model, is straightforward and will be discussed below.

In what follows, we use the following definitions:

- Input uncertainty: it is defined as the uncertainty in the data input to the model, which is quantified by an underlying probability distribution. It is related to observation methods and networks.

- Parameter uncertainty: it is defined as the uncertainty in the model parameter vector. It is mainly related to model structure, calibration method and consistency of the underlying data base.

- Model error: for a given model, input data and parameter vector, it is defined as the difference between model simulation and the corresponding data value. It is mainly related to model inability to reproduce the related real processes (model structural error). Here, it is assumed to resemble all the uncertainties that are not included in input data and parameter uncertainty.

- Prediction uncertainty: for a given model (or set of models in a multi-model framework), it is defined as the uncertainty in the prediction of the true value of a given hydrological variable. It is quantified by the probability distribution of the variable to be predicted and is typically expressed also in the form of prediction limits of the simulation. These latter quantify the range for the variable within which the true value falls with probability equal to the confidence level. Prediction uncertainty is formed up by input and data uncertainty and model error.

The analytical procedure to convert a deterministic model into a stochastic framework is rather technical and is expressed by equations (1) to (6) below. We would like to introduce the new blueprint with a fully comprehensible treatment for those who are not acquainted with (or do not like) stochastics. Therefore, the presentation is structured to allow the reader who is interested in the application only to directly jump to equations (7) and (8) without any loss of practical meaning.

Hydrological models are typically expressed through a deterministic formulation, namely, a single valued transformation. In general, it can be written as

$$Q = S\left(\Theta, \mathrm{X}\right) \tag{1}$$

where Q is the model prediction which, in a deterministic framework, is implicitly assumed to equal the true value of the variable to be predicted. The mathematical relationship S represents the model structure, X indicates the input data vector and Θ the parameter vector. In the stochastic framework, the hydrological model is expressed in stochastic terms, namely [*Koutsoyiannis*, 2010],

$$f_Q\left(Q\right) = K_S f_{\Theta,\mathrm{X}}\left(\Theta, \mathrm{X}\right) \tag{2}$$

where f indicates a probability density function, and K_S is a transfer operator that is related to, and generalizes in a stochastic context, the deterministic model S. Within this context, Q indicates the true variable to be predicted, which is unknown at the prediction time and therefore is treated as a random variable. K_S can be generalized to represent a so-called stochastic operator, which implements a shift from one to many transformations S.

Note that, by starting from eq. (1) and (2) above, we assume that Q depends on input data X and parameters Θ through the model S. Therefore, f_Q (and thus uncertainty of Q) depends on $f_{X,\Theta}$ (and thus uncertainty of X and Θ) and model S uncertainty (through the operator K_S). It follows that the model error is assumed to resemble all the uncertainties that are not included in input data and parameter uncertainty, as it was noted in the definitions above. In principle, other uncertainty sources could be considered explicitly. For instance, dependence on the initial conditions and therefore their uncertainty can be easily included in eq.(1) and (2). In what follows it is omitted to simplify notation (note that initial conditions can be included in the input data vector X).

A stochastic operator can be defined by using a stochastic kernel $k\left(e, \Theta, \mathrm{X}\right)$, with e reflecting a deviation from a single valued transformation. Here we will assume that e is a stochastic process, with marginal probability density $f_e(e)$, representing the model error according to the additive relationship

$$Q = S\left(\Theta, \mathrm{X}\right) + e \ . \tag{3}$$

Note that alternative structures for the model error can be defined, for instance by introducing multiplicative terms. The term e accounts for the model uncertainty that was discussed above.

The stochastic kernel introduced above must satisfy the following conditions:

$$k\left(e,\Theta,\mathrm{X}\right)\geq 0 \text{ and } \int_{e} k\left(e,\Theta,\mathrm{X}\right)de = 1\ , \tag{4}$$

which are met if $k\left(e,\Theta,\mathrm{X}\right)$ is a probability density function with respect to e.

Specifically, the operator K_S applying on $f_{\Theta,\mathrm{X}}\left(\Theta,\mathrm{X}\right)$ is then defined as [*Lasota and Mackey*, 1985, p. 101]

$$K_S f_{\Theta,\mathrm{X}}\left(\Theta,\mathrm{X}\right)=\int_{\Theta}\int_{\mathrm{X}} k\left(e,\Theta,\mathrm{X}\right) f_{\Theta,\mathrm{X}}\left(\Theta,\mathrm{X}\right)d\Theta \mathrm{dX}\ . \tag{5}$$

Under the assumption that parameter uncertainty is independent of data uncertainty, for the purpose of estimating the probability density $f_Q\left(Q\right)$ the joint probability distribution $f_{\Theta,\mathrm{X}}\left(\Theta,\mathrm{X}\right)$ can be substituted by the product of the two marginal distributions $f_{\Theta}\left(\Theta\right)f_{\mathrm{X}}\left(\mathrm{X}\right)$. Note that we are not excluding dependence among the single elements of the input data as well as parameter vector, and also note that this assumption can be removed and therefore does not affect the generality of the approach, as we discuss in Section 5.

In view of this latter result, by combining eq. (2) and eq. (5), in which the model error can be written as $e = Q - S(\Theta,\mathrm{X})$ according to eq. (3), we obtain

$$f_Q\left(Q\right)=\int_{\Theta}\int_{\mathrm{X}} k\left(Q-S\left(\Theta,\mathrm{X}\right),\Theta,\mathrm{X}\right) f_{\Theta}\left(\Theta\right) f_{\mathrm{X}}\left(\mathrm{X}\right)d\Theta d\mathrm{X}\ . \tag{6}$$

At this stage we need to identify a suitable expression for $k\left(Q-S\left(\Theta,\mathrm{X}\right),\Theta,\mathrm{X}\right)$. Upon substituting eq. (3) in eq. (6) and remembering that k is a probability density function with respect to the model error e, we recognize that the kernel is none other than the conditional density function of e for the given Θ and X, i.e., $f_e\left(Q-S\left(\Theta,\mathrm{X}\right)|\Theta,\mathrm{X}\right)$.

To summarise the whole set of analytical derivations expressed by equations (1) to (6) we may see that we passed from the deterministic formulation of the hydrological model expressed by eq. (1), i.e. (to replicate it for clarity),

$$Q = S(\Theta, \mathrm{X}) \tag{7}$$

to the stochastic formulation expressed by

$$f_Q(Q) = \int_\Theta \int_\mathrm{X} f_e(Q - S(\Theta, \mathrm{X})|\Theta, \mathrm{X}) f_\Theta(\Theta) f_\mathrm{X}(\mathrm{X}) \, d\Theta d\mathrm{X} \tag{8}$$

with the following meaning of the symbols:

- $f_Q(Q)$: probability density function of the true value of the hydrological variable to be predicted;

- $S(\Theta, \mathrm{X})$: deterministic part of the hydrological model;

- $f_e(Q - S(\Theta, \mathrm{X})|\Theta, \mathrm{X})$: conditional probability density function of the model error. According to eq. (2) it can also be written as $f_e(e|\Theta, \mathrm{X})$;

- Θ: model parameter vector;

- $f_\Theta(\Theta)$: probability density function of model parameter vector;

- X: input data;

- $f_\mathrm{X}(\mathrm{X})$: probability density function of input data.

In eq. (8) the conditional probability distribution of the model error $f_e(Q - S(\Theta, \mathrm{X})|\Theta, \mathrm{X})$ is conditioned on input data X and parameter vector Θ. Such formulation would be useful if we needed to account for changes in time of the conditional statistics of the model error (like, for instance, those originated by heteroscedasticity). On the other hand, if we assumed that the model error is independent of X and Θ, then eq. (8) can be written in the simplified form

$$f_Q(Q) = \int_\Theta \int_\mathrm{X} f_e(Q - S(\Theta, \mathrm{X})) f_\Theta(\Theta) f_\mathrm{X}(\mathrm{X}) \, d\Theta d\mathrm{X} \, . \tag{9}$$

The presence of a double integral in eq. (8) and eq. (9) may induce the feeling to the reader that the practical application of the proposed framework is cumbersome. Actually, the double integral can be easily computed through numerical integration, namely, by applying a Monte Carlo simulation procedure that is well known and already used in hydrology (see *Koutsoyiannis* [2010]). The only problem is related to the computational requirement, which might become significant when dealing with complex models and large basins. We explain the numerical integration in the next section of the paper.

