

---

# Hydrological catchment modelling: past, present and future

E. Todini

University of Bologna, Italy

Email: [todini@geomin.unibo.it](mailto:todini@geomin.unibo.it)

---

## Abstract

This paper discusses basic issues in hydrological modelling and flood forecasting, ranging from the roles of physically-based and data-driven rainfall runoff models, to the concepts of predictive uncertainty and equifinality and their implications. The evolution of a wide range of hydrological catchment models employing the physically meaningful and data-driven approaches introduces the need for objective test beds or benchmarks to assess the merits of the different models in reconciling the alternative approaches. In addition, the paper analyses uncertainty in models and predictions by clarifying the meaning of uncertainty, by distinguishing between parameter and predictive uncertainty and by demonstrating how the concept of equifinality must be addressed by appropriate and robust inference approaches. Finally, the importance of predictive uncertainty in the decision making process is highlighted together with possible approaches aimed at overcoming the diffidence of end-users.

**Keywords:** catchment models, physically-based models, data-driven input-output models, equifinality, predictive uncertainty

## Introduction

The history of hydrological modelling ranges from the Rational Method (Mulvaney, 1850) to recent distributed physically-meaningful models (Abbott *et al.*, 1986a,b; Wigmosta *et al.*, 1994; Refsgaard and Storm, 1995; Ewen *et al.*, 2000; Kouwen, 2000; De Roo *et al.*, 1998, 2000; Liu and Todini, 2002; Vivoni, 2003). Over the same period, starting from the simple Unit Hydrograph (Sherman, 1932), input-output models, now called data-driven models, have evolved into ANN models (Garcia-Bartual, 2002) and Data Based Mechanistic (DBM) models (Young, 2001, 2002). From the wide range of models available, the choice of the one most appropriate for any specific task is difficult, particularly as each modeller tends to promote the merits of his/her own approach. Moreover, apart from the intercomparison of conceptual models conducted in the seventies, (WMO, 1975), no objective comparisons using benchmarks, or test beds using standard data sets, have been proposed or effected.

Hydrological models serve many purposes, one of the most important applications being flood forecasting in which uncertainty plays a major role. Unfortunately, the implications of using uncertainty in the decision-making

process and even the concept of uncertainty seem to deter hydrologists from addressing the problem. Indeed, many hydrologists do not appear to appreciate the need to quantify predictive uncertainty and tend to describe the uncertainty implicit in the model rather than in possible future values of the quantity to be predicted or forecast, conditional upon the output of the model (Krzysztofowicz, 1999). To this confusion is added a lack of appreciation of the difference between parameter uncertainty (and the possible need to estimate the parameter values) and predictive uncertainty. Moreover, the introduction of the ‘equifinality’ principle (Beven and Binley, 1992) has not served to clarify this distinction.. Hence, in delineating possible future lines of research, the uncertainty problem must be viewed in the correct perspective.

After a brief history of the development of hydrological models, this paper proposes a reconciliation of the different approaches, by developing objective benchmarks or test-beds relevant to the many diverse potential applications. The problem of uncertainty in hydrological models is addressed as well as possible approaches to describing a real-world case study.

## Brief history of hydrological models

### FROM EMPIRICAL TO PHYSICALLY MEANINGFUL MODELS

The Rational Method proposed by Mulvaney (1850) is a clear exposition of the concept of time of concentration and its relation to the maximum runoff; it estimates peak flow but not flood volume and is physically meaningful only in small impervious catchments in which flow is effectively a purely kinematic process. Applications of the method to the design of sewers appeared in the literature from the end of the 19<sup>th</sup> century (Kuichling, 1889; Lloyd-Davies, 1906).

The concept of the unit hydrograph (UH) (Sherman, 1932) was based on the principle of superposition of effects; it enabled the complete flood hydrograph to be predicted from rainfall sampled at constant intervals. With the introduction of systems theory, the unit hydrograph was then interpreted as the response of a linear, causative, dynamic stationary system and two forms of the unit hydrograph were then considered. The first one, the continuous time impulse response of a linear system, is known in hydrology as the instantaneous unit hydrograph (IUH) and the second one, the response to a time discretised input, is known as the finite period unit hydrograph (TUH) (O'Donnell, 1966). Indeed, the introduction of the IUH can be viewed as the starting point that led to the separation of physically meaningful and data driven models. If the 'shape' of the IUH is defined *a priori* by the modeller as the integral solution, a set of linear or linearised differential equations and the parameter values are estimated not from the input-output historical data but computed as a function of the physical characteristic quantities of the phenomenon, then the IUH is a physical interpretation of the phenomenon. Examples can be found easily in flood routing models. For instance, Kalinin and Milyukov (1957) demonstrated that, by linearising the unsteady flow equations, the integral solution is a Gamma density function, namely a Nash cascade (1958, 1960) with parameters  $n$  and  $k$ , where the parameter  $n$  is now extended to the domain of real numbers, which can be expressed in terms of the Froude number, the bed slope, the velocity, etc. (Dooge, 1973). Furthermore, Hayami (1951) showed how to derive an IUH from the linear diffusion equation, while Todini and Bossi (1986) derived a TUH from the linear parabolic approximation of the unsteady flow equations, with the two parameters, celerity and diffusivity, which are recomputed at each integration time interval in terms of the hydrodynamic characteristics of the reach (discharge, the friction slope, etc.).

However, if the shape of the IUH/TUH cannot be defined *a priori* on physical grounds, both the shape and the relevant parameters must be derived from the measurements, so

clearly, the result is a data-driven model (Natale and Todini, 1976a,b).

As well as difficulties of physical interpretation, the extension of the IUH/TUH approach to larger, not necessarily impervious, catchments presented problems requiring subjective choices, such as:

- separation of storm runoff from base flow;
- the determination of 'effective' rainfall, namely that portion of the rainfall that is not lost through replenishing soil moisture etc;
- the actual derivation of the IUH/TUH shape and/or of the IUH/TUH parameters from the measurements available.

To overcome these problems, research into non-linear or threshold-type systems led to representations based on:

- (i) Volterra integrals of an order greater than the first,
- (ii) orthogonal polynomials (Amorocho and Orlob, 1961) or
- (iii) piecewise linearisations (Todini and Wallis, 1977; Todini, 2002b), reproducing the consequences of threshold effects introduced by soil saturation.

To achieve a better physical interpretation of catchment response, the 1960s saw the development of models in which individual components in the hydrological cycle were represented by interconnected conceptual elements; each of these represented, in the hydrological model, the response of a particular subsystem: Dawdy and O'Donnell, 1965, Crawford and Linsley, 1966 – Stanford Watershed IV; Burnash *et al.*, 1973 – Sacramento; Rockwood, 1964 – SSARR; Sugawara, 1967, 1995 – Tank, etc. All these models represented in different ways the responses of, and the interconnections between, the various sub-systems; at the time, they were regarded as the very best that could be achieved with the then current data and computational resources. At that time the modellers strongly believed that the parameters of their models, such as the storage coefficients, roughness coefficients or the different thresholds, were physical entities which could be inferred from the physiographic characteristics of the catchments. Due to the need to obviate a time-consuming trial and error approach in parameterising these models, model parameter optimisation was introduced (Dawdy and O'Donnell, 1965). As a result, when the estimates were made on the basis of objective functions to be minimised (for example, the sum of squares criterion), the resulting parameter values were generally unrealistic, perhaps because they incorporated errors of measurements as well as those of the model itself.

Also, the conditions of observability (the need for sufficient information content in the data to determine the parameter values) were not always guaranteed, particularly for multiple input–output hydrological models (Gupta and Sorooshian, 1983; Sorooshian and Gupta, 1983; Singh and Woolhiser, 2002). In essence, these models became data-driven.

At the end of the 1970s, a new type of lumped models was introduced, based on the idea that the rainfall runoff process is mainly dominated by the dynamics of saturated areas, which can be related to the soil moisture storage using a simple monotone function, thus leading to the variable contributing area models. These models generally employed the Dunne assumption that all precipitation enters the soil and that surface runoff originates by saturation of the upper soil layer. These variable contributing area models, the Xinanjiang (Zhao, 1977) and the Probability Distribution (PDM) (Moore and Clarke, 1981) were characterised by few significant parameters: although expressing the physical concepts of continuity of mass they were still not entirely meaningful in their dynamics. Thereafter, Juemou *et al.* (1987) combined the Xinanjiang soil moisture distribution function with the Constrained Linear Systems (CLS) model (Natale and Todini, 1976a,b; Todini and Wallis, 1977; Todini, 2002b) into the Synthesized Constrained Linear Systems model (SCLS). Later, by modifying the Xinanjiang soil moisture distribution function, Todini (1996, 2002a) developed the ARNO model, from which Wood *et al.* (1992) derived the VIC model by increasing the number of soil layers (Liang *et al.*, 1994, 1996a,b). The core of all these models is a two parameter distribution function curve representing the relation between the total volume of water stored in the soil and the extension of the saturated areas. Unfortunately, the parameterisation of this curve, as well as of the other processes represented (drainage, percolation, groundwater flow, etc), was based on empirical parameters to be estimated from the data. Beven and Kirkby (1979) originated a more physically-meaningful distribution function model, TOPMODEL, based on the distribution function of a topographic index. This assumes that the accumulation of soil moisture can be approximated by successive steady states of the water table originating in the upper soil layer. They derived a new relation between the volume of water stored in the soil and the extent of saturated areas (the topographic index distribution function) on the basis of physically-meaningful parameters. Unfortunately this proved to be true only for very small hill-slope catchments represented with extremely fine meshes (Franchini *et al.*, 1996).

