7-1-2010

Improving Robustness of Hydrologic Parameter Estimation by the Use of Moving Block Bootstrap Resampling

Hamid Moradkhani
Portland State University

Mohammad Ebtehaj

Hoshin V. Gupta

Citation Details

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

Improving robustness of hydrologic parameter estimation by the use of moving block bootstrap resampling

Mohammad Ebtehaj,[1] Hamid Moradkhani,[2] and Hoshin V. Gupta[3]

Received 14 March 2009; revised 20 January 2010; accepted 22 February 2010; published 16 July 2010.

[1] Modeling of natural systems typically involves conceptualization and parameterization to simplify the representations of the underlying process. Objective methods for estimation of the model parameters then require optimization of a cost function, representing a measure of distance between the observations and the corresponding model predictions, typically by calibration in a static batch mode and/or via some dynamic recursive optimization approach. Recently, there has been a focus on the development of parameter estimation methods that appropriately account for different sources of uncertainty. In this context, we introduce an approach to sample the optimal parameter space that uses nonparametric block bootstrapping coupled with global optimization. We demonstrate the applicability of this procedure via a case study, in which we estimate the parameter uncertainty resulting from uncertainty in the forcing data and evaluate its impacts on the resulting streamflow simulations.

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

1. Introduction

[2] To estimate the streamflow response of a catchment, the conventional approach has been to rely on human expertise to manually calibrate a hydrologic model, exploiting the knowledge and experience of the forecaster while taking into account local hydroclimatic conditions. During the past several decades, various methods have been developed to automate and systematize the process, helping to enhance the accuracy and quality of such predictions (e.g., *Sorooshian and Dracup* [1980], *Duan et al.* [1992], *Beven and Binley* [1992], *Sorooshian et al.* [1993], *Gupta et al.* [1998], *Vrugt et al.* [2003], *Moradkhani et al.* [2005a, 2005b], *Kavetski et al.* [2006], *Tolson and Shoemaker* [2007], among many others). While considerable progress has been made, critics of the automated parameter estimation approach suggest that such procedures do not suitably take into account the coherence of the underlying physical processes, as a consequence of which the estimated parameters may not fall within acceptable, physically meaningful, ranges [e.g., *Burnash*, 1995]. Nonetheless, as new and more complex hydrologic models continue to be developed and computational resources become more powerful, the need for robust statistical and mathematical approaches that help to increase the accuracy of forecasts will continue to grow.

[3] By treating parameter estimation as a problem of constrained deterministic optimization, a number of well-known algorithms have been developed. Popular methods for the calibration of hydrologic models include the Shuffled Complex Evolution method (SCE-UA) [*Duan et al.*, 1992], the Bayesian Total Error Analysis framework (BATEA) [*Kavetski et al.*, 2006], and the Dynamically Dimensioned Search algorithm (DDS) [*Tolson and Shoemaker*, 2007]. *Vrugt et al.* [2003] extended the context to acknowledge the probabilistic nature of the parameter estimation problem, and introduced the Shuffled Complex Evolution Metropolis (SCEM-UA) algorithm that provides an estimate of the posterior joint probability density function of the parameters. The robustness of these algorithms has been investigated extensively in the hydrologic literature, and numerous successful applications to calibration of hydrologic models have been reported (e.g., *Yapo et al.* [1996], *Blasone et al.* [2008], *Tolson and Shoemaker* [2007], among many others).

[4] Regardless of the approach used for solving the optimization problem, several questions arise concerning the reliability of operational forecasts derived using the resulting parameter estimates. These include (1) what confidence bounds can be placed on the estimated parameters, given the different sources of uncertainty arising from observational and model parameterization errors? (2) How can one estimate the impact that use of a limited amount of calibration data will have on the confidence bounds attributed to the estimated parameters? (3) How do these various sources of uncertainty affect the hydrologic forecasts?

[5] Furthermore, although the optimization context typically assumes that the model parameters represent time invariant properties of the catchment, the various sources of uncertainty mentioned above will tend to cause the parameter estimates to deviate (sometimes quite significantly) from their underlying "true" or population values. Several probabilistic approaches have been proposed to explain this

[1]Saint Anthony Falls Laboratory, National Center for Earth Surface Dynamics, Department of Civil Engineering, University of Minnesota, Minneapolis, Minnesota, USA.

[2]Department of Civil and Environmental Engineering, Maseeh College of Engineering and Computer Science, Portland State University, Portland, Oregon, USA.