The above theoretical scheme is very general yet its formalism has been given in terms of converting a single deterministic model into a stochastic model. However, the generality and flexibility of the approach allow for an extension to a multi-model framework. Multi-modeling schemes allow to test multiple working hypotheses and model structures thereby testing individual components of process-based models (for an extensive discussion see *Clark et al.* [2011]; for an example of application see *Krueger et al.* [2010]). A preliminary estimation of the weight w_i of each i-th model is necessary, which quantifies the importance of each model in the simulation process. The weight is related to model performances with respect to other candidate models (*Neuman* [2003], page 297, defines the weight as the probability that the model is correct), and can be estimated by using prior judgemental information. Uniform probability across different models is a reasonable working hypothesis, which should however be supported by expert knowledge to avoid that the same importance is given to models with much dissimilar predicting capabilities. The multi-model probability density function $f_Q(Q)$ can be written as

$$f_Q\left(Q\right) = \sum_{i=1}^{M} w_i f_Q^{(i)}\left(Q\right) \ , \tag{10}$$

where M is the number of the considered models and $f_Q^{(i)}(Q)$ is the probability distribution derived through eq. (8) or (9) for each single model i, $1 < i \leq M$.

It can be seen that methodological scheme introduced above does not require computation of the model likelihood, therefore avoiding to introduce any related assumption. However, the statistical properties of the model error still need to be deciphered, although in a less detailed (and perhaps non-parametrical) manner, to compute the integrals in eq. (8) and (9) (see Section 4 for details on application). In the applications presented in this paper we use the meta-Gaussian approach by *Montanari and Brath* [2004] to this end. Its robustness notwithstanding, further research studies are needed to provide an efficient statistical characterisation of the errors from hydrological models, which are often heteroscedastic and affected by several forms of dependence not easy to decipher (see, for instance, *Refsgaard et al.* [2006], *Kavetski et al.* [2006] and *Beven* [2006]). The interested reader is also referred to *Montanari and Grossi* [2008] for an additional discussion on the meta-Gaussian approach and error dependency. It is important to note that the model error should be representative of model performances in validation.

4. Application of the Proposed Framework: Integrating Hydrological Model Implementation and Uncertainty Assessment

Estimating the probability distribution of the true value of the variable to be predicted by a hydrological model (prediction uncertainty) is equivalent to simultaneously carrying out model implementation and uncertainty assessment. The framework for estimating the probability density function of model prediction, $f_Q(Q)$, was described in Section 3. Here we show how eq. (8) can be applied in practice.

We assume that the probability density functions of model parameters, input data and model error are known, for instance because they were already estimated by using procedures such as those found in the hydrological literature (see, for instance, *Clark and Slater* [2006], *McMillan et al.* [2011], *Renard et al.* [2011] and *Di Baldassarre and Montanari* [2009] for data uncertainty; *Vrugt et al.* [2007], *Ebtehaj et al.* [2010] and *Srikanthan et al.* [2009] for parameter uncertainty; and *Montanari and Brath* [2004], *Montanari and Grossi* [2008] and *Krzysztofowicz* [2002] for model error). A practical demonstration showing how this can be determined is contained in Section 6 below. Some of the above mentioned techniques require the estimation of model likelihood, which may imply approximations in the definition of the related uncertainties. For instance, likelihood assessment is required by DREAM, which is used in the applications presented in Section 6 to estimate parameter uncertainty. However, methods are available to avoid the use of the likelihood to define the above densities. For instance, bootstrap and resampling methods can be used for parameter uncertainty [*Ebtehaj et al.*, 2010; *Srikanthan et al.*, 2009]. It is important to note that definition of the above densities is a key task as an imperfect estimation of input and parameter uncertainty would propagate through the simulation chain thus inducing lack of fit. It is well known that the identification of such distributions is still a challenging task in hydrology. In particular, for the data the problem still remains a relevant research issue. Information on observation error, and the related probability distribution, can be used to this end (see, for instance, *Pelletier* [1987]).

Under the above premises the double integral in eq. (8) can be easily computed through a Monte Carlo simulation procedure, which can be carried out in practice by performing many implementations of the deterministic hydrological model $S(\Theta, \mathrm{X})$.

In detail the simulation procedure is carried out through the following steps that refer to the flowchart in Figure 1:

1. A parameter vector for the hydrological model is picked up at random from the model parameter space according to the probability density $f_{\Theta}(\Theta)$.

2. An input data vector for the hydrological model is picked up at random from the input data space according to the probability density $f_{\mathrm{X}}(\mathrm{X})$.

3. The hydrological model is run and a model prediction (or a vector of predictions) $S(\Theta, \mathrm{X})$ is computed.

4. An outcome of the model error (or vectors of errors) is picked up at random from the model error population according to the probability density $f_e(e)$ and added to the model prediction $S(\Theta, \mathrm{X})$.

5. The simulation described by items from 1 to 4 is repeated j times. Therefore we obtain j (vectors of) outcomes of the prediction Q.

6. Finally the probability density $f_Q\,(Q)$ is inferred through the realizations mentioned in item 5.

It is important to note that j needs to be sufficiently large, in order to accurately estimate the probability density $f_Q(Q)$. Generally, a good compromise of accuracy and computational efficiency to find an optimal j value should be evaluated case by case.

Once the probability distribution of the true value to be predicted Q is known we obtain a best estimate for the prediction along with an assessment of uncertainty for a given confidence level, under the above assumptions that are further discussed in the next Section of the paper.

For the practical application of the multi-model approach, the whole simulation procedure is to be repeated for any subsequent candidate model therefore obtaining the individual estimates for $f_Q^{(i)}(Q)$ to be averaged according to eq. (10).

In principle, the above framework allows one to estimate the single contribution of each uncertainty source. For instance, if we are interested in the impact of parameter uncertainty, the simulation procedure can be carried out by skipping items 2 and 4, therefore neglecting the impact of data uncertainty and model error. However, we should be fully aware that neither in the proposed framework nor in the real world are uncertainties necessarily additive. Thus, even if assessment of individual impacts is possible, these latter cannot be summed up a posteriori to assess the overall prediction uncertainty.

The algorithm presented above has some similarity with the operational flow chart of other simulation methods like GLUE or multi-model approaches. The relevant difference is the use of the model error to summarize uncertainties other than those induced by imprecise input data and parameters. In this way likelihood computation can be avoided. We stress once again that, in order to preserve the statistical consistency that is ensured by the underlying theoretical development, the probability distribution of the model error must be reliably inferred with the support of statistical tests.

5. Discussion of the Underlying Assumptions

Like any scientific method, the blueprint proposed in Sections 3 and 4 is based on assumptions in order to ensure applicability. When dealing with uncertainty assessment in hydrology, assumptions are often treated with suspicion, because it is felt that they undermine the effectiveness of the method and therefore its efficiency and credibility with respect to users. We stress here that assumptions (typically simplifying ones) are a means to reach a better understanding of the behaviors of natural processes and allow science to be effectively put into practice. As a matter of fact, assumptions are unavoidably needed to set up models, calibrate their parameters and estimate their reliability, whatever approach is used. Evidently, flawed assumptions may falsify statistical inference as well as any alternative model of uncertain and deterministic systems. Therefore the target of the researcher should not be to avoid assumptions, but rather discuss them transparently, evaluate their effects and, when possible, check them, for instance through statistical testing.

In order to discuss the assumptions conditioning the blueprint we introduced above, first we note that we assumed that the uncertainty of model outputs only depends on input data uncertainty, parameter uncertainty and model error, according to eq. (1) which states that model output itself only depends on input data, parameters and model structure S. Therefore, some sources of uncertainty are not taken explicitly into account, like for instance operation uncertainty [*Montanari et al.*, 2009] and discretisation errors when dealing with daily data. Indeed, several uncertainties are not explicitly accounted for in any uncertainty assessment method. To this regard, we would like to point out that the model error, for a given model and given input data and parameters, implicitly takes into account, in an aggregated and very practical form, all the sources of uncertainty that make the model output different with respect to observed values. However, other uncertainties sources, like for instance uncertainty in the initial conditions and state variables, could be included explicitly, provided the related probability distributions are quantified and random outcomes for their values are randomly picked up at each simulation step. We did not include additional uncertainties to simplify notation.

Second, we assumed that parameter uncertainty is independent of input uncertainty. If the data are sufficient, this assumption is reasonable, because parameters of a given model depend on statistics of the input data and not their particular values [*Casella and Berger*, 1994]. A practical demonstration of the limited sensitivity of rainfall-runoff model output to artificially induced input errors was recently given by *Montanari and Di Baldassarre* [2012]. We must note, however, that the data are seldom of sufficient size when fitting hydrological models. As a result, parameters often turn out to be dependent on the input uncertainty (so that changes in input data result in changes of parameters).