In a further step towards a physical representation of the rainfall-runoff process, Wooding (1965a,b, 1966), and Woolhiser and Liggett (1967) used kinematic models for

the study of small urban basins, while Freeze and Harlan (1969) proposed, albeit only as a future project, the creation of a mathematical model based on distributed physical knowledge of surface and subsurface phenomena. By numerical integration of the coupled sub-systems of partial differential equations describing surface flow and flow in the unsaturated and saturated zones, and by matching the solutions of each sub-system with the boundary conditions of another, catchment scale predictions could be produced. This concept was developed into SHE (Système Hydrologique Européen), by the Danish Hydraulic Institute (DK), the Institute of Hydrology at Wallingford (UK) and SOGREAH (France) (Abbott *et al.*, 1986a,b). SHE has since evolved into a robust physically-based model, available as MIKE-SHE (Refsgaard and Storm, 1995) and SHETRAN (Ewen *et al.*, 2000). The limitation to its practical use is the large requirement for data and computational time which restrict its use to small, extensively instrumented catchments.

More recently, the wider availability of distributed information, ranging from soil types and land use to radar rainfall, have facilitated the production of simplified physically-meaningful distributed hydrological models. These models, based on simplifying assumptions, with simpler and more parsimonious parameterisations than those employed in MIKE-SHE and SHETRAN, can be applied successfully to flood forecasting. Such models are: WATFLOOD (Kouwen, 2000), DHSVM (Wigmosta *et al.*, 1994), TOPKAPI (Todini, 1995; Todini and Ciarapica, 2002; Liu and Todini, 2002), FEWS NET

hydrological behaviour of a catchment, although with a larger loss of physical interpretation. This loss of physicality increased further with Artificial Neural Network (ANN) approaches, which can be viewed as non-linear analogues of the original linear transfer function models; unfortunately, forecasts may be poor when the events are larger than those in the training set (Cameron *et al.*, 2002, Gaume and Gosset, 2003). Although Dawson and Wilby (2001) and Shamseldin (1997) review applications of ANN to rainfall-runoff modelling, few operational forecasting systems are presently based on ANN (Garcia-Bartual, 2002); as already noted, outside the range of the training set, the ANN may be less robust and may sometimes diverge (Gaume and Gosset, 2003). More recently, a Data Based Mechanistic (DBM) modelling approach, introduced by Young (2002), derived the model structure and the parameter values from the input and output data using system engineering identification and parameter estimation techniques that attempted to go beyond the black-box concept by selecting those (not necessarily linear) model structures that are considered physically meaningful (Young, 2001, 2002). Although the DBM modelling approach recognises the importance of the physical coherence of the identified model structure, it derives it from the observations, thus disregarding *de facto* the results of at least 50 years of research efforts aimed at specifying the physical hydrological mechanisms that generate floods. This contrasts with the Bayes principle which would combine the observations with all possible *a priori* knowledge on the hydrological processes and possibly on the parameter values to obtain less uncertain *a posteriori* forecasts.

#### THE NEED TO RECONCILE THE TWO APPROACHES

Unfortunately, hydrological modellers emphasise the merits of their own approaches while disregarding those of others. In particular, physical process-oriented modellers have no confidence in the capabilities of data-driven models' outputs with their heavy dependence on training sets, while the more system engineering-oriented modellers claim that data-driven models produce better forecasts than complex physically-based models. Implicit in this controversy is the basic question: should 50 years of research by scientists seeking better representations of hydrological processes be jettisoned?

In this new century, there is the need, for the sake of hydrology as a science, to reconcile the two streams of thought by (a) combining the advantages of both approaches (Klemes, 1983) and (b) designing extensive benchmarks or test beds, to determine the role, validity and fields of application of the different models.

Krzysztofowicz (1999) has already proposed estimating the predictive probability by combining a generic model of unknown nature — but presumably physically-meaningful — with an autoregressive model by means of a Bayesian approach. Recent experiments have shown that the resulting Bayesian processor works well over a short forecasting horizon when the autoregressive model output is highly correlated with the observations, but it decays rapidly as soon as the required length of the forecasting horizon becomes larger; moreover, when dealing with water stages or with discharges, the autoregressive model is successful in the recession but less so in the rising limb of the hydrograph. Therefore, to reconcile the different approaches, as an improved alternative to the Krzysztofowicz approach, the Bayesian processor could be used to combine a physically-based to a data driven model. This approach would in fact benefit from the robustness of the former and the adaptability of the latter.

At the same time, appropriate benchmarks or test beds might be made freely available for modellers to assess, objectively, the qualities of their models as well as their limitations and the possible fields of application. The task is not simple but, once problems have been characterised, along with high quality data sets from appropriate catchments, a set of test cases, reflecting the various fields of application, could be devised. Thus, models could then be tested on the benchmarks or test beds, and quality assurance certificates could be issued.

#### Uncertainty in models and parameters

Uncertainty plays an increasing role in modelling and forecasting but has yet to reach a consensus among hydrologists, partly due to the statistical complexity and also because end users have no clear ideas on the use of the relevant information.

To provide a logical introduction to the discussion below, the reader is provided with a simple representation of certainty and uncertainty. For a line from  $-\infty$  to  $+\infty$ , full knowledge or certainty can be represented by a Dirac delta over a specific real number. For instance the certainty that the value of a given quantity  $x$  is 3, can be represented by  $\delta_{x=3}$ . At the same time, total uncertainty can be represented mathematically by a uniform distribution between  $-\infty$  and  $+\infty$ . Between these limits, partial knowledge will always be represented by a probability density, which will tend to be more and more peaky, as a function of the increasing knowledge level, around the imperfectly known quantity (here  $x = 3$ ) and, in the limit, when all the necessary information is acquired, it may converge onto the Dirac delta, for instance  $\delta_{x=3}$ .

Therefore, either a quantity is known perfectly and probabilistic concepts need not be invoked, or a quantity is not known perfectly and the most comprehensive information available is either its probability density function (pdf) or its integral, the Probability Distribution Function (PDF). This point must be clear when dealing with uncertainty in mathematical models based on a number of model parameters. In this case, either the model parameter values are known perfectly, in the sense that they do not need to be estimated because they are known on physical grounds (or it is believed so), or alternatively, if we are not sure and they are uncertain, we must not estimate ONLY their expected values but rather try to obtain their entire pdf, because the pdf is the information necessary for dealing with uncertain quantities.

When dealing with uncertainty, the issues to be clarified are:

1. How is predictive uncertainty defined and how is it derived?
2. Can the conditional predictive uncertainty be used instead of the unconditional?
3. What is the implication of using model uncertainty, as in the ‘ensemble forecasts’ instead of predictive uncertainty?
4. What is the implication of focussing on parameter estimation and estimation of their uncertainty, as opposed to prediction and the estimation of predictive uncertainty?
5. Should formal Bayesian inference be used or less formal likelihood-based approaches such as GLUE?
6. What is the benefit of using predictive probability in operational decision making?

#### PREDICTIVE UNCERTAINTY

In clarifying to hydrologists the real meaning of predictive uncertainty, Krzysztofowicz (1999), points out that “*Rational decision making (for flood warning, navigation, or reservoir systems) requires that the total uncertainty about a hydrologic predictand (such as river stage, discharge, or runoff volume) be quantified in terms of a probability distribution, conditional on all available information and knowledge.*” and that “*Hydrologic knowledge is typically embodied in a deterministic catchment model*”.

These statements underline two aspects usually not clearly understood by hydrologists. The first is that the objective of forecasting is to describe the uncertainty of actual future values of water stage, discharge, runoff volume, etc. rather than the uncertainty of predictions generated by the hydrological forecasting models. The second is that this

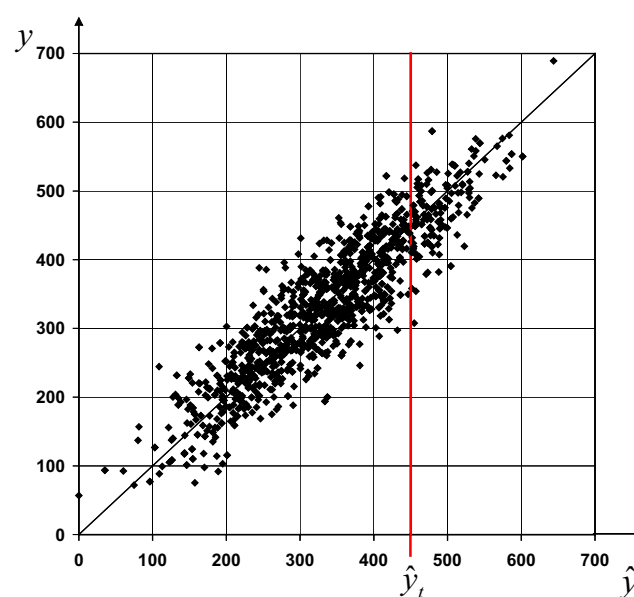
uncertainty, generally expressed in terms of a probability density (or probability distribution) function, is ‘conditional’ upon the forecasting model prediction, which is now seen as the available, although uncertain, knowledge of the future. In other words, the forecasting model prediction is now a function in the decision making process and not the provider of deterministic (and therefore ‘certain’) future levels, flows, etc.

