A stochastic approach for assessing the uncertainty of rainfall-runoff simulations

Alberto Montanari and Armando Brath
Faculty of Engineering, University of Bologna, Bologna, Italy

Received 31 July 2003; revised 2 November 2003; accepted 26 November 2003; published 31 January 2004.

[1] Rainfall-runoff models have received a great deal of attention by researchers in the last decades. However, the analysis of their reliability and uncertainty has not been treated as thoroughly. In the present study, a technique for assessing the uncertainty of rainfall-runoff simulations is presented that makes use of a meta-Gaussian approach in order to estimate the probability distribution of the model error conditioned by the simulated river flow. The proposed technique is applied to the case study of an Italian river basin, for which the confidence limits of simulated river flows are derived and compared with the respective actual observations. *INDEX TERMS:* 1860 Hydrology: Runoff and streamflow; 3210 Mathematical Geophysics: Modeling; 1821 Hydrology: Floods; 1894 Hydrology: Instruments and techniques; *KEYWORDS:* uncertainty, rainfall-runoff, hydrological model, confidence limits, meta-Gaussian

Citation: Montanari, A., and A. Brath (2004), A stochastic approach for assessing the uncertainty of rainfall-runoff simulations, *Water Resour. Res.*, *40*, W01106, doi:10.1029/2003WR002540.

1. Introduction

[2] Rainfall-runoff (R-R) models are increasingly used in real world applications. Indeed, the growing availability of both computing power and hydrological data observed at fine spatial and temporal scales makes the application of these models an attractive option for answering many of the questions which are frequently posed to hydrologists.

[3] Intensive research efforts have been devoted to R-R models in the recent past. However, in spite of the promising results of many studies it often appears to the practitioner that there is still significant uncertainty associated with R-R simulations. Nevertheless, almost all existing river flow simulation techniques are conceived to provide a point estimation, since most research in operational hydrology has been dedicated to finding the best estimate rather than quantifying the uncertainty of model predictions [*Brath and Rosso*, 1993; *Singh and Woolhiser*, 2002].

[4] The uncertainty in R-R models output is originated by several causes, such as input uncertainty, parameter uncertainty and model uncertainty. The latter is due to the intrinsic inability of the R-R model to provide a perfect simulation of the hydrological processes involved in the R-R transformation. Furthermore, a potential source of uncertainty relies in the model output, which may be affected by measurement errors that are influent on model parameters and evaluation of model performances.

[5] Uncertainty can be estimated using different approaches, the choice among them depending also on the nature of the R-R model. A first option is to structure the latter as a probability model. The corresponding optimal simulation, along with its confidence interval (CI) can be computed probabilistically [*Montanari et al.*, 1997]. A second option is to estimate uncertainty by analyzing the statistical properties of the R-R model error series that occurred in reproducing observed historical river flow data. This approach has been followed by many authors in statistics [*Yar and Chatfield*, 1990] and hydrology [see, e.g., *Loukas et al.*, 2002]. A third option is to use simulation and re-sampling techniques, thus applying the so-called Monte Carlo method [*Metropolis*, 1987]. There are two essentially different types of re-sampling approaches, namely those that use importance sampling [see *Kuczera and Parent*, 1998] and those that define the response surface using weights (for instance the generalised likelihood uncertainty estimator described here below).

[6] In recent years, the awareness of the key role that reliability assessment can play in R-R modeling has stimulated intensive research on this subject. The most known method for inferring the global R-R uncertainty is probably the generalized likelihood uncertainty estimator (GLUE) proposed by *Beven and Binley* [1992]. This approach was used in a number of studies presented in the hydrologic literature [see, e.g., *Cameron et al.*, 1999, 2000; *Blaszkova and Beven*, 2002] and is increasingly used today. GLUE rejects the concept of an optimum model and parameter set and assumes that, prior to input of data into a model, all model structures and parameter sets have an equal likelihood of being acceptable. The existence of multiple likely models and parameter sets has been called equifinality [*Beven*, 1993, 2001, 2002], to suggest that this should be accepted as a generic problem in hydrological modeling rather than simply reflecting the problem of identifying the true model in the face of uncertainty. Once different candidate models are identified, GLUE is performed by first identifying, for each model, the parameters which most affect the output. Then, a high number of parameter sets is generated via uniform sampling, or incorporating prior knowledge about the distribution of parameters. The R-R models are then run for each of the sets and the model output is compared to a record of observed data (e.g., observed hydrographs or annual maximum peak flows [see *Cameron et al.*, 1999]). The performance of each trial

0043-1397/04/2003WR002540$09.00

is assessed through likelihood measures. This includes rejecting some models and/or parameter sets as non behavioral. For instance, *Cameron et al.* [1999] used the *Nash and Sutcliffe* [1970] efficiency to evaluate the likelihood of the simulation of a continuous hydrograph. All models (and corresponding parameter sets) that provide a likelihood measure reaching a minimum threshold are retained. The likelihood weighted uncertainty bounds can be calculated using a standard procedure [*Freer et al.*, 1996]. First, the calculated likelihoods are rescaled to produce a cumulative sum of 1.0. A cumulative distribution function of simulated discharges is then constructed using the rescaled weights. This allows uncertainty bounds to be derived, in addition to a median simulation. The possibility to include in the analysis different contending models is a key feature of the alternative blueprint for a hydrologic response modeling system recently introduced by *Beven* [2002].

[7] *Krzysztofowicz* [2002] proposed a Bayesian technique for the estimation of uncertainty when the R-R model is used in forecasting, namely the Bayesian system for probabilistic river stage forecasting (Bayesian Forecasting System, BFS) [see also *Krzysztofowicz*, 1999, 2001; *Krzysztofowicz and Kelly*, 2000; *Kelly and Krzysztofowicz*, 2000; *Krzysztofowicz and Herr*, 2001]. The purpose is to produce a probabilistic river stage (or river discharge) forecast based on a probabilistic quantitative precipitation forecasting as an input and a deterministic hydrological model as a means of simulating the response of a river basin to precipitation. The system can work with any hydrological model and aims at estimating the total uncertainty of the forecast, which is considered to be caused by precipitation uncertainty, which is dominant, and hydrologic uncertainty, which is the aggregate of all uncertainties arising from sources other than precipitation uncertainty. The final product of the BFS is the probability distribution of the forecasted river stage (or river flow).

[8] While the above mentioned approaches allow one to estimate the total uncertainty in the model output, several studies were developed to estimate confidence limits of the R-R model parameters, therefore allowing a quantification of parameter uncertainty. For instance, we may mention the NLFIT [*Kuczera*, 1994], PEST [*Brockwell and Davis*, 1991] and BARE [*Thiemann et al.*, 2001] techniques and the recently introduced SCEM-UA method [*Vrugt et al.*, 2003]. This latter allows one to estimate the best parameter set along with its confidence bands by merging the strengths of the Metropolis annealing algorithm, controlled random search, competitive evolution, and complex shuffling.

[9] The aim of this study is to propose an approach for estimating the uncertainty associated with R-R models when these are used in simulation (and hence not in forecasting). A stochastic technique is proposed with low computational demands, which can therefore be easily applied when dealing with highly computer intensive R-R approaches (the spatially distributed models are a typical example). The proposed method can be used in conjunction with any kind of R-R model. Once a model has been selected, the uncertainty assessment procedure requires the prior identification of an optimal parameter set and the computation of the simulation errors in the reproduction of observed river flow data. Hence the technique presented here relies on the concept of optimality instead of equifinality. Model uncertainty is estimated by inferring the statistical properties of the R-R model errors and hence the estimated CIs are related to the selected model structure and parameter set. Since the model error results from the aggregation of different causes of approximation, the technique presented here aims at estimating their total effect without attempting at separating each individual contribution.