Further assumptions may be needed to estimate the probability density f_e of model error, which might be non-Gaussian and affected by heteroscedasticity. For instance, in the application presented in Section 6 the meta-Gaussian approach by *Montanari and Brath* [2004] is applied. In brief, the method recognizes the dependence on model prediction of the conditional probability distribution of model error. In this way change of the statistical properties during time is efficiently modeled and therefore the marginal probability distribution of the model error is allowed to be heteroscedastic (see Section 6 and *Montanari and Brath* [2004]). The meta-Gaussian approach assumes that the model error does not depend on parameter and data uncertainty. In this case also, if the data are sufficient the assumption is reasonable. We checked it with extended simulation experiments that are independently presented by *Montanari and Di Baldassarre* [2012].

In principle the above assumptions of independence could be removed by conditioning the model error and parameter uncertainty on data. Parameter uncertainty can be conditioned by calibrating the hydrological model for different outcomes of the input data from the related probability distribution. The model error can be conditioned by estimating its probability distribution at each step of the simulation procedure described in Section 4, therefore obtaining different error probability densities for different input data and parameters. The main problem with this solution is given by the increased computational requirements. We further discuss this issue in Section 6.4.

If other approaches are used to derive the probability distribution of the model error, different assumptions would be introduced depending on the (possibly informal) approach that is adopted. No matter which method is used, any additional assumption introduced for inferring f_e should be appropriately checked.

Another relevant issue has been pointed out by some researchers (see, for instance, *Beven et al.* [2011]) who are convinced that epistemic errors arising from hydrological models might be not aleatory and therefore are difficult (or impossible) to model by using stochastic approaches. In our view, variables are either deterministic or random. That is, if they cannot be described deterministically, then they can be modeled by using stochastics, no matter if their stochastic dynamics are driven by epistemic uncertainty or natural variability. Another issue that is frequently raised is that epistemic errors are affected by non-stationarity and therefore cannot be efficiently modeled by using stochastics [*Beven and Westerberg*, 2011]. Actually, such a view neglects the fact that even the definition of stationarity and nonstationarity relies on the theory of stochastic processes [*Koutsoyiannis*, 2006, 2011] and thus dealing with it is necessarily an issue of applying stochastics. In our opinion, non-stationarity might be necessary to enrol when environmental changes are present, but it is not induced by epistemic uncertainty. Irrespective of its origin, non-stationarity can be efficiently dealt with by using non-stationary stochastic processes [*Brockwell and Davis*, 1987], and by introducing and checking suitable assumptions. For instance, the stochastic kernel introduced in eq. (3) and (4) is conditioned on the input data and therefore its marginal statistical properties are changing in time. In addition, in the case studies we present here the model error is allowed to be heteroscedastic and correlated. Indeed, as we mentioned above, the meta-Gaussian approach provides a conditional probability distribution of the model error that is changing in time depending on the simulated river flow [*Montanari and Brath*, 2004].

It is important to note that the blueprint proposed here relies much on data. Although probability distributions of input data, model parameter and model error could be estimated according to expert knowledge, data analysis is a fundamental requirement for assessing uncertainty. Therefore, particular attention should be paid to data collection and checking, to avoid as much as possible the use of disinformative observations (for a detailed discussion on this issue the interested reader is referred to *Beven and Westerberg* [2011]).

One may be concerned by the computational requirements of the proposed framework, especially when dealing with complex modeling approaches. For instance, spatially distributed models might require sampling from several parameters and might involve significantly longer computational times. This issue is indeed a matter of concern for any numerical integration procedure. It needs to be carefully considered in view of the length of the simulation period and the minimal number of simulated data points that is required to reliably infer the probability distribution of the model output.

One reviewer of this paper asserted that, strictly speaking, none of the above assumptions is satisfied. We believe that if we accept such an assertion in its generality, we would convict all models and perhaps all scientific disciplines except pure mathematics, because all sciences that describe Nature seek to provide approximations of reality. It is well known that all models are wrong, and, likewise, all assumptions are never strictly satisfied. The purpose in modeling is to produce approximations of reality, which are tested whether or not they are satisfactory. If they are not, then the model should be changed, trying different model structures or perhaps relaxing some assumptions.

However, our practical experience suggests that the assumptions we introduced are reasonable. For instance, we believe that data uncertainty is indeed playing a negligible effect on parameter uncertainty in most real world applications (see, for instance, *Montanari and Di Baldassarre* [2012]).

6. Examples of Application

In order to illustrate the proposed blueprint with practical examples, two applications are presented here below that refer to different rainfall-runoff models applied to catchments located in Italy. In detail, the catchments are those of the Secchia River at Bacchello Bridge and the Leo River at Fanano, in the Emilia-Romagna region, in Northern Italy. Figure 2 shows their locations.

6.1. The case study of the Secchia River

The Secchia River is located in northern Italy and is a tributary to the Po River. The catchment area is 1214 km^2 at the Bacchello Bridge river cross section that is located about 62 km upstream of the confluence in the Po River. The maximum altitude is 2121 m above sea level (a.s.l.) at Mount Cusna. The main stream length up to Bacchello Bridge is about 98 km and the climate over the region is continental.

Hourly rainfall and temperature data are available for the years 1972 and 1973 in five raingauges. For the same period, hourly river flow data at Bacchello Bridge were collected.

To test the blueprint proposed here over an extended data set with controlled uncertainty, we used synthetic hourly rainfall, temperature and river flow data that cover a 50-year observation period. The same data set was used by *Montanari* [2005] who gives additional details. Synthetic data simulation is briefly described here below.

Rainfall data, for the 5 raingauges mentioned above, were generated using the generalized multivariate Neyman-Scott rectangular pulses model [*Cowpertwait*, 1995] that was calibrated using the observed data. Mean areal rainfall was then computed as a weighted sum of the rainfall in each raingauge, where weights were estimated by using the Thiessen polygons. Rainfall uncertainty was introduced through weight corruption by randomly picking up their value, at each time step, from a uniform distribution in the range ± 20% of the related uncorrupted value. The obtained weights were rescaled so that their cumulative sum is equal to one.

Synthetic hourly temperature data were generated by applying a fractionally differenced ARIMA model (FARIMA; see *Montanari et al.* [1997]). A mean areal value for temperature was obtained by rescaling the synthetic observations to the mean altitude of the basin area, by adopting a standard temperature gradient. Temperature data were not corrupted, in view of their limited uncertainty with respect to rainfall and river flow.

Synthetic river flow data were generated by using the previously generated synthetic rainfall and temperature records as input to the lumped rainfall-runoff model ADM [*Franchini*, 1996]. The ADM model is a nine-parameter lumped conceptual scheme that was calibrated against historical data obtaining a Nash efficiency [*Nash and Sutcliffe*, 1970] of

0.81 in validation (see *Montanari* [2005] for more details). Table 1 presents the model parameters. River flow data were corrupted by multiplying each observation by a coefficient that was picked up, at each time step, from a uniform distribution in the range 0.8−1.2. The coefficient of determination of the linear regression of corrupted versus uncorrupted river flow data is 0.86.

The observations included in the first 30 years of the synthetic record were used to calibrate a rainfall-runoff model that has reduced complexity with respect to ADM, therefore introducing model structural uncertainty (see below for a detailed description). Years from 31 to 40 were used to validate the model itself and to infer the probability distribution of the model error by using the meta-Gaussian approach by *Montanari and Brath* [2004]. The goodness-of-fit was checked by using the statistical tests described in *Montanari and Brath* [2004], which were satisfied over the whole range of river flows.

Finally, data for the years 41-50 of the observation period were used to test, in full validation mode, the proposed blueprint (rainfall-runoff modeling and uncertainty assessment).

The rainfall-runoff model we used for the Secchia River is HyMod, namely, the same 5-parameter lumped and conceptual rainfall-runoff model that was used by *Montanari* [2005]. HyMod was introduced by *Boyle* [2000] and extensively used thereafter. Model parameters are shown in Table 1. Evapotranspiration is accounted for by using the radiation method [*Doorembos et al.*, 1984]. With a total of only five parameters, HyMod can be considered an approach of reduced complexity with respect to ADM and therefore model structural uncertainty is introduced.