To clarify these aspects, let us introduce the concept of the joint probability distribution of the real quantity of interest  $y$ , the predictand (namely the discharge, the water level in a specific cross section, etc.), and the model forecast  $\hat{y}$ . Unless the model is so exceptionally accurate, thus perfectly matching the observations, a scatter will always be observed in the  $y - \hat{y}$  plane as in Fig. 1. This scatter is a representation of the joint sample frequency of  $y$  and  $\hat{y}$  that can be used to estimate the joint probability density.

For any given model, the model forecast  $\hat{y}_t$ , where  $t$  is time, is a function of a specific value  $\hat{g}$  of the parameter set and of the input forcing  $x_t$  (the covariate); thus the joint probability density can be expressed as in Eqn.(1):

$$f(y_t, (\hat{y}_t | x_t, \hat{g})) \quad (1)$$

which, for the sake of clarity, is written in a more explicit way than in the classical statistical notation, by explicitly writing the model output  $\hat{y}_t$  conditional on the covariate and the parameters.





If there is no scatter and  $y_t = \hat{y}_t \forall t$ , then and only then can  $\hat{y}_t$  be used as a forecast of  $y_t$ . In all the other, more generally occurring, cases where there is inherent uncertainty, the conditional probability of  $y_t$  given  $\hat{y}_t$  must be derived in order to predict  $y_t$ . This is easily done by cutting for a given  $\hat{y}_t$  the previously mentioned joint probability density (Fig. 1) and renormalising it, which can be formalised as:

$$f(y_t | (\hat{y}_t | x_t, \hat{\theta})) = \frac{f(y_t, (\hat{y}_t | x_t, \hat{\theta}))}{\int_0^\infty f(y_t, (\hat{y}_t | x_t, \hat{\theta})) dy} \quad (2)$$

It is important that the reader understands that the conditional uncertainty given in Eqn. (2) expresses the predictive uncertainty of a ‘given’ model under a ‘given’ input forcing and a ‘given’ set of parameter values. This has nothing to do with the uncertainty induced by the model choice, and/or by the parameter values and/or by the input and output measurement errors. If it is believed that these additional uncertainties may strongly affect the forecast and its uncertainty, then they have to be assessed and marginalised out. Following the Bayesian approach, all the previously mentioned additional uncertainty (namely model, parameters and measurements) is concentrated in a number of ‘dummy’ parameters, the uncertainty of which is described via a posterior probability density and successively marginalised out.

In the notation in Mantovan and Todini (2006), a formal definition of such predictive uncertainty is represented by:

$$f(y_t | x_t, Y_n, X_n) = \int_{\Theta} f(y_t | x_t, \theta) g(\theta | Y_n, X_n) d\theta \quad (3)$$

This can also be written in a more explicit way, as:

$$f(y_t | (\hat{y}_t | x_t, Y_n, X_n)) = \int_{\Theta} f(y_t | (\hat{y}_t | x_t, \theta)) g(\theta | Y_n, X_n) d\theta \quad (4)$$

where the predictand  $y_t$  is explicitly written conditionally upon the model output  $\hat{y}_t$ , which is in turn conditional on the covariate and the parameters.

In Eqns. (3) and (4):

$$f(y_t | x_t, Y_n, X_n) \text{ or}$$

$f(y_t | (\hat{y}_t | x_t, Y_n, X_n))$  is the probability density of the predictand conditional upon the historical observations and the covariate after marginalising the uncertainty due to the parameters.

$Y_n$  is the set of historical predictand observations (for instance water levels, discharges, etc.) and  $n$  is the record length;

$X_n$  is the set of historical covariates (for instance rainfall, upstream inflows, etc.)

$y_t$  is the predictand value of interest

$x_t$  is the corresponding value of the covariate

$\theta$  is a given parameter vector

$\Theta$  is the ensemble of all possible parameter realisations

$f(y_t | x_t, \theta)$  or  $f(y_t | (\hat{y}_t | x_t, \theta))$  is the probability density of the predictand value of interest conditional upon the covariate and a generic set of parameters  $\theta$

$g(\theta | Y_n, X_n)$  is the posterior density of parameters given the historical observations

Equation (4) shows that, when all the different (model, parameters and measurements) uncertainties are account for, the predictive uncertainty is obtained by marginalising our dummy parameters effect from the conditional probability density expressed by Eqn. (2). Therefore, the posterior probability density for the parameter vector  $\theta$ , namely  $g(\theta | Y_n, X_n)$ , can be derived by means of a Bayesian inferential process (see Mantovan and Todini, 2006) and plays an essential role in the derivation of the predictive uncertainty. The Bayesian inference process allows this posterior probability density to be derived from the historical observations  $Y_n, X_n$ , starting from a prior density expressing our subjective knowledge about the parameters. The posterior density is used to marginalise out the conditionality on the parameters; this involves integrating over  $\Theta$ , the entire domain of existence of the parameters, its product with the conditional probability density function  $f(y_t | (\hat{y}_t | x_t, \theta))$  of the predictand  $y_t$  conditioned on the covariate  $x_t$  and the parameter vector  $\theta$  that identifies the specific model. Equation (1) is then the basis for deriving the predictive probability density for the  $t^{\text{th}}$  observation, and can be used to describe the predictive uncertainty both in ‘hindcast’ mode, when  $t \leq n$ , and in ‘forecast’ mode for  $t > n$ .

The predictive uncertainty expressed by Eqn. (4) is unconditional on the parameters, since parameter uncertainty has been integrated out. This means that, following Eqn. (4), to estimate the most likely outcome (the expected value of the predictand) one has to take into account and use all possible model predictions (namely one per parameter vector realisation) and not only the one relevant to the most likely parameter values, and all the predictions have then to be averaged using the derived posterior probability function to marginalise out all the uncertainty which is now encapsulated in the model parameters (De Finetti, 1975).

Therefore, in dealing with the derivation of the unconditional predictive probability, which is one of the main scopes of predictive modelling (and in particular of hydrological modelling), the estimation of the parameter

values is not the main goal, since the full predictive uncertainty information about the parameters will be incorporated in the derived posterior probability density of the parameters, given the observations.

Hence, the essential steps in evaluating the predictive uncertainty using a Bayesian inferential approach are to:

- define a subjective prior density for the parameters;
- assume an appropriate likelihood for the parameters, namely a probability density of the observations given the parameters coherent with the Bayes theorem (Mantovan and Todini, 2006);
- derive a pdf for the parameters from the observations (the posterior density);
- compute the probability density of the predictand conditional on the parameters; and
- marginalise the parameter uncertainty by integrating, over the parameter space, the derived parameters posterior pdf times the probability density of the predictand conditional upon the parameters.

The above explanation has addressed Issue 1; Issue 2 is now addressed. When dealing with modelling and prediction, it is common practice to estimate a ‘best set’ of parameter values to make predictions and forecasts (Sorooshian and Gupta, 1983; Thiemann *et al.*, 2001; Vrugt *et al.*, 2003). In this case the estimated predictive uncertainty

is referred to as ‘conditional upon the parameters’ and can be expressed as:

$$f(y_t | (\hat{y}_t | x_t, Y_n, X_n)) = f(y_t | (\hat{y}_t | x_t, \hat{g}(Y_n, X_n))) \quad (5)$$

where:

$\hat{g}$  represents the given parameter values which can be estimated as expected values, as maximum likelihood (ML) estimates, as modal values, as trial and error estimates, etc. or just simply assigned on the basis of physical considerations.

