Advances in Water Resources 33 (2010) 1235–1242

Contents lists available at ScienceDirect

Advances in Water Resources

journal homepage: www.elsevier.com/locate/advwatres

Calibration of rainfall-runoff models in ungauged basins: A regional maximum likelihood approach

Simone Castiglioni, Laura Lombardi, Elena Toth, Attilio Castellarin, Alberto Montanari *

Faculty of Engineering, DICAM — University of Bologna, Viale del Risorgimento 2, Bologna, 40136, Italy

ARTICLE INFO

Article history:
Received 30 November 2009
Received in revised form 12 April 2010
Accepted 15 April 2010
Available online 26 May 2010

Keywords:
Rainfall-runoff
Ungauged basins
PUB
Regional calibration
Spectral calibration

ABSTRACT

Parameter estimation for rainfall-runoff models in ungauged basins is a challenging task that is receiving significant attention by the scientific community. In fact, many practical applications suffer from problems induced by data scarcity, given that hydrological observations are often sparse or unavailable. This study focuses on regional calibration for a generic rainfall-runoff model. The maximum likelihood function in the spectral domain proposed by Whittle [40] is approximated in the time domain by maximising the fit of selected statistics of the river flow process, with the aim to propose a calibration procedure that can be applied at regional scale. Accordingly, the statistics above are related to the dominant climate and catchment characteristics, through regional regression relationships. The proposed technique is applied to the case study of 4 catchments located in central Italy, which are treated as ungauged and are located in a region where detailed hydrological, as well as geomorphologic and climatic information, is available. The results obtained with the regional calibration are compared with those provided by a classical least squares calibration in the time domain. The outcomes of the analysis confirm the potential of the proposed methodology and show that regional information can be very effective for setting up hydrological models.

1. Introduction

Calibration of rainfall-runoff (R-R) models in ungauged or scarcely gauged basins is the subject of increasing attention by the hydrologic community, especially after the introduction of the Prediction in Ungauged Basins (PUB) initiative of the International Association of Hydrological Sciences [30]. Indeed, the emerging need for efficient water resources assessment and flood frequency estimation makes the real world application of R-R models an attractive perspective, with the potential problem that practical applications are almost always concerned with ungauged or scarcely gauged catchments. This is the reason why the scientific community recently proposed a number of innovative approaches for R-R model calibration. Regional calibration is an interesting option, based on the idea of linking R-R model parameters to macro-scale properties of the contributing catchment. So far, the contributions focused on the possibility to directly estimate R-R model parameters with regional relationships (see, for instance, [29,18,44,42,27,38,26,1]). This approach requires the preliminary identification of a suitable hydrological model for the study area. On the one hand it might be reasonable to assume that a single R-R model is adequate for approximating the hydrological regime of a whole region, especially if it is characterised by fairly similar runoff generation processes. On the other hand, different data and information availability, as well as different scales, may dictate the necessity for the end users to consider different models in the same region, therefore suggesting the need for a more flexible procedure.

This paper presents a regional approach for the calibration of any type of R-R model at regional scale. The basic concept underlying our idea is that the way forward for modelling ungauged catchments is the integration of any useful information in a parsimonious and flexible model structure, whose complexity is to be adapted to the consistency of the available data base. Within this approach regional information is a valuable opportunity and the flexibility of the model structure is an essential requirement.

The idea is to derive through regional relationships quantitative signatures of the river flow regime like, for instance, selected flow statistics, and then carry out model calibration by searching the parameter sets that better reproduce the signatures themselves. The literature proposes several approaches to regional estimation of river flow statistics (traditional approaches are illustrated for instance in [32,17,25,13]) and presents applications to different hydrological problems and contexts such as peak flow estimation (see e.g., [12,6,7,20]) or frequency analysis of low flows (see e.g., [34,15,31,8,19]).

Key research questions are (1) the selection of the signatures and (2) how to search for the optimal parameter set. To address question (1), we refer to the maximum likelihood estimation procedure proposed by [40] in the context of time series models, which appears

* Corresponding author. Tel.: +39 051 2093356.
E-mail addresses: simone.castiglioni@mail.ing.unibo.it (S. Castiglioni), laura.lombardi6@unibo.it (L. Lombardi), elena.toth@unibo.it (E. Toth), attilio.castellarin@unibo.it (A. Castellarin), alberto.montanari@unibo.it (A. Montanari).

0309-1708/$ – see front matter
doi:10.1016/j.advwatres.2010.04.009

to be a promising tool in the PUB framework [22]. The Whittle's maximum likelihood is basically based on matching the mean value and the periodogram of observed and simulated time series (see Section 2.1). Given that the periodogram can be expressed as a function of the standard deviation and autocorrelogram of the river flows [5], we select these statistics, together with the mean value, as leading signatures for the calibration procedure. The reason for selecting the above likelihood function is that the related signatures, under certain assumptions, can be estimated by using regional information, as we will show below. Moreover, Whittle's likelihood has valuable statistical properties, namely, under mild assumptions it provides asymptotically consistent and normally distributed estimates. Even though the latter assumptions are generally not strictly satisfied for rainfall-runoff models and the PUB context, the estimator provides good approximations [22] and therefore appears to be a useful guide for the selection of the above signatures. To address question (2) above we apply a multiobjective calibration procedure to simultaneously maximise the fit of the selected signatures.

The proposed approach is applied in validation mode (that is, by assuming fully ungauged conditions) to 4 catchments located in a fairly hydrologically homogeneous region of central Italy, for which enough information is available to estimate the above statistics at the regional level. The results show that (1) the proposed approach to R-R model calibration is viable and flexible (i.e., it is not dependent on the model structure), and (2) the use of the available regional information enables the user to reduce the uncertainty of R-R models in applications to ungauged basins.

2. An approximate maximum likelihood procedure for the parsimonious calibration of rainfall-runoff models

Montanari and Toth [22] remarked that parameter calibration in the spectral domain can be an interesting opportunity for hydrological applications. Spectral methods were widely applied in the past for the statistical analysis of hydrologic data (see, for instance, [4,14]) but only recently they were considered for hydrological model calibration, by also focusing on wavelets analysis [28]. Spectral calibration is particularly interesting for ungauged basins because the statistical properties of river flows can be robustly inferred in the spectral domain by using sparse information like fragmented and old river flow records [22].