[3]Department of Hydrology and Water Resources, University of Arizona, Tucson, Arizona, USA.

0043-1397/10/2009WR007981

variability, using the framework of statistical inference [e.g., *Beven and Binley*, 1992; *Thiemann et al.*, 2001; *Vrugt et al.*, 2003, 2008; *Moradkhani et al.*, 2005b; *Blasone et al.*, 2008; *Kavetski et al.*, 2006]. In general, these studies pose the problem in the context of Bayesian inference, so that observational data are used to update a "prior" hypothesis (presented as a probability density on the parameter estimates). As the number of sample observations is increased (as evidence is accumulated over time) the Bayesian framework enables the degree of belief attributed to a given hypothesis to evolve.

[6] In this study we explore the use of bootstrap resampling to estimate the uncertainty in "optimal" parameters estimates that arises due to uncertainties in the forcing data. While bootstrap resampling has typically been used to generate synthetic streamflow time series having certain desirable statistical properties [e.g., *Vogel and Shallcross*, 1996; *Sharma et al.*, 1997; *Srinivas and Srinivasan*, 2005, 2006], we employ it here in the context of Monte Carlo sampling to arrive at an estimate of the unknown true parameter hypothesis (the probabilistic structure of the optimal parameter estimates) on the basis of the sampled data.

[7] The paper is organized as follows. Section 2 presents the statistical context for nonparametric bootstrap resampling of independent and dependent data sets. Section 3 presents an overview of the proposed procedure and discusses the theoretical conditions to be met for proper implementation of block bootstrap resampling with respect to the statistical characterization of the hydrologic time series. Section 4 presents a case study using a well-known data set to investigate the applicability of the approach. Section 5 couples the block bootstrap method with the SCE-UA global optimization method to obtain an estimate of the distribution of optimal parameters, and to compute the parameter and associated forecast confidence intervals. Section 6 discusses the computational cost of the proposed method and section 7 presents some concluding remarks. A detailed description of the independent bootstrap sampling method is presented in Appendix A.

2. Review of the Bootstrap Methods

2.1. Bootstrapping of Independent Data

[8] Inferential statistical methods are used to draw conclusions about particular properties of a given population on the basis of finite sample size. Most conventional parametric hypotheses testing methods (e.g., t-test, F-test, chi-square, and ANOVA) rely on normality assumptions, and are designed for simple statistics such as the mean. Deviations from normality often cause the results of such tests to be inaccurate. Further, closed-form mathematical expressions often do not exist for the sample distribution of various statistics of interest (e.g., median, higher order moments, quantiles), and parametric construction of the confidence intervals may not be achievable. Bootstrapping is a computationally intensive statistical method that allows us to relax some of the conditions and assumptions required by traditional parametric inference. As a branch of the broader class of resampling techniques, Bootstrapping can be used to estimate the statistical properties of a population by sampling from an approximate or empirically constructed distribution without needing to make prior assumptions regarding its mathematical form. Initially proposed by *Efron* [1979], the method forms a nonparametric procedure for population inference from independent identically distributed (*iid*) samples.

[9] Let $\boldsymbol{X} = \{X_1, X_2, \ldots, X_n\}$ represent a set of *iid* random samples of size n from an unknown distribution $\boldsymbol{F}$ (denoted as $X_i \sim_{iid} \boldsymbol{F}$), and let $\boldsymbol{x} = \{x_1, x_2, \ldots, x_n\}$ be an observed realization. Our goal is to approximate the sample distribution of a pivotal statistic $\boldsymbol{R}_n(\boldsymbol{X}, \boldsymbol{F})$, which possibly depends on both $\boldsymbol{X}$ and $\boldsymbol{F}$, on the basis of the observed sample data $\boldsymbol{x}$ [*Efron*, 1979]. By definition, the pivotal form of a statistic is a transformation of that statistic such that its sample distribution holds a mathematically tractable form (e.g., normal distribution) and its shape is independent of the population distribution. For example, $\boldsymbol{R}_n(\boldsymbol{X}, \boldsymbol{F})$ might be of the form $\sqrt{n}[\bar{x}_n - \mu(\boldsymbol{F})]$ or possibly a studentized version, where $\bar{x}_n$ is the sample mean and $\mu(\boldsymbol{F})$ denotes the population mean.