HyMod was calibrated by using DREAM [*Vrugt et al.*, 2007], in which a standard Gaussian likelihood function was used. DREAM is a modified SCEM-UA global optimisation algorithm [*Vrugt et al.*, 2003]. It makes use of population evolution like a genetic algorithm together with a selection rule to assess whether a candidate parameter set is to be retained. The sample of retained sets after convergence can be used to infer the probability distribution of model parameters. Herein, a number of 6000 parameter sets were retained, which indirectly determine the density function $f_{\Theta}(\Theta)$ of the parameter vector in a non-parametric empirical manner, fully respecting the dependencies between different parameters. HyMod explained about 81% and 82% of the river flow variance in calibration and validation, respectively, with the best DREAM parameter combination from the joint Markov chains. The values of the corresponding Nash efficiency are 0.81 - 0.82. These are feasible values in real world applications. Figure 3 reports a comparison over a 1500 hour window of the full validation period (years 41-50) between observed and simulated hydrographs.

6.2. The simulation procedure for the Secchia River

As we mentioned above, the simulation procedure refers to the years 41-50 of the observation period. HyMod was run 5000 times, by randomly picking up parameter sets from those retained by DREAM and accounting for input uncertainty by corrupting, for each simulation, the rainfall data as described in Section 6.1 (that is, by reproducing the same type of error that was introduced in the synthetic data set. Namely, a perfect uncertainty assessment for the rainfall data was assumed). A random outcome from the probability distribution of the model error was added to each observation, therefore obtaining, for each simulated river flow, a sample of 5000 points that allows to infer the related probability distribution.

Figure 4 shows the 95% prediction limits for the same 1500 hour window of the full validation period mentioned above, along with the corresponding observations. By looking at the overall prediction, one notes that 5.4% and 4.3% of the observations are located above the upper and below the lower limit, respectively, against theoretical values of 2.5% (at the 95% confidence level). These results indicate a slight underestimation of the band widths. Figure 5 shows a coverage probability plot (CPP), which gives information on the accuracy of the uncertainty estimation. A placement of the points along the 1:1 line is expected. For more details on drawing and interpreting the CPP plot (which is sometimes referred to as probability plot or Q-Q plot) see *Laio and Tamea* [2007]. For the present case, we note that Figure 5 confirms the underestimation of the band width, which nevertheless is scarcely significant in practice.

6.3. The case study of the Leo River

The catchment area of the Leo River basin at the closure section of Fanano is 64.4 km^2 and the main stream length is about 10 km. The maximum elevation in the catchment is the Mount Cimone (2165 m a.s.l.), which is the highest peak in the northern part of the Apennine Mountains. The climate is continental.

Daily river flow data at Fanano are available for the period January 1st, 2003 - October 26th, 2008, for a total of 2126 observations. For the same period, daily mean areal rainfall and temperature data over the catchment are available, as estimated by the Italian National Hydrographic Service based on observations collected in nearby stations.

The observations collected from January 1st 2003 to December 31st 2006 were used for calibrating the rainfall-runoff model, while the period January 1st 2007 - October 26th 2008 was reserved for its validation. We estimated the probability distribution of the model error by referring to the first year of the validation period (2007), in order to obtain an assessment of $f_e(e|\Theta, \mathrm{X})$ in a real world application. Finally, the period January 1st 2008 - October 26th 2008 was reserved for testing, in full validation mode, of the proposed blueprint (rainfall-runoff modeling and uncertainty assessment).

The rainfall-runoff model is AFFDEF [*Moretti and Montanari*, 2007], a spatially-distributed grid-based approach where hydrological processes are described with physically-based and conceptual equations. AFFDEF counts 8 calibrated parameters, which are described in Table 1. In order to limit the computational requirements, and in view of the limited catchment area, the Leo river basin was described by using only one grid cell, therefore applying a lumped representation. AFFDEF was calibrated by using DREAM [*Vrugt et al.*, 2007], again by using a standard Gaussian likelihood function. Herein, a number of 32000 parameter sets was retained. Figure 6 shows the probability density function for the model parameters. They all appear to be unimodal and well defined. Parameter values are in agreement with what one would expect from AFFDEF applications to similar catchments [*Moretti and Montanari*, 2007].

AFFDEF explained about 57% and 47% of the river flow variance in calibration and validation, respectively. The values of the corresponding Nash efficiency are 0.59 - 0.36. Figure 7 depicts a comparison during the validation period (2007 and 2008) between observed and simulated hydrographs. This latter was obtained by using the best DREAM parameter set. We can see that a significant uncertainty affects the model performance, which is unlikely merely due to lumping the model at catchment scale. We are interested in checking whether the proposed blueprint provides a consistent assessment of such uncertainty. The probability distribution of the model error was again inferred by using the meta-Gaussian approach, which provided a satisfactory fit for river flows greater than 0.5 m^3/s.

6.4. The simulation procedure for the Leo River

No information is available about input uncertainty. This is an important limitation in many practical applications. In particular, input uncertainty is usually dominant in real time flash-flood forecasting, where input rainfall to a rainfall-runoff model is usually predicted to increase the lead time of the river flow forecasting. If a probabilistic prediction for rainfall is available then input uncertainty can be efficiently taken into account in the blueprint proposed above. Alternatively, input uncertainty can be estimated using expert knowledge, Bayesian procedures like BATEA [*Kavetski et al.*, 2006] or conditional simulation methods [*Clark and Slater*, 2006; *Götzinger and Bárdossy*, 2008; *Renard et al.*, 2011]. For the present application, in absence of any information and similarly to *Vrugt et al.* [2008] and *Renard et al.* [2010], we introduced in each simulation and each data point a rainfall multiplier that was picked up from a Gaussian distribution with unit mean and standard deviation equal to 0.1.

A number $j = 5000$ of AFFDEF simulations were run during the 300-day full validation period January 1st 2008 - October 26, 2008. A random outcome from the probability distribution of the model error was added to each observation, therefore obtaining, for each simulated river flow, a sample of 5000 points.

Figure 8 shows the 95% prediction limits for the full validation period, along with the corresponding observations. The results confirm the relevant uncertainty that is anticipated by model performance. In fact, the prediction bands cover a large range of river flows. The observations located above the upper and below the lower limit are 7.3% and 9.1%, respectively, for river flows greater than 0.5 m^3/s. In this case also, the width of the prediction limits appears to be slightly underestimated. Figure 9 shows the CPP plot for the 300-day full validation period, for river flows greater than 0.5 m^3/s. The slight underestimation of the band width is confirmed.

It is interesting to inspect the reasons for the underestimation of the prediction limits width. In fact, one may note in Figure 8 that the lower prediction limit is satisfactorily estimated, with the only exception of the final part of the validation period where observations systematically fall sligthly outside the limit itself. On the contrary, the upper limit seems to be too large and too narrow for low and high flows, respectively. Further information can be gained by comparing the probability density functions of observed and simulated data. Given that such distribution, for the simulated data, depends on the magnitude of the model prediction according to the assumptions of the meta-Gaussian approach (see Section 5), the above comparison must be carried out by focusing on a restricted range of river flows, for which distribution changes are negligible. Figure 10 reports the result of the comparison for the validation period and the low flow range between 2 m^3/s (the observed mean) and 5 m^3/s. It can be seen that the overestimation of the upper limit is confirmed. To further inspect this issue, the comparison was also performed between the probability density functions of actual and simulated model error for the validation period and the same low flow range. Results are shown in Figure 11.

One can see that the variability of the model error is overestimated as well. Therefore, it appears that the unsatisfactorily assessment of the upper prediction limit for low flows is mainly due to inefficient representation of the statistical properties of the model error by the meta-Gaussian approach, which is induced by the limited extension of the calibration period (the year 2007 only) that makes statistical analysis and testing scarcely efficient. In practice, the data base is not extended enough for the method to recognize the variability of the band width depending on the river flow magnitude. Then, the method tends to predict constant band width thus resulting in overestimation and underestimation for low and high flows, respectively.

Other reasons for the lack of fit could be improper characterisation of input and parameter uncertainty as well as failure of the fulfilment of the underlying assumptions and in particular that of independence between the model error and parameter/data uncertainty that is adopted by the meta-Gaussian approach. In fact, the statistical properties of the model error were estimated by referring to the simulation obtained with the best DREAM parameter combination from the joint Markov chains. Actually, the error behaviors inferred in this way are not fully representative of suboptimal input data and parameter vectors that are picked up randomly in the simulation procedure which may induce larger errors for the given data set. To avoid this problem, two solutions can be used: (a) to estimate the model error by referring to a parameter and data set that provides average performances instead of the best ones. This approach is computationally efficient and therefore preferable when computing resources are a matter of concern. (b) To infer the statistical properties of the model error at each simulation step, therefore significantly increasing computational requirements.