It is important to point out that the main difference between spectral calibration and traditional procedures is not merely the application in the spectral domain. In fact, Whittle's likelihood can be translated in the time domain, as the periodogram can be estimated as a function of river flow statistics. What makes Whittle's likelihood different is the nature of the above statistics and the weight attributed to them. Although Whittle's likelihood is conditioned by assumptions that are difficult to fully meet in the context of PUB (see, for instance, [11]), we believe it is particularly suitable for this very context in view of the possibility to regionalise its computation. The purpose of this paper is to provide a practical demonstration to this end. Within the present contribution we assume that the input data for the selected R-R model are available to produce model simulations for trial parameter sets, whereas no hydrometric observations are used in the ungauged calibrations.

2.1. The Whittle likelihood measure

The likelihood measure proposed by Whittle [40] for the parameters of a generic model will be denoted 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. For a stationary time series $L(\theta)$ is computed on the spectral density (see, for instance, [2,21]).

In order to introduce $L(\theta)$ for a R-R model, let us first focus on the gauged basin case, for which N river flow observations are available to calibrate the model. Once a model simulation of the N observations is available for a trial parameter set θ, with the same mean value for observed and simulated records, i.e. a zero mean simulation error series, the Whittle's likelihood can be computed through the approximate relationship

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

where $\lambda_j = 2\pi j/N$ are the Fourier frequencies; J is the periodogram of the series of the N observed river flows; J_M is the periodogram of the hydrological model output, that depends on the parameter vector θ, and J_e is the periodogram of the simulation error which depends on the standard deviation of the error itself. Under hypotheses on the spectral densities [11], Whittle's likelihood provides asymptotically consistent and normally distributed estimates in both the Gaussian and non-Gaussian, linear and non-linear cases. However, Eq. (1) is approximated because the likelihood should be computed on the spectral densities of the simulation and simulation error rather than on the periodograms (which are estimates for the corresponding spectral densities) and, moreover, the periodograms J_M and J_e are assumed to be mutually independent. For more details about the Whittle's likelihood the interested reader may refer to [2] and [11].

Note that maximisation of Eq. (1) is equivalent to minimisation of the absolute value of the summation at the right hand side. Moreover, in the ungauged case the spectral density J_e of the simulation error is unknown and therefore it is assumed to be identically null. Therefore, one has to make sure that the model simulation is unbiased (that is, the mean value of the river flows is satisfactorily reproduced) and the following objective function is then minimized in order to apply Whittle's approximate likelihood [22]:

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

in which the periodogram $J(\lambda_j)$ is estimated in the present study by using other information than the time series of observed flows. According to this procedure, observed river flows are not needed to perform the model calibration, but nevertheless the capability to carry out model simulations is required and therefore one needs input data for the model itself. Usually this is not a very limiting requirement because historical meteorological data (at least at coarse time scale) and macro-scale information on the catchment are often available in most parts of the world.

Minimising Eq. (2) implies that model calibration is carried out by essentially matching the mean value and the periodogram of model output and river flow. Given that any time series can be decomposed in a sum of periodic components through a harmonic analysis, the periodogram describes the variability of the time series that is explained by each component. It is essential to note that the periodogram can be derived from the autocovariance function of the data. Therefore one may say that Whittle's likelihood performs model calibration by essentially matching mean value and autocovariance function of model output and observed river flow record.

2.2. Approximating the Whittle likelihood measure in the time domain

In order to apply the proposed calibrating approach in ungauged basins one needs to estimate the mean value and periodogram in the absence of observed data. In particular, if one aims to use regional methods, the periodogram cannot be estimated by directly referring to the spectral domain, because it would be impossible to infer the river flow variability explained by each of the Fourier frequencies. However, a regional estimation of the periodogram is possible if one shifts to the time domain and estimates with a regional procedure mean value, standard deviation and autocorrelogram of the river flows. There is an extensive literature about regional estimation of mean value and variability while, to our knowledge, regionalization of the autocorrelogram was not inspected so far. A possible solution is to fit the autocorrelogram with a parametric function and then regionalise the related parameters. Accordingly, we assumed that the memory structure of the river flows can be well approximated with that of an Auto-Regressive first-order stochastic process (AR1, see [5]). We are aware of the physical limitations of this latter simplifying assumption, which was introduced for the sake of seeking a robust estimation of the river flow autocorrelation properties, and nevertheless it leads to a good fit for the short-term memory properties. The same does not hold for the long term autocorrelation, but in view of the uncertainty affecting PUB the approximation is believed to be acceptable. A more refined regionalization procedure, considering a two-parameter approximation for the autocorrelogram, is the subject of on-going work.

Therefore, under the above assumptions model calibration is equivalent to match the mean value $\mu(Q)$, standard deviation $\sigma(Q)$ and lag-1 autocorrelation coefficient $\rho_1(Q)$ of observed and simulated streamflows, where the weight attributed to each statistic depends on the river flow magnitude. We postulate that the statistics above can be derived by applying the regional procedure described in Section 5.1.

However, instead of identifying a single optimal solution we preferred to apply a 3-objective calibration procedure therefore identifying the full Pareto set of the non-dominated parameter vectors θ_n for the R-R model [43], in order to inspect a wider range of behavioral simulations in the presence of uncertainty in each of the above statistics. In this way the user might also choose the fitting accuracy for mean value, variability and persistence properties, depending on the scopes of the problem at hand and the specific behaviors of the case study. The envelope of the R-R model outputs obtained by running the model with all the θ_n provides an indication of the simulation uncertainty induced by fitting with different accuracy the individual patterns of $\mu(Q)$ and the periodogram resembled by $\sigma(Q)$ and $\rho_1(Q)$.

The multiobjective optimisation was carried out by using a recently developed multimethod evolutionary search, named AMALGAM [36]. AMALGAM (A Multi Algorithm Genetically Adaptive Method) runs simultaneously, for population evolution, a set of different optimisation methods (namely NSGA-II, Differential Evolution, Adaptive Metropolis Search and Particle Swarm Optimisation), resulting in a combination of the respective strengths by adaptively updating the weights of these individual methods based on their reproductive success (for more details see [36] and references therein). This ensures a fast, reliable and computationally efficient solution to multiobjective optimisation problems. The self-adaptive search properties of AMALGAM make it able to quickly adjust to the specific peculiarities and difficulties of each search problem and, in addition, it reduces the need for tuning of the algorithmic parameters.