[10] The main idea of bootstrap estimation of the sample distribution of $\boldsymbol{R}_n(\boldsymbol{X}, \boldsymbol{F})$ can be explained as follows:

[11] 1. Construct an empirical population distribution $\hat{\boldsymbol{F}}$ (Bootstrap population) by uniform sampling with replacement from the observed sample $\boldsymbol{x} = \{x_1, x_2, \ldots, x_n\}$ with the probability of x_i set equal to 1/n.

[12] 2. Draw *iid* random samples X_i^* of size n, from $\hat{\boldsymbol{F}}$, and call them bootstrap samples:

$$X_i^* = x_i^*, \qquad X_i^* \sim_{iid} \hat{\boldsymbol{F}}, \qquad i = 1, 2, \ldots, n. \tag{1}$$

[13] 3. Approximate the sampling distribution of the statistic $\boldsymbol{R}_n(\boldsymbol{X}, \boldsymbol{F})$, using the bootstrap distribution of $\boldsymbol{R}_n(\boldsymbol{X}^*, \hat{\boldsymbol{F}})$.

[14] To supplement this explanation, the procedure for bootstrap estimation of the distribution of the sample mean is explained in the Appendix.

2.2. Bootstrapping of Dependent Stationary Data

[15] For dependent processes, construction of the Bootstrap population is more complicated and far less obvious than for the independent setup presented by *Efron* [1979]. Block bootstrapping (BB) for weakly dependent stationary time series was proposed independently by *Künsch* [1989] and *Liu and Singh* [1992]. In this method, the sample distribution of the estimator is calculated using pseudo time series obtained by sampling with replacement from whole blocks of consecutive observations.

[16] The notion of weakly dependent data structures was first introduced by *Rosenblatt* [1965] to prove the central limit theorem. Considering A and B to be two subsets of a bounded sequence of stationary random numbers $\{\boldsymbol{X}_n, n \in \mathrm{N}\}$ and defining a distance measure between those sets as $\boldsymbol{d}(A, B)$, *Rosenblatt* [1965] argued that the strict weak dependence condition is satisfied for all subsets A and B of $\{X_n\}$, when the following holds in the space of probability measures:

$$|\boldsymbol{P}(A \cap B) - \boldsymbol{P}(A).\boldsymbol{P}(B)| < f[\boldsymbol{d}(A, B)], \tag{2}$$

where the function $f[\boldsymbol{d}(A, B)] \to 0$ as $\boldsymbol{d}(A, B) \to \infty$. By relaxing the Rosenblatt conditions in terms of the second-order statistics, $\{\boldsymbol{X}_n, n \in \mathbb{N}\}$, we get what is known as a second-order weakly dependent or asymptotically uncorrelated process, in which the covariance $Cov(X_n, X_{n+d}) \to 0$ as $d \to \infty$.

[17] In recent years, several different block bootstrap techniques have been developed for reproducing the structural characteristics of dependent stationary time series [*Carlstein*, 1986; *Künsch*, 1989; *Liu and Singh*, 1992; *Politis and Romano*, 1994; *Lahiri*, 1995, 1999; *Horowitz*, 2003; *Zoubir and Iskander*, 2004]. *Carlstein* [1986] proposed a nonoverlapping block bootstrap method to estimate the distribution of sample variance of a stationary process. *Künsch* [1989] and *Liu and Singh* [1992] independently suggested overlapping moving block bootstrap (MBB) and proved asymptotic convergence of the method for a broad range of statistics of a stationary process. *Politis et al.* [1992] generalized the sampling method by presenting a circular block-sampling method, which allows the blocks to start near the end of the data and wrap around the starting point. Further to the development of circular block bootstrapping, *Politis and Romano* [1994] proposed a random base block length method, the so-called stationary bootstrap (SB). *Politis* [2001] and *Synowiecki* [2007] showed that by restricting the sampled block length the original idea of the block bootstrap can also be applied consistently for cyclostationary time series.

[18] MBB and SB are the two algorithms most widely used for bootstrap estimation of weakly dependent discrete stochastic processes. In this section, the principal similarities and differences between these two methods of block resampling are explained in brief.