It should be noted that, unless the relation between the parameters and the predictand is linear, the conditional probability density derived from Eqn. (5) will not coincide with that given by Eqn. (4). Therefore, this approach may lead to less robust predictions and uncertainty estimates and its use should be justified either by a need for less computationally demanding approaches, such as in the case of real-time flood forecasting (as in Krzysztofowicz, 1999), or by proving that the distortion generated by disregarding the parameter uncertainty (which may result from a relatively small variability of the parameter values or by weak non-linear structures in the hydrological model) does not seriously affect the predictive probabilities. Following the example in Liu *et al.* (2005), Fig. 2 shows how in the case of the River Xixian, the difference between the unconditional predictive probability distribution function, computed by

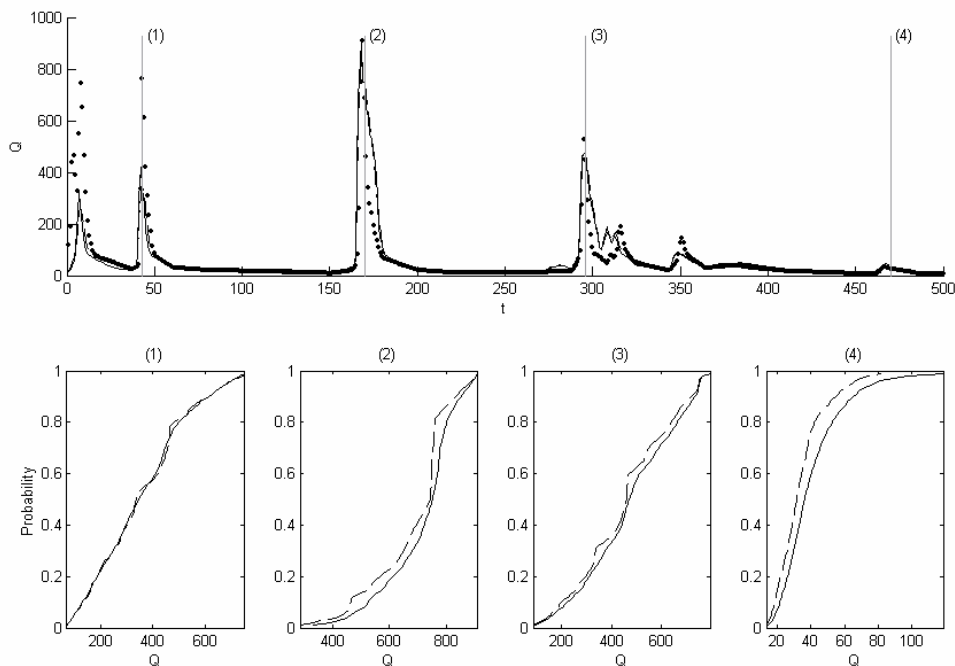


Fig. 2. Cumulative predictive probabilities estimated at various points in time. The solid line represent the unconditional (Eqn. (4)) while the dashed line represents the conditional (Eqn. (5)).

marginalising the 50 000 distributions obtained per each set of parameters (solid line) does not differ too much from the conditional (dashed line) obtained using the maximum likelihood parameter value estimates. This may happen either when the forecast is linearly or quasi-linearly related to the parameters, or when the forecast is relatively insensitive to the parameter variation around the estimated value, or, alternatively, the variance of the parameters is small. Therefore, this second and less formal approach may be acceptable in certain circumstances.

Issue 3 is now addressed which relates to a tendency among hydrologists and meteorologists to confuse predictive uncertainty with model uncertainty (Stephenson *et al.*, 2005). Model uncertainty is either expressed in the form of an ensemble of forecasts by varying model parameters as well as the initial and boundary conditions or by mimicking the predictive probability estimation, using an equation similar to Eqn. (6):

$$f(\hat{y}_t | x_t, Y_n, X_n) \triangleq g(\mathcal{Y}_n, X_n) \quad (6)$$

Clearly the density of Eqn. (6) differs from that represented by Eqns. (4) and (5) since it does not include the conditional density of the predictand, given the model forecast, as expressed by Eqn. (2) and therefore has nothing to do with the definition of predictive probability.

Although the use of ‘ensemble forecasts’ to represent the predictive uncertainty has been increasingly proposed by

meteorologists, it is not difficult to show, particularly when the members of the ensemble are used to produce flood forecasts, that very rarely will the measured value of the predictand lie between the 0.05 and the 0.95 probability bounds derived from the ensemble. Figure 3 compares the observed flow (thick solid line) in the Po river in Italy at the gauging station of Ponte Becca (~35 000 km<sup>2</sup>) with the modelled one using the actually measured rain over the catchment (dashed line) and the forecasts obtained using all the ensemble members (thin lines). Both the observed flows and the ones modelled using the measured rainfall are most of the time completely outside the uncertainty band described by the ensemble members.

As previously mentioned, many hydrologists also make the mistake of not including the conditional probability expressed by Eqn. (2) in the computation of the predictive uncertainty. This confusion also occurs in many GLUE papers where the predictive probability is defined as the cumulative probability deriving from Eqn. (6) instead of from Eqn. (4). To clarify this point, a few lines of the Beven and Freer (2001) paper are reproduced here.

*“Given a large enough sample of Monte Carlo simulations, the range of likelihood weighted predictions may be evaluated to obtain prediction quantiles at any time step. This is most easily done if the likelihood values are renormalized such as  $\sum_{i=1}^B L[M(\Theta_i)] = 1$ , where  $M(\Theta_i)$  now indicates the  $i$ th behavioural Monte Carlo sample, so that*

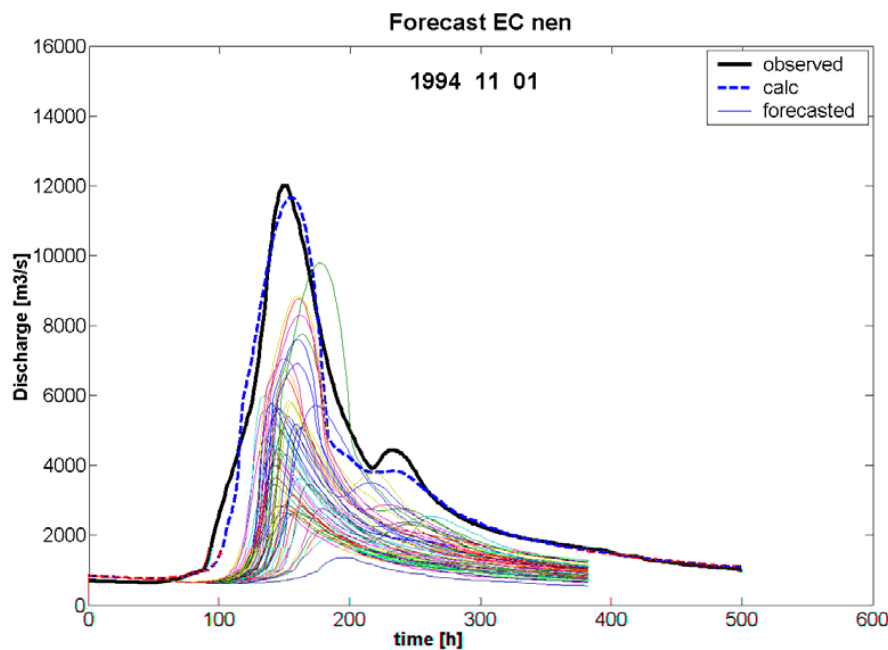


Fig. 3. Up to 10 days flow forecasts in the Po river in Italy at the gauging station of Ponte Becca (~35 000 km<sup>2</sup>). Observed flow (thick solid line); modelled flow using observed rainfall (dashed line); forecasts obtained using all the ensemble members (thin lines).



at any time step  $t$

$$P(\hat{Z}_t < z) = \sum_{i=1}^B L[M(\Theta_i) | \hat{Z}_{t,i} < z] \quad (7)$$

where  $\hat{Z}_{t,i}$  is the value of variable  $z$  at time  $t$  by model  $M(\Theta_i)$ .”

Noting that, in the Beven and Freer (2001) paper, the notations  $Z_t$ ,  $\hat{Z}_t$  are equivalent to  $y_t$ ,  $\hat{y}_t$  used in the present paper, it is not difficult to understand that in GLUE the quantiles essentially relate to the ‘model’ predicted variable  $\hat{Z}_t$ , not to the possible true value  $Z_t$  conditional on the predicted value  $\hat{Z}_t$  as a correct definition of predictive probability (Krzysztofowicz, 1999; Mantovan and Todini, 2006) implies, namely  $P(Z_t \leq z | \hat{Z}_t)$ . The reason for this definition is that decisions must be taken on the basis of what will possibly happen, not on what the model is predicting: the model is not the reality, it is only a tool to reduce uncertainty on future outcomes, by conditioning the predictand uncertainty on the model forecast. Therefore, in the rest of this paper the predictive probability used will be that given by Eqn. (4).

#### EQUIFINALITY VS INEQUIFINALITY

Beven and Binley (1992) introduced the concept that alternative models (generally the same hydrological model with different values of the parameters) may have the same predictive value, which was subsequently formalised as ‘equifinality’. The equifinality principle reflects, up to a certain point, the concepts expressed above where a distinction was made between a prediction based on a unique set of parameters, and that derived by assuming a prior range within which the parameters may vary and deriving the posterior parameter density from the observations. In principle, when dealing with non-linear models, the typical hydrological modelling with fixed parameter values may lead to large predictive biases and to an incorrect evaluation of the predictive uncertainty. If uncertain parameters are assumed in accounting for the presence of uncertainties, then, after expressing Eqn. (4) in discrete form, an ensemble of model predictions (one for each parameter vector realisation) must be averaged with the derived posterior parameter probability mass (the marginalisation) which will reflect the knowledge gained from the observations (the Bayesian learning process). Therefore, the objective of any Bayesian inference process is to obtain peakier posterior parameter densities from the observations, to reach ‘inequifinality’, namely that some models (linked to some of the parameter vector values) and their predictions are more likely than others (Mantovan and Todini, 2006).

The GLUE approach proposed by Beven and Binley (1992) uses Eqn. (7) rather than Eqn. (4), resulting in extremely ‘flat’ posterior probability distributions and unrealistically wide predictive uncertainties; by contrast, the application of Bayesian inference leads to peakier posterior densities and smaller predictive uncertainties.