In detail, we optimise with AMALGAM the following objective functions:

$$\phi_1(\theta) = \left[\frac{\mu[Q_s(\theta)] - \hat{\mu}(Q)}{\hat{\mu}(Q)}\right]^2 \tag{3}$$

$$\phi_2(\theta) = \left[\frac{\sigma[Q_s(\theta)] - \hat{\sigma}(Q)}{\hat{\sigma}(Q)}\right]^2 \tag{4}$$

$$\phi_3(\theta) = \left[\frac{\rho_1[Q_s(\theta)] - \hat{\rho}_1(Q)}{\hat{\rho}_1(Q)}\right]^2 \tag{5}$$

where $\mu[Q_s(\theta)]$, $\sigma[Q_s(\theta)]$ and $\rho_1[Q_s(\theta)]$ are the mean, standard deviation and lag-one autocorrelation coefficient of the simulated runoff series, and $\hat{\mu}(Q)$, $\hat{\sigma}(Q)$ and $\rho_1(Q)$ are the corresponding values obtained with regional relationships (see Section 5.1).

3. Rainfall-runoff model

The calibration method introduced in Section 2 is in principle applicable to any R-R model. However, in view of the application to the ungauged case it is advisable to restrain parameter uncertainty by using a parsimonious model. In this study we use HYMOD, a five-parameter lumped and conceptual model that was proposed by [3] and recently used by [37], [35] and [22] among others. This model consists of a relatively simple rainfall excess model that is connected to two series of linear reservoirs. In detail, three identical reservoirs in series are used to simulate the quick response and a single reservoir is adopted for the slow response. HYMOD input data are precipitation and evapotranspiration time series in addition to catchment area. The five parameters of the model are the maximum water storage capacity of the soil over the catchment, C [mm], a parameter representing the degree of spatial variability of the water storage capacity itself over the catchment, b [–], the fraction of rainfall excess that flows downstream as quick response, a [–], and the fractions of quick, R_q [–], and slow, R_s [–], reservoirs that empty at each time step.

4. Study area

The study area includes 52 catchments located in northern central Italy (Fig. 1) for which at least five years of daily series of streamflow records are available. It was checked that the river flow data are not significantly affected by water storages and withdrawals. This set of watersheds covers a wide range of hydrologic conditions. Table 1 reports the minimum, mean and maximum values of all selected geomorphological and climatic and river network attributes. In detail, the attributes are: drainage area, A (Km^2); percentage of permeable area, P (%); maximum, mean and minimum elevations, $H_{\max}$, H_{mean} and $H_{\min}$ (m above the sea level, m asl); average elevation relative to $H_{\min}$, $\Delta H = H_{\text{mean}} - H_{\min}$ (m); main channel length, L (km); concentration time, t_c (h); mean annual precipitation, MAP (m); mean annual temperature, TAM (C°) (see also [10]). The concentration time, t_c (h), was computed through Giandotti's equation (see [16] and also [10]) As shown in Table 1, mountainous and hilly, as well as pervious and practically impervious basins are included (see also [8], Table 1 on p. 958 and Section 6.3). The high variability of perviousness can be ascribed to the different geological units that are present in the study area, which range from clayey formation (northern portion) to fractured limestone (southern portion). The catchments are scarcely urbanized, with the upper part covered by pastures and broad leaved woods.

For a subset of 4 catchments (Candigliano River at Acqualagna, Metauro River at Barco di Bellaguardia, Esino River at Moie and Potenza River at Cannucciaro) simultaneous daily series of streamflow, precipitation and evapotranspiration are available with a minimum continuous record length of three and a half years. The observation periods are comprised between the years 1951 and 1978. Mean areal precipitation over the catchment was computed by applying the Thiessen polygons technique [8]. Daily evapotranspiration series at catchment scale were derived from the Thornthwaite's [33] equation. The quality of the data was tested and certified by the former National Hydrographic Service of Italy.

The 4 catchments (see Fig. 1) were selected for testing the proposed calibration procedure by assuming ungauged conditions. This means that the related river flow records were only used for verifying the model performances (see ALUC calibration in Section 5.3).

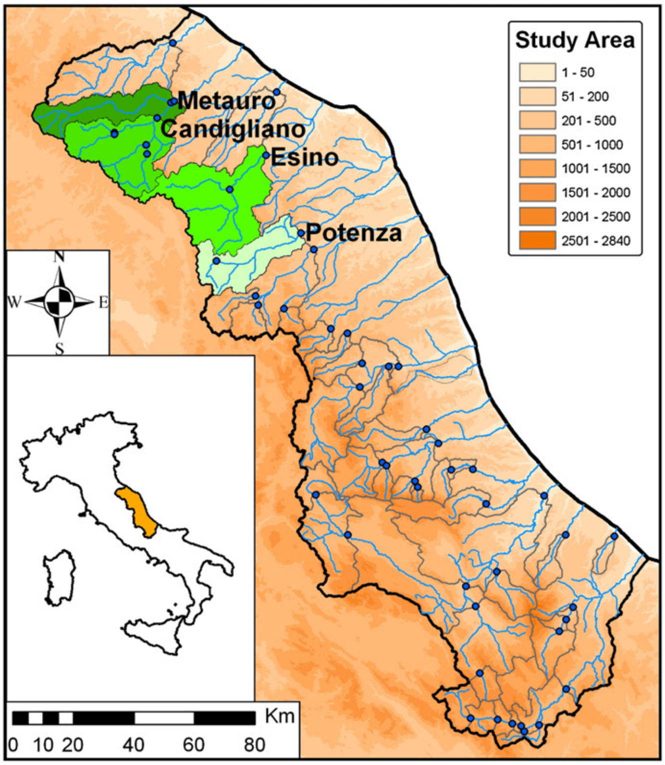

Fig. 1. Study area: 52 basins utilized in the regional analysis; 4 study catchments where the calibration procedure was tested in ungauged mode.