We believe that performances like those we obtained in the two case studies above are sufficiently accurate for real world decision making in view of the consistency of the related data base. For other cases the most appropriate solution should be decided after considering the related practical needs.

7. Process-based stochastic modeling as a learning process

Uncertainty estimation allows one to quantitatively assess model reliability. If the model is process-based, the correctness of the underlying schematizations can be effectively checked by looking at the obtained prediction limits. In fact, these latter provide a comprehensive picture of the probability distribution of the prediction error, for a given confidence level and different river regimes. Therefore, prediction limits are a possible mean to check the correctness of our understanding of the hydrological processes at a given place. A closer look at the full probability distribution of the model error, for different flow regimes (an example that refers to low flows is presented in Figure 11) allows one to complete the information about model failures in different hydrological situations, therefore providing useful indications on possible adjustments at the model structure (for a recent discussion on model structural adequacy see *Gupta et al.* [2012]). In particular, the above distribution indicates when the model fails and the type of failure that is occurring, so that its impact can be evaluated. The prediction error should be assessed by also considering input and parameter uncertainty, to better understand whether the weakness is related to model structure rather than calibration information. Analysis of parameter uncertainty, as depicted by the parameter distributions shown in Figure 6, allows one to assess whether the parameters themselves are well defined and what is their impact on the results. A flat distribution may correspond to a poorly defined parameter and/or a scarce impact of the related process on the results. Such analysis is particularly useful when adopting a flexible model structure, to identify the relevant model components (for applications, see *Montanari et al.* [2006]; *Fenicia et al.* [2008]; *Schoups et al.* [2008]).

8. Conclusions and discussion

A blueprint is presented to introduce a novel methodological scheme for estimating the uncertainty of the output from a process-based (deterministic) hydrological model through the estimation of the related probability distribution. The scheme is obtained by developing a theoretical formulation to convert a deterministic hydrological model into a stochastic one, therefore incorporating randomness in hydrological modeling as a fundamental component. The scheme shows that to include an arbitrarily complex deterministic model within a stochastic framework, where randomness is a fundamental part of the system, is in principle possible. Although we explicitly focused on process-based approaches, the blueprint that we are proposing is applicable to any deterministic scheme. The relevant feature of the approach herein proposed, which can be applied to models of arbitrary complexity, is that model likelihood computation can be avoided. In fact, the approach proposed replaces the single output of a deterministic model with the probability distribution thereof which is estimated by stochastic simulation. A comprehensive discussion of the underlying simplifying assumptions and how they can be removed is presented, therefore allowing to structure modeling in a objective setting. The proposed method allows hydrological modeling and uncertainty assessment to be jointly carried out.

Two applications are presented for illustrating the introduced blueprint. One of them makes use of synthetic data. Although simplifying assumptions are introduced to reduce the computational effort, the case studies show that the proposed approach is efficient and provides a consistent uncertainty assessment. However, the results show that the opportunity of removing some of the above assumptions should be considered depending on the user needs.

We believe the theoretical framework introduced here may open new perspectives regarding modeling of uncertain hydrological systems. In fact, analyzing randomness within process-based system representations is an invaluable opportunity to improve system understanding therefore increasing predictability, according to the "models of everywhere" concept [*Beven*, 2007]. In particular, it is possible to analyse (possibly) changing or shifting behaviors and their reaction to (human induced) changes. Moreover, we believe that the proposed procedure is very useful for educational purposes, putting the basis for developing a unified theoretical basis for uncertainty assessment in hydrology.

A successful application of the proposed blueprint requires a reliable estimation of input, parameter and model uncertainty. The latter is obtained through the estimation of the probability density $f_e(e)$ of the model error. The meta-Gaussian model [*Montanari and Brath*, 2004; *Montanari and Grossi*, 2008] was herein used. However, in condition of data scarcity it may be scarcely efficient, as we show in Section 6.4. Data assimilation methods can also be considered, like machine learning and nearest neighbor techniques [*Shrestha and Solomatine*, 2006]. All the above methods rely on limiting assumptions and some of them are also computer intensive. We believe that estimating model uncertainty in hydrology is still a difficult problem for which more focused research is needed [*Montanari*, 2011]. The proposed framework may facilitate streamlining of this research and linking it with other components within an holistic modeling approach.

Finally, as mentioned above, estimation of parameter and input uncertainty is a relevant challenge as well which has an impact on model prediction. Possibilities are the GLUE method [*Beven and Binley*, 1992] and the DREAM algorithm [*Vrugt et al.*, 2007] for parameter uncertainty, which nevertheless are computer intensive as well and may turn out to be impractical with spatially-distributed models applied to fine time scale at large catchments. Information on observation error, and the related probability distribution, can be used to estimate input uncertainty. Additional and focused research is needed to improve the above techniques therefore ensuring a more practical application of the framework herein proposed.

Acknowledgments. To authors are grateful to Hessel Winsemius, Stacey Archfield, Keith Beven, Nataliya Bulygina, Martyn Clark, Hoshin Gupta, Francesco Laio, Elena Montosi, Jasper Vrugt, and two other anonymous referees who provided very useful comments and help. A.M. was partially supported by the Italian government through the grant "Uncertainty estimation for precipitation and river discharge data. Effects on water resources planning and flood risk management".

References

Abbot, M. B., J. C. Bathurst, J. A. Cunge, P. E. O'Connell, and J. Rasmussen (1986), An introduction to the European Hydrological System - Systme Hydrologique Europen, SHE. 2. Structure of a physically-based, distributed modeling system., *J. Hydrol.*, *102*, 87–77.

Beven, K. J. (1989), Changing ideas in hydrology The case of physically-based models, *J. Hydrol.*, *105*, 157–172.

Beven, K. J., and A. M. Binley (1992), The future of distributed models: Model calibration and uncertainty prediction, *Hydrol. Proc.*, *6*, 279–298.

Beven, K. J. (2001), How far can we go in distributed hydrological modeling, *Hydrol. and Earth System Sci.*, *5*, 1–12.

Beven, K. J. (2002), Towards an alternative blueprint for a physically based digitally simulated hydrologic response modeling system, *Hydrol. Proc.*, *16*, 189–206.

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

Beven, K. J. (2007), Towards integrated environmental models of everywhere: uncertainty, data and modelling as a learning process, *Hydrol. and Earth System Sci.*, *11*, 460–467.

Beven, K. J. (2009), *Environmental modelling: an uncertain future? An introduction to techniques for uncertainty estimation in environmental prediction*, Routledge.

Beven, K. J., P. J. Smith, and A. Wood (2011), On the colour and spin of epistemic error (and what we might do about it), *Hydrol. Earth Syst. Sci. Discuss.*, *8*, 5355–5386.

Beven, K. J., and I. Westerberg (2011), On red herrings and real herrings: disinformation and information in hydrological inference, *Hydrol. Proc.*, *10*, 1676–1680.

Box, G. E. P., and G. C. Tiao (1973), *Bayesian Inference in Statistical Analysis*, Addis-onWesley, Boston, USA.

Boyle, D. P. (2000), Multicriteria calibration of hydrological models, *Ph.D. Dissertation*, Dep. of Hydrol. and Water Resour., Univ. of Ariz.

Brockwell, P. J., and R. A. Davis (1973), *Time series: theory and methods*, Springer, New York, USA.

Bulygina, N., and H. Gupta (2009), Estimating the uncertain mathematical structure of a water balance model via Bayesian data assimilation, *Water Resour. Res.*, *45*, doi:10.1029/2007WR006749.

Bulygina, N., and H. Gupta (2010), How Bayesian data assimilation can be used to estimate the mathematical structure of a model, *Stoch. Environ. Res. Risk Assess.*, *24*, 925–937, doi:10.1007/s00477-010-0387-y.

Bulygina, N., and H. Gupta (2011), Correcting the mathematical structure of a hydrological model via Bayesian data assimilation, *Water Resour. Res.*, *47*, doi:10.1029/2010WR009614.

Casella, G., and R. L. Berger (1990), *Statistical inference*, Brooks/Cole Pub., Pacific Grove, USA.

Clark, M. P., D. Kavetski, and F. Fenicia (2011), Pursuing the method of multiple working hypotheses for hydrological modeling, *Water Resour. Res.*, *47*, doi:10.1029/2010WR009827.

Clark, M. P., and A. G. Slater (2006), Probabilistic Quantitative Precipitation Estimation in Complex Terrain, *J. Hydrometeor.*, *7*, 3-22, doi: http://dx.doi.org/10.1175/JHM474.1

Cowpertwait, P. S. P. (1995), A generalized spatial-temporal model of rainfall based on a clustered point process, *Proc. of the R. Soc. of London*, Series A, *450*, 163175.