Although the Bayesian statisticians would prefer to start with informative priors on the parameters (namely by using densities with a mode expressing prior knowledge), the hypothesis of complete ‘equifinality’ may be accepted at the outset of the Bayesian inference process, thus expressing the idea that all the models, one per parameter vector realisation, have the same informative value, due to prior lack of knowledge. This is why the real scope of the Bayesian inference process should be associated with ‘inequifinality’, to produce peakier posterior densities where some of the models are more likely to be correct than others (Mantovan and Todini, 2006).

#### PARAMETER ESTIMATION VERSUS PREDICTIVE UNCERTAINTY

This section addresses Issue 4. In the classical Bayesian approach, the parameters of a model do not necessarily represent physically meaningful quantities which have true (albeit unknown) values, but rather temporary ‘dummy’, ‘convenient’ or ‘nuisance’ quantities of uncertain nature, over which all uncertainty in the model, observations, boundary conditions, etc. is projected, to be marginalised out by their ‘posterior probability density’, obtained from observations via the Bayesian inference process (De Finetti, 1975, Chapters 11 and 12). Therefore, it is necessary to clarify the objective of this uncertainty assessment: Is ‘parameter estimation’ the main objective, or is it ‘prediction’?

If the objective is ‘parameter estimation’, it is assumed, implicitly, that the parameters have a true, albeit unknown, value to be found. Note that now the scope is not to identify dummy values for the parameters aiming at estimating the predictive density, but rather to determine the true, physically meaningful parameter values. In this case, if the observability conditions (Gelb, 1974) are met and if all the uncertainties in the model, the input, the output, etc. can be reflected, fully and correctly, in an appropriate likelihood function (which is not easy), only then will it be possible to estimate physically meaningful parameter values after deriving the posterior parameter probability density function (which coincides with the likelihood function if one takes the uniform as the *a priori* on the parameters), either using an ML approach or as an expected value. This is an extremely complex problem that in real cases can rarely be

solved to produce physically meaningful parameter values. The alternative ‘prediction’ problem is less complex and more feasible because, in the Bayesian approach, the estimation of the ‘true’ parameter values is not required; it is rather their entire ‘posterior probability density’ that expresses their uncertainty after sampling the observations. This posterior density may not be that associated with the ‘true’ value of the parameters, since in Bayesian inference the parameters become ‘convenient’ quantities, used to absorb and reflect all the sources of uncertainty; these are finally marginalised out at the end of the process. It is true that, if the different types of uncertainties, such as input and output uncertainties, model structural uncertainties, parameter uncertainties, etc. can be defined and described through pdfs, this could be beneficial, in the sense that they could then be marginalised out to obtain the predictive uncertainty, but as mentioned earlier, this is not essential when the aim is specifically ‘predictive uncertainty’.

#### LESS FORMAL LIKELIHOODS AS AN ALTERNATIVE TO FORMAL BAYESIAN INFERENCE

The point raised by Beven and Binley (1992), who advocated the use of less formal likelihoods to represent the likelihood of parameters when dealing with complex models and several types of errors (input, output, model, parameters, etc.) showing non-normal or asymmetrical densities, correlation and heteroscedasticity, is good. This descends from the fact that to be successful, a Bayesian inference process requires stringent assumptions on the shape of the probability density of errors, which is difficult to define correctly in the case of complex hydrological models. Unfortunately the proposed solution, GLUE, is based on a number of non-formal likelihoods that do not satisfy Bayes theorem and which lead to incoherent and paradoxical Bayesian inferential properties with a reduced capability of extracting information from the observations, and a consequent over-estimation of predictive uncertainty (Mantovan and Todini, 2006).

Presently, research aims, on the one hand, at finding coherent less formal likelihoods and, on the other hand, to transform the errors in ‘convenient’ spaces, where formal Bayesian inference can be used to demonstrate the robustness of the derived approaches with respect to their complexity and different typologies of errors. Liu *et al.*, (2006), have recently used the Normal Quantile Transform (NQT) (Van der Waerden, 1952, 1953a,b) to convert discharge observations and model predictions into a multivariate normal space where the Bayesian approach could be applied successfully. Recently, using the a,b,c model, an extremely simplified hydrological model

introduced by Fiering (1967) for didactic purposes, both GLUE- and NQT-based Bayesian inference were applied to synthetic data with complex asymmetric, heteroscedastic and time correlated errors, to describe the improvement obtainable—in terms of predictive uncertainty—from the prior uncertainty expressed by a multi-uniform distribution on the parameters. Figure 4 compares the results that can be obtained; they will be described more completely in a forthcoming paper with Mantovan and Martina. Figure 4(a) shows the expected value, the 0.05 and the 0.95 quantiles deriving from the discretisation of Eqn. (8):

$$f(y_i | (\hat{y}_i | x_i)) = \int_{\Theta} f(y_i | (\hat{y}_i | x_i, \theta)) g_0(\theta) d\theta \quad (8)$$

written in discretised form, where  $g_0(\theta)$  represents the multi-uniform prior density on the parameters; this represents the prior predictive uncertainty before Bayes theorem and the NQT have been applied to derive the posterior pdf of the parameters. Figure 4(c) displays the expected value, the 0.05 and the 0.95 quantiles obtained from Eqn. (4) in discretised form (i.e. using the posterior density of the parameters), while Fig. 4(b) reproduces the 0.05, the 0.5 (not the expected value, since this is how GLUE results are typically shown) and the 0.95 quantiles obtained using the GLUE approach. It is interesting to see that, while the formal Bayesian inference approach applied to the NQT transformed data, which does not require assumptions on the probability distribution of errors in deriving the posterior pdf of the parameters, largely reduces the prior predictive uncertainty expressed by Eqn. (4), GLUE reduces it much less and in places, for instance around the peaks at the 120 and the 160 time steps, the posterior predictive uncertainty appears to be larger than the *a priori* uncertainty. Therefore, in response to Issue 5, it appears that formal Bayesian inference should be employed in preference to GLUE.

One of the issues that presently enriches the debate about uncertainty among hydrologists is how to show the benefits arising from the operational use of predictive uncertainty, a corollary of which is how to communicate uncertainty to the end-users, namely the decision-makers. Indeed, the end-users such as water managers, emergency managers, etc. have difficulty in perceiving the benefits arising from the operational use of predictive uncertainty. What is certain is that hydrologists must not make statements such as: “the probability of flooding in the next 12 hours is 67.5%”. This is meaningless to an end-user. What he/she would like to hear is the answer to the basic question: “what are the expected benefits and drawbacks of issuing a flood alert for the next 12 hours?”. Therefore, hydrologists must define, in dialogue with end-users, subjective utility functions, which can be used to compute the expected benefits or the

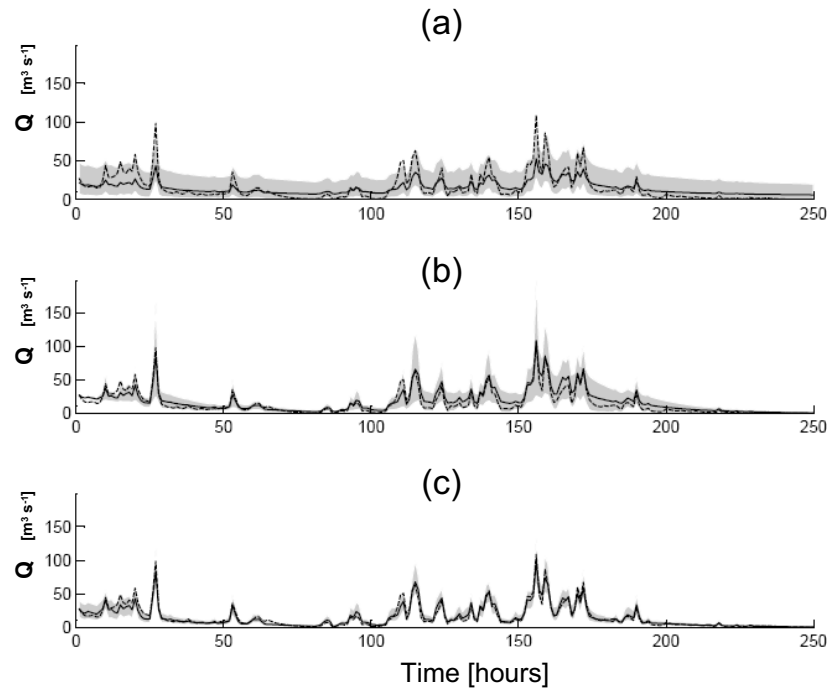


Fig. 4. Comparison of predictive uncertainty estimates. (a) expected value, 0.05 and 0.95 quantiles obtained using the prior parameter density by discretising Eqn. (8); (b) 0.05, 0.50, 0.95 quantiles obtained using the GLUE approach; (c) expected value, 0.05 and 0.95 quantiles obtained using the posterior parameter density obtained via the NQT transform and Bayesian inference by discretising Eqn. (4).

expected damages contingent on the predictive density of the quantity of interest.