[19] Suppose that $\boldsymbol{x} = \{x_1, x_2, \ldots, x_n\}$ is a sample realization of size n from a real valued stationary and weakly dependent process $\{\boldsymbol{X}_n\}$ and $\boldsymbol{R}_n(\boldsymbol{X}, \boldsymbol{F})$ is a pivotal statistic with some functional dependence on the observed data and the process of joint distribution $\boldsymbol{F}$. Because this is a dependent process, *Künsch* [1989] proposed the use of block resampling for construction of an empirical estimate of the bootstrap joint distribution of $\boldsymbol{R}_n(\boldsymbol{X}, \boldsymbol{F})$. To describe the algorithm, assume a constant block size l, such that $l/n \to 0$ as $n,l \to \infty$. Let, $B_{i,l} = \{x_i, x_{i+1}, \ldots, x_{i+l-1}\}$ denote all possible overlapping blocks of observations with length l that can be sampled from $\{x_n\}$, where $i = 1, \ldots, n - l + 1$. In the original MBB, the blocks are randomly sampled with replacement $b_1 = \lceil n/l \rceil$ times, where $\lceil \bullet \rceil$ denotes the integer part, from $\{B_{1,l}, B_{2,l}, \ldots, B_{(n-l+1),l}\}$ and concatenated to generate a pseudo time series of bootstrap versions $\boldsymbol{X}_b^* = \{x_1^*, x_2^*, \ldots, x_{n1}^*\}$ of the observation $\{x_n\}$, where, $n_1 = b_1 l$. Accordingly, the sample distribution of a particular statistic can be estimated by repeating the above process several times, say, $b = 1, \ldots, B$.

[20] In the SB method, instead of using fixed block lengths, the lengths of the sampled blocks are randomized using a geometric distribution. A circular sampling method proposed by *Politis et al.* [1992] and *Politis and Romano* [1994] can be used to preserve the stationarity of the generated pseudo series in a strict sense. Let $B_{i,l} = \{x_i, x_{i+1}, \ldots, x_j, \ldots, x_{i+l-1}\}$ be the data block with a random length of l starting from x_i. The block length l is sampled randomly from the geometric distribution $P_L(l) = (1 - p)^{l-1}p$, where $p \in (0,1]$. The index i of the block starting point $x_i \in x$, is sampled from a sequence of *iid* random numbers distributed uniformly on $1 \le i \le n$. When $j > n$, x_j is defined to be x_i, where $i = j (\mathrm{mod}\ n)$ and $x_0 = x_n$. By repeating this block construction procedure and by concatenating the sequences, a bootstrap pseudo time series $\boldsymbol{X}_b^* = \{x_1^*, x_2^*, \ldots, x_N^*\}$ can be generated that preserves the original stationarity characteristics of the observations in a strict sense. Of course the concatenation process stops once the size of the bootstrap pseudo time series N exceeds n. As before, by constructing a large number of bootstrap pseudo time series (i.e., $b = 1, \ldots, B$), an approximate sample distribution for $\boldsymbol{R}_n(\boldsymbol{X}, \boldsymbol{F})$ can be computed. In this paper, we will focus on the MBB method, and leave comparisons of the MBB and SB methods and further investigation regarding the optimum block length for future work.

2.3. Smoothed Bootstrap Resampling

[21] In bootstrap estimation, owing to the size and properties of the original sample, it is possible for the samples to contain repeated values, thereby making the bootstrap empirical distribution $\hat{\boldsymbol{F}}$ discrete and irregular. The smoothed bootstrap [*Efron*, 1979, 1982] modifies the original bootstrap approach to improve the properties of the bootstrap samples. Instead of resampling directly from the empirical distribution, the smoothed bootstrapapproach uses a smoothed version $\hat{\boldsymbol{f}}$ of $\hat{\boldsymbol{F}}$ to generate the bootstrap samples. *Efron* [1982] showed, by direct simulation, that smoothed bootstrap resampling can result in a smaller standard error of estimation of the correlation structure of a data set. *Young* [1990] also studied the estimation of quadratic population functionals (i.e., $\int \psi(x)^2 d\boldsymbol{F}(x)$), where $\psi(x)$ is a specified function) and argued that the smoothed bootstrapcan provide better results than the standard nonparametric version. *Efron* [1979, 1982] and *Silverman and Young* [1987] defined the smoothed bootstrapdistribution as

$$\hat{\boldsymbol{f}}(t) = n^{-1}h^{-1}\sum_{i=1}^{n}\boldsymbol{K}\left(\frac{t - x_i}{h}\right), \tag{3}$$