Di Baldassarre, G., and A. Montanari (2009), Uncertainty in river discharge observations: A quantitative analysis, *Hydrol. and Earth Sys. Sci.*, *1*3, 913–921.

Doorembos, J., W. O. Pruitt, A. Aboukhaled, J. Damagnez, N. G. Dastane, C. van der Berg, P. E. Rijtema, O. M. Ashford, and M. Frere, (1984), Guidelines for predicting crop water requirements, *FAO Irrig. Drainage Pap.*.

Ebtehaj, M., H. Moradkhani, and H.V. Gupta (2010), Improving robustness of hydrologic parameter estimation by the use of moving block bootstrap resampling, *Water Resour. Res.*, *46*, W07515, doi:10.1029/2009WR007981.

Fenicia, F., H. H. G. Savenije, and L. Pfister, (2008), Understanding catchment behavior through stepwise model concept improvement, *Water Resour. Res.*, *44*, doi:10.1029/2006WR005563.

Franchini, M. (1996), Use of a genetic algorithm combined with a local search method for the automatic calibration of conceptual rainfall runoff models, *Hydrol. Sci. J.*, *41*, 2139.

Freeze, R. A., and R. L. Harlan (2008), Blueprint for a physically-based, digitally- simulated hydrologic response model, *J. Hydrol.*, *9*, 237–258.

Grayson, R. B., I. D. Moore, and T. A. McMahon, (1992), Physically-based hydrologic modeling. 2. Is the concept realistic?, *Water Resour. Res.*, *28*, 2659–2666.

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

Gupta, H.V., M. P. Clark, J. A. Vrugt, G. Abramowitz, and M. Ye (2012), Towards a Comprehensive Assessment of Model Structural Adequacy, *Water Resour. Res.*, in press.

Koutsoyiannis, D. (2006), Nonstationarity versus scaling in hydrology, *J. of Hydrol.*, *324*, 239–254.

Koutsoyiannis, D., C. Makropoulos, A. Langousis, S. Baki, A. Efstratiadis, A. Christofides, G. Karavokiros, and N. Mamassis (2009), HESS Opinions: "Climate, hydrology, energy, water: recognizing uncertainty and seeking sustainability", *Hydrol. Earth Syst. Sci.*, *13*, 247–257.

Koutsoyiannis, D. (2010), HESS Opinions “A random walk on water”, *Hydrol. Earth Syst. Sci.*, *14*, 585–601.

Koutsoyiannis, D. (2011), Hurst-Kolmogorov dynamics and uncertainty, *Journal of the American Water Resources Association*, *47*, 481–495.

Krzysztofowicz, R. (2002), Bayesian system for probabilistic river stage forecasting, J. of Hydrol., *268*, *16–40*.

Krueger, T., J. Freer, J. N. Quinton, C. J. A. Macleod, G. S. Bilotta, R. E. Brazier, P. Butler, and P. M. Haygarth (2010), Ensemble evaluation of hydrological model hypotheses, *Water Resour. Res.*, *46*, doi:10.1029/2009WR007845.

Kavetski D., G. Kuczera, and S.W. Franks (2006), Bayesian analysis of input uncertainty in hydrological modeling: 1. Theory, *Water Resour. Res.*, *42*, doi:10.1029/2005WR004368.

Kundzewicz, Z. W. (2007), Predictions in Ungauged Basins: PUB Kick-off, Proceedings of the PUB Kick-off meeting held in Brasilia, 2022 November 2002, *IAHS Publ.* 309, 38–44.

Laio, F., and S. Tamea (2007), Verification tools for probabilistic forecasts of continuous hydrological variables, *Hydrol. Earth Syst. Sci.*, *11*, 1267–1277.

Lasota, D. A., and M. C. Mackey (1985), Probabilistic properties of deterministic systems, Cambridge University Press.

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

Liu, Y., and H. V. Gupta (2007), Uncertainty in hydrologic modeling: Toward an integrated data assimilation framework, *Water Resour. Res.*, *43*, doi:10.1029/2006WR005756.

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

Merz, R., and G. Blöschl (2008a), Rainfall uncertainty in hydrological modelling: an evaluation of multiplicative error models, *J. of Hydrol.*, *400*, 83–94.

Merz, R., and G. Blöschl (2008a), Flood frequency hydrology: 2. Combining data evidence, *Water Resour. Res.*, *44*, doi:10.1029/2007WR006745.

Montanari, A., R. Rosso, and M. S. Taqqu (1997), Fractionally differenced ARIMA models applied to hydrologic time series: identification, estimation and simulation, *Water Resour. Res.*, *33*, 10351044.

Montanari, A., and A. Brath (2004), A stochastic approach for assessing the uncertainty of rainfall-runoff simulations, *Water Resour. Res.*, *40*, doi:10.1029/2003WR002540.

Montanari, A., (2005), Large sample behaviors of the generalized likelihood uncertainty estimation (GLUE) in assessing the uncertainty of rainfall-runoff simulations, *Water Resour. Res.*, *41*, doi:10.1029/2004WR003826.

Montanari, A., What do we mean by uncertainty? The need for a consistent wording about uncertainty assessment in hydrology (2007), *Hydrol. Proc.*, *2*1, 841–845.

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

Montanari, A. , C. A. Shoemaker, and N. van de Giesen (2009), Introduction to special section on Uncertainty Assessment in Surface and Subsurface Hydrology: An overview of issues and challenges, *Water Resour. Res.*, *45*, doi:10.1029/2009WR008471.

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

Montanari, A. and G. Di Baldassarre (2012), Data errors and hydrological modelling: the role of model structure to propagate observation uncertainty, submitted to *Adv. Water Resour.*.

Montanari, L., M. Sivapalan, and A. Montanari (2006), Investigation of dominant hydrological processes in a tropical catchment in a monsoonal climate via the downward approach, *Hydrol. and Earth Sys. Sci.*, *1*0, 769–782.

Moretti., G., and A. Montanari (2007), AFFDEF: A spatially distributed grid based rainfall-runoff model for continuous time simulations of river discharge, *Environ. Mod. & Soft.*, *22*, 823–836.

Nash, J. E., and J. V. Sutcliffe, River flow forecasting through conceptual models 1: a discussion of principles (1970), *J. Hydrol.*, *10*, 282290.

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

Pelletier, M. P. (1987), Uncertainties in the determination of river discharge: a literature review, *Can. J. Civil Eng.*, *41*, 834-850.

Pianosi, F., and L. Raso (2012), Dynamic modeling of predictive uncertainty by regression on absolute errors, *Water Resour. Res.*, *48*, doi:10.1029/2011WR010603.

Refsgaard, J. C., J. P. van Der Sluijs, J. Brown, and P. van der Keur (2006), A framework for dealing with uncertainty due to model structure error, *Adv. Water Resour.*, *29*, 1586–1597.

Reggiani, P., M. Sivapalan and S. M. Hassanizadeh (1998), A unifying framework for watershed thermodynamics: balance equations for mass, momentum, energy and entropy and the second law of thermodynamics, *Adv. Water Resour.*, *22*, 367–398.

Reggiani, P., S. M. Hassanizadeh, M. Sivapalan and W. G. Gray (1999), A unifying framework for watershed thermodynamics: constitutive relationships, *Adv. Water Resour.*, *23*, 15–39.

Renard, B., D. Kavetski, G. Kuczera, M. Thyer, and S.W. Franks (2010), Understanding predictive uncertainty in hydrological modeling: The challenge of identifying input and structural errors, *Water Resour. Res.*, *46*, doi:10.1029/2009WR008328.

Renard, B., D. Kavetski, E. Leblois, M. Thyer, G. Kuczera, and S. W. Franks (2011), Toward a reliable decomposition of predictive uncertainty in hydrological modeling: Characterizing rainfall errors using conditional simulation, *Water Resour. Res.*, *47*, doi:10.1029/2011WR010643.

Savenije, H. H. G. (2009), HESS Opinions "The art of hydrology", *Hydrol. Earth Syst. Sci.*, *13*, 157–161.

Schoups, G., N.C. van de Giesen, and H.H.G. Savenije (2008), Model complexity control for hydrologic prediction, *Water Resour. Res.*, *44*, doi:10.1029/2008WR006836.

Schoups, G., and J. A. Vrugt (2010), A formal likelihood function for parameter and predictive inference of hydrologic models with correlated, heteroscedastic, and nonGaussian errors, *Water Resour. Res.*, *46*, doi:10.1029/2009WR008933.