[10] The section 2 introduces the proposed uncertainty estimation technique. Section 3 describes the application to the case study. Section 4 is devoted to comparing the obtained CIs with those computed through the GLUE approach. Section 5 outlines the conclusions of the study.

2. A Meta-Gaussian Model for the Estimation of Rainfall-Runoff Simulation Uncertainty

[11] The uncertainty of synthetic data generated by R-R models has to be estimated using a different approach from that used when the model is applied for performing real time forecasting of future river flows. Indeed, in the latter case the user knows the actual river flow observations up to the forecast time t and therefore the uncertainty assessment can take advantage of the knowledge of the past R-R model errors occurred until time t. This is not the case when performing synthetic river flow generation, since the only known variables are those included in the R-R model input.

[12] An approach to quantifying the uncertainty of R-R models, when these are used in simulation (and hence not in forecasting), is proposed here. The aim is to estimate the probability distribution of the model error, conditioned by the value of the contemporary simulated river flow. Once this distribution is known, the user can estimate the upper and lower CI of the simulation, by adding to the synthetic river discharge the estimated upper and lower value of the model error for the given confidence level.

[13] The technique proposed here is based on fitting a probability model to the R-R error series and can therefore be considered as an approach of the second type with respect to the classification of the uncertainty assessment methods mentioned in section 1.

2.1. Estimation of R-R Model Uncertainty

[14] Let us denote with h_t, $t = 1, \ldots n$ an observed time series of river flow containing n observations, which is supposed to be a realization of a stochastic process $H(t)$, and with s_t the corresponding R-R model simulations. Let us refer to the case where s_t is the model output resulting from validation, that is carried out by simulating observed river flows which were not used in calibration. The simulation error is defined as

$$e_t = h_t - s_t, \tag{1}$$

with $t = 1, \ldots n$. Let us suppose that e_t and s_t are realizations from two stochastic processes, denoted with the symbols $E(t)$ and $S(t)$, respectively. The statistical properties of such stochastic processes of course vary depending on the case study, the R-R model and its parameter set. The mean and standard deviation of $E(t)$ and $S(t)$ are indicated with the symbols μ_E, σ_E, μ_S and σ_S, respectively.

[15] The estimation of the probability distribution of $E(t)$, conditioned by the contemporary value of $S(t)$, is

carried out here by using a meta-Gaussian model, which is fully described by *Kelly and Krzysztofowicz* [1997]. Let us denote with $P(E \leq e_t)$ and $P(S \leq s_t)$ the arbitrarily specified marginal cumulative probability distributions of $E(t)$ and $S(t)$, respectively, which are assumed to be strictly increasing and continuous. First, a standard normal quantile transform (NQT) is applied in order to make $P(E \leq e_t)$ and $P(S \leq s_t)$ Gaussian. The NQT is fully described by *Kelly and Krzysztofowicz* [1997]; it can be found in the literature of nonparametric statistics, where it is called the inverse normal score, or simply the normal score. The NQT has been used in hydrological simulation studies as well [see, e.g., *Moran*, 1970; *Hosking and Wallis*, 1988].

[16] In order to explain the NQT, let us refer to $S(t)$. The composition of the inverse Q^{-1} of the standard normal distribution and the marginal probability distribution $P(S \leq s_t)$ defines the NQT of the original variate $S(t)$, which will be referred to as $\mathrm{NQT_S}$,

$$NS(t) = Q^{-1}[P(S \leq s_t)], \tag{2}$$

in which N indicates that the variables are referred to the normalized space.

[17] A marginal probability distribution has to be chosen for $S(t)$. The scientific literature reports numerous examples where the lognormal probability distribution provides a good fit for the frequency of occurrence of observed river flow data [see, e.g., *Kottegoda*, 1980]. However, we observed that in many cases it is difficult to identify a suitable probability model for river flow sequences observed at short time step. Therefore the cumulative probability $P(S \leq s_t)$ of each observation of $S(t)$ is approximated here with the corresponding sample frequency $F(s_t)$, which is estimated using the Weibull plotting position [*Stedinger et al.*, 1993], that is, $F(s_t) = j_t/(n + 1)$. Here j_t is the position occupied by s_t in the sample rearranged in ascending order.

[18] Therefore the $\mathrm{NQT_S}$ involves the following steps: (1) for each s_t the cumulative frequency $F(s_t)$ is computed; (2) for each $F(s_t)$ the standard normal quantile Ns_t is computed and it is associated with the corresponding s_t. Thus a discrete mapping of (2), which gives the $\mathrm{NQT_S}$, is obtained. In order to be able to apply the inverse of the $\mathrm{NQT_S}$, that is, $\mathrm{NQT_S}^{-1}$ for any $NS(t) \in \Re$, linear interpolation is used to connect the points of the discrete mapping previously obtained. The region beyond the minimum and maximum available Ns_t values is covered by linear extrapolation.

[19] Once the normalized series $NS(t)$ and $NE(t)$ are derived by applying the respective transformations $\mathrm{NQT_S}$ and $\mathrm{NQT_E}$, let us make the following hypotheses: (1) $NS(t)$ and $NE(t)$ are stationary and ergodic; and (2) the cross dependence between $NS(t)$ and $NE(t)$ is governed by the normal linear equation

$$Ne_t = \rho_{NE,NS}(0)Ns_t + N\varepsilon_t, \tag{3}$$

where $\rho_{NE,NS}(0)$ is the lag zero Pearson's cross correlation coefficient between $NS(t)$ and $NE(t)$ and $N\varepsilon_t$ is an outcome of the stochastic process $N\Theta(t)$ which is stochastically independent of Ns_t and normally distributed with mean zero and variance $1 - \rho^2_{NE,NS}(0)$. Consequently, the conditional mean and variance of the R-R model error in the normalized space are

$$\mu(NE(t)|Ns_t) = \rho_{NE,NS}(0)Ns_t \tag{4}$$

$$\sigma^2(NE(t)|Ns_t) = \left[1 - \rho^2_{NE,NS}(0)\right] \tag{5}$$

and the transition density is still Gaussian and it is given by

$$r(NE(t)|Ns_t) = \frac{1}{\left[1 - \rho^2_{NE,NS}(0)\right]^{0.5}}\, q\left\{\frac{Ne_t - \rho_{NE,NS}(0)Ns_t}{\left[1 - \rho^2_{NE,NS}(0)\right]^{0.5}}\right\}, \tag{6}$$

where q denotes the standard normal density.

[20] In principle one could consider more complex linear dependence structures than the one given in (3), for instance by allowing a dependence of Ne_t on previous values of $NS(t)$. However, such formulations did not provide significant gains in the analyses performed within this study and therefore were not applied.

[21] In applications, assumption 2 is subjected to testing, as described in section 2.2. The implications of assumptions 1 and 2 are discussed in sections 2.2 and 2.3.

[22] From (4) and (5), remembering that the random variable $E(t)$ is the model error, one can compute the 95% CI for the simulated river flow s_t by

$$s_t^+ = s_t + \mathrm{NQT_E^{-1}}[\mu(NE(t)|Ns_t) + 1.96\sigma(NE(t)|Ns_t)], \tag{7a}$$

$$s_t^- = s_t + \mathrm{NQT_E^{-1}}[\mu(NE(t)|Ns_t) - 1.96\sigma(NE(t)|Ns_t)], \tag{7b}$$

thus obtaining a quantification of the simulation uncertainty.