where the density $\hat{\boldsymbol{f}}$ is a kernel estimator of the population density $\boldsymbol{F}$ of interest, given a smoothing bandwidth parameter, h, that needs to be determined. The function $\boldsymbol{K}(\bullet)$ is assumed to be a symmetric zero mean density function with unit variance. Accordingly, using a relevant kernel function such as a Gaussian kernel to generate smoothed bootstrap samples $\hat{\boldsymbol{x}}^* = \{\hat{x}_1^*, \hat{x}_2^*, \ldots, \hat{x}_N^*\}$ from the $\hat{\boldsymbol{f}}$, it suffices to uniformly sample with replacement from the available observations $\boldsymbol{X} = \{x_1, x_2, \ldots, x_n\}$ and then perturb the sample by adding random noise as follows [*Young*, 1990]:

$$\hat{x}_i^* = x_i + \varepsilon_i, \tag{4}$$

where, $i = 1, \ldots, n$ and ε_i are *iid* random variables with the density $f_\varepsilon(t) = h^{-1}\boldsymbol{K}(t/h)$.

3. Block Bootstrapping for Hydrologic Parameter Estimation

[22] From a statistical standpoint, available observations of streamflow, precipitation, and other hydrological data, should be treated as finite-length samples from a population of very large size, one that is not entirely observable. In the context of the estimation of hydrologic model parameters, it is typical for the computed optimal estimates to be nonlinear functionals of the sample observations. Therefore the true population distribution of the optimal hydrologic parameters $\boldsymbol{F}_{Hp}$ can be approximated by the empirical bootstrap distribution $\hat{\boldsymbol{F}}_{Hp}$ as long as the assumptions of the method are satisfied. Because the data time series used for model cali-

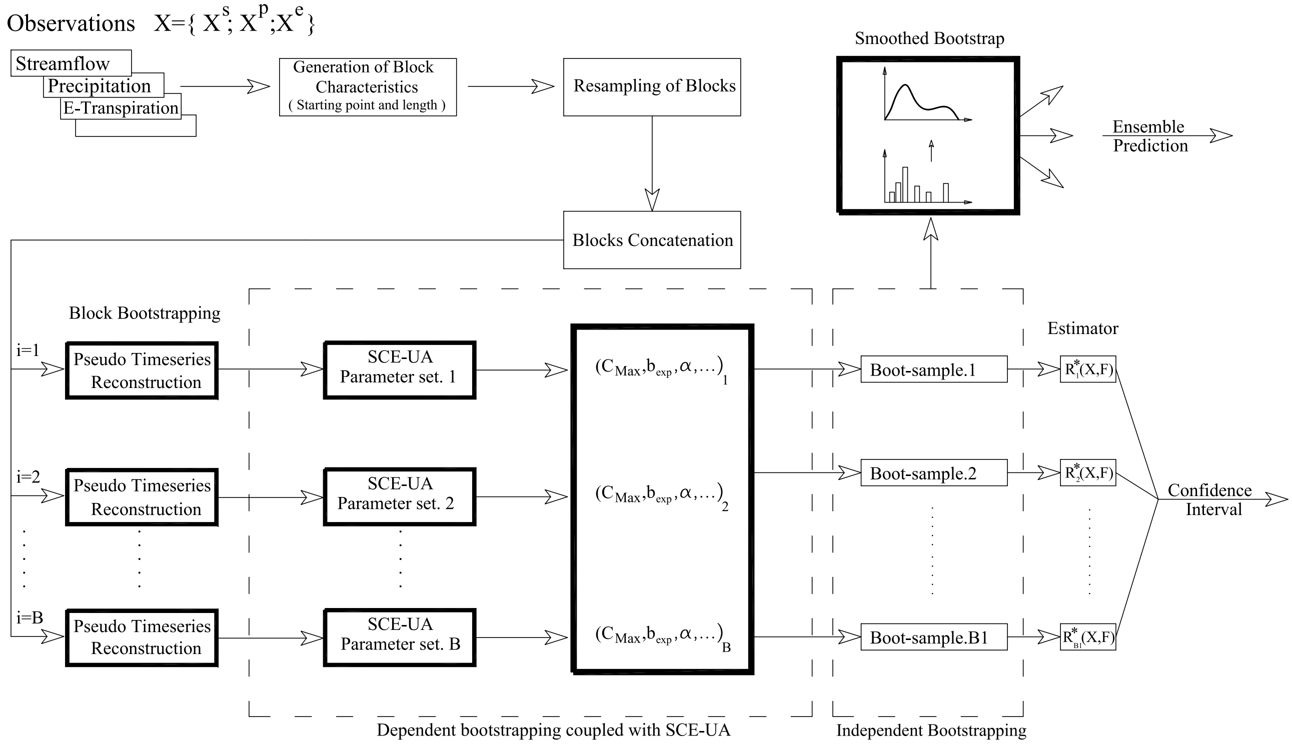

Figure 1. The proposed framework to couple block bootstrapping with global optimization for estimating the distribution of the optimal hydrologic parameters by calibration of a hydrologic model.