Shrestha, D. L., and D. P. Solomatine (2006), Machine learning approaches for estimation of prediction interval for the model output, *Neural Networks*, *19*, 225–235.

Soil Conservation Service (1972), *National Engineering Handbook*, section 4, Hydrology, U.S. Dept. of Agriculture, Washington D.C.

Sivapalan, M., K. Takeuchi, S.W. Franks, V.K. Gupta, H. Karambiri, V. Lakshmi, X. Liang, J.J. McDonnell, E.M. Mendiondo, P.E. O'Connell, T. Oki, J.W. Pomeroy, D. Schertzer, S. Uhlenbrook, and E. Zehe (2003), IAHS Decade on Predictions in Ungauged Basins (PUB), 2003-2012: Shaping an exciting future for the hydrological sciences, *Hydrol. Sci. J.*, *48*, 857–880.

Srikanthan, R, G. Kuczera, and M. Thyer (2009), 'Calibrate It Twice': A Simple Resampling Method for Incorporating Parameter Uncertainty in Stochastic Data Generation. In: Proceedings of H2O 09: 32nd Hydrology and Water Resources Symposium, Newcastle: Adapting to Change. Barton, A.C.T.: Engineers Australia, 2009: 1028-1039. Availability: http://search.informit.com.au/documentSummary;dn=757766656276837;res=IELENG, ISBN 97808258259461

Stedinger, J. R., Vogel, R. M., Lee, S. U., and R. Batchelder (2008), Appraisal of the generalized likelihood uncertainty estimation (GLUE) method, *Water Resour. Res.*, *44*, doi:10.1029/2008WR006822.

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

Vrugt J. A., H. V. Gupta, W. Bouten, and S. Sorooshian (2003), A Shuffled Complex Evolution Metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters, *Water Resour. Res.*, *39*, doi:10.1029/2002WR001642.

Vrugt, J. A., and B. A. Robinson (2007), Improved evolutionary optimization from genetically adaptive multimethod search, *Proc. of the Natl. Acad. of Sci. of USA*, *1*04, 708–711.

Vrugt, J., C. ter Braak, M. Clark, J. Hyman, and B. Robinson (2008), Treatment of input uncertainty in hydrologic modeling: Doing hydrology backward with Markov chain Monte Carlo simulation, Water Resour. Res., 44, W00B09, doi:10.1029/2007WR006720.

Winsemius, H. C., B. Schaefli, A. Montanari, and H. H. G. Savenije (2009), On the calibration of hydrological models in ungauged basins: A framework for integrating hard and soft hydrological information, *Water Resour. Res.*, *45*, doi: 10.1029/2009WR007706, 2009.

Ye, M., P. D. Meyer, and S. P. Neuman (2008), On model selection criteria in multimodel analysis, *Water Resour. Res.*, *44*, doi:10.1029/2008WR006803.

Table 1. Parameters of the ADM, HyMod and AFFDEF rainfall-runoff models. Symbols are given for the parameters of AFFDEF which refer to Figure 6.

Parameter	Unit	ADM	HyMod	AFFDEF	Symbol
Maximum soil storage capacity	[cm]	X	X		
Shape parameter of the storage capacity curve	[-]	X	X		
Surface/subsurface flow partition factor	[-]		X		
Residence time quick flow reservoir	[h]		X		
Residence time low flow reservoir	[h]	X	X	X	K
Shape parameter of the drainage curve	[-]	X			
Shape parameter of the percolation curve	[-]	X			
Maximum drainage rate	[cm/s]	X			
Maximum percolation rate	[cm/s]	X			
Convectivity	[cm/s]	X			
Diffusivity	[cm^2/s]	X			
Multiplying factor for soil storativity	[-]			X	H
Multiplying factor for interception storage	[-]			X	C_{int}
Residence time soil water	[h]			X	H_s
Threshold temperature for snow accumulation	[°C]			X	T_s
Threshold temperature for snow melting	[°C]			X	T_{melt}
Snow conversion factor	[−]			X	SCF
Melting factor	[mm/(°C d)]			X	M_f

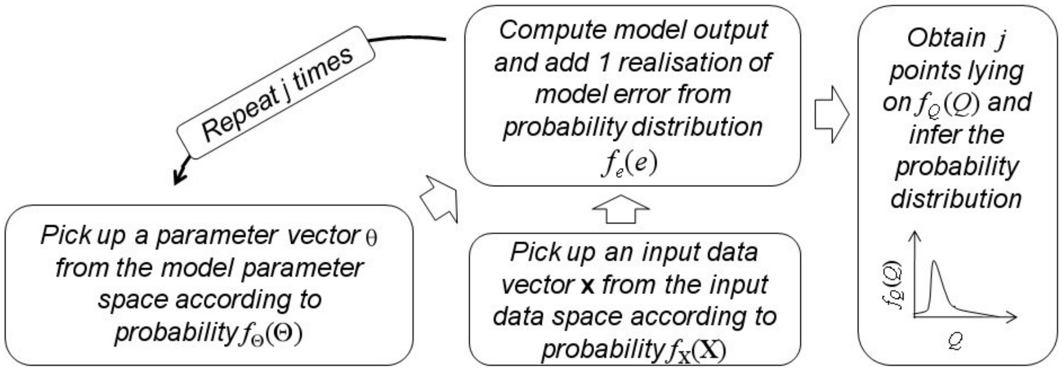

Figure 1. Flowchart of the Monte Carlo simulation procedure for performing the numerical integration in eq. (8) and (9).

Figure 2. Location of the case study basins (Italy). A and B indicate the positions of Bacchello Bridge and Fanano, respectively.

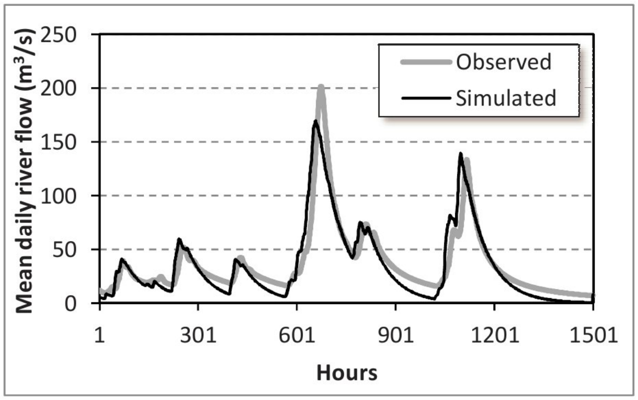

Figure 3. Case study of Secchia River. Comparison between observed and simulated hydrographs during a 1500-hour window included in the full validation period (years 41-50 of the synthetic record). The simulated hydrograph was obtained by using the best DREAM parameter set during the calibration period.

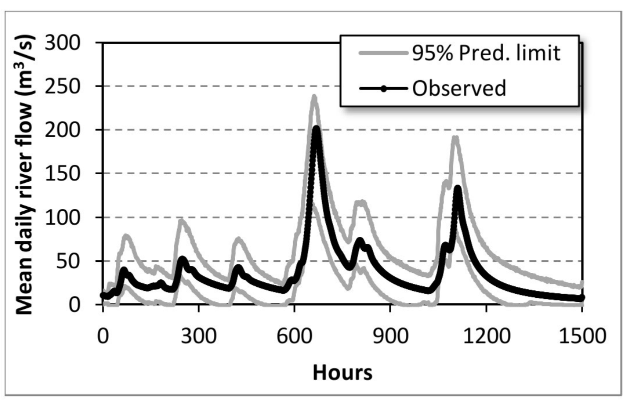

Figure 4. Case study of Secchia River. 95% prediction limits provided by HyMod during a 1500-hour window of the full validation period (years 41-50 of the synthetic record).

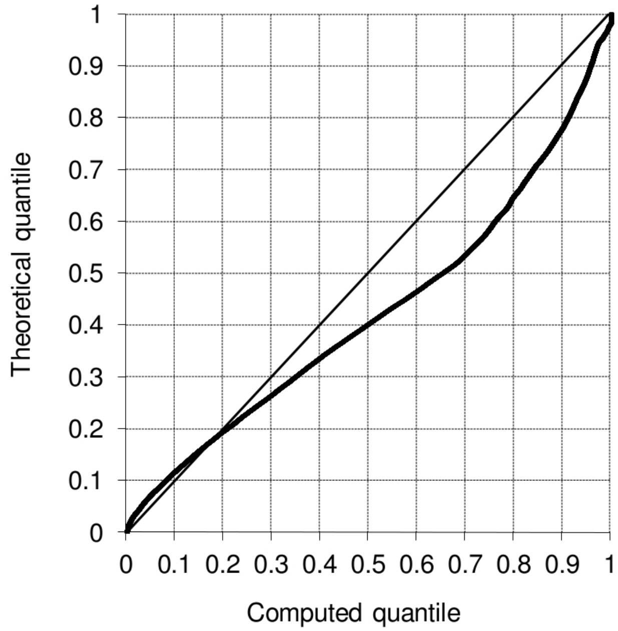

Figure 5. Case study of Secchia River. CPP plot of the prediction provided by HyMod during the full validation period (years 41-50 of the synthetic record).