2.2. Goodness of Fit Checking

[23] The meta-Gaussian approach, described in the previous section, presents some relevant statistical properties which are well summarized by *Kelly and Krzysztofowicz* [1997]. In particular, whereas the regression of $NE(t)$ on $NS(t)$ is Gaussian, linear and with constant dispersion, the regression of $E(t)$ on $S(t)$ may be non Gaussian, non linear and with varying dispersion. This property can provide a first means for verifying the goodness of the fit provided by the meta-Gaussian model, which consists in verifying the hypotheses that condition the validity of the linear regression model of $NE(t)$ on $NS(t)$. The residuals of such a model are given by

$$N\varepsilon_t = Ne_t - \rho_{NE,NS}(0)Ns_t, \tag{8}$$

and should be Gaussian, homoscedastic, with mean 0 and variance $1 - \rho^2_{NE,NS}(0)$. Gaussianity is verified here by applying the Kolmogorov-Smirnov (K-S), Anderson-Darling (A-D) and probability plot correlation coefficient (C-C) tests [*D'Agostino and Stephens*, 1986; *Stedinger et al.*, 1993] and by drawing the normal probability chart. For checking the linearity and homoscedasticity assumptions, the graphical approach proposed by *Cook and Weisberg*

[1994] and used, in a similar context as here, by *Krzysztofowicz and Kelly* [2000] is applied, which consists of drawing the residual plot of $N\varepsilon_t$ versus $\rho_{NE,NS}(0)Ns_t$. Under the target model, the residual plot will have no systematic features. If it is curved, then the regression function is non linear. If the plot is fan shaped or otherwise shows systematic changes in variation across the plot, then the variance of $N\varepsilon_t$ is not constant.

[24] In practical applications the goodness of fit checking may suggest that some of the basic assumptions of the meta-Gaussian model are not satisfied. The lack of fit can be resolved by preliminarily transforming the model error accordingly to the indications given in section 2.3.

2.3. Resolving the Lack of Fit

[25] Among the assumptions which condition the application of the meta-Gaussian approach introduced above, the conditions of normality and heteroscedasticity of the residuals given by (8) are the ones which result more frequently violated. It is well known that non normality and unequal variances are often related and that transformations to stabilize the variance often help to correct non normality as well. This type of lack of fit may occur, for instance, when σ_E is dependent on $S(t)$ but independent of μ_E, since the NQT is ineffective in reducing this kind of heteroscedasticity, which, as a matter of fact, is often present in the samples of R-R model errors. In order to overcome this problem, a preliminary variance stabilizing transformation can be performed on the $E(t)$ series, which is structured accordingly to the following steps.

[26] First of all, let us compute the absolute value of the deviation from the mean value of the $E(t)$ record,

$$e_t^* = |e_t - \mu_E|. \tag{9}$$

thus obtaining a realization of the stochastic process $E^*(t)$. Since the variability of $E(t)$ is increasing with s_t, the mean value μ_{E^*} of $E^*(t)$ is expected to increase with s_t as well, accordingly to the relationship

$$\mu_{E^*} = f(s_t), \tag{10}$$

in which $f(\cdot)$ is supposed to be a strictly monotonic function. In order to estimate the shape of $f(\cdot)$, the NQT is applied to $E^*(t)$ and $S(t)$ following the same procedure described in section 2.1 and a linear regression is performed in the normalized space, thus obtaining

$$f(s_t) = \mathrm{NQT}^{-1}_{\mathrm{E}^*}[k + cNs_t], \tag{11}$$

in which k and c are the coefficients of the above regression of $NE^*(t)$ on $NS(t)$.

[27] Finally, the error series $E(t)$ is transformed accordingly to the variance stabilizing transformation (VST) expressed by

$$e'_t = \frac{e_t - \mu_E}{f(s_t)}, \tag{12}$$

thus obtaining a realization of the stochastic process $E'(t)$. Reapplying the meta-Gaussian approach to $E'(t)$ and $S(t)$ allows one to verify the effectiveness of the above variance stabilizing transformation, by checking that the goodness of fit tests described in section 2.2 are now satisfied. The confidence limits of the simulated river flow s_t can be derived by applying the inverse of the above VST, that is,

$$s_t^+ = s_t + \mathrm{NQT}^{-1}_{\mathrm{E}'}[\mu(NE'(t)|Ns_t) + 1.96\sigma(NE'(t)|Ns_t)]f(s_t) + \mu_E, \tag{13a}$$

$$s_t^- = s_t + \mathrm{NQT}^{-1}_{\mathrm{E}'}[\mu(NE'(t)|Ns_t) - 1.96\sigma(NE'(t)|Ns_t)]f(s_t) + \mu_E. \tag{13b}$$

[28] Whenever the VST described above is ineffective, the lack of fit may be due, for instance, to the presence of non stationarity in the dependence structure between $NE(t)$ and $NS(t)$. In this case, one may try to fit the meta-Gaussian model by referring to a limited sample of model errors. Since in most practical applications one is interested in estimating the CI for peak discharges, it might be advisable to focus on the higher flows only, therefore considering a reduced joint sample of $NE(t)$ and $NS(t)$. This procedure may lead to improving the goodness of fit. Of course the user has to make sure that the size of the reduced joint sample is large enough to provide a reliable estimation.

2.4. Meta-Gaussian Model: Some Relevant Features and Limitations

[29] The meta-Gaussian model proposed in section 2.1 does not impose any restriction on the marginal dependence structure of $NE(t)$ and $NS(t)$ and therefore it is not influenced by the presence of serial correlation. The reliability of the meta-Gaussian approach might be affected by the hypothesis of stationarity and ergodicity for $NE(t)$ and $NS(t)$, which is necessarily introduced in order to be able to estimate the statistical properties of such random variables on the basis of a finite and perhaps limited sample. In order to limit the resulting approximation, it is advisable to fit the meta-Gaussian model on an extended data set, which should include a wide variety of hydrological scenarios.

[30] The presence of non stationarity in the model errors can be qualitatively checked by carrying out a validation of the meta-Gaussian model results (see section 3.3), that allows one to check their reliability in real world applications. It is significant to point out that non stationarity might be due to R-R model inadequacy or unreliability of input and output data. Therefore in these circumstances the lack of fit could be resolved by changing the R-R modeling approach or by reconsidering the hydrological data set. However, this is not always the case since non stationarity might originate from other reason, for instance a variation of river banks management or a change in the watershed land-use. The meta-Gaussian approach does not allow one to identify the reasons for the presence of non stationarity and therefore it is not a means for verifying the suitability of a given R-R model. It gives an estimation of the uncertainty associated to the simulations provided by an assigned model and parameter set. Its lack of fit, which might be due to the presence of non stationarity and therefore even to model inadequacy, only indicates that the R-R uncertainty cannot be reliably estimated.

[31] It is worth remarking that different model structures affect the analysis only through their simulation errors.

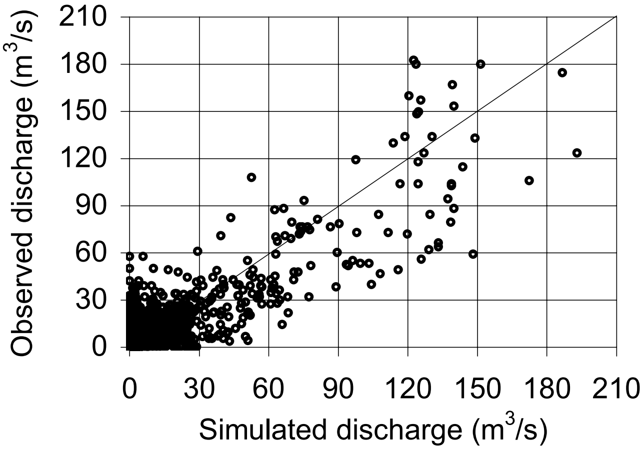

Figure 1. Dispersion diagram of observed versus simulated by the distributed model hourly discharges (validation mode, 1994–1997).

Unsatisfactory R-R models are expected to be characterized by wider CIs but do not compromise the applicability of the proposed uncertainty estimation technique, unless nonstationarity is induced in the model errors.

[32] Particular care has to be taken when trying to apply the meta-Gaussian approach outside the range of s_t covered by the (s_t, e_t) couples used in the calibration phase. Like any extrapolation, this procedure may result in unsatisfactory approximations if non stationarity is present in the dependence structure and probability distribution of $E(t)$ and $S(t)$.