bration typically exhibit a correlation structure, we must use the *dependent bootstrap* approach to sample the data space when estimating the hydrologic parameters. Furthermore, we will need to account for the periodicities that are characteristic of most natural time series; these arise from the axial rotation of the Earth and its revolution around the Sun, introducing strong periodic components at different time scales into the overall variability of the natural time series and making them nonstationary in a strict sense.

[23] Broadly speaking, a real valued discrete stochastic process $\{\boldsymbol{X}_n, n \in \mathbb{N}\}$ is called cyclostationary with period T_0, when the joint density of $\{X_n, X_{n+1},\ldots, X_{n+kT_0}\}$ $\forall n, k \in \mathbb{N}$, is independent of n for a constant value of k (see comprehensive explanation of *Gardner et al.* [1985, 2005]). In the context of block bootstrap estimation of cyclostationary dependent processes, *Politis* [2001] has shown that when the block size is restricted to be an integer multiplier of the process period T_0, the method of block bootstrapping can be applied for estimating the process statistics. To apply this to bootstrap estimation of hydrologic parameters, it is sufficient to restrict the block sampling lengths of the data time series to coincide with their annual periodic structures.

[24] Accordingly, we propose the modulated framework illustrated in Figure 1. For a selected model calibration period, a number (say, B) of pseudo time series can be generated using either the moving block bootstrap or the stationary bootstrap method, taking into account the periodic structure of the time series. For each pseudo time series, the model is calibrated to obtain a corresponding optimal estimate for the parameters. Consequently, we obtain $b = 1,\ldots, B$ sets of Bootstrapped Optimal Hydrologic Parameter (BOHP) estimates $\theta^*_b = \{\theta^*_1, \theta^*_2,\ldots, \theta^*_{n_h}\}$ (where n_h denotes the number of model parameters) constituting samples from the empirical population distribution $\hat{\boldsymbol{F}}_\theta$. To relax the inferential constraints arising from the limited sample size of the BOHP, we use independent and smoothed bootstrap sampling. By repeated sampling ($b_1 = 1,\ldots, B_1$) with replacement from the BOHP, the approximate sample distribution of various estimators of the parameters (e.g., mean, median, quantiles, and so on), along with their associated confidence intervals, can be computed without the need to make prior assumptions regarding their underlying distribution. Moreover, by employing smoothed bootstrap sampling, a nonparametric smoothed density model can be derived for the parameters, obviating the need to rely on the histogram estimator; this smoothed density model enables us to inexpensively generate larger sample sizes of the parameter estimates for the generation of ensemble simulations by propagation through the hydrologic model.

4. Case Study

4.1. Study Area and Data

[25] The Leaf River Basin, a subbasin of the Pascagoula watershed, originates in west-central Mississippi and drains nearly 9200 km^2. The basin (more than 70% forested) has an approximate total length of 170 km and a maximum width of 90 km, and comprises low rolling hills and ridges, characteristic of the coastal plains. Because of the hydrometeorological and geomorphic characteristics of the watershed, it has experienced several devastating floods [*U.S. Army Corps of Engineers Report*, 1975] and has been studied extensively (e.g., *Yapo et al.* [1996], *Vrugt et al.* [2003], *Moradkhani et al.* [2005a, 2005b], among many others). Forty years (i.e., October 1948 to September 1988) of observed hydrological time series are available for the Collins