A schematic example of such utility functions is shown in Fig. 5, redrawn from Martina *et al.* (2006), for the case of a flood alert (please note that in this simple schematic example casualties are not taken into account). The dashed line represents the end-user perception of damage (not necessarily the real one) that will occur if the dykes are

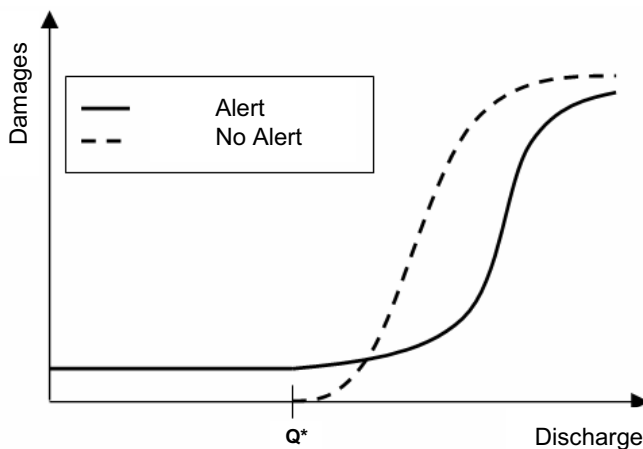


Fig. 5. The utility functions deriving from a flood alert problem (Redrawn from Martina *et al.*, 2006). Solid line represents perceived cost and damage if an alert is issued; dashed line represents perceived damage if alert is not issued.

overtopped, namely  $Q > Q^*$  where  $Q^*$  is the maximum discharge that may safely flow in the river. The solid line represents the perception of cost plus damages when an alert has been issued. As can be seen from Fig. 5, if an alert is issued a cost must be inevitably incurred for mobilising civil protection agents, alerting the population, laying sandbags, etc., but the damage in that case will be smaller than in the previous case due to the raised awareness of the incoming flood. The decision on whether or not to issue an alert will then descend from the comparison of the ‘expected damage’ for the two options, obtained by integrating the product of the cost function times the predictive uncertainty pdf over all possible values of future discharge. It should be noted that the ‘expected damages’ are a function of the actual future discharge that will happen, not of the discharge predicted by the model. By using the expected value of damage instead of the ‘model forecast’, the probability of false alarms as well as of missed alarms should be much reduced, as the uncertainty about the future discharge is taken into account. In addition, the peakier the predictive density is, the more reliable will be the resulting decision, so that improvements in forecasting, rather than looking for a better ‘deterministic’ forecast, must essentially aim at reducing predictive uncertainty by whatever means are available.

To show how one can use predictive uncertainty in

operation (Issue 6), the Lake Como real-time management decision support system is considered here as one of the few successful examples of the operational use of forecast uncertainty (Todini, 1999). Lake Como is a natural lake in northern Italy closed at its exit and managed as a multi-purpose lake for flood control, irrigation and electric power production. Using a stochastic dynamic programming approach, a standard operating rule was developed on a ten-day basis to optimise long term irrigation and energy production. However, when a flood is forecast, the reservoir manager would like to modify the standard operating rule to deal with the incoming flood. To achieve this goal, a utility function describing the damage perception of the manager was developed; every morning an incoming flood forecast, together with its predictive uncertainty, is issued and an optimal release, computed by minimising the expected damage using the inflow predictive uncertainty, is then proposed. Note that all this process is totally hidden from the water manager who is aware only of the suggested optimal release and of its expected consequences (Fig. 6).

The performance of the system was assessed on the basis

of a hindcast simulation for the 15 year period January 1<sup>st</sup>, 1981 to December 31<sup>st</sup>, 1995; the results are presented in Table 1. When applying the optimised rule, the lake level never falls below the lower acceptable limit of  $-0.40$  m, while historically this was observed on 214 days. In terms of Como flooding, over the 15 years, the lake level was

Table 1. Summary of results. A comparison between recorded water level occurrences (historical) and the results of the operation rule based on the forecasting uncertainty (optimised) for the period January 1<sup>st</sup>, 1981 to December 31<sup>st</sup>, 1995.

Water level	No. of days	
	Historical	Optimised
$> -40$ cm	214	0
$> 120$ cm	133	75
$\geq 140$ cm	71	52
$\geq 173$ cm	35	34
Water deficit	$890.27 \cdot 10^6 \text{ m}^3$	$788.59 \cdot 10^6 \text{ m}^3$

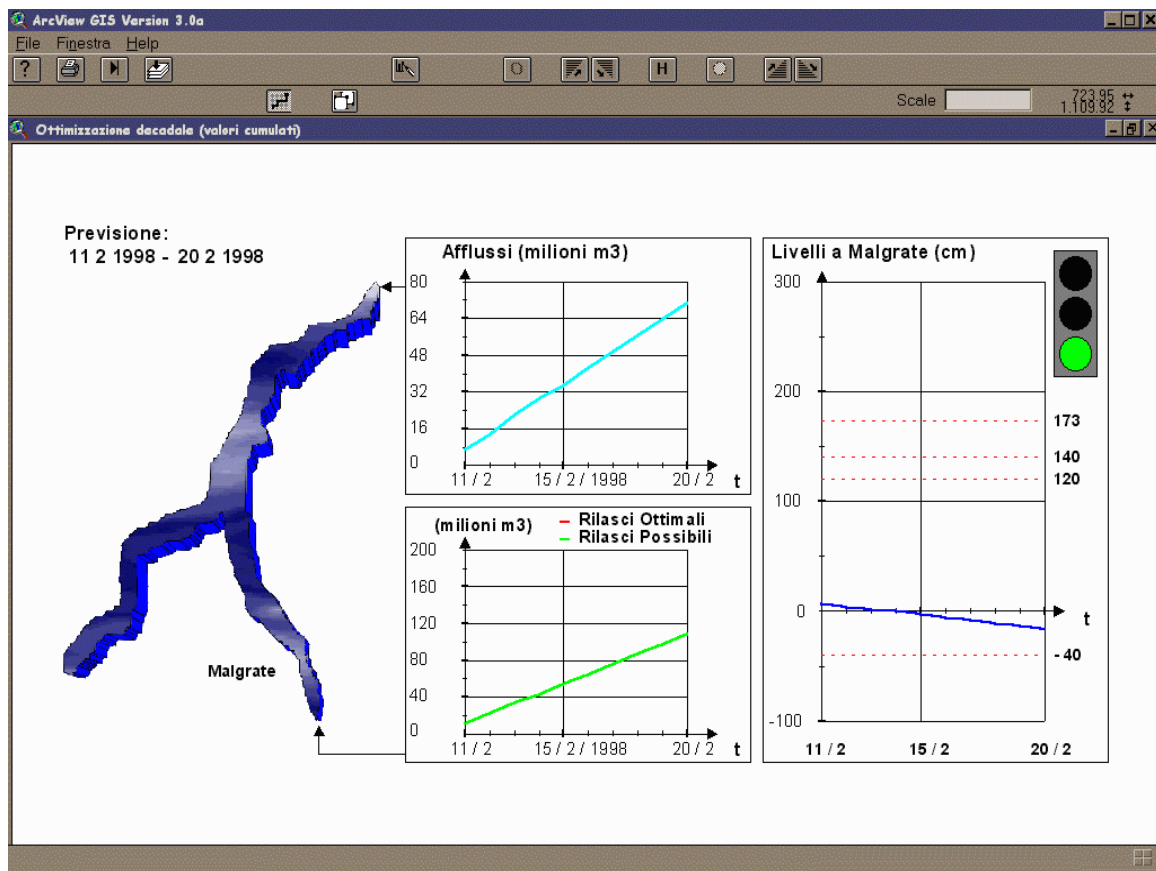


Fig. 6. The Lake Como operational decision support system. The system, on the basis of the expected value of inflows to the lake (light blue line) and its uncertainty (not shown on the screen, but used in the process) suggests to the water manager the optimal release (green line) which minimises the expected damage and shows the consequent expected lake level (blue line) for the following 10 days.

historically recorded to be above the lower flood limit of 1.20 m on 133 days, whereas the optimised rule reduced it to 75 days. A noticeable reduction also appears at higher lake levels: at 1.40 m, when the traffic must stop in the main square of Como, the reduction is from 71 to 52 days and at 1.73, the legal definition of ‘normal flood’ when people can claim compensation for their damage, the reduction is from 35 to 34 days. At the same time, the irrigation water deficit decreases by an average of almost  $100 \times 10^6 \text{ m}^3 \text{ yr}^{-1}$ . This result is exceptional, given that meeting irrigation demand implies higher lake levels, an objective conflicting with the need to reduce the frequency of flooding.

It is quite interesting how the system was accepted by the end-user. At the end of 1997, the system was installed operationally and the Director of Consorzio dell’Adda, who is in charge of lake management, was invited to look at it but not to use it until he had confidence in its effectiveness. After six months the Director admitted that he had been beaten four to nil. Every time he took a decision different from that suggested by the Decision Support System (DSS), he was wrong. Ever since, the system has been in operation and used successfully; it has produced not only a reduction in the number, frequency and magnitude of Como flooding events, but also a 3% increase in energy production and a large volume of extra water for irrigation.

This example shows that, if appropriately briefed and involved, the end-users will quickly become aware of the benefits arising from the use of predictive uncertainty, provided they are not asked to interpret the forecasting in statistical terms or the stochastic computation and optimisation frequently required in problems in this type.