[33] Finally, it has to be considered that the meta-Gaussian approach leads to estimating the uncertainty in an aggregated solution. Input and output uncertainty, as well as parameter uncertainty, are accounted for implicitly, without separating their individual contribution. Therefore the meta-Gaussian approach reliability is expected to be conditioned by changes in the different uncertainty sources. Variations in the reliability of input and output variables, for instance those due to modification of the gauging methods, are likely to affect the R-R model performances, as well as using synthetic input data instead of observed ones, as it is often done when estimating design variables. In all these cases the effects of non stationarity in the uncertainty sources should be carefully evaluated.

3. Application of the Proposed Technique to an Italian Case Study

[34] A detailed description of the application of the meta-Gaussian approach to the case study of the Samoggia River, located in northern Italy, is provided in this section. The R-R transformation was simulated by a lumped and a spatially distributed R-R model.

[35] Calibration of the R-R models, and the calibration and validation of the meta-Gaussian approach were performed referring to three different observation periods, in order to be able to assess the capability of the proposed uncertainty assessment method with respect to a real world application. There is a slight overlap between the calibration period of the R-R and meta-Gaussian models, which is limited to a single flood event.

3.1. Study Watershed

[36] The Samoggia River flows northward across the Apennines Mountains. The watershed is mostly mountainous, with a maximum and minimum altitudes of 850 and 50 m above sea level (asl) respectively, a mainstream length of 60 km and a basin extension at the cross river section of Calcara of 158 km^2. The Samoggia River basin is monitored by rain gauges and hydrometric stations managed by the Italian National Hydrographic Service. For the purposes of the present analysis, historical data of hourly river discharges at Calcara and hourly rainfall depths over the basin have been collected, from 1 January 1994 through 31 December 1997. The rainfall data have been observed in 3 rain gauges located in the basin or in the immediate vicinity. These are placed at Monte San Pietro (317 m asl), Montepastore (596 m asl) and Monteombraro (727 m asl). Hourly temperature data recorded at Monteombraro in the period 1994–1997 are also available. An extensive database of soil texture, soil type and soil use at the local scale is available as well, retrieved from surveys carried out in the last 10 years [*Brath et al.*, 2002, 2003].

[37] Rainfall and river stage are monitored by continuous time, automatically recording gauges. The conversion from water level to river discharge is obtained by means of a rating curve derived on the basis of flow velocity measures and field surveys of the cross river section geometry. The rating curve is periodically verified and did not change during the observation period considered here. It is not expected to undergo significant variations since it refers to cross-river section with stable geometry.

3.2. Rainfall-Runoff Models

[38] The simulation procedure is based on the use of a spatially distributed and a lumped R-R model, which provided a continuous simulation at a hourly time step.

3.2.1. Spatially Distributed Model

[39] The spatially distributed model applied here was developed at the University of Bologna and can be considered a mixed conceptual and physically based approach. The model divides the basin in square cells, coinciding with the pixels of the digital elevation model (DEM) which describes the basin topography. The river network is auto-

Table 1. Application of the Meta-Gaussian Approach for Estimating the Uncertainty of ADM and Distributed Models Applied to the Case Study of the Samoggia River Basin[a]

Model	$\rho_{NE,NS}(0)$	Residual Mean	Residual Variance	K-S Statistic	A-D Statistic	C-C Statistic
ADM	0.166	0.00 (0.00)	0.9722 (0.9725)	0.030 (0.042)	0.690 (2.492)	0.9985 (0.9982)
Distributed	−0.144	0.00 (0.00)	0.9788 (0.9781)	0.025 (0.042)	0.460 (2.492)	0.9990 (0.9982)

[a]Coefficient of correlation $\rho_{NE,NS}(0)$, residual mean and variance (against their theoretical values reported in parentheses), and K-S, A-D, and C-C statistics on the residual series (against their 95% limit values reported in parentheses; note that the tests were carried out by using theoretical values of distribution parameters) are shown.

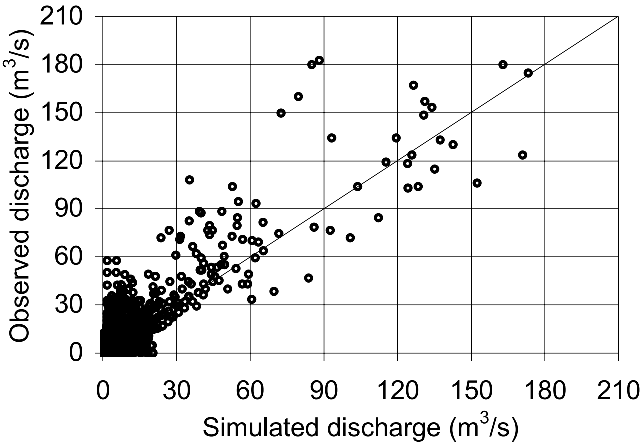

Figure 2. Dispersion diagram of observed versus simulated by the ADM model hourly discharges (validation mode, 1994–1997).

matically extracted from the DEM itself by applying the D-8 method [*Band*, 1986], which makes it possible to estimate the flow paths. Each cell receives water from its up-slope neighbors and discharges to its down-slope neighbor. Distinction between hillslope rill and network channel is based on the concept of constant critical support area [*Montgomery and Foufoula-Georgiou*, 1993]. Accordingly, rill flow is assumed to occur in each cell where the upstream drainage area does not exceed 0.5 km^2, while channel flow occurs otherwise.

[40] The interaction between soil, vegetation and atmosphere is modeled by applying a modified CN approach in each DEM cell [*Soil Conservation Service*, 1972], which allows one to compute infiltration and surface runoff continuously in time [*Brath et al.*, 2003]. Effective evapotranspiration is computed at a local scale by applying the radiation method [*Doorembos et al.*, 1984].

[41] Surface and subsurface flows are propagated downstream by applying the variable parameters Muskingum-Cunge model. Extensive details are given by *Cunge* [1969] and *Orlandini et al.* [1999], for the surface and subsurface propagation, respectively.

[42] Some of the model parameters, such as the roughness of the river network, have a physical meaning at a point scale and were estimated on the basis of in situ surveys, therefore assuming that they are still valid at the grid scale. It was necessary to optimize six parameters by means of a trial and error procedure, through manual calibration performed by comparing observed and simulated hourly river flows of the flood event that occurred on 8 October 1996. The model was subsequently validated by simulating the hourly discharges of the whole 1994–1997 period. A dispersion diagram of the observed versus simulated hourly flows is reported in Figure 1. The *Nash and Sutcliffe* [1970] efficiency for the simulation of the whole period is 0.68. For the periods 1994–1995 and 1996–1997 the efficiency amounts to 0.60 and 0.81 respectively.

3.2.2. Lumped Model

[43] The ADM model [*Franchini*, 1996] was used in order to provide a lumped schematization of the R-R transformation. The ADM model is derived from the Xinanjiang model [*Zhao et al.*, 1980], and it is based on the same concept of probability distributed soil moisture storage capacity. The model is divided into two main blocks: the first represents the water balance at soil level, that is, the balance between the moisture content and the incoming (precipitation) and outgoing (evapotranspiration, surface runoff, interflow and base flow) water flows, and it is characterized by seven parameters. The second block represents the transfer of runoff production to the basin outlet and involves four parameters. The soil, in turn, is divided into two zones: the upper zone produces surface and subsurface runoff (interflow), while the lower zone produces base flow runoff. The transfer of these components to the outlet section takes place in two distinct stages, the first representing the flow along the hillslopes toward the channel network, while the second represents the flow along the channel network toward the basin outlet. Surface runoff and interflow are summed and transferred along the hillslopes, with a transfer function obtained as the solution of the convective diffusive flow equation when a lateral, uniformly distributed input is considered, which involves two parameters. An analogous function is used for the transfer of the total runoff along the river network to the closure section, which involves two additional parameters.