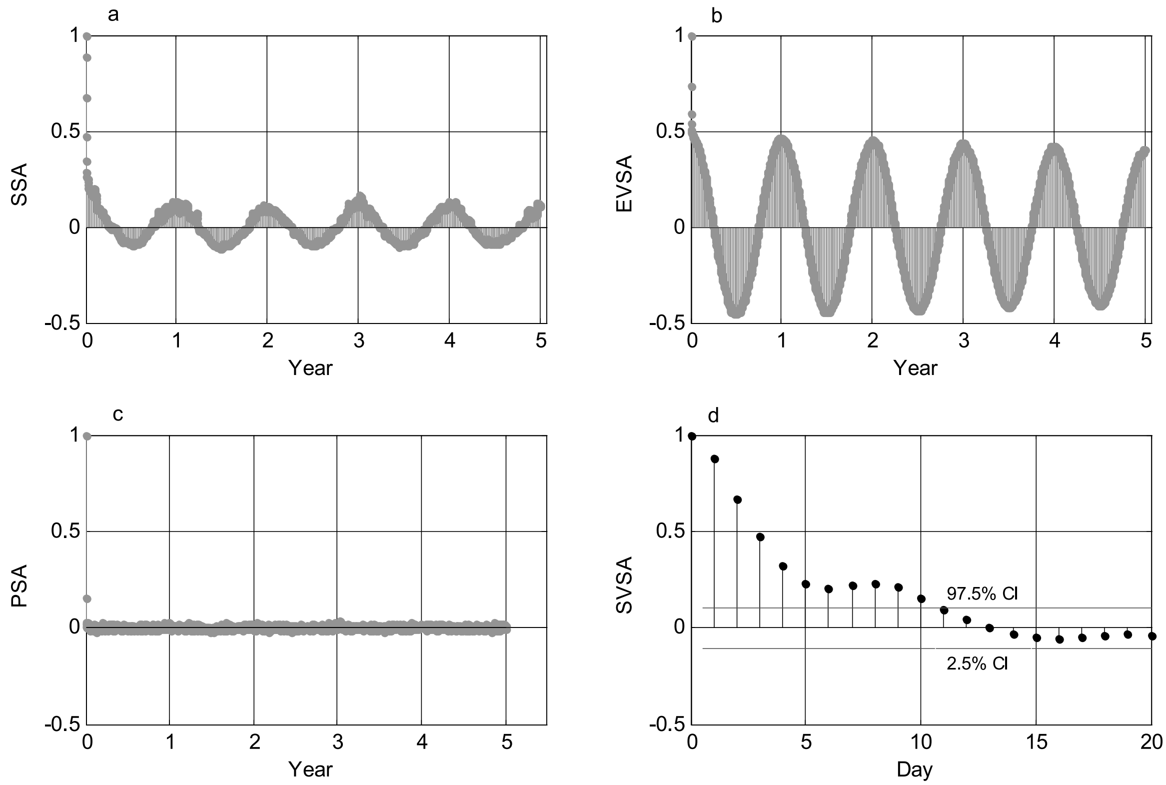

Figure 2. Display multiannual autocorrelation functions of the observed (a) streamflow (SSA), (b) evapotranspiration (EVSA), and (c) precipitation (PSA), respectively. (d) The interannual autocorrelation function of the observed streamflow in the water year of 1948 implies short-term persistence of the time series.

station (1944 km^2, drainage area) located approximately 28 miles northwest of Hattiesburg, Mississippi.

[26] We have examined the conditions that must be met for applying the proposed block bootstrap parameter estimation scheme including the cyclostationary and weak dependence structure of the observations. Analysis of the sample autocorrelation function (ACF) of the daily observations (Figure 2) shows that the autocorrelation functions of daily streamflow (SSA) and evapotranspiration (EVSA) exhibit annual sinusoidal-like periodic structures, while the precipitation time series do not show pronounced periodicities (Figure 2c). The interannual autocorrelation function of streamflow (Figure 2d shows an example for water year 1948) decays exponentially and dies off gradually; beyond a lag of 10 days, the ACF is effectively zero with 95% confidence. This indicates that when the annual periodic component is removed from the time series, the process exhibits only short-term memory and weak dependence properties.

[27] Although the scope of this study does not include the detection, removal, and merging of nonstationary components, we investigated the second-order cyclostationary of the time series by looking for linear trends in the streamflow annual statistics over the period of observation. Figures 3a and 3b show the annual mean and variance of the streamflow time series indicating no significant nonstationary trend. However, a slightly positive sample autocorrelation was detected for both the mean and the variance. A *t*-test, conducted to check the null hypothesis that the correlations are equal to zero (i.e., H_0: $\rho(Y, X) = 0$) indicated that the null hypothesis cannot be strongly rejected; therefore, the correlations do not seem to be statistically significant.

[28] To examine the weak dependence structure of the observations, we computed the variance of the sample mean for different sample sizes at the annual scale. If the observations $\{X_1, X_2,\ldots, X_m\}$ belong to a weak memory process, we expect the variance of the sample mean, $X_{\mathrm{bar}}(n) = \sum_n X_i/n$, $n \in (1, m)$, to be proportional to the sample size n [*Beran*, 1994] such that

$$\mathrm{Var}[X_{\mathrm{bar}}(n)] \propto n^{-1}. \tag{5}$$

[29] Figure 4a shows the variance of the sample mean for streamflow as a function of different sample sizes at the annual scale, $\{n = kT_0, k = 1,\ldots, 20; T_0 = 1 \text{ year}\}$. The slope of the least-square regression line in the log-space indicates that the variance of the process decays proportionally with n^{-1}, a characteristic signature of weak memory properties. Similar results were obtained for the precipitation and evapotranspiration time series; therefore, they were treated similarly during block bootstrap sampling.

4.2. Hydrologic Model

[30] The HyMod model [*Boyle et al.*, 2000] is a parsimonious conceptual representation of catchment scale precipitation-runoff dynamics (Figure 5) that has been used in numerous research studies (e.g., *Vrugt et al.* [2003],

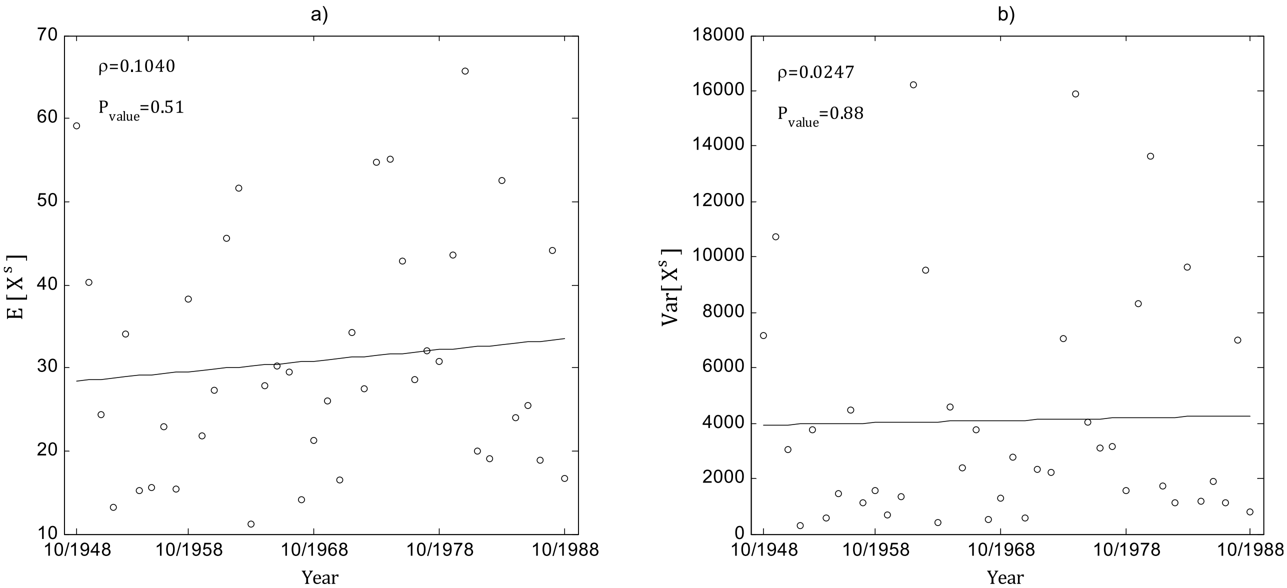

Figure 3. The annual (a) mean and (b) variance of the observed streamflow in cubic meters per second is illustrated. A *t*-test conducted on the least-square fitted line to the sample data indicates that no statistically meaningful nonstationarity exists in the first- and second-order moments.

Moradkhani et al. [2005a, 2005b], *Blasone et al.* [2008], among others). The model has a simple structure and is based on the probabilistic principle of the spatial distribution of catchment runoff in response to rainfall [*Moore*, 1985]. A nonlinear tank represents soil moisture variability, while two parallel chains of linear reservoirs simulate the fast- and slow-routing processes. The model requires specification of five parameters: (1) maximum storage capacity of the watershed ($C_{\max}$); (2) spatial variability of soil moisture capacity ($B_{\exp}$); (3) a factor to distribute flow between fast- and slow-flow-routing processes (Alpha); (4) a parameter associated with slow-flow routing (k_s); and (5) a parameter associated with quick flow (k_q). Feasible ranges for the model parameters are presented in Table 1.