It is the author’s personal belief that considerable effort is required to inform the end-users of the improvements obtainable without burdening them with the computational complexity. In this way, they can appreciate and receive the full benefits of an approach aimed at improving the success of their decision-making.

## Conclusions and future perspectives

From the Rational Method to the distributed hydrological models presently available, a vast evolution of conceptualisation and parameterisation of hydrological processes has evolved in catchment models over several decades of research. This effort in describing hydrological phenomena in physical terms from the pixel to the catchment scale (the bottom-up approach), is not fully acknowledged by the proponents of data-driven or data-mechanistic models (the top-down approach) who do not require the derivation of rigid model structures from the physical balance equations. Similarly, the proponents of the physically

meaningful models do not acknowledge the merits of the data-driven approach.

It is difficult to demonstrate the superiority of either approach but, although recognising the merit of the data-driven models, a danger exists in broad philosophical terms: namely, that all the work aimed at gaining a better physical understanding of the runoff generation and routing processes and their representation at different space and time scales can be dismissed as unnecessary. Therefore, the hydrological research community would do well to follow the advice of Klemes from over 20 years ago (Klemes, 1983) and combine the advantages of both approaches while designing appropriate benchmarks or test-beds to evaluate the roles and fields of application of the different types of models. Furthermore, the recent introduction of the ‘equifinality’ principle, instead of leading to possible solutions, has amplified the dissatisfaction with physically-based models because their parameter uncertainty is deemed to be enormous. Again, in respect of physically-based models, research hydrologists, with knowledgeable statisticians, must agree on the principles of ‘predictive uncertainty’ and on the counter-principle of ‘inequifinality’; they must construct formally correct and less diffuse posterior parameter distribution functions to reflect the quantity of data available as indeed they do for the data-driven models. These can then be used to marginalise the parameter uncertainty to deliver more appropriate measures of predictive uncertainty. Only by sharing the framework for estimating predictive uncertainty for both data-driven and physically-based models can benchmarks or test-beds be established to determine their predictive merits and roles.

Finally, water and emergency managers must be made aware of the potential benefits of correct estimates of predictive uncertainty. For their part, hydrologists must understand the actual problems of stakeholders rather than retreating into a haze of equations and complex statistical and mathematical representations.

## References

- Abbott, M.B., Bathurst, J.C., Cunge, J.A., O’Connell, P.E. and Rasmussen, J., 1986a. An introduction to the European Hydrological System - *Système Hydrologique Européen*, “SHE”, 1: History and philosophy of a physically-based, distributed modelling system. *J. Hydrol.*, **87**, 45–59.
- Abbott, M.B., Bathurst, J.C., Cunge, J.A., O’Connell, P.E. and Rasmussen, J., 1986b. An introduction to the European Hydrological System - *Système Hydrologique Européen*, “SHE”, 2: Structure of physically-based, distributed modelling system. *J. Hydrol.*, **87**, 61–77.
- Amorocho, J. and Orlob, G.T., 1961. *Nonlinear analysis of hydrologic systems*. Water Resources Centre, Contribution 40. University of California, Berkeley, USA.



- Beven, K.J. and Kirkby, M.J., 1979. A physically based, variable contributing area model of basin hydrology. *Hydrolog. Sci. Bull.*, **24**, 43–69.
- Beven, K.J. and Binley, A.M., 1992. The future of distributed models: model calibration and uncertainty prediction. *Hydrol. Process.*, **6**, 279–298.
- Beven, K.J. and Freer, J., 2001. Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. *J. Hydrol.*, **249**, 11–29.
- Box, G.E.P. and Jenkins, G.M., 1970. *Time Series Analysis Forecasting and Control*. Holden Day, San Francisco, USA.
- Burnash, R.J.C., Ferral, R.L. and McGuire, R.A., 1973. *A General Streamflow Simulation System - Conceptual Modelling for Digital Computers*. Report by the Joint Federal State River Forecasts Center, Sacramento, USA.
- Cameron, D., Kneale, P. and See, L., 2002. An evaluation of a traditional and a neural net modeling approach to flood forecasting for an upland catchment. *Hydrol. Process.*, **16**, 1033–1046.
- Crawford, N.H. and Linsley, R.K., 1966. *Digital simulation in Hydrology, Stanford Watershed model IV*. Tech. Rep. 39. Dept. Civil Eng. Stanford University, USA.
- Dawdy, D.R. and O'Donnell, T., 1965. Mathematical models of catchment behavior. *J. Hydraul. Div.- ASCE, HY4*, **91**, 123–137.
- Dawson, C.W. and Wilby, R.L., 2001. Hydrological modelling using artificial neural networks. *Prog. Phys. Geogr.*, **25**, 80–108.
- de Finetti, B., 1975. *Theory of Probability*, vol. 2. Wiley, Chichester, UK.
- De Roo, A.P.J., Wesseling, C.G. and Van Deursen, W.P.A., 1998. Physically based river modelling within a GIS. The LISFLOOD model. Proc. 3rd Int. Conf. on Geo-Computation. In: *Geo-Computation CD-ROM* produced by R.J. Abrahart. ISBN 0-9533477-0-2. [http://www.geocomputation.org/1998/06/gc\\_06.htm](http://www.geocomputation.org/1998/06/gc_06.htm)
- De Roo, A.P.J., Wesseling, C.G. and Van Deursen, W.P.A., 2000. Physically-based river basin modelling within a GIS: the LISFLOOD model. *Hydrol. Process.*, **14**, 1981–1992.
- Dooge, J.C.I., 1973. *Linear Theory of Hydrologic Systems*. Technical Bull. No. 1468. U.S. Department of Agriculture. Washington, USA.
- Ewen, J., Parkin, G. and O'Connell, P.E., 2000. SHETRAN: Distributed river basin flow and transport modeling system. *J. Hydrolog. Eng.*, **5**, 250–258.
- Fiering, M.B., 1967. *Streamflow synthesis*. Harvard University Press, Cambridge, MA, USA.
- Franchini, M., Wendling, J., Obled, Ch. and Todini, E., 1996. Physical interpretation and sensitivity analysis of the TOPMODEL. *J. Hydrol.*, **175**, 293–338.
- Freeze, R.A. and Harlan R.L., 1969. Blueprint for a physically-based digitally-simulated hydrologic response model. *J. Hydrol.*, **9**, 237–258.
- Garcia-Bartual, R., 2002. Short-term river forecasting with Neural Networks, Integrated Assessment and Decision Support. *Proc. 1st biennial meeting Int. Environmental Modelling and Software Society*, **2**, 160–165. (ISBN: 88-900787-0-7)..
- Gaume, E. and Gosset, R., 2003. Over-parameterisation, a major obstacle to the use of artificial neural networks in hydrology? *Hydrol. Earth Syst. Sci.*, **7**, 693–706.
- Gelb, A. (Ed.), 1974. *Applied optimal estimation*. M.I.T. Press, Cambridge, MA, USA.
- Gupta, V.K. and Sorooshian, S., 1983. Uniqueness and observability of conceptual rainfall-runoff model parameters: the percolation process examined. *Water Resour. Res.*, **19**, 269–276.
- Hayami, S., 1951. On the propagation of flood waves. *Disaster Prevention Research Institute Bul.*, **1**, Kyoto University, Japan. 1–16.
- Juemou, W., Ruifang, Z. and Guanwu, X., 1987. Synthesised Constrained Linear System (SCLS). *J. Hydraul. Eng.*, no. 7. Beijing, China.
- Kalinin, G.P. and Milyukov, P.I., 1957. O raskete neustanovivshegosya dvizhenia vody v otkrytykh ruslakh (On the computation of unsteady flow in open channels). *Meteorologiya i gidologiya zhurnal* **10**, 10–18. (in Russian).
- Klemes, V.K., 1983. Conceptualization and scale in hydrology. *J. Hydrol.*, **65**, 1–23.
- Kowen, N., 2000. *WATFLOOD/SPL: Hydrological model and flood forecasting system*. Dept. Civil Engineering, University of Waterloo, Waterloo, Ont., Canada
- Krzysztofowicz, R., 1999. Bayesian theory of probabilistic forecasting via deterministic hydrologic model. *Water Resour. Res.*, **35**, 2739–2750.
- Kuichling, E., 1889. The relation between the rainfall and the discharge of sewers in populous districts. *Amer. Soc. Civil Eng. Trans.*, **20**, 1–56.
- Liang, X., Lettenmaier, D.P. and Wood, E.F., 1996a. One-dimensional Statistical Dynamic Representation of Subgrid Spatial Variability of Precipitation in the Two-Layer Variable Infiltration Capacity Model. *J. Geophys. Res.*, **101**(D16), 21,403–21,422.
- Liang, X., Wood, E.F. and Lettenmaier, D.P., 1996b. Surface soil moisture parameterization of the VIC-2L model: Evaluation and modifications. *Global Planet. Change*, **13**, 195–206.
- Liang, X., Lettenmaier, D.P. Wood, E.F. and Burges, S.J., 1994. A Simple hydrologically Based Model of Land Surface Water and Energy Fluxes for GSMs. *J. Geophys. Res.*, **99**(D7), 14,415–14,428.
- Liu, Z. and Todini, E., 2002. Towards a comprehensive physically-based rainfall-runoff model. *Hydrol. Earth Syst. Sci.*, **6**, 859–881.
- Liu, Z., Martina, M.V.L. and Todini, E., 2005. Flood forecasting using a fully distributed model: application of the TOPKAPI model to the Upper Xixian Catchment. *Hydrol. Earth Syst. Sci.*, **9**, 347–364.
- Lloyd-Davies, D.E., 1906. The elimination of stormwater from sewerage systems. *Inst. Civil Eng. Proc. (London)*, **164**, 41–67.
- Mantovan, P. and Todini, E., 2006. Hydrological forecasting uncertainty assessment: Incoherence of the GLUE methodology. *J. Hydrol.*, **330**, 368–381.
- Martina, M.L.V., Todini, E. and Libralon, A., 2006. A Bayesian decision approach to rainfall thresholds based flood warning. *Hydrol. Earth Syst. Sci.*, **10**, 413–426.
- Moore, R.J. and Clarke, R.T., 1981. A Distribution Function Approach to Rainfall-Runoff Modelling. *Water Resour. Res.*, **17**, 1367–1382.
- Mulvany, T.J., 1850. On the use of self registering rain and flood gauges. *Inst. Civ. Eng. Proc. (Dublin)*, **4**, 1–8.
- Nash, J.E., 1958. The form of the instantaneous unit hydrograph. IUGG General Assembly of Toronto, Vol. III. *IAHS Publication no. 45*, 114–121.
- Nash, J.E., 1960. A unit hydrograph study with particular reference to British catchments. *Proc. Inst. Civil. Eng.*, **17**, 249–282.
- Natale, L. and Todini, E., 1976a. A stable estimation for linear models - 1. Theoretical development and Monte-Carlo experiments. *Water Resour. Res.*, **12**, 667–671.
- Natale L. and Todini E., 1976b. A stable estimator for linear models - 2. Real world hydrologic applications. *Water Resour. Res.*, **12**, 672–675.