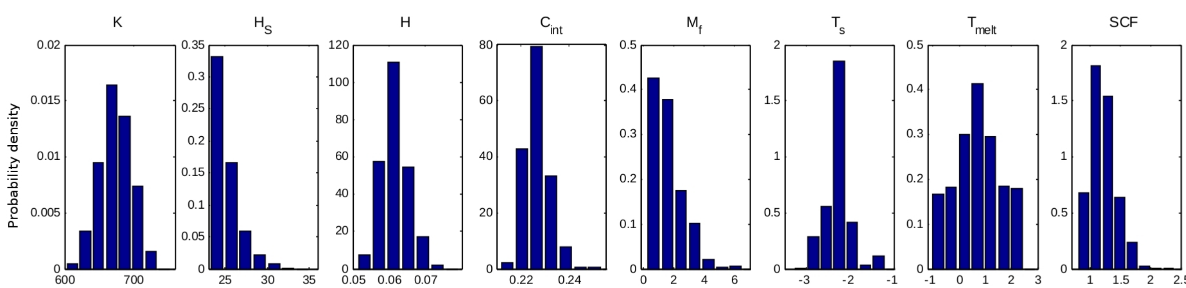

Figure 6. Case study of Leo River. Probability density functions for the AFFDEF parameters obtained with DREAM. Symbol meanings are given in Table 1.

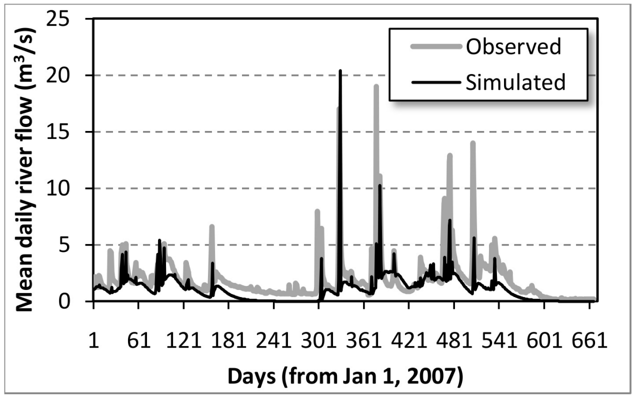

Figure 7. Case study of Leo River. Comparison between observed and simulated hydrographs during the validation period (Jan 1st 2007 - October 26, 2008). The simulated hydrograph was obtained by using the best DREAM parameter set during the calibration period.

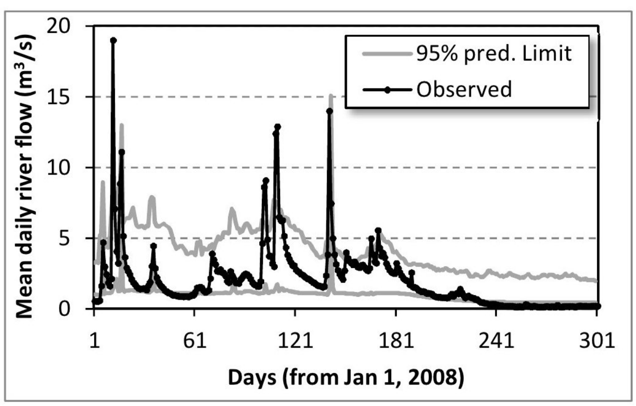

Figure 8. Case study of Leo River. 95% prediction limits provided by AFFDEF during the full validation period (Jan 1st, 2008 - Oct 26th, 2008), along with the corresponding observed values.

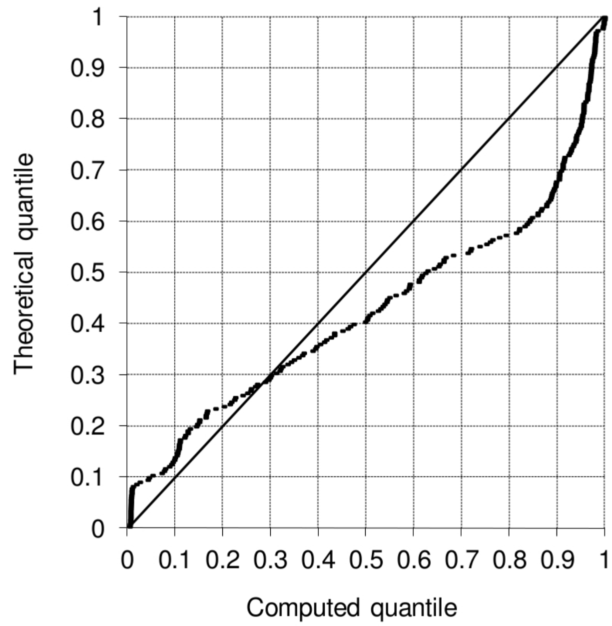

Figure 9. Case study of Leo River. CPP plot of the prediction provided by AFFDEF during the full validation period (Jan 1st, 2008, Oct 26th, 2008).

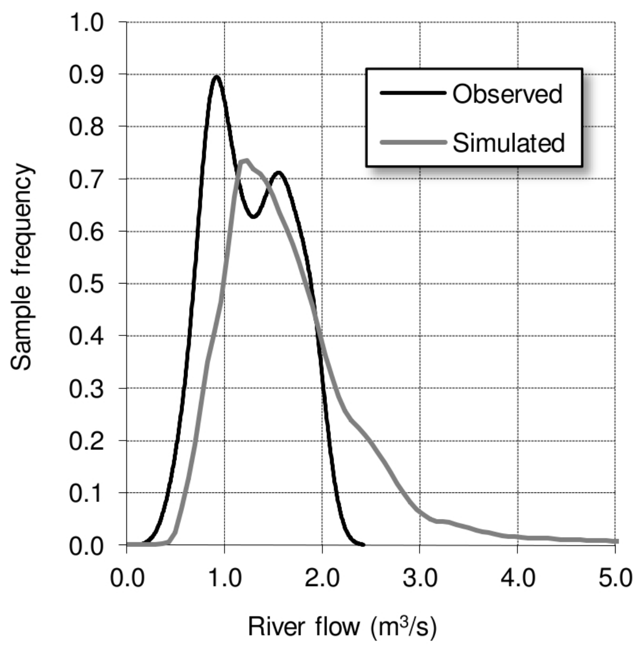

Figure 10. Case study of Leo River. Comparison between the probability density functions of observed and simulated data during the full validation period (Jan 1st, 2008, Oct 26th, 2008) for the river flow range between 2 m^3/s and 5 m^3/s.

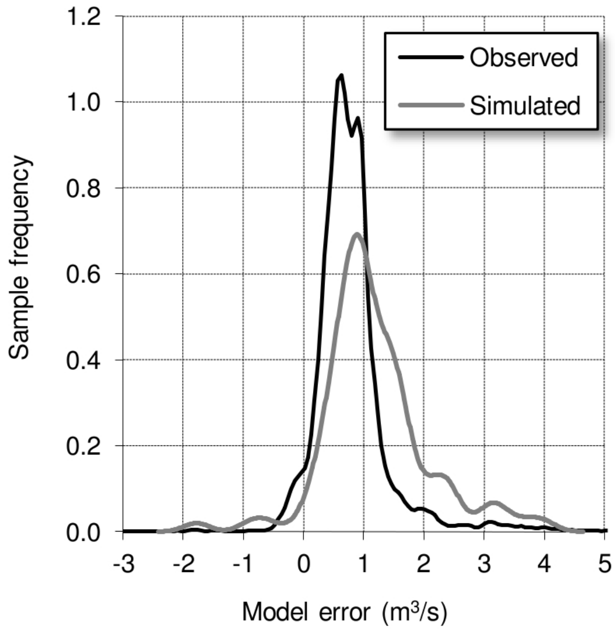

Figure 11. Case study of Leo River. Comparison between the probability density functions of actual and simulated model error during the full validation period (Jan 1st, 2008, Oct 26th, 2008) for the river flow range between 2 m^3/s and 5 m^3/s.