[44] The model has a total of eleven parameters, which were estimated by automatically optimizing the simulation of river flows of the flood event that occurred on 8 October 1996 (the same data set used for calibrating the spatially distributed model). Automatic calibration was performed by using the SCE-UA method [*Duan et al.*, 1992]. The model was validated by simulating the 1994–1997 flows. The efficiency was 0.67 for the whole validation period and 0.69 and 0.65 for the periods 1994–95 and 1996–97, respectively. Figure 2 reports the dispersion diagram of the observed versus ADM-simulated hourly flows.

[45] The R-R model efficiencies are to be evaluated considering that both models were calibrated by using only the data from one flood event. Calibrating the model over a longer period would undoubtedly lead to an improvement of the validation efficiency, but this option would have restricted the sample size of the model errors available for calibrating and validating the meta-Gaussian approach. It should be noted that the present analysis is not aimed at obtaining a good fit of the observed river flows, while the

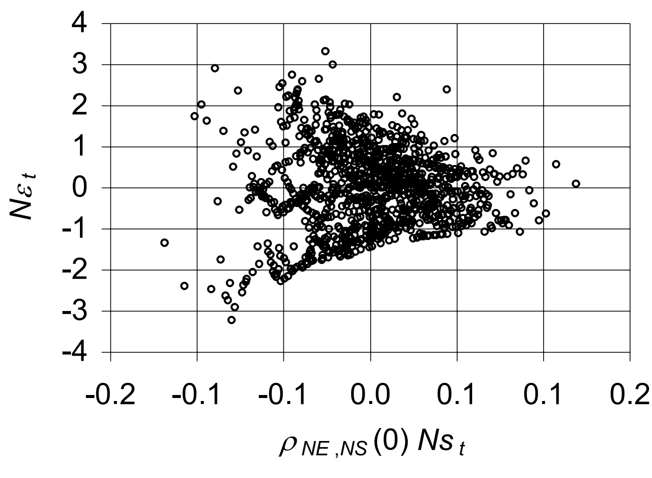

Figure 3. ADM model, years 1996 and 1997. Residual plot of the linear regression of $NE(t)$ on $NS(t)$.

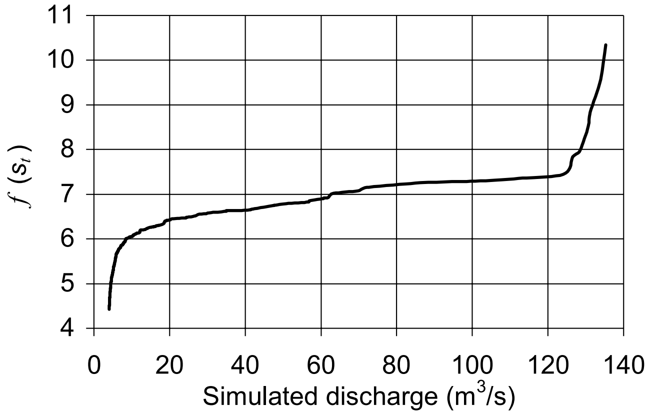

Figure 4. ADM model, years 1996 and 1997. Plot of the function $f(s_t)$ used in the variance stabilizing transformation.

Samoggia case study is not aimed at performing a comparison between the performances of the two R-R models (which were comparable in terms of model efficiency). The case study is developed to prove the capability of the meta-Gaussian approach to infer the uncertainty of the R-R models simulations, regardless of its actual magnitude.

3.3. Analysis of the Rainfall-Runoff Model Uncertainty

[46] The uncertainty in the simulation of the river flow data was evaluated by applying the meta-Gaussian approach described in section 2.1, which was calibrated and validated using separate records of hourly river flow data, as described here below. Calibration and observation periods are different but close in time and there seems to be no evidence of non stationarity in the observable uncertainty (observed input and output data), since the gauging instruments and methods were not altered. Therefore a comparison of the estimated CI with actual river flow data in validation mode may provide indications about the reliability of the uncertainty estimation procedure.

[47] From the validation phase of the distributed and ADM models applied to the Samoggia River basin, two samples of 35064 observations of simulated river flows and corresponding model errors were obtained, for the years 1994, 1995, 1996 and 1997. The samples referring to the years 1996 and 1997 were used for calibrating the respective meta-Gaussian models, which were in both cases (distributed and ADM models) validated using the data of the years 1994 and 1995. To limit the possible effects of non stationarity (see section 2.3), the meta-Gaussian model was applied on the couples (s_t, e_t) characterized by simulated river flows values greater than 4.00 and 3.45 m^3/s, for the simulations provided by distributed and ADM models, respectively, thus concentrating on samples containing 1037 and 1033 data.

[48] Looking at the residual plot of $N\varepsilon_t$ versus $\rho_{NE,NS}(0)Ns_t$ (see Figure 3 for the residuals of the ADM model), one can note a fan shaped pattern, which highlights the presence of residual heteroscedasticity for both models, which compromises the validity of the meta-Gaussian assumption. Therefore according to the guidelines provided in section 2.3, the VST was applied on the model errors in order to stabilize their variance. Figure 4 depicts the shape of the estimated function $f(s_t)$ given by (11), within the range of the s_t values comprised in the analyzed (s_t, e_t) couples. Figure 5 shows an example of application of the NQT to the transformed (after the VST) ADM model error series.

[49] Application of the meta-Gaussian model after the VST led to obtaining values of $\rho_{NE,NS}$ (0), residual mean and variance, K-S, A-D and C-C tests statistics reported in Table 1.

[50] It can be seen that all the tests resulted in not rejecting normality at the 95% confidence level for the residuals given by (8). Referring to the ADM model, the residual plot, showed in Figure 6, provides convincing support to the hypothesis of a linear and homoscedastic regression of $NE'(t)$ on $NS(t)$. Figure 7 reports the normal probability plot for the residual series, which confirms the good agreement with the hypothesis of Gaussian residuals. The goodness of the fit was also confirmed for the spatially distributed model.

[51] The meta-Gaussian model was validated using the 1994 and 1995 hourly discharge simulations. It has to be noted that the distributed model performed significantly worse, and hence with a greater uncertainty, in 1994–1995 than in 1996–1997 (see section 3.2.1). Validation of the meta-Gaussian approach is therefore performed in a non favorable context in this case, since the assumption of stationarity of model performances appears to be weak. This is probably a consequence of having used a limited data set for R-R model calibration (see also section 3.2.2).

[52] Figures 8 and 9 show dispersion diagrams of observed versus simulated river flows, referring to validation periods of both models, along with the respective estimated 95% CIs. If the pattern of the model error was well captured by the meta-Gaussian approach, approximately 95% of the observed data should lie inside the computed CIs, while about 5% should lie outside and should be with good approximation uniformly distributed above the upper and below the lower CI and along the range of discharges the CIs were computed for. Therefore checking the number and distribution of the observed data lying outside the CIs allows one to obtain a qualitative evaluation of the CI reliability. A substantial violation of the above prerequisites points out an underestimation or overestimation of the CI width and may be a significant clue of the presence of non stationarity in $E(t)$.

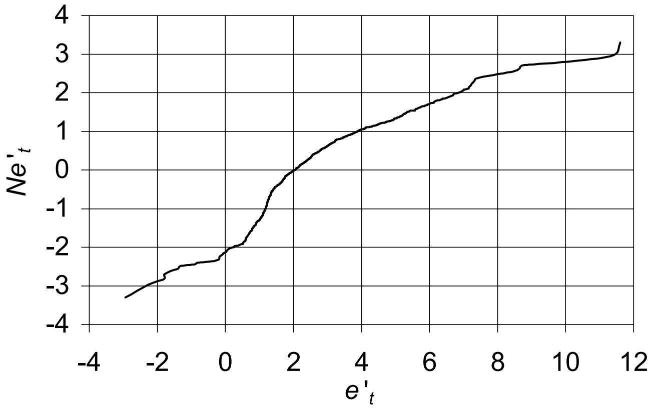

Figure 5. ADM model, years 1996 and 1997. Normal quantile transform for the model error series (after application of the variance stabilizing transformation).

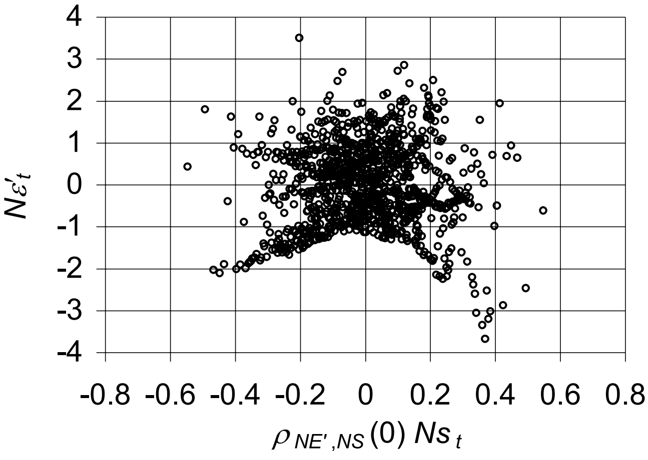

Figure 6. ADM model, years 1996 and 1997. Residual plot of the linear regression of NE' (t) on $NS(t)$, after performing the preliminary variance stabilizing transformation.

[53] It can be noted that the meta-Gaussian approach successfully captures the bias of the simulation, which has opposite signs for the two models. This is clearly displayed by the progress of the respective confidence bands. The pattern of the model error seems to be correctly captured as well; in particular, the meta-Gaussian approach well represents the tendency of the ADM error to increase less significantly for increasing river flows.

[54] Looking at the validation plot of the distributed model, one sees that about 15% of the validation data fall outside the computed 95% CIs (11% and 4% below the lower and above the upper CI, respectively); these points are quite uniformly distributed along the range of the observed flows. Therefore there appears to be an underestimation of the width of the confidence bands, in particular of the lower one. This may be caused by the higher magnitude of the model errors which occurred in 1994 and 1995, compared to those that occurred in 1996 and 1997. It is therefore recognized that the distributed model performed significantly worse in the years 1994 and 1995 with respect to 1996 and 1997. As a matter of fact, it is confirmed that non stationarity of the R-R model performances may affect the uncertainty estimation.

[55] Looking at the validation plot of the ADM model one sees that 10% of the validation data fall outside the computed 95% CIs (7% and 3% below the lower and above the upper CI, respectively). Therefore the width of the upper CI appears to be satisfactorily estimated. As far as the lower CI is concerned, a more detailed analysis reveals that almost all the outside data are in the range 0–15 m^3/s. Hence the width of the lower CI seems to be underestimated only for the very low river flows.

[56] Figure 10 depicts a focus on the flood event occurring on 23 June 1995 along with the respective 95% CIs. This event was unsatisfactorily simulated by both ADM and distributed models, which produced a significant underestimation and overestimation of the peak flows, respectively. The confidence bands of the distributed model appear to be wider for the higher river flows; this is a consequence of the higher magnitude of the distributed model errors, which is confirmed by looking at the validation plots in Figures 8 and 9. Looking for the observations where the meta-Gaussian approach failed in inferring the uncertainty at a 95% confidence level, it can be noticed that a part of the observed rising limb is not comprised within the distributed model confidence bands, and the peak flow is not included in the ADM CIs. However, the percentage of observations which fall within the estimated CIs is satisfactory, even for this event which was poorly reproduced by both models.

4. Comparison With the GLUE Approach

[57] A further analysis was made in order to compare the confidence bands estimated through the meta-Gaussian approach with those provided by the GLUE technique [*Beven and Binley*, 1992]. The comparison was carried out referring to the Samoggia River basin, for which GLUE confidence bands were computed for the simulation of the flood event occurring on 23 June 1995. The likelihood of the different models and parameter sets considered within GLUE should be computed on the basis of a statistically significant data set, in order to minimize the effects due to sampling variability and possible non stationarity. However in the context of the present analysis, in order to provide a homogeneous comparison, a situation of historical data scarcity is emulated and the likelihood measure used within GLUE is computed by only using the continuous hydrograph for the flood event of 8 October 1996. This is the same data set that was used for finding the optimal parameter set of the R-R models, as described in section 3.2. In order to account for model uncertainty, two different models (ADM and distributed, see section 3.2) were incorporated in the GLUE approach. Uncertainty estimation was performed accordingly to the following steps: (1) a sensitivity analysis was carried out for both models in order to identify the parameters which are more influent on the magnitude of the peak river flows. Accordingly, six parameters were selected for both models. (2) Plausible ranges for each parameter were selected, by allowing a variation of ±50% around the optimal values identified when calibrating the models as described in section 3.2. (3) Uniform sampling was performed in order to generate candidate parameter sets, which were used to simulate the October 1996 flood. (4) The *Nash*

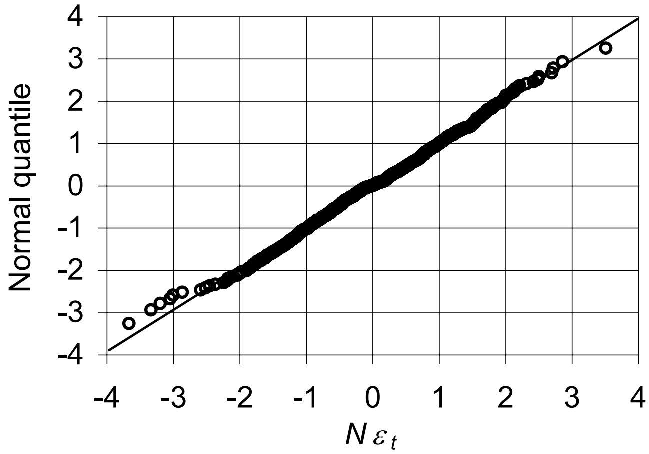

Figure 7. ADM model, years 1996 and 1997. Normal probability plot of the residual of the regression of NE' (t) on $NS(t)$.

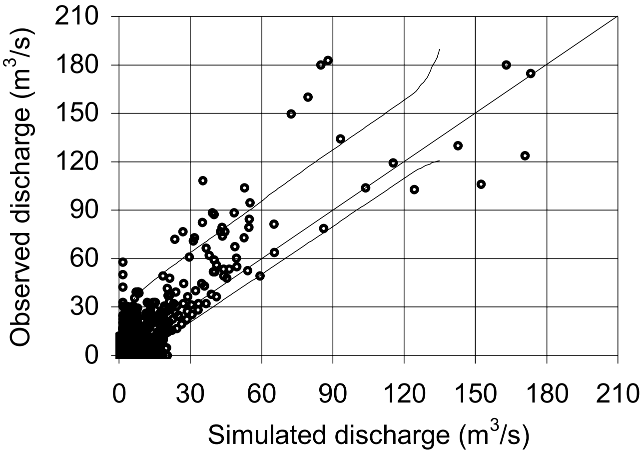

Figure 8. Dispersion diagram of observed versus simulated by the ADM model hourly discharges, along with the 95% confidence bands estimated by applying the meta-Gaussian approach (validation mode, 1994 and 1995).