- O'Donnell, T., 1966. Methods of computation in hydrograph analysis and synthesis, Recent trends in hydrograph synthesis, *Proc. Tech. Meeting no. 21, T.N.O.*, The Hague, The Netherlands. 65–102.
- Refsgaard, J.C. and Storm, B., 1995. MIKE SHE (Chapter 23). In: *Computer models of watershed hydrology*, V.P. Singh (Ed.), Water Resources Publications, Littleton, Colorado, USA.
- Rockwood, D.M., 1964. *Streamflow synthesis and reservoir regulation*. U.S. Army Engineer Division, North Pacific, Portland, Oregon, Engineering Studies Project 171, Technical Bulletin No. 22.
- Shamseldin, A.Y., 1997. Application of neural network technique to rainfall-runoff modelling. *J. Hydrol.*, **199**, 272–294.
- Sherman, L.K., 1932. Streamflow from rainfall by the unit graph method. *Engineering News Record*, **108**, 501–505.
- Singh, V.P. and Woolhiser, D.A., 2002. Mathematical modeling of watershed hydrology. *J. Hydrol. Eng.*, **7**, 270–292.
- Sorooshian, S. and Gupta, V.K., 1983. Automatic calibration of conceptual rainfall-runoff models: The question of parameter observability and uniqueness. *Water Resour. Res.*, **19**, 260–268.
- Stephenson, D.B., Coelho C.A.S., Balmaseda M., and Doblas-Reyes F.J., 2005: Forecast Assimilation: A unified framework for the combination of multi-model weather and climate predictions. *Tellus A*, **57A**, 253–264.
- Sugawara, M., 1967. The flood forecasting by a series storage type model. *IAHS Publication no. 85*, 1–6.
- Sugawara, M., 1995. Chapter 6: Tank model. In: *Computer models of watershed hydrology*, V. P. Singh (Ed.), Water Resources Publications, Littleton, Colo., USA.
- Thiemann, M., Trosset, M., Gupta, H. and Sorooshian, S., 2001. Bayesian recursive parameter estimation for hydrologic models. *Water Resour. Res.*, **37**, 2521–2535.
- Todini E., 1978. Using a desk-top computer for an on-line flood warning system. *IBM J. Res. Develop.*, **22**, 464–471.
- Todini, E., 1995. New trends in modeling soil processes from hillslope to GCM scales. In: *The Role of Water and the Hydrological Cycle in Global Change*, H.R. and S.A Oliver, (Eds). NATO ASI Series, Series I: Global Environmental Change, 31: 317–347.
- Todini E., 1996. The ARNO Rainfall-Runoff model. *J. Hydrol.*, **175**, 339–382.
- Todini, E., 1999. Using phase-space modeling for inferring forecasting uncertainty in non-linear stochastic decision schemes. *J. Hydroinformatics*, **01.2**, 75–82.
- Todini, E., 2002a. The ARNO model. (Chapter 16). In: *Mathematical Models of Large Watershed Hydrology*, V.P. Singh, D.K. Frevert and S.P. Meyer (Eds.), Water Resources Publications, Littleton, Colorado, USA. 687–716.
- Todini, E., 2002b. The CLS model. (Chapter 20). In: *Mathematical Models of Large Watershed Hydrology*. V. P. Singh, D. K. Frevert and S. P. Meyer (Eds.), Water Resources Publications, Littleton, Colorado, USA. 861–886.
- Todini, E. and Wallis, J.R., 1977. Using CLS for daily or longer period rainfall-runoff modelling. In: *Mathematical Models for Surface Water Hydrology*, T.A. Ciriani, U. Maione and J.R. Wallis (Eds.), Wiley, Chichester, UK. 149–168.
- Todini, E. and Bossi, A., 1986. PAB (parabolic and backwater) an unconditionally stable flood routing scheme suited for real-time forecasting and control. *J. Hydraul. Res.*, **24**, 405–424.
- Todini, E. and Ciarapica, L., 2002. The TOPKAPI model. (Chapter 12). In: *Mathematical Models of Large Watershed Hydrology*. V.P. Singh, D.K. Frevert and S.P. Meyer (Eds.), Water Resources Publications, Littleton, Colorado, USA. 471–506.
- Van der Waerden, B.L., 1952. Order tests for two-sample problem and their power: I. *Indagationes Mathematicae*, **14**, 453–458.
- Van der Waerden, B.L., 1953a. Order tests for two-sample problem and their power: II. *Indagationes Mathematicae*, **15**, 303–310.
- Van der Waerden, B.L., 1953b. Order Tests for the two-sample problem and their Power: III. *Indagationes Mathematicae*, **15**, 311–316.
- Verdin, J., and Klaver, R., 2002. Grid-cell-based crop water accounting for the Famine Early Warning System. *Hydrol. Process.*, **16**, 1617–1630.
- Vivoni, E.R., 2003. *Hydrologic Modeling using triangulated irregular networks: terrain representation, flood forecasting and catchment response*. PhD Thesis, MIT, Cambridge, Mass., USA.
- Vrugt, J.A., Gupta, H.V., Bouten, W. and Sorooshian, S., 2003. A Shuffled Complex Evolution Metropolis Algorithm for optimization and uncertainty assessment of hydrological model parameters. *Water Resour. Res.*, **39**, 1201, doi:10.1029/2002WR001642.
- Wigmosta, M.S., Vail, L.W., and Lettenmaier, D.P., 1994. A distributed hydrology-vegetation model for complex terrain. *Water Resour. Res.*, **30**, 1665–1679.
- WMO, 1975. *Intercomparison of conceptual models used in operational hydrological forecasting*. Operational Hydrology Rep. 7, WMO 429. WMO, Geneva, Switzerland.
- Wood, E.P., Lettenmaier, D.P. and Zartarian, V.G., 1992. A land-surface hydrology parameterization with subgrid variability for general circulation models. *J. Geophys. Res.*, **97** (D3), 2717–2728.
- Wooding, R.A., 1965a. A hydraulic model for the catchment-stream problem, I. Kinematic wave theory. *J. Hydrol.*, **3**, 254–267.
- Wooding, R.A., 1965b. A hydraulic model for the catchment-stream problem, II. Numerical solutions. *J. Hydrol.*, **3**, 268–282.
- Wooding, R.A., 1966. A hydraulic model for the catchment-stream problem, III. Comparison with runoff observations. *J. Hydrol.*, **4**, 21–37.
- Woolhiser, D.A. and Liggett, J.A., 1967. Unsteady, one-dimensional flow over a plane – the rising hydrograph. *Water Resour. Res.*, **3**, 753–771.
- Young, P.C., 2001. Data-based mechanistic modelling and validation of rainfall-flow processes. In: *Model validation: perspectives in hydrological science*, M.G. Anderson and P.D. Bates, (Eds.), Wiley, Chichester, UK. 117–161.
- Young, P.C., 2002. Advances in real-time flood forecasting. *Phil. Trans. Roy. Soc. London, A*, **360**, 1433–1450.
- Zhao, R.J., 1977. *Flood forecasting method for humid regions of China*. Report, East China College of Hydraulic Engineering, Nanjing, China.