WATER RESOURCES RESEARCH, VOL. 43, W05434, doi:10.1029/2006WR005184, 2007

Calibration of hydrological models in the spectral domain: An opportunity for scarcely gauged basins?

Alberto Montanari[1] and Elena Toth[1]

Received 9 May 2006; revised 1 November 2006; accepted 15 November 2006; published 24 May 2007.

[1] This study considers the use of the maximum likelihood estimator proposed by Whittle for calibrating the parameters of hydrological models. Whittle's likelihood provides asymptotically consistent estimates for Gaussian and non-Gaussian data, even in the presence of long-range dependence. This method may represent a valuable opportunity in the context of ungauged or scarcely gauged catchments. In fact, the only information required for model parameterization is essentially the spectral density function of the actual process simulated by the model. When long series of calibration data are not available, the spectral density can be inferred by using old and sparse records, regionalization methods, or information on the correlation properties of the process itself. The proposed procedure is applied to the case studies of two Italian river basins, for which a lumped rainfall-runoff model has been calibrated by emulating scarcely gauged situations. It is shown that the Whittle estimator can be applied in such context with satisfactory results.

Citation: Montanari, A., and E. Toth (2007), Calibration of hydrological models in the spectral domain: An opportunity for scarcely gauged basins?, *Water Resour. Res.*, *43*, W05434, doi:10.1029/2006WR005184.

1. Introduction

[2] Parameter calibration plays an important role in the use of hydrological models of any type. In fact, even when using physically based approaches, there is nevertheless the need to calibrate at least a part of the model parameters in practical applications. This operational procedure is requested as any hydrological model provides only an approximation of reality and therefore one needs to adjust the parameter values in order to improve the goodness of the fit [see, e.g., *Beven*, 1989*,* 2006].

[3] Calibration involves the processing of input data through the model and the adjustment of parameter values in order to produce a reliable simulation. The goodness of the fit is usually evaluated by comparing the data simulated by the model with the corresponding observed variables. This direct comparison can be performed only when referring to catchments where sufficiently long and simultaneous records of input and output variables are available. When the catchment or the location of interest is poorly gauged or ungauged, that is, when hydrological data are limited or unavailable, alternative methods for identifying model parameters are needed.

[4] Calibration of hydrological models in ungauged basins is the subject of increasing attention by the scientific community. In fact, this is one of the main goals of the Prediction in Ungauged Basins (PUB) initiative promoted by the International Association of Hydrological Sciences [*Sivapalan et al.*, 2003]. Despite of being widely investigated nowadays, this topic is considered by many scientists more a utopia than a scientific goal. Indeed, without the support of some observed data a reliable model calibration is still impossible. At the present state of our knowledge, the challenge is to develop techniques in order to be able to apply models when only a limited information is available, at local or regional scale.

[5] Numerous recent contributions proposed by the literature have explored the possibility to regionalize hydrological model parameters, therefore obtaining a tool for inferring their values when dealing with ungauged basins [see, e.g., *Bloeschl and Sivapalan*, 1995; *Post and Jakeman*, 1999; *Seibert*, 1999; *Uhlenbrook et al.*, 1999; *Hundecha and Bardossy*, 2004; *Merz and Bloeschl*, 2004; *McIntyre et al.*, 2005; *Parajka et al.*, 2005]. Other authors pointed out the operational advantages that can be gained by using the information provided by soft data, such as sparse measurements or qualitative observations [*Seibert and McDonnell*, 2002].

[6] The purpose of this paper is to propose the use of the maximum likelihood estimator introduced in the context of time series analysis by *Whittle* [1953] for calibrating hydrological model parameters. The estimator has good statistical properties, as it is asymptotically consistent. It was successfully used in other hydrological applications [e.g., *Montanari et al.*, 2000; *Montanari*, 2003]. In the context of scarcely gauged/ungauged basins it has significant advantages with respect to traditional calibration procedures, as in principle the calibration can be carried out even when observed output data are not available. In fact, the parameter estimation, rather than attempting to fit observed and simulated output time series, is carried out essentially by matching the spectral densities of the model simulation and the actual process. In absence of observed data, it is argued here that the spectral density of the process can be derived

[1]Dipartimento di Ingegneria delle Strutture, dei Trasporti, delle Acque, del Rilevamento e del Territorio, Faculty of Engineering, University of Bologna, Bologna, Italy.

0043-1397/07/2006WR005184$09.00

from different types of information that could be available in poorly gauged situations.

[7] Three applications of the proposed estimator for calibrating a lumped rainfall-runoff model to the Reno and Secchia river basins (Italy) are shown herein, referring to a real data and a synthetic data case study, respectively. In the first two cases the model is parametrized by using input and output data referring to different observation periods and, in the real data case, measured at a different time step. In the third case, the parameterization is carried out by exploiting some basic statistical properties of the output river flows that could be, for example, inferred at regional scale. The results show that the Whittle's likelihood can allow the user to profit from information that could not be used with traditional calibration methods.

2. Approximation Proposed by Whittle to the Gaussian Maximum Likelihood Function

[8] The likelihood measure proposed by *Whittle* [1953] for the parameters of a generic model will be denoted in the following as $L(\theta)$, where θ is the model parameter vector. Note that the likelihood of a parameter set θ is proportional to the probability of obtaining a correct model simulation when the model parameter set is θ. Therefore maximizing the likelihood for varying θ allows one to identify the optimal parameter set.

[9] For a stationary time series $L(\theta)$ is computed on the spectral density [see, e.g., *Beran*, 1994; *Montanari et al.*, 2000; *Chouduri et al.*, 2004]. In practice, the model calibration is carried out essentially by matching the spectral densities of model output and river flow process. Given that any time series can be decomposed in a sum of periodic components through a harmonic analysis, the spectral density describes the variability of the time series that is explained by each component. It is essential to note that the spectral density can be derived from the autocovariance function of the data. Therefore one may say that Whittle's likelihood performs model calibration essentially by matching the autocovariance functions of model output and observed river flow record.

[10] Whittle's likelihood has been widely used in the time series literature for constructing estimators. Referring to the case of rainfall-runoff (R-R) models, for the purpose of introducing $L(\theta)$ let us first focus on the gauged basin case, for which N river flow observations are available to calibrate the model (the scarcely gauged basin case will be introduced in section 2.3). In this case, the approximation of the observed streamflow, $Q_{obs}(t)$, performed by the rainfall-runoff transformation can be written as:

$$Q_{obs}(t) = M[\theta, \mathrm{I}(t)] + e(t) = M[\theta, \mathrm{I}(t)] + \Phi^{-1}(B)\varepsilon(t), \quad t = 1, \ldots N, \tag{1}$$

in which it is assumed that the R-R model residuals $e(t)$ can be modeled by an autoregressive stochastic process [*Brockwell and Davis*, 1987]. In (1) $M[\theta,\mathrm{I}(t)]$ is the transformation operated by the hydrological model, that is assumed to be stationary, and $\mathrm{I}(t)$ is the input vector (for instance rainfall and temperature or evapotranspiration at time t); $\Phi(B)$ is the autoregressive operator, which is in charge of taking into account the correlation in $e(t)$; and $\varepsilon(t)$ represents a zero mean, independent and identically distributed (IID) random variable. The assumption of zero mean for $\varepsilon(t)$ is essential and will be thoroughly discussed below. In general, the autoregressive process is conditional on the hydrological model output $M[\theta,\mathrm{I}(t)]$. An extensive study carried out by the *World Meteorological Organization* [1992] proved that a first order autoregressive process is sufficient to account for correlation in the residual series of many practical applications of rainfall-runoff models.

[11] The Whittle's likelihood $L(\theta)$ for the model given by (1) can be computed through the relationship

$$L(\theta) = \exp\left[-\sum_{j=1}^{N/2}\left\{\log\left[f_M(\lambda_j,\theta) + f_e(\lambda_j,\Phi)\right] + \frac{J(\lambda_j)}{f_M(\lambda_j,\theta) + f_e(\lambda_j,\Phi)}\right\}\right], \tag{2}$$

where $\lambda_j = 2\pi j/N$ are the Fourier frequencies; J is the periodogram (which is an estimate for the spectral density) of the series of the N observed river flows; f_M is the spectral density of the hydrological model output that depends on the parameter vector θ and f_e is the spectral density of the autoregressive operator Φ, that is,

$$f_e(\lambda_j, \Phi) = \frac{\sigma_\varepsilon^2}{2\pi\left|\Phi(e^{-i\lambda_j})\right|^2}, \tag{3}$$

where σ_ε is the standard deviation of $\varepsilon(t)$.

[12] The periodogram $J(\lambda_j)$ of an observed time series at the frequency λ_j can be computed as [*Brockwell and Davis*, 1987]

$$J(\lambda_j) = \sum_{|k|<N} \gamma(k)e^{-ik\lambda_j}, \tag{4}$$

where $\gamma(k)$ is the sample autocovariance coefficient at lag k and i is the imaginary unit. For more details about the derivation of the approximation proposed by *Whittle* [1953] to the Gaussian maximum likelihood function the interested reader is referred to the book by *Beran* [1994] which provides a detailed explanation.

[13] To derive analytically the spectral density $f_M(\lambda_j,\theta)$ of a hydrological model might be not possible in many cases. To solve this problem the spectral density can be approximated numerically, by computing the periodogram, $J_M(\lambda_j,\theta)$, of a sufficiently long model simulation.

[14] Given that maximizing $L(\theta)$ as given by (2) is equivalent to minimizing the absolute value of the quantity between square brackets in the right-hand side of (2), model calibration can be carried out by minimizing

$$W(\theta) = \sum_{j=1}^{N/2}\left\{\log\left[J_M(\lambda_j,\theta) + f_e(\lambda_j,\Phi)\right] + \frac{J(\lambda_j)}{J_M(\lambda_j,\theta) + f_e(\lambda_j,\Phi)}\right\}. \tag{5}$$

[15] In the present study the minimization of (5) was performed numerically (see section 3.2). Whittle's likelihood provides asymptotically consistent and normally distributed estimates for the hydrological model parameters in

the case of Gaussian and linear models even in the presence of long-range dependence (which induces the Hurst effect) [*Fox and Taqqu*, 1986; *Montanari*, 2003]. This means that for Gaussian and linear models the estimator is capable of providing the probability distribution of the parameters, which allows one to derive their confidence limits. Asymptotic normality and consistency is assured under mild conditions also in the non-Gaussian case [*Giraitis and Surgailis*, 1990], while asymptotic normality is no more guaranteed for nonlinear models, as it is the case in many practical applications in hydrology. Nonetheless, the absence of asymptotic normality does not affect the consistency of the estimator but only the computation of the confidence limits of the parameters. Therefore, when applying the Whittle's estimator to nonlinear models alternative methods for computing the confidence limits are to be sought (see section 2.2) [*Giraitis and Taqqu*, 1999].

[16] When contemporaneous input records and river flow data are available over a time period [0, T], the estimator given by (5) can be used to perform a classical model calibration based on the analysis of the goodness of the fit of observed data. In this case the fit is evaluated by comparing the frequency behaviors of simulated and actual records. We will refer to this procedure with the term "Whittle direct calibration." Once a trial value for the parameter vector θ is selected, the periodogram $J_M(\lambda_j,\theta)$ of the simulation run over [0, T] can be computed. Subsequently, the model residuals over [0, T] can be calculated depending on θ so that the autoregressive model $\Phi(B)$ can be identified and its parameters estimated. Therefore it is possible to compute $f_e(\lambda_j,\Phi)$ and, finally, $W(\theta)$ accordingly to (5), that can be minimized in order to identify the best θ.

2.1. Some Relevant Properties of the Estimator and Their Effect in Practical Applications

[17] One of the hypotheses of the Whittle's likelihood is the stationarity of the process. Note that this is not in contrast with the seasonal structure of many hydrological records. In fact, seasonality is reflected in the spectral density and is fully compatible with a stationary process.

[18] Attention should also be focused on the i.i.d. assumption for $\varepsilon(t)$. In fact, in real world applications the $\varepsilon(t)$ are often heteroscedastic and therefore are not identically distributed. In fact, during peak flow periods the errors $e(t)$ of the hydrological model are usually higher, therefore implying an amplification (and a corresponding heteroscedasticity) of the $\varepsilon(t)$. Consequently, a bias can be induced in the parameter estimates. Note that if the quantities $Q_{obs}(t)$ and $M[\theta,\mathrm{I}(t)]$ were taken to be log transformed observations and hydrological model output, then the effect of a residual variance that changes with the magnitude of the observations would be accounted for. Transformations, such as the normal quantile transform, may also be used to stabilize the variances [*Montanari and Brath*, 2004]. In this study these transformations were not applied because the analysis is focused on ungauged basins. Given the high uncertainty of the parameter estimates in this case, the approximation induced by residual heteroscedasticity can be considered negligible.

[19] The spectral density of a process (like river flows at an assigned location) does not convey any information about the mean value and phase of the process itself. In practice, the periodogram of a time series does not change after a shift in the mean or a shift of the whole time series in time (this latter shift is equivalent to a change in phase). The estimator provided by equation (5) preserves the mean of the process through the assumption of zero mean for $\varepsilon(t)$. This condition implies that the hydrological model error $e(t)$ has zero mean as well, which corresponds to imposing equality for the mean values of observed and simulated data.

[20] Preservation of the process phase, therefore avoiding possible shifts of the model output in time, is attained through the term f_e, i.e., the spectral density of the hydrological model error $e(t)$. In fact, f_e and the phase of the hydrological model are strictly connected. A change in phase of the hydrological model simulation (or in other words a shift in time of its simulation) implies a change in the variance (and therefore the spectral density) of $e(t)$. The results is that minimizing (5) implies an optimization of the phase of the simulated record.

[21] The considerations above imply two important consequences on the use of the Whittle's likelihood. First, those parameter sets that lead to significant differences $\Delta\mu_{obs,sim}$ in the mean values of observed and simulated river flows should be rejected, as in this case the assumption of zero mean for $\varepsilon(t)$ would be violated. Strictly speaking one should exclude the parameter sets that do not allow one not to reject the hypothesis of null value of $|\Delta\mu_{obs,sim}|$ at an assigned confidence level. Second, one should note that neglecting the term f_e in the estimation procedure may imply an imprecision in fitting the phase of the signal. The consequences of this potential drawback will be discussed in section 2.3.

2.2. Assessment of the Calibration Uncertainty

[22] It was mentioned above that in the case of linear models (either Gaussian or not) Whittle's likelihood provides normally distributed estimates even in the presence of long-range dependence. Therefore in this case it is straightforward to compute confidence limits for the estimated parameters by using the asymptotical results summarized by *Beran* [1994]. Details of such computation are not reported here, as the linear case is not frequent in hydrology. Asymptotic normality is no more guaranteed for nonlinear models and therefore an alternative procedure for parameter uncertainty assessment is to be used.

[23] Being $L(\theta)$ an estimate of the likelihood of the model parameters, an interesting possibility for approximating the probability distribution of the estimated parameters is offered by Markov Chain Monte Carlo (MCMC) methods. In particular, the Whittle's likelihood can be employed within the Shuffled Complex Evolution Metropolis algorithm (SCEM-UA) optimization algorithm by *Vrugt et al.* [2003]. The SCEM-UA algorithm works by integrating the Metropolis algorithm, controlled random search, competitive evolution, and complex shuffling to drive the sampler of the model parameters toward their target distribution. In detail, $L(\theta)$ can be used as a likelihood measure that is alternative to the one provided by equation (2) of *Vrugt et al.* [2003]. Once the probability distribution of the model parameters is approximately known, we can derive an estimate of the optimal parameter set along with the related confidence limits for an assigned confidence level.

[24] It should be noted that the procedure suggested above provides an estimate of the parameter uncertainty under the hypothesis that the information used for calibrating the model is correct. Therefore an estimate of the uncertainty given by model structural inadequacy only is obtained. Actually, the information used in model calibration is always uncertain in real word applications. A significant source of uncertainty relies in the definition of the spectral density of the process f_M. Note that in the case of ungauged basins this latter uncertainty is relevant and most likely dominant with respect to other sources of uncertainty, included that given by model structural inadequacy. Therefore the evaluation of the uncertainty of the calibrated parameters should be carried out by taking into account the reliability of f_M. A relevant problem is that in the case of ungauged basins to derive an estimate of the uncertainty of the spectral density may be extremely difficult, if not impossible. However, in some cases the user may be capable of inferring a confidence range for f_M (or, alternatively, the autocovariance function of the signal), at least on a perceptual basis. In this case, Monte Carlo simulation methods (like the GLUE approach [*Beven and Binley*, 1992]) may allow one to derive sensitivity envelops for model parameters and output. These techniques will not be applied here.

2.3. Application to Ungauged or Scarcely Gauged Catchments

[25] When contemporaneous input and output records are not available, the direct calibration is not applicable and indirect estimation methods for ungauged basins need to be worked out. In this case, the Whittle's likelihood can be an interesting opportunity, because the streamflow periodogram might also be obtained in an indirect way. We will refer to this procedure with the term ‘‘Whittle indirect calibration’’. For example, since the periodogram of a time series can be computed from its autocovariance function (see equation (4)), the model estimation may be based on information on the autocorrelation structure of streamflow values, that may be derived also from regionalization procedures, under assumptions that will be discussed here below.

[26] In the remainder of the paper we will consider some examples of indirect calibrations. The first situation refers to a river basin whose data set consists of input observations, that is precipitation and temperature data, and streamflow measures that are not simultaneous. In this case, the availability of an old, and possibly sparse, streamflow series may allow a successful application of the Whittle method. In fact, once a trial value for the parameter vector θ is selected, the spectral density function of the model, f_M, can be estimated by computing the periodogram J_M of a simulation run obtained by using the available input data. If the process is assumed to be stationary, the periodogram J of the actual river flow process can be estimated by using old records or sparse measurements. Assuming that the role of the spectral density of the hydrological model residuals, f_e, is negligible, the estimator given by equation (5) may be applied without the availability of contemporaneous input and output data.

[27] The second situation refers to a river basin for which observed river discharges are not available but information are at disposal about the mean value, the variability and the autocorrelation structure of the river flows. This type of knowledge could be derived by using regionalization procedures and allows one to infer the periodogram J of the river flow record. In this case also, f_e is assumed to play a negligible role.

[28] As mentioned in section 2.1, neglecting f_e implies that the estimator is no more able to fit the mean and phase of the signal, as only the correct fitting of the spectral density of the model is assured. Therefore, when referring to scarcely gauged catchments the users is required to specify a value for the mean, μ_{obs}, of the river flows, in order to be able to reject the parameter sets that lead a difference $|\Delta\mu_{obs,sim}|$ exceeding a given threshold. Such mean value may be problematic to identify. However, given the high uncertainty involved in the analysis of ungauged catchments, the approximation due to the choice of a perceptual value for μ_{obs} may be considered scarcely influent.

[29] The lack of fitting of the phase of the signal is a minor problem in the case of many modern hydrological models, whose spectral density significantly depends on the values of parameters describing hydrographs characteristics that are related to the phase of the signal, like the base time or time to peak. This evidence is confirmed by the applications presented here below, where a satisfactory fit is reached after neglecting *fe*. However, the user should be aware that neglecting *fe* may induce rough approximations for those models whose spectral density is not related to the displacement in time of the hydrograph, like for instance the Nash model [*Nash*, 1958].

3. Description of the Case Studies: Application to Scarcely Gauged Catchments

[30] The performances of the Whittle estimator were tested by developing three case studies, for which a situation of data scarcity was emulated. In detail, for the first two case studies we tested the possibility to apply the Whittle's likelihood in order to calibrate a lumped R-R model when the hydrological information consists of (1) recent records of rainfall and temperature and (2) old records of river flows, possibly observed at a different time step. The possibility to calibrate a R-R model by using not contemporaneous records of input and output variables constitutes an interesting opportunity. In fact, the rainfall and temperature monitoring network is generally quite dense in many countries, while the information about the river flows is much more sparse, fragmentary and sometimes obsolete. For instance in Italy the economic costs related to river flow monitoring have caused a decrease of discharge data availability in the latter decades. Therefore the situation in which one may need to exploit the information given by old and sparse records of river flows is quite common.

[31] The third case study considers the application of the Whittle's likelihood in order to calibrate a lumped R-R model on the basis of (1) recent records of rainfall and temperature and (2) mean value, standard deviation and lag one autocorrelation coefficient of the river flow process. The possibility to estimate the model parameters on the basis of this type of information would constitute an interesting options when the above statistics can be inferred at regional scale [*Castellarin et al.*, 2004].

[32] The test sites herein considered refer to two medium sized watersheds located in the Apennines Mountains in north central Italy: the Secchia River and the Reno River. The Secchia River was used as reference for performing a synthetic data extensive test while the Reno River was used for performing real data tests.

[33] The Secchia River flows northward across the Apennine Mountains and is a right tributary to the Po River. The contributing area is 1214 km^2 at the Bacchello Bridge river cross section that is located about 62 km upstream of the confluence in the Po River. The maximum altitude is 2121 m above sea level at Mount Cusna. The mainstream length up to Bacchello Bridge is about 98 km and the basin concentration time is about 15 hours. The mean annual rainfall depth ranges between 700 and more than 2000 mm/year over the basin area. The maximum peak discharge observed at Bacchello Bridge in the period 1923–1981 was 823 m^3/s (20 April 1960).

[34] The period of observation is formed by two years of hourly hydrometeorological data, from 1 January 1972 to 31 December 1973. The hourly data set consists of streamflow measures at Bacchello Bridge, the temperatures recorded in 1 gauge and the precipitation depths measured in 5 rain gauges.

[35] The Reno River at the closure section of Casalecchio, just upstream the city of Bologna, has a drainage area of the watershed of 1050 km^2. The mainstream is around 76 km long. The average elevation is 635 m above sea level, the highest peak and the outlet being at an altitude of 1900 and 63 m above sea level respectively. The majority of precipitation events occur from October to April, November being the wettest month and the runoff regime follows closely the precipitation trend. The 10-year return period discharge is estimated as around 1100 m^3/s, from the sample cumulative distribution function of 70 observed annual maxima.

[36] The observed data set is formed by eight years of hydrometeorological hourly data, from 1 January 1993 to 31 December 2000, and 50 years of daily streamflows at Casalecchio, from 1 January 1930 to the 31 December 1979. The hourly data set consists of streamflow measures at Casalecchio, the temperatures recorded in 7 gauges and the precipitation depths measured in 13 rain gauges, even if not all contemporarily operative for the whole observation period.

3.1. Conceptual Rainfall-Runoff Models

[37] Two lumped conceptual rainfall-runoff models with a different degree of complexity were used for developing the case studies. The first model, which was also used for generating the synthetic data referred to the Secchia River, is the ADM rainfall-runoff model [*Franchini*, 1996]. The model is formed by two main blocks: the first represents the water balance at soil and subsoil level, while the second represents the transfer of runoff production at the basin outlet. The basin, in turn, is divided into two zones: the upper zone produces surface and subsurface runoff, having as inputs precipitation and potential evapotranspiration, while the lower zone (whose input from the first one is the percolation flow) produces base 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 and the second the flow along the channel network toward the basin outlet. The surface runoff generation is based on the concept of probability distributed soil moisture storage capacity of the Xinanjiang model [*Zhao et al.*, 1980], where the total surface runoff is the spatial integral of the infinitesimal contribution deriving from the saturated elementary areas. The model has a total of 11 parameters to be estimated: 2 for the computation of surface runoff, 4 for interflow and percolation, 1 for the base flow and four parameters for the transfer components.

[38] HYMOD is a five-parameter conceptual rainfall-runoff model that was introduced by *Boyle* [2000] and recently used by *Wagener et al.* [2001], *Vrugt et al.* [2003], and *Montanari* [2005]. HYMOD consists of a relatively simple rainfall excess model, described in detail by *Boyle* [2000], that is connected with two series of linear reservoirs: three identical reservoirs for the quick response and a single reservoir for the slow response. This model requires the optimization of five parameters: the maximum storage capacity in the catchment, C, the degree of spatial variability of the soil moisture capacity within the catchment, b, the factor distributing the flow between the two series of reservoirs, α, and the residence time of the linear quick and slow reservoirs, K_s and K_l, respectively. With a total of only five parameters, HYMOD can be considered an approach of reduced complexity with respect to ADM.

[39] The estimates of potential evapotranspiration, the second meteorological variable provided as input to both models, is obtained from temperature data through a simplified form of the radiation method [*Doorenbos et al.*, 1984]. Empirical radiation-based equations for estimating potential evaporation are generally based on the energy balance. The method used here is a simplified form of the Penman-Monteith equation in which the terms related to vapor pressure and wind velocity are neglected.

3.2. Numerical Maximization of the Whittle's Likelihood Function and Reference Baselines

[40] In this study we used a genetic algorithm in order to perform numerical optimization, namely, the GENOUD method (genetic optimization using derivatives [see *Sekhon and Mebane*, 1998]). GENOUD combines evolutionary algorithm methods with a derivative based, quasi-Newton method to solve optimization problems. A relevant feature of GENOUD is its computing time, which is sustainable even when working with high dimensional parameter sets.

[41] As a reference baseline, the performances of the Whittle direct estimations were compared with those obtained by maximizing the traditional Nash-Sutcliffe efficiency [*Nash and Sutcliffe*, 1970] of the simulated river flows. To obtain a term of comparison for appraising the performances of the indirect calibration we performed an extensive test of models characterized by random parameters. In detail, for each calibration exercise we computed the Nash-Sutcliffe efficiency obtained by running the model with 10,000 random parameter sets. These were generated accordingly to a uniform distribution within the same parameter range that was utilized when running the GENOUD optimization algorithm for indirect calibration. The resulting median value of the efficiency obtained with random parameters (median efficiency with random param-

eters) was assumed as a reference performance for a calibration that is carried out without using any type of optimization information but the predefined parameter range.

3.3. Generation of Synthetic Rainfall and Temperature Data for the Secchia River

[42] The Secchia River case study was developed by using synthetic data, in order to be able to perform an extensive test of the Whittle's likelihood. Accordingly, artificial rainfall, temperature and river flow data were generated in order to obtain a 100-year long sample of input and output time series to be used to calibrate the R-R model.

[43] The synthetic rainfall data were generated by using the generalized multivariate Neyman-Scott rectangular pulses model. For more details the interested reader is referred to *Cowpertwait* [1995] and *Montanari* [2005]. The rainfall model was applied to the Secchia River basin by generating data for five locations where rain gauges are present. By using the method of moments, historical hourly rainfall data observed in the two year period 1972–1973 were used to calibrate the Neyman-Scott model parameters. 100 years of hourly rainfall data were subsequently generated for the five rain gauges. The depth duration frequency curves in each of them were well simulated; in particular, the percentage error in the simulation of the 12-hour (a time span comparable to the concentration time of the basin) cumulated rainfall with return period of 100 years was always lower than 6% in all the rain gauges. The mean areal hourly rainfall over the basin was then estimated through a weighted average of the hourly rainfall data in the five locations. The weights w_j, $j = 1, \ldots, 5$ are estimated with the Thiessen polygons method.

[44] The rainfall data were subsequently corrupted in order to emulate a real situation, where the observed records are always affected by uncertainty. In detail, the weights w_j used to compute the mean areal rainfall over the basin were perturbed by randomly generating, at each time step, each w_j accordingly to a uniform distribution in the range ±20% of their actual value. Then the w_j obtained at each time step are rescaled to make their cumulative sum equal to one.

[45] The rainfall data corruption introduced an uncertainty that can be quantified by the coefficient of determination of the linear regression of corrupted versus uncorrupted mean areal rainfall depths that resulted equal to 0.76. The type of uncertainty introduced when corrupting the rainfall data is not unlikely in practical applications. In fact, the weights attributed to each raingauge for the computation of the mean areal rainfall over the basin can change in time, depending on the spatial variability of the rainfall field. Finally, the rainfall generation allowed to obtain a 100-year record of lumped and uncertain hourly rainfall $P_m(t)$ over the basin that was used as input to the rainfall-runoff simulation model.

[46] The generation of the synthetic hourly temperature record was performed by referring to a location where historical data are available and was carried out by applying a fractionally differenced ARIMA model (FARIMA). On many occasions, this class of models turned out to be able to fit the autocorrelation structure of temperature series, that is very often affected by a slow decay, which may suggest the existence of long-term persistence, implying this way the presence of the Hurst effect [*Montanari*, 2003]. More details on FARIMA models and the simulation procedure herein applied are given by *Montanari et al.* [1997]. A mean areal value of the hourly temperature data was obtained by rescaling the synthetic observations to the mean altitude of the basin area, by adopting an elevation thermal gradient.

3.4. Generation of Synthetic Streamflow Data for the Secchia River

[47] Synthetic river flow data were obtained by using the previously generated synthetic rainfall and temperature records as input data to the lumped rainfall-runoff model ADM [*Franchini*, 1996], which is described in the section 3.1.

[48] Using the GENOUD optimization algorithm (see section 3.2) the parameters of the ADM model were estimated by automatically optimizing the simulation of historical hourly river flow data that were observed at Bacchello Bridge in the year 1972 and part of the year 1973. The model was validated by simulating the fraction of the 1973 flows that was not used for calibration. The resulting Nash-Sutcliffe efficiency was 0.81 for the validation period. The ADM model was subsequently applied to generate a 100-year long record of hourly synthetic river flows of the Secchia River.

[49] The synthetic river flows were subsequently corrupted by multiplying each observation by a coefficient that assumed different values at each hourly time step. The time series of the coefficient values was obtained by generating outcomes from a uniform distribution in the range 0.8–1.2. The coefficient of determination of the linear regression of corrupted versus uncorrupted river flow data is equal to 0.86. Clearly, there are many possible ways for corrupting a river flow series. The procedure used here is representative of a situation where each river flow measurement is affected by a random inaccuracy, like the one that can be caused by measurement errors.

4. Results of Model Calibration

4.1. Secchia River, Synthetic Data: Calibration by Using Past River Flows

[50] The years 51–100 of the synthetic meteorological and streamflow data were divided into 10 nonoverlapping 5-year long subsamples and 10 separate calibrations of the HYMOD model for the Secchia river basin, using both the Whittle estimator (Whittle direct calibration) and the Nash-Sutcliffe efficiency maximization (Nash-Sutcliffe direct calibration) were performed. In these direct calibrations, simultaneous time series of synthetic hourly meteorological and streamflow data were used, here assumed as true data. In the Whittle direct calibrations, $W(\theta)$ given by (5) was minimized, whereas in the Nash-Sutcliffe direct calibrations the efficiency was maximized.

[51] In addition to the direct calibrations, the Whittle estimator was used also in indirect mode. Here HYMOD calibration was repeated for each 5-year long subsample mentioned above, therefore performing 10 indirect optimizations of model parameters. In this case we used as input data the rainfall and temperature of each 5-year nonoverlapping period of the second half (time span between years 51 and 100) of the synthetic record, while the streamflow periodogram J and the mean value μ_{obs} were estimated

Table 1. Secchia River Basin: Nash-Sutcliffe Efficiency Obtained for the 10 Periods of 5 years Belonging to the Years 51–100 of the Synthetic Record[a]

Years of the Synthetic Record	Nash-Sutcliffe Direct Calibration	Whittle Direct Calibration	Whittle Indirect Calibration	Median Efficiency With Random Parameters
51–55	0.77	0.75	0.60	−0.45
56–60	0.79	0.74	0.51	−0.93
61–65	0.80	0.78	0.57	−0.09
66–70	0.83	0.63	0.64	−2.66
71–75	0.76	0.77	0.57	−1.06
76–80	0.79	0.70	0.75	−1.62
81–85	0.82	0.63	0.56	−0.86
86–90	0.72	0.62	0.26	−1.80
91–95	0.84	0.71	0.57	−1.21
96–100	0.79	0.72	0.48	−2.12
Mean values	0.79	0.71	0.55	−1.28

[a]Efficiencies are reported for Nash-Sutcliffe direct calibration, Whittle direct and indirect calibrations. The median efficiency with random parameters is also shown.

using the river flow observations of the first half record (time span between the years 1–50). In such Whittle indirect calibrations the streamflow data corresponding to the 5-year subsamples (between years 51–100) are considered unknown and are used only for validation purposes. Model calibration was performed by optimizing θ in order to minimize equation (5) for each 5-year period, by applying the GENOUD method.

[52] The performances of the direct and indirect calibrations for each 5-year period are analyzed through the evaluation of the Nash-Sutcliffe efficiency. Table 1 indicates the goodness of fit measures obtained for the ten 5-year periods of the years 51–100 of the synthetic record with the Whittle indirect calibration, Whittle direct calibration and Nash-Sutcliffe direct calibration. The median efficiency with random parameters is also shown. It may be seen that the performances of the two direct procedures are good and comparable, thus allowing one to see that the Whittle estimator provides satisfactory performances as calibration procedure. Of course, the efficiencies of the direct calibrations obtained by directly maximizing E itself are slightly greater than the ones obtained by applying the Whittle method directly, as this latter optimizes a different objective function.

[53] As far as the indirect calibration is concerned, which is the most interesting result in the context of scarcely gauged basins, it can be seen that the results are more dependent on the calibration period and, as was to be expected, always worse than direct calibration. In fact, when an indirect calibration is carried out a loss of performances is to be expected and depends on the amount of the information that is still exploited in the indirect mode. It should be noted that the indirect calibration efficiency is evaluated over data that were not used in the calibration phase, thus referring to a validation test. This is a further cause of performance loss.

[54] However, the efficiency of the Whittle indirect calibration is always remarkably better than the median efficiency with random parameters. This proves the considerable value of the Whittle's method in this case.

[55] In view of the considerations above, the results may be considered satisfactory, and prove that the Whittle's likelihood can be profitably used when sparse and old records of river flows are only available along with model input data not necessarily referred to the same period.

4.2. Reno River, Real Data: Calibration by Using Past River Flows

[56] A real data case study was implemented referring to the Reno river basin by emulating a situation with scarcity of observed river flow data. The ADM model was calibrated in this case.

[57] For the Reno catchment, the spatial average of hourly rainfall depths was estimated with an inverse squared distance weighting of the rain gauges observations, accordingly to what was suggested by previous studies [*Brath et al.*, 2004]. Spatial average of temperature values was based on an inverse squared distance weighting and an elevation thermal gradient. In this case the direct calibrations, making use of simultaneous input and output data, were performed on the eight years of hourly data, both with the Whittle estimator and the Nash-Sutcliffe efficiency maximization.

[58] The indirect calibration was carried out by exploiting the daily streamflow measured in the 1930–1979 period for estimating the periodogram J in equation (5), along with the hourly precipitation and temperature data relative to the years 1993–2000 for estimating the periodogram of the hydrological model J_M. The 1993–2000 hourly streamflow measures were not used in the indirect calibration.

[59] Table 2 compares the values of the efficiency of the simulation of the 8-year hourly flow data obtained with the Whittle indirect calibration, Whittle direct calibration and Nash-Sutcliffe direct calibration. The median efficiency with random parameters is also shown. Figure 1 shows the scatterplot of simulated versus observed hourly streamflows for the Nash-Sutcliffe direct calibration and the Whittle indirect calibration. Figure 2 shows the mean annual flow duration curves obtained with the observed river flow and the simulated flow provided by the ADM model directly (Nash-Sutcliffe) and indirectly (Whittle) calibrated. The flow duration curves are plotted in logarithmic scale

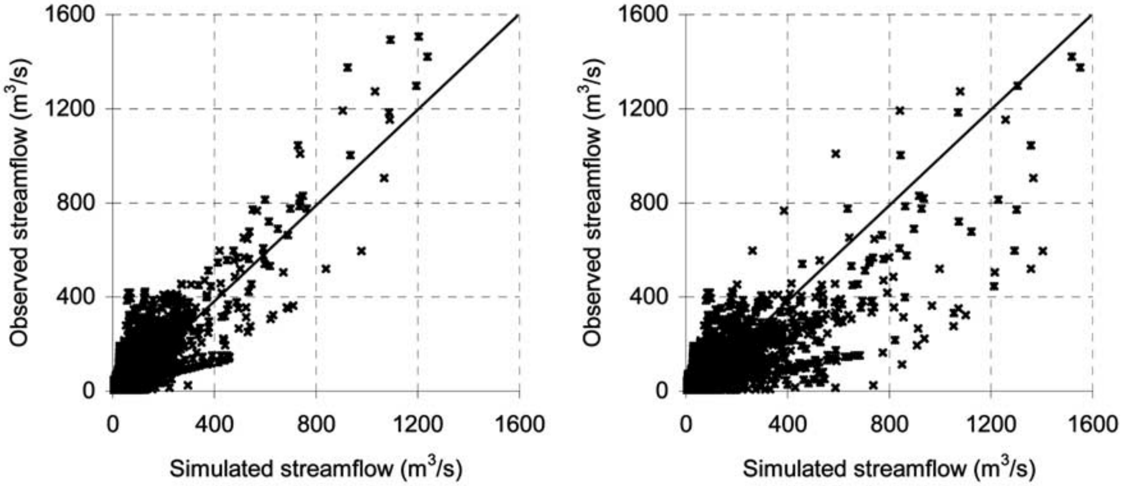

Figure 1. Reno River basin (1993–2000). Scatterplot of simulated versus observed hourly streamflow data for the (left) Nash-Sutcliffe direct calibration and (right) indirect Whittle calibration.

and limited to river flows greater than 0.5 m^3/s. Table 2 suggests that the two direct calibrations perform similarly, thus confirming the efficiency of the Whittle estimator. The simulation obtained with the Whittle indirect calibration, even if obviously less accurate than those allowed by the direct procedures, is satisfactory (see Figures 1 and 2), especially when compared with the performances obtained with random parameters (Table 2). In addition, one should consider also in this case that the simulation is carried out in validation mode and that the calibration was based on a kind of information that is not exploitable with a traditional calibration procedure.

4.3. Reno River, Real Data: Calibration by Using River Flow Statistics

[60] A second case study was developed for the Reno River, by applying again the ADM model. In detail, in this case it is assumed that the river flow process can be described by utilizing a first order autoregressive process. Obviously, assuming that the dynamics of the river flow can be approximated by a linear model provides a crude approximation of reality, that can be accepted for the purpose of calibrating the model in condition of data scarcity, by exploiting basic statistical information of the actual process. Under this assumption, the autocovariance function (and therefore the periodogram J in equation (5)) of the process can be derived depending on the mean, variance and lag one autocorrelation coefficient of the river flows (see also equation (3), which refers to a zero mean, autoregressive process of generic order). In the present application the above statistics were inferred by computing their value in the river flow record of the year 1993 alone. In a more general context, these statistics could be derived through a regional analysis [*Castellarin et al.*, 2004]. The periodogram of the hydrological model J_M was derived by using the same precipitation and temperature data used in the application presented in section 4.2.

[61] Table 2 shows the value of the efficiency of the simulation obtained over the 8-year hourly data using this indirect calibration based on river flow statistics. In this case also, the efficiency of the Whittle indirect calibration is much better than the median efficiency with random parameters, that is obtained without any calibration, and shows another way for taking into account, in a statistical framework under given assumptions, a type of information that could be derived on regional basis and is not exploitable with a traditional calibration procedure.

5. Concluding Remarks

[62] The proposed calibration procedure, that uses the approximation proposed by *Whittle* [1953] to the Gaussian likelihood function, can be an interesting approach for estimating the parameters of hydrological models in the context of ungauged or scarcely gauged catchments. As an example, three case studies are presented in the paper that could be representative of application in scarcely gauged situations.

[63] The main feature of the proposed method, which is based on the comparison between the spectral density of the

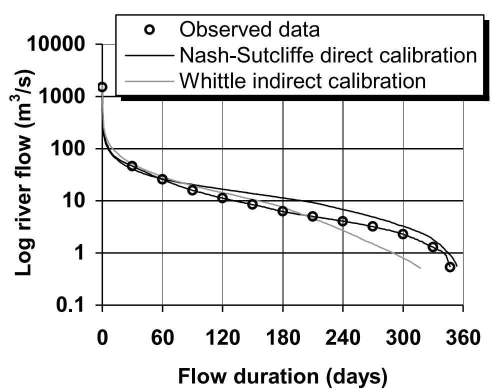

Figure 2. Reno River basin (1993–2000). Mean annual flow duration curves (for river flow higher than 0.5 m^3/s) obtained with observed data, Nash-Sutcliffe direct calibration, and indirect Whittle calibration. River flows are in logarithmic scale.

Table 2. Reno River Basin: Nash-Sutcliffe Efficiency Obtained for the 8 years of Hourly Data (1993–2000) With the Nash-Sutcliffe Direct Calibration and Whittle Direct and Indirect Calibrations[a]

	Value
Nash-Sutcliffe direct calibration	0.77
Whittle direct calibration	0.73
Whittle indirect calibration (past river flows)	0.48
Whittle indirect calibration (river flow statistics)	0.51
Median efficiency with random parameters	−0.66

[a]The median efficiency with random parameters is also shown.

simulated and actual process, lies in the possibility to perform the model calibration also in absence of the output measurements, since the spectral density may be inferred also from an alternative information. The applications illustrated in the present work are only possible examples of the practical use of the Whittle estimator, which may be applied also by inferring the spectral density of the process from perceptual knowledge.

[64] Of course, when applying an indirect calibration procedure the uncertainty may be relevant, whatever technique is chosen; nonetheless, the Whittle method may be useful to reduce the feasible parameter space in order to identify a set of behavioral models in the framework of equifinality [*Beven and Binley*, 1992; *Gupta et al.*, 1998; *Beven and Freer*, 2001]. This operation can be carried out by exploiting a type of knowledge that could not be used within traditional parameterization procedures. For instance, when a direct calibration provides more than one parameter set equally acceptable, the proposed technique may help to identify the parameterization that better represents old or sparse measurements or information of different type that may be available for inferring the spectral density of the simulated process.

[65] It is worth noting that in the case of sufficient data availability the proposed method may have advantages in preserving statistical characteristics of the river flow records such as variance, autocovariance structure and spectral properties across selected Fourier frequencies. Moreover, the focus on signatures of observed streamflow variability, such as the spectral signature, rather than on the raw time series may help the hydrologist to understand the signal behavior beyond what can be learned from curve fitting.

[66] A software for the application of the Whittle estimator, in direct mode, which runs under the R package is available at http://www.costruzioni-idrauliche.ing.unibo.it/people/alberto/indexeng.html.

[67] **Acknowledgments.** The work presented here has been carried out in the framework of the activity of the Working Group at the University of Bologna of the Prediction in Ungauged Basins (PUB) initiative of the International Association of Hydrological Sciences. The study has been partially supported by the Italian Ministry of University and Research in Science and Technology (MURST) through its national grant to the program on "Advanced techniques for estimating the magnitude and forecasting extreme hydrological events, with uncertainty analysis." A special acknowledgement for providing very useful comments is owed to Liudas Giraitis, Paul Torfs, the Associate Editor Murugesu Sivapalan, the referees Demetris Koutsoyiannis and Ralf Merz, and an anonymous referee.

References

Beran, J. (1994), *Statistics for Long-Memory Processes*, CRC Press, Boca Raton, Fla.

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

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

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

Beven, K. J., and J. Freer (2001), Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology, *J. Hydrol.*, *249*, 11–29.

Bloeschl, G., and M. Sivapalan (1995), Scale issues in hydrological modelling: A review, *Hydrol. Processes*, *9*, 251–290.

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

Brath, A., A. Montanari, and E. Toth (2004), Analysis of the effects of different scenarios of historical data availability on the calibration of a spatially-distributed hydrological model, *J. Hydrol.*, *291*, 232–253.

Brockwell, P. J., and R. A. Davis (1987), *Time Series: Theory and Methods,* Springer, New York.

Castellarin, A., G. Galeati, L. Brandimarte, A. Brath, and A. Montanari (2004), Regional flow-duration curves: Reliability for ungauged basins, *Adv. Water Resour.*, *27*, 953–965.

Chouduri, N., S. Ghosal, and A. Roy (2004), Contiguity of the Whittle measure for a Gaussian time series, *Biometrika*, *91*, 211–218.

Cowpertwait, P. S. P. (1995), A generalized spatial-temporal model of rainfall based on a clustered point process, *Proc. R. Soc. London, Ser. A*, *450*, 163–175.

Doorenbos, 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, *Irrig. Drain. Pap.*, *24*, 156 pp.

Fox, R., and M. S. Taqqu (1986), Large sample properties of parameter estimates for strongly dependent stationary Gaussian time series, *Ann. Stat.*, *14*, 517–532.

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.

Giraitis, L., and D. Surgailis (1990), A central limit theorem for quadratic forms in strongly dependent linear variables and application to asymptotical normality of Whittle's estimate, *Probab. Theory Related Fields*, *86*, 87–104.

Giraitis, L., and M. S. Taqqu (1999), Whittle estimator for finite-variance non-Gaussian time series with long memory, *Ann. Stat.*, *27*, 178–203.

Gupta, H. V., S. Sorooshian, and P. Yapo (1998), Toward improved calibration of hydrologic models: Multiple and noncommensurable measures of information, *Water Resour. Res.*, *34*, 751–763.

Hundecha, Y., and A. Bardossy (2004), Modeling the effect of land use changes on the runoff generation of a river basin through parameter regionalization of a watershed model, *J. Hydrol.*, *292*, 281–295.

McIntyre, N., H. Lee, H. Wheater, A. Young, and T. Wagener (2005), Ensemble predictions of runoff in ungauged catchments, *Water Resour. Res.*, *41*, W12434, doi:10.1029/2005WR004289.

Merz, R., and G. Bloeschl (2004), Regionalisation of catchment model parameters, *J. Hydrol.*, *287*, 95–123.

Montanari, A. (2003), Long range dependence in hydrology, in *Theory and Applications of Long-Range Dependence*, edited by P. Doukhan, G. Oppenheim, and M. S. Taqqu, pp. 461–472, Springer, New York.

Montanari, A. (2005), Large sample behaviors of the generalized likelihood uncertainty estimation (GLUE) in assessing the uncertainty of rainfall-runoff simulations, *Water Resour. Res.*, *41*, W08406, doi:10.1029/2004WR003826.

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.

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.

Montanari, A., R. Rosso, and M. S. Taqqu (2000), A seasonal fractional ARIMA model applied to the Nile River monthly flows at Aswan, *Water Resour. Res.*, *36*, 1249–1259.

Nash, J. E. (1958), The form of the instantaneous unit hydrograph, *IAHS Publ.*, *42*, 114–118.

Nash, J. E., and J. V. Sutcliffe (1970), River flow forecasting through conceptual models 1: A discussion of principles, *J. Hydrol.*, *10*, 282–290.

Parajka, J., R. Merz, and G. Bloeschl (2005), A comparison of regionalisation methods for catchment model parameters, *Hydrol. Earth Syst. Sci.*, *9*, 157–171.

Post, D. A., and A. J. Jakeman (1999), Relationships between catchment attributes and hydrological response characteristics in small Australian mountain ash catchments, *Hydrol. Processes*, *10*, 877–892.

Seibert, J. (1999), Regionalisation of parameters for a conceptual rainfall-runoff model, *Agric. For. Meteorol.*, *98–99*, 279–293.

Seibert, J., and J. J. McDonnell (2002), On the dialog between experimentalist and modeler in catchment hydrology: Use of soft data for multicriteria model calibration, *Water Resour. Res.*, *38*(11), 1241, doi:10.1029/2001WR000978.

Sekhon, J. S., and W. R. Mebane (1998), Genetic optimization using derivatives, *Polit. Anal.*, *7*, 187–210.

Sivapalan, M., et al. (2003), IAHS Decade on Predictions in Ungauged Basins (PUB), 2003–2012: Shaping an exciting future for the hydrological sciences, *Hydrol. Sci. J.*, *48*, 857–880.

Uhlenbrook, S., J. Seibert, C. Lebundgut, and A. Rodhe (1999), Prediction uncertainty of conceptual rainfall-runoff models caused by problems in identifying model parameters and structure, *Hydrol. Sci. J.*, *44*, 779–797.

Vrugt, J. A., H. V. Gupta, W. Bouten, and S. Sorooshian (2003), A Shuffled Complex Evolution Metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters, *Water Resour. Res.*, *39*(8), 1201, doi:10.1029/2002WR001642.

Wagener, T., D. P. Boyle, M. J. Lees, H. S. Wheater, H. V. Gupta, and S. Sorooshian (2001), A framework for development and application of hydrological models, *Hydrol. Earth Syst. Sci.*, *5*, 13–26.

Whittle, P. (1953), Estimation and information in stationary time series, *Ark. Mat.*, *2*, 423–434.

World Meteorological Organization (1992), Simulated real-time intercomparison of hydrological models, *Oper. Hydrol. Rep. 38,* Geneva, Switzerland.

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. Montanari and E. Toth, DISTART, Faculty of Engineering, University of Bologna, Via del Risorgimento 2, I-40136 Bologna, Italy. (alberto.montanari@unibo.it)