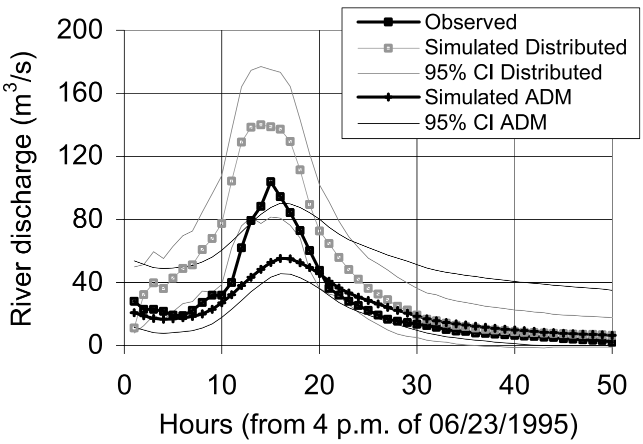

Figure 10. ADM and distributed model. Observed and simulated flows of the flood event of 23 June 1995, along with the respective 95% CIs computed with the meta-Gaussian approach.

and Sutcliffe [1970] efficiency was computed for each simulation. All the parameter sets which did not provide an efficiency of at least 90% were rejected as non behavioral. Uniform sampling described in step 3 was repeated until a collection of 3000 behavioral parameter sets was reached; 1920 and 1080 of the related simulations were obtained through the lumped and distributed models, respectively. (5) Likelihood weights were computed by applying the procedure used, among others, by *Freer et al.* [1996] and *Cameron et al.* [1999] and described in section 1. (6) Likelihood weighted uncertainty bounds, as well as a median simulation, were then computed referring to the flood event occurring on 23 June 1995. These are reported in Figure 11, along with the corresponding observed river flows.

[58] Comparing Figures 10 and 11, one may assess the differences among the CIs estimated with the meta-Gaussian and GLUE approaches. Before attempting a comparison among the results, one should note that the two techniques are very different in nature. The most relevant difference is perhaps that GLUE, as performed here, accounts for model uncertainty by considering only two candidate models. It is quite possible that considering additional model structures would lead to different CIs. Moreover, GLUE, as applied here, does not account for input uncertainty, which is instead implicitly considered by the meta-Gaussian approach, which treats all the sources of uncertainty in an aggregated form. Therefore one would expect that the width of the GLUE CI, as computed here, is smaller with respect to the meta-Gaussian CI. In fact, looking at Figures 10 and 11, it can be noticed that the width of the GLUE CI, referring to the peak river flows, is slightly smaller with respect to the one related to the distributed model. It results fairly comparable to the one provided by the meta-Gaussian approach for the ADM simulation of the peak flows, while it becomes significantly smaller for decreasing river flows. The GLUE CI do not include the peak river flow, thus confirming that the considered flood event is indeed difficult to simulate using the considered R-R models when calibration (in the case of the meta-Gaussian approach) and computation of the likelihood measures (for the GLUE approach) are performed using the selected data set, which refers to only one flood event.

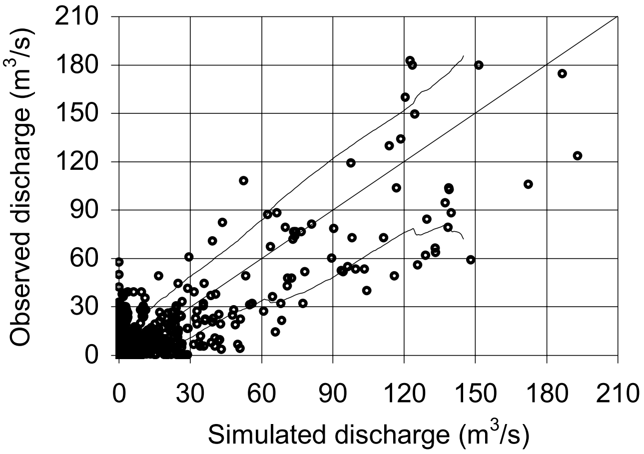

Figure 9. Dispersion diagram of observed versus simulated by the distributed model hourly discharges, along with the 95% confidence bands estimated by applying the meta-Gaussian approach (validation mode, 1994 and 1995).

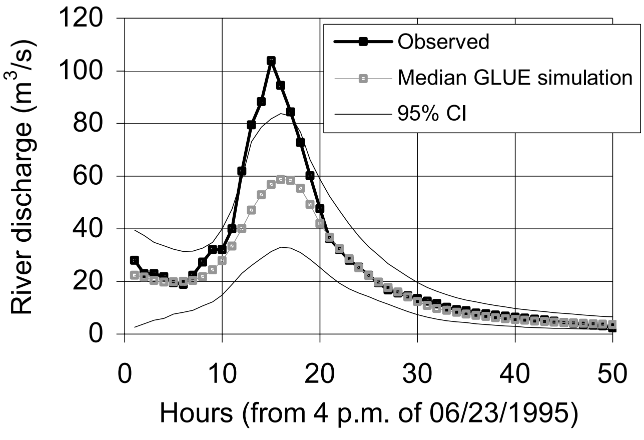

Figure 11. ADM and distributed model. Median simulation and 95% CI of the flood event of 23 June 1995, computed by means of the GLUE approach.

[59] Overall, even if a comparison among the GLUE and meta-Gaussian CIs is indeed difficult to carry out due to their different nature, the coherence among them seems to support the reproducibility of the estimated R-R model uncertainty.

5. Conclusions

[60] This study presents a technique for assessing the uncertainty of rainfall-runoff simulations, which is estimated in an aggregated solution, without attempting to separate the contribution given by the different uncertainty sources. Model uncertainty estimated here may arise from misspecification of the R-R model or inappropriateness of its mathematical scheme to represent the hydrological processes which take part at the R-R transformation, as well as from input and output uncertainty.

[61] The uncertainty estimation is carried out by analyzing the properties of the sample frequency distribution of the model errors in reproducing observed historical river flow data, and makes use of a meta-Gaussian model in order to estimate the probability distribution of the R-R model error conditioned by the value of the simulated river flow. The standard normal quantile transform is used in order to make the marginal probability distributions of both river flows and model errors Gaussian. Straightforward application and low computer times are the main prerogatives of the meta-Gaussian approach.

[62] The proposed technique is statistically based and goodness of fit tests have been discussed which can be applied in order to verify whether a reliable estimation is provided. However, even if the fit turns out to be satisfactory, one must be aware that the computed uncertainty bounds may be too optimistic, in the sense that they might be too narrow. This is a common problem when estimating the uncertainty associated with rainfall-runoff models and forecasting models in general. As long as 30 years ago, *Wallis* [1974] noted that it is "a common experience for models to have worse error variances than they should when used in forecasting outside the period of fit." Subsequent empirical studies have generally confirmed this statement [see, e.g., *Gardner*, 1988]. An interesting discussion about this issue is provided by *Chatfield* [1993], who identified some possible reasons which may, in general, cause the estimated CIs to be too narrow. Some of these apply to the case of rainfall-runoff modeling. For instance, it is worth noting that the prediction model and its parameters may change, either during the period of fit or in the future, especially if calibration is performed by using a limited data set. The meta-Gaussian approach indeed rests on the assumption that the model error is ergodic which, as a matter of fact, is rarely the case (see section 3.2.2). Moreover, there may be outliers or error in the data.

[63] The meta-Gaussian approach has been applied in order to estimate the confidence limits of the simulated river flows referring to the case study of an Italian river basin, the discharges of which were simulated by two different rainfall-runoff models. The results show that the derived confidence limits are consistent when compared with the order of magnitude of the model errors in validation mode, thus showing that the proposed approach may be valuable for quantifying the R-R models uncertainty with a reasonable degree of approximation.

[64] **Acknowledgments.** The authors wish to thank Roman Krzysztofowicz for providing suggestions about the goodness of fit tests for meta-Gaussian models. Insightful comments provided by two anonymous referees and the Associate Editor of the journal were greatly appreciated and lead to a substantial improvement of the framework of the analysis and readability of the manuscript. The work presented here has been partially supported by the Italian National Research Council - Group for the Prevention of Hydrogeological Disasters (contract 01.01014.PF42) and by the Italian Government through the national grant to the research program on "Hydrological Safety of Impounded Rivers".

References

Band, L. E. (1986), Topographic partition of watersheds with digital elevation models, *Water Resour. Res.*, *22*, 15–24.
Beven, K. J. (1993), Prophesy, reality and uncertainty in distributed hydrological modeling, *Adv. Water Resour.*, *16*, 41–51.
Beven, K. J. (2001), How far can we go in distributed hydrological modeling, *Hydrol. 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. Processes*, *16*, 189–206.
Beven, K. J., and A. Binley (1992), The future of distributed models: Model calibration and uncertainty prediction, *Hydrol. Processes*, *6*, 279–298.
Blaszkova, S., and K. Beven (2002), Flood frequency estimation by continuous simulation for a catchment treated as ungauged (with uncertainty), *Water Resour. Res.*, *38*, 1–14.
Brath, A., and R. Rosso (1993), Adaptive calibration of a conceptual model for flash flood forecasting, *Water Resour. Res.*, *29*, 2561–2572.
Brath, A., A. Montanari, and G. Moretti (2002), On the use of simulation techniques for the estimation of peak river flows, paper presented at the International Conference on Flood Estimation, Int. Comm. for the Hydrol. of the Rhine Basin, Lelystad, Netherlands.
Brath, A., A. Montanari, and G. Moretti (2003), Assessing the effects on the flood risk of the land-use changes in the last five decades: An Italian case study, *IAHS Publ.*, *278*, 435–441.
Brockwell, P. J., and R. A. Davis (1991), *Time Series: Theory and Method*, 2nd ed., Springer-Verlag, New York.
Cameron, D. S., K. J. Beven, J. Tawn, S. Blazkova, and P. Naden (1999), Flood frequency estimation by continuous simulation for a gauged upland catchment (with uncertainty), *J. Hydrol.*, *219*, 169–187.
Cameron, D. S., K. J. Beven, and P. Naden (2000), Flood frequency estimation by continuous simulation under climate change (with uncertainty), *Hydrol. Earth Syst. Sci.*, *4*, 393–405.
Chatfield, C. (1993), Calculating interval forecast, *J. Bus. Econ. Stat.*, *11*, 121–144.
Cook, R. D., and S. Weisberg (1994), *An Introduction to Regression Graphics*, John Wiley, Hoboken, N. J.
Cunge, J. A. (1969), On the subject of a flood propagation computation method (Muskingum Method), *J. Hydraul. Res.*, *7*, 205–230.
D'Agostino, R. B., and M. A. Stephens (1986), *Goodness-Of-Fit Techniques*, Chapman and Hall, New York.
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. Drain. Pap.*, *24*, 144 pp.
Duan, Q., S. Sorooshian, and V. K. Gupta (1992), Effective and efficient global optimization for conceptual rainfall-runoff models, *Water Resour. Res.*, *28*, 1015–1031.
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*, 21–39.
Freer, J., K. J. Beven, and B. Ambroise (1996), Bayesian uncertainty in runoff prediction and the value of data: An application of the GLUE approach, *Water Resour. Res.*, *32*, 2163–2173.
Gardner, E. S. (1988), A simple method for computing prediction intervals for time-series forecasts, *Manage. Sci.*, *34*, 541–546.
Hosking, J. R. M., and J. R. Wallis (1988), The effect of intersite dependence on regional flood frequency analysis, *Water Resour. Res.*, *24*, 588–600.
Kelly, K. S., and R. Krzysztofowicz (1997), A bivariate meta-Gaussian density for use in hydrology, *Stochastic Hydrol. Hydraul.*, *11*, 17–31.
Kelly, K. S., and R. Krzysztofowicz (2000), Precipitation uncertainty processor for probabilistic river stage forecasting, *Water Resour. Res.*, *36*, 2643–2653.
Kottegoda, N. T. (1980), *Stochastic Water Resources Technology*, McMillan, New York.

Krzysztofowicz, R. (1999), Bayesian theory of probabilistic forecasting via deterministic hydrologic model, *Water Resour. Res.*, *35*, 2739–2750.
Krzysztofowicz, R. (2001), The case for probabilistic forecasting in hydrology, *J. Hydrol.*, *249*, 2–9.
Krzysztofowicz, R. (2002), Bayesian system for probabilistic river stage forecasting, *J. Hydrol.*, *268*, 16–40.
Krzysztofowicz, R., and H. D. Herr (2001), Hydrologic uncertainty processor for probabilistic river stage forecasting: Precipitation-dependent model, *J. Hydrol.*, *249*, 46–68.
Krzysztofowicz, R., and K. S. Kelly (2000), Hydrologic uncertainty processor for probabilistic river stage forecasting, *Water Resour. Res.*, *36*, 3265–3277.
Kuczera, G. (1994), NLFIT: A Bayesian nonlinear regression program suite, users manual, Dep. of Civ. Eng. and Surv., Univ. of Newcastle, Callaghan, Australia.
Kuczera, G., and E. Parent (1998), Monte Carlo assessment of parameter uncertainty in conceptual catchment models: the Metropolis algorithm, *J. Hydrol.*, *211*, 69–85.
Loukas, A., L. Vasiliades, and N. R. Dalezios (2002), Climatic impacts on the runoff generation processes in British Columbia, Canada, *Hydrol. Earth Syst. Sci.*, *6*, 211–228.
Metropolis, N. (1987), The beginning of the Monte Carlo method, *Los Alamos Sci.*, *15*, 125–130.
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*, 1035–1044.
Montgomery, D. L., and E. Foufoula-Georgiou (1993), Channel network source representation using digital elevation models, *Water Resour. Res.*, *29*, 3925–3934.
Moran, P. A. P. (1970), Simulation and evaluation of complex water system operations, *Water Resour. Res.*, *6*, 1737–1742.
Nash, J. E., and J. V. Sutcliffe (1970), River flow forecasting through conceptual models 1: A discussion of principles, *J. Hydrol.*, *10*, 282–290.
Orlandini, S., A. Perotti, G. Sfondrini, and A. Bianchi (1999), On the storm flow response of upland Alpine catchments, *Hydrol. Processes*, *13*, 549–562.
Singh, V. P., and D. A. Woolhiser (2002), Mathematical modeling of watershed hydrology, *J. Hydrol. Eng.*, *7*, 270–292.
Soil Conservation Service (1972), *National Engineering Handbook*, sect. 4, *Hydrology*, U.S. Dep. of Agric., Washington, D. C.
Stedinger, J. R., R. M. Vogel, and E. Foufoula-Georgiou (1993), Frequency analysis of extreme events, in *Handbook of Hydrology*, edited by D. R. Maidment, pp. 18.1–18.66, McGraw-Hill, New York.
Thiemann, M., M. Trosset, H. Gupta, and S. Sorooshian (2001), Bayesian recursive parameter estimation for hydrologic models, *Water Resour. Res.*, *37*, 2521–2536.
Vrugt, J. A., H. V. Gupta, W. Bouten, and S. Sorooshian (2003), A shuffled complex evolution Metropolis algorithm for optimization and uncertainty assessment of hydrological model parameters, *Water Resour. Res.*, *39*(6), 1201, doi:10.1029/2002WR001642.
Wallis, K. F. (1974), Discussion of "Experience with forecasting univariate time-series and the combination of forecasts" by P. Newbold and C. W. J. Granger, *J. R. Stat. Soc., Ser. A*, *137*, 155–156.
Yar, M., and C. Chatfield (1990), Prediction intervals for Holt-Winters forecasting procedure, *Int. J. Forecast.*, *6*, 1–11.
Zhao, R. J., Y. L. Zhuang, L. R. Fang, X. R. Liu, and Q. S. Zhang (1980), The Xinanjiang model, *IAHS Publ.*, *129*, 351–356.

A. Brath and A. Montanari, Faculty of Engineering, University of Bologna, Via del Risorgimento 2, I-40136 Bologna, Italy. (alberto.montanari@mail.ing.unibo.it)