5. Rainfall-runoff model calibration in ungauged conditions

To optimise the HYMOD parameters a regional analysis was first developed in order to estimate $\mu(Q)$, $\sigma(Q)$ and $\rho_1(Q)$ depending on the geomorphological and climatic attributes.

5.1. Regional analysis

Three regional predictive models for $\mu(Q)$, $\sigma(Q)$ and $\rho_1(Q)$ were developed. To this aim, the geomorphological and climatic indexes for the 52 sites, whose ranges of variation are reported in Table 1, were regressed against the corresponding empirical values of $\mu(Q)$, $\sigma(Q)$ and $\rho_1(Q)$ through a multivariate stepwise regression analysis [39,8,9].

In detail, we considered linear, logarithmic and exponential regression relationships that, in this order, can be written as,

$$\gamma(Q) = a_0 + a_1x_1 + a_2x_2 + \ldots + a_nx_n \tag{6}$$

$$\gamma(Q) = a_0 + a_1\log(x_1) + a_2\log(x_2) + \ldots + a_n\log(x_n) \tag{7}$$

$$\gamma(Q) = a_0x_1^{a_1}x_2^{a_2}\ldots x_n^{a_n} \tag{8}$$

where $\gamma(Q)$ is either $\mu(Q)$, $\sigma(Q)$ or $\rho_1(Q)$, x_i for $i = 1, 2, \ldots, n$, is the i-th explanatory variable of the model (i.e., a selected geomorphological or climatic index), and a_i, for $i = 0, 1, \ldots, n$, is the i-th regression coefficient. The optimal subset of explanatory variables and the coefficients a_i were estimated through the stepwise analysis. This latter is carried out by adding at each step an explanatory variable, namely the one minimising the Mallows statistic [39]. Then, simpler regression models are tried by dropping in turn each one of the explanatory variables that were previously included. The model minimising the Mallows statistic is retained and then the procedure adds one more explanatory variable, until no further reductions in Mallows statistics can be obtained.

The following regression equations were finally identified:

$$\begin{aligned}\mu(Q) = &-3.907 + 0.016A - 0.003H_{\min} + \\ &-0.118L - 0.422t_c + 5.543MAP; \\ &E = 0.95; E_j = 0.94\end{aligned} \tag{9}$$

$$\begin{aligned}\sigma(Q) &= 0.0167P^{-0.133}\Delta H^{0.289}L^{1.200}MAP^{1.480}; \\ E &= 0.65; E_j = 0.51\end{aligned} \tag{10}$$

$$\begin{aligned}\rho_1(Q) &= 0.761 + 10^{-3}(1.6P - 0.03H_{\max} + 0.04\Delta H - 69.41MAP); \\ E &= 0.67; E_j = 0.59\end{aligned} \tag{11}$$

In the relationships above the values of the Nash efficiency [24] obtained in calibration, E, and in a jack-knife cross-validation, E_j are indicated. The efficiency varies between $-\infty$ and 1, where 1 indicates the perfect fit, and 0 stands for a model that performs as efficiently as a mean regional value. The jack-knife cross-validation procedure was carried out as in [8] and [9]. In detail, one catchment is supposed to be ungauged and therefore the related data set is eliminated. Then, Eqs. (9), (10) and (11) are recalibrated against the reduced data base. The obtained regional model is then used to estimate the statistics for the ungauged site. The procedure is repeated by excluding all stations in turn, thereby obtaining a set of 52 jack-knifed estimates of each statistic in ungauged mode. The small differences between E and corresponding E_j point out a satisfactory robustness for the considered regional models, which are therefore suitable for application in ungauged basins (see also [8,9]). The results of the jack-knife validation exercise are presented in Fig. 2.

5.2. Baseline calibrations

In order to better assess the performances of the calibration procedure in the ungauged case, we first carried out two baseline calibrations. The first is a classical calibration of HYMOD to the 4 considered basins by minimising the model residuals through a least squares procedure. Hereafter we will call it "classical calibration" (CC). This exercise aims at obtaining indications about the HYMOD performances in the considered catchments in gauged conditions. The second calibration was carried out by minimising with AMALGAM the objective functions (3), (4) and (5), but $\hat{\mu}(Q)$, $\hat{\sigma}(Q)$ and $\hat{\rho}_1(Q)$ were computed by using the observed record instead of the regional procedure described in Section 5.1. Given that the multiobjective search leads to the identification of a Pareto set and not to a single optimal solution, an average simulation was computed by taking at each time step the mean value of the simulations given by all the non-

Table 1
Minimum, mean and maximum values of geomorphological and climatic attributes of the 52 basins utilized in the regional analysis.

	A (km^2)	P (%)	H_{max} (m asl)	H_{mean} (m asl)	H_{min} (m asl)	ΔH (m)	L (km)	t_c (h)	MAP (m)	TAM (C°)
Minimum	14.4	0.1	279.5	178.0	3.0	158.3	5.3	0.9	0.820	8.3
Average	351.1	49.1	2085.7	959.3	363.8	595.4	36.2	6.4	1.099	11.7
Maximum	3082.0	99.0	2914.0	1950.0	1103.1	1543.2	159.9	18.9	1.530	15.3

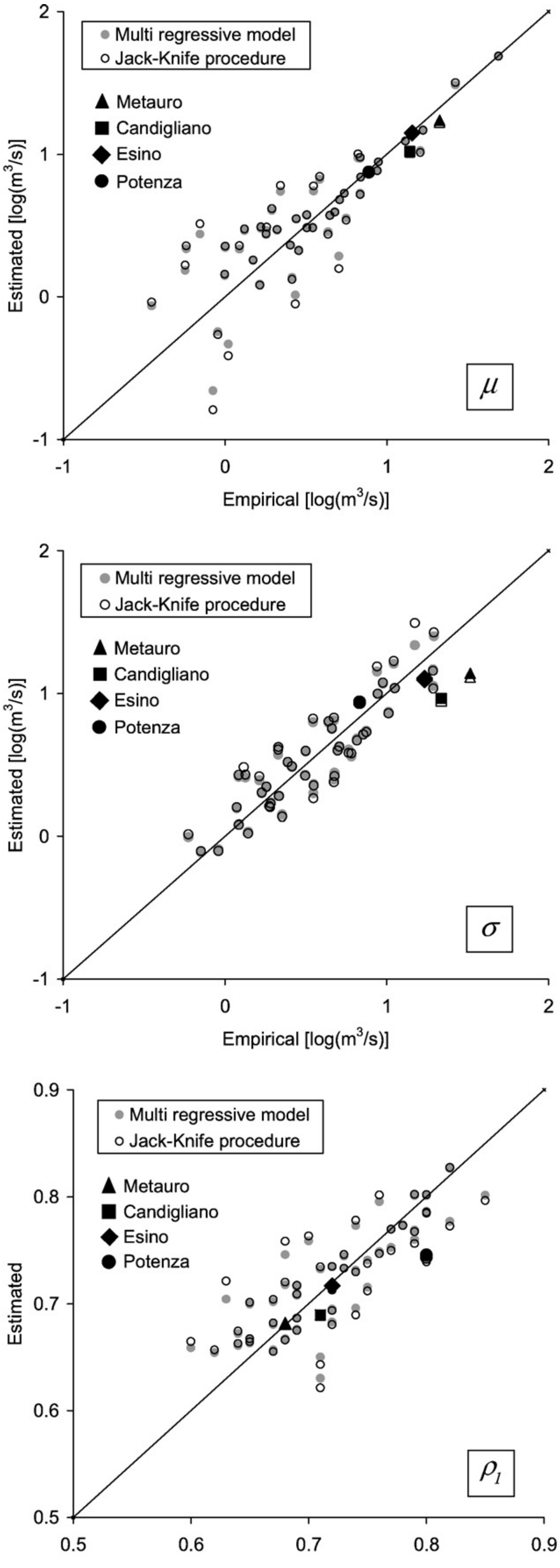

Fig. 2. Scatter-plots reporting empirical versus regional estimates of $\mu(Q)$, $\sigma(Q)$ and $\rho_1(Q)$; regional estimates obtained by multiregression models (9)–(11) are represented as dots (Multiregressive model), regional estimates obtained through cross-validation are represented as circles (Jack-knife procedure); empirical and estimated values of the statistics for the 4 test sites are also shown.

dominated parameter vectors. Hereafter we will call it "approximate likelihood direct calibration" (ALDC). This latter exercise aims first to quantify the approximation introduced by using the approximate likelihood instead of least squares. Of course a deterioration of the performances is expected because the long term variability of the river flow record is not described by neglecting the low frequencies of the periodogram. Second, a comparison with the ALDC procedure will allow quantifying the additional uncertainty induced by using regional statistics instead of those computed on the observed record. Table 2 reports the Nash efficiencies [24] obtained with the classical calibration and approximate likelihood direct calibration.

5.3. Approximate likelihood in the ungauged case

We identified the Pareto set of the non-dominated solutions for the test catchments by using AMALGAM to minimise the differences computed through (3), (4) and (5) between the statistics of the simulated data and the corresponding ones estimated through the regional regression equations. In particular, since we simulated the ungauged conditions for the four test sites, we used the jack-knife estimates for the three considered statistics (see Section 5.1). Therefore we were able to compute the envelope of the simulations provided by the non-dominated sets as well as an average simulation. This latter was used to compute the Nash efficiency with respect to the observed river flow record, which is reported in Table 2. Figs. 3–6 show the scatter-plots of observed river flows versus the average simulation. Figs. 7 and 8 report the obtained simulation envelopes and the respective observed records for the Candigliano and Metauro rivers, for a time window of 150 days around the event characterised by the highest peak flow during the observation period.

6. Discussion of results

Table 2 shows that a deterioration of the performances is experienced when using the approximate likelihood (ALDC) instead of least squares (CC) for all river basins. Nevertheless, keeping in mind that the approximate likelihood cannot take into account the low frequency behaviors (long term autocorrelation) of the river flow process, the results of this preliminary analysis are encouraging, showing the potential of the proposed approach. Further experiments are currently in progress, which also consider a more accurate description of the persistence properties instead of referring to $\rho_1(Q)$ only.

Concerning the calibration in ungauged mode (ALUC), one can see that a slight decrease of the model reliability is experienced for the Metauro River, for which the efficiency reduces from 0.58 to 0.51.

For the Candigliano and Esino catchments one notes that the average simulation in the ungauged case (ALUC) is unexpectedly characterised by a slightly improved efficiency with respect to the direct calibration (ALDC). This outcome may be justified by the presence of a significant model structural uncertainty, as shown by the efficiencies of HYMOD. When additional uncertainty is introduced by the ungauged calibration a compensation may occur therefore hiding the deterioration of the performances.

Table 2
Nash efficiencies obtained with classical calibration (CC), approximate likelihood direct calibration (ALDC) and approximate likelihood ungauged calibration (ALUC) for the test catchments.

Basin	CC	ALDC	ALUC
Candigliano	0.71	0.59	0.67
Metauro	0.63	0.58	0.51
Esino	0.61	0.51	0.55
Potenza	0.72	0.71	0.22

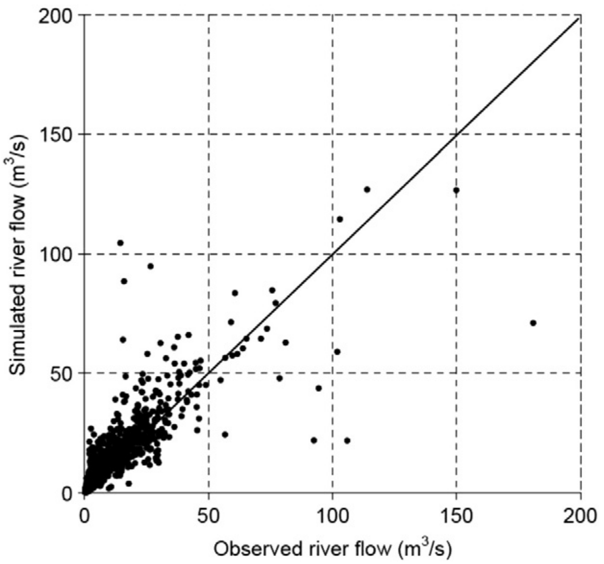

Fig. 3. Candigliano River at Acqualagna. Scatterplot of observed versus simulated (average simulation) river flows obtained with the non-dominated parameter sets identified by AMALGAM in ungauged conditions (ALUC).

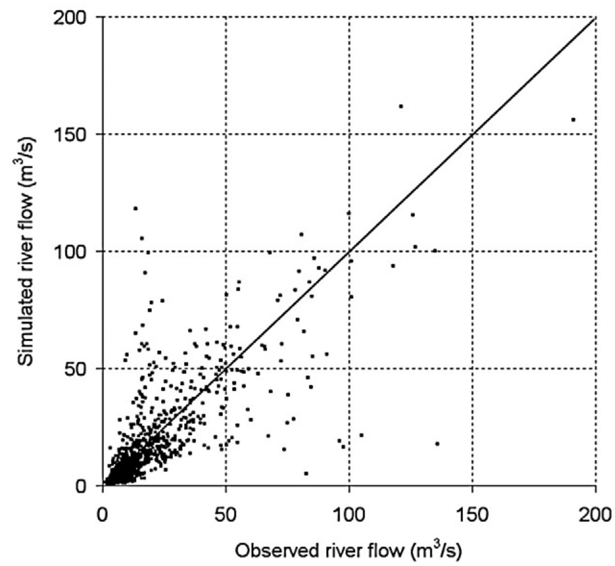

Fig. 5. Esino River at Moie. Scatterplot of observed versus simulated (average simulation) river flows obtained with the non-dominated parameter sets identified by AMALGAM in ungauged conditions (ALUC).

Figs. 3–6 show that the average simulations are fairly unbiased, with the exception of the Potenza River. For the Candigliano catchment the simulation of the highest river flows is generally satisfactory with a few exceptions. For the Metauro case the high flows are generally underestimated, but the shape of the hydrograph is overall satisfactorily reproduced, as confirmed by Fig. 8, where one can see that most of the observations are included in the simulation envelope. The Esino simulation shows no significant bias but the presence of some unreliability in the reproduction of low to medium discharges.

Even if these results point out that the model could be inaccurate for some practical applications, it is nevertheless interesting to note the significance of the information provided by the regional calibration, which could also enable the user to restrict the feasible space for the parameters.

As Table 2 shows, Nash efficiency decreases significantly in the ALUC experiment for the Potenza River, resulting equal to 0.22. This deterioration of performance is mainly due to a significant overestimation (see Fig. 6). Given the good ALDC performance for this very catchment, these poor results are probably to be ascribed to the regionalization procedure: in particular, a significant discrepancy between the empirical value and jack-knife estimate of $\rho_1(Q)$ should be noted in Fig. 2.

We believe it is worth discussing the value and the significance of the simulation envelopes shown in Figs. 7 and 8. As a matter of fact the terminology related to uncertainty used in the hydrological literature is not always clear [23]. Therefore it is necessary to stress how each envelope was obtained, that is, by encompassing all the simulations provided by the non-dominated (lying on the Pareto front) parameter sets identified by AMALGAM. Thus, the simulation envelope represents the uncertainty induced by the impossibility to identify a unique

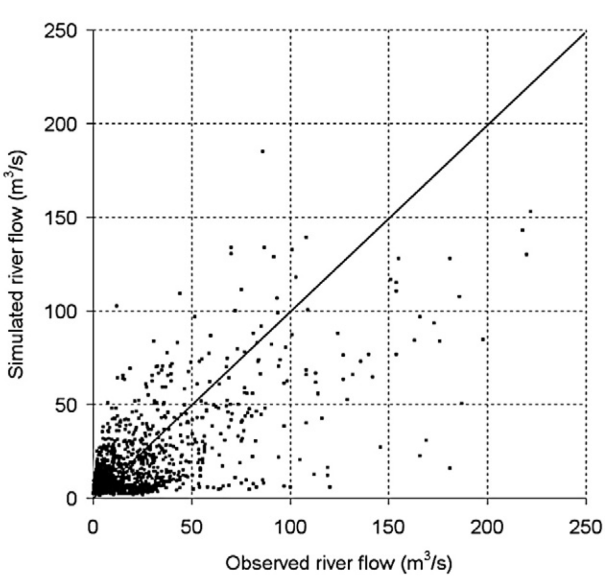

Fig. 4. Metauro River at Barco di Bellaguardia. Scatterplot of observed versus simulated (average simulation) river flows obtained with the non-dominated parameter sets identified by AMALGAM in ungauged conditions (ALUC).

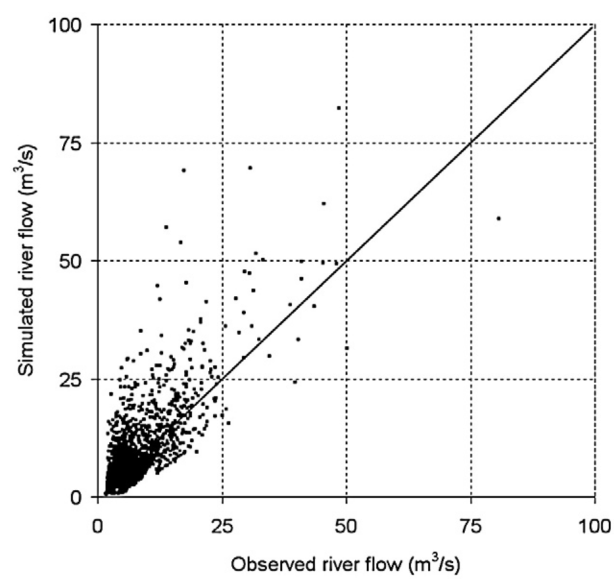

Fig. 6. Potenza River at Cannucciaro. Scatterplot of observed versus simulated (average simulation) river flows obtained with the non-dominated parameter sets identified by AMALGAM in ungauged conditions (ALUC).

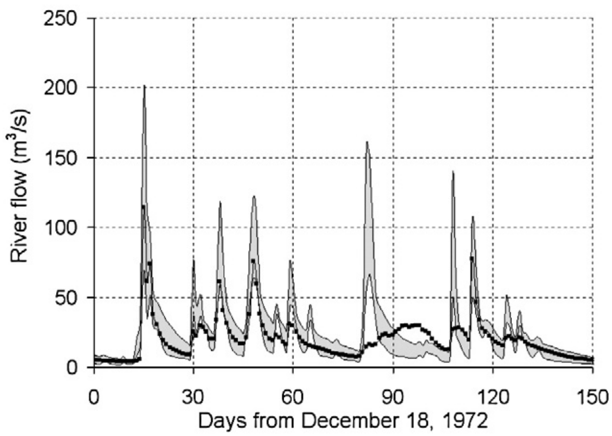

Fig. 7. Candigliano River at Acqualagna. Simulation envelope (in light grey) obtained with the non-dominated parameter sets identified by AMALGAM in ungauged conditions (ALUC), along with the observed record (thick line), for the flood event with the highest peak flow during the observation period.

solution that optimises all the three calibration objectives adopted in the study (i.e., $\mu(Q)$, $\sigma(Q)$ and $\rho_1(Q)$). This results in the non existence of a unique candidate for the maximum likelihood parameter set, but rather multiple candidates that originate a simulation envelope instead of one optimal simulation. It follows that the simulation envelope is not strictly representative of the actual discrepancy one would expect between the model output and the observations but just of the simulation variability one obtains by travelling along the Pareto front. The global uncertainty (i.e., the actual discrepancy between the model output and the observations) is expected to be larger, for the contribution given by additional sources of uncertainty like structural inadequacy of the model and limited reliability of the regional estimates. Accordingly, the obtained envelope is likely to represent a lower limit for the global uncertainty and, in any case, cannot be associated with an actual probability for an observation to fall within it.

7. Conclusion

The idea proposed in this paper is to carry out the R-R model calibration, in ungauged conditions, through an approximate regional maximum likelihood approach. This latter is obtained by maximising, through a multiobjective calibration procedure, the fit between the mean, standard deviation and lag-one autocorrelation coefficient of river flows simulated by the R-R model and the regional estimates of the same statistics. We tested the proposed procedure with respect to the case study of a wide geographical region in Italy. Within this region we selected 4 catchments for which the HYMOD R-R model was calibrated by assuming that no observations of river flows are available.

The results show that the regional calibration procedure is potentially able to convey useful information. In particular, regional information may be useful for constraining the feasible parameter space.

It is unlikely that the regional information is enough to calibrate a rainfall-runoff model with the reliability that is required in real world applications, especially if one takes into account that it might be problematic to estimate the global uncertainty affecting the model simulations. We propose an approach for estimating an envelope for the model output that provides an idea of simulation uncertainty, which was derived by encompassing the outputs provided by different parameter sets lying on the Pareto front. However, this envelope does not quantitatively estimate the actual discrepancy that one would expect between model and reality, which might be impossible to quantify in ungauged conditions.

The regional maximum likelihood may be an important piece of knowledge to aid model calibration, in view of the idea that in ungauged conditions the integration of different information is the way forward to obtain indications to reduce the feasible space for the model parameters. Indeed, we believe that integrating different types of soft data, for instance by following the approach recently proposed by [41], is an attractive perspective for process understanding and hydrological modelling in ungauged locations. Regional information could represent an avenue for future research, especially since it can be exploited through numerous methods hydrologists are familiar with. The integration of different competencies and skills may be a compelling and fascinating feature for better understanding and modelling ungauged basins.

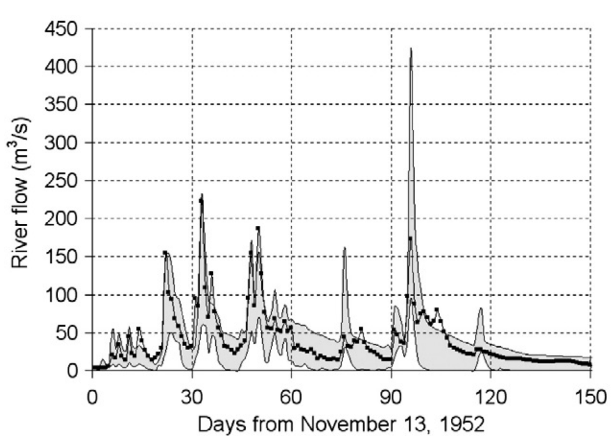

Fig. 8. Metauro River at Barco di Bellaguardia. Simulation envelope (in light grey) obtained with the non-dominated parameter sets identified by AMALGAM in ungauged conditions (ALUC), along with the observed record (thick line), for the flood event with the highest peak flow during the observation period.

Acknowledgements

The authors are grateful to Jasper Vrugt for providing the software for the application of AMALGAM. Three anonymous referees are also acknowledged. 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 Government through its national grants to the programme on “Relations between hydrological processes, climate, and physical attributes of the landscape at the regional and basin scales”.

References

[1] Bárdossy A. Calibration of hydrological model parameters for ungauged catchments. Hydrol Earth Syst Sci 2007;11:703–10.
[2] Beran J. Statistics for long-memory processes. New York: Chapman & Hall; 1994.
[3] Boyle, D.P., Multicriteria calibration of hydrological models, *Ph.D. dissertation*, Dep. of Hydrol. and Water Resour., Univ. of Ariz., Tempe., 2000.
[4] Bras RL, Rodriguez-Iturbe I. Random functions and hydrology. New York: Dover Publications; 1993.
[5] Brockwell PJ, Davis RA. Time series: theory and methods. New York: Springer; 1987.
[6] Burn DH. Evaluation of regional flood frequency analysis with a region of influence approach. Water Resour Res 1990;26(10):2257–65.
[7] Castellarin A, Burn DH, Brath A. Assessing the effectiveness of hydrological similarity measures for flood frequency analysis. J Hydrol 2001;241:270–85.
[8] Castellarin A, Galeati G, Brandimarte L, Montanari A, Brath A. Regional flow-duration curves: reliability for ungauged basins. Adv Water Resour 2004;27(10):953–65.
[9] Castellarin A, Camorani G, Brath A. Predicting annual and long-term flow-duration curves in ungauged basins. Adv Water Resour 2007;30(4):937–53.
[10] Castiglioni S, Castellarin A, Montanari A. Prediction of low-flow indices in ungauged basins through physiographical space-based interpolation. J Hydrol 2009;378:272–80.

[11] Contreras–Cristán A, Gutiérrez–Peña E, Walker SG. A note on Whittle's likelihood. Commun Stat–Simul Comput 2006;35:857–75.
[12] Dalrymple T. Flood frequency analyses. Reston (VA, USA): USGS Wat. Supply Paper 1543-A; 1960.
[13] FEH, Flood Estimation Handbook (FEH). Wallingford (UK) Institute of Hydrology, 1999.
[14] Fleming SW, Lavenue AM, Aly AH, Adams A. Practical applications of spectral analysis to hydrologic time series. Hydrol Processes 2001;16:565–74.
[15] Furey PR, Gupta VA. Space-time variability of low streamflows in river networks. Water Resour Res 2000;36(9):2679–90.
[16] Giandotti M. Previsione delle piene e e delle magre dei corsi d'acqua. Min. LL.PP., Servizio Idrografico Italiano, Memorie e studi idrografici, vol. 8, (Rep. No. 2), Roma, 1934.
[17] Hosking JRM, Wallis JR. Regional frequency analysis: an approach based on L-moments. Cambridge (UK: Cambridge University Press; 1997.
[18] Hundecha Y, Ouarda TBMJ, Bárdossy A. Regional estimation of parameters of a rainfall-runoff model at ungauged watersheds using the spatialstructures of the parameters within a canonical physiographic-climatic space. Water Resour Res 2008;44:W01427, doi:10.1029/2006WR005439.
[19] Laaha G, Blöschl G. Seasonality indices for regionalizing low flows. Hydrol Process 2006;20:3851–78.
[20] Merz R, Blöschl G. Flood frequency regionalisation-spatial proximity vs. catchment attributes. J Hydrol 2005;302(1–4):283–306.
[21] Montanari A, Rosso R, Taqqu MS. A seasonal fractional ARIMA model applied to the Nile River monthly flows at Aswan. Water Resour Res 2000:1249–59.
[22] Montanari A, Toth E. Calibration of hydrological models in the spectral domain: an opportunity for scarcely gauged basins? Water Resour Res 2007;43:W05434, doi:10.1029/2006WR005184.
[23] Montanari A. What do we mean by uncertainty? The need for a consistent wording about uncertainty assessment in hydrology. Hydrol Processes 2007;21: 841–5.
[24] Nash J, Sutcliffe J. River flow forecasting through conceptual models, Part 1: a discussion of principles. J Hydrol 1970;10:282–90.
[25] Pandey GR, Nguyen VTV. A comparative study of regression based methods in regional flood frequency analysis. J Hydrol 1999;225(1–2):92–101.
[26] Parajka J, Merz R, Blöschl G. A comparison of regionalisation methods for catchment model parameters. Hydrol Earth Syst Sci 2005;9:157–71.
[27] Pokhrel P, Gupta HV, Wagener T. A spatial regularization approach to parameter estimation for a distributed watershed model. Water Resour Res 2008;44: W12419, doi:10.1029/2007WR006615.
[28] Schaefli B, Zehe E. Hydrological model performance and parameter estimation in the wavelet-domain. Hydrol Earth Syst Sci 2009;13:1921–36.
[29] Seibert J, Uhlenbrook S, Leibundgut C, Halldin S. Multiscale calibration and validation of a conceptual rainfall-runoff model. Phys Chem Earth, Part B: Hydrol Oceans Atmosphere 1998;25:59–64.
[30] Sivapalan M. Prediction in ungauged basins: a grand challenge for theoretical hydrology. Hydrol Process 2003;17:3163–70.
[31] Smakhtin VU. Low flow in hydrology: a review. J Hydrol 2001;240:147–86.
[32] Stedinger JR, Vogel RM, Foufoula-Georgiou E. Chp. 18 Frequency analysis of extreme events. In: Maidment DR, editor. Handbook of hydrology. York (NY, USA): McGraw-Hill Inc.; 1993.
[33] Thornthwaite CW. An approach toward a rational classification of climate. Trans Am Geophysys Union 1946;27(1).
[34] Vogel RM, Kroll CN. Regional geohydrologic–geomorphic relationships for the estimation of low-flow statistics. Water Resour Res 1992;28(9):2451–8.
[35] Vrugt JA, Gupta HV, Bouten W, Sorooshian S, Shuffled A. Complex Evolution Metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters. Water Resour Res 2003;39:1201, doi:10.1029/2002WR001642.
[36] Vrugt JA, Robinson BA. Improved evolutionary optimization from genetically adaptive multimethod search. Proc Natl Acad Sci USA 2007;104:708–11.
[37] Wagener T, Boyle DP, Lees MJ, Wheater HS, Gupta HV, Sorooshian S. A framework for development and application of hydrological models. Hydrol Earth Syst Sci 2001;5:13–26.
[38] Wagener T, Wheater HS. Parameter estimation and regionalization for continuous rainfall-runoff models including uncertainty. J Hydrol 2006;320:132–54.
[39] Weisberg S. Applied linear regression2nd ed. . New York: John Wiley; 1985.
[40] Whittle P. Estimation and information in stationary time series. Ark Mat 1953;2: 423–34.
[41] Winsemius HC, Schaefli B, Montanari A, Savenije HHG. On the calibration of hydrological models in ungauged basins: A framework for integrating hard and soft hydrological information, Water Resour Res 2009;45:W12422, doi:10.1029/2009WR007706.
[42] Yadav M, Wagener T, Gupta H. Regionalization of constraints on expected watershed response behavior for improved predictions in ungauged basins. Adv Water Resour 2007;30(8):1756–74.
[43] Yapo PO, Gupta HV, Sorooshian S. Multi-objective global optimization for hydrologic models. J Hydrol 1998;204:83–97.
[44] Zhang Z, Wagener T, Reed P, Bhushan R. Reducing uncertainty in predictions in ungauged basins by combining hydrologic indices regionalization and multiobjective optimization. Water Resour Res 2008;44:W00B04, doi:10.1029/2008WR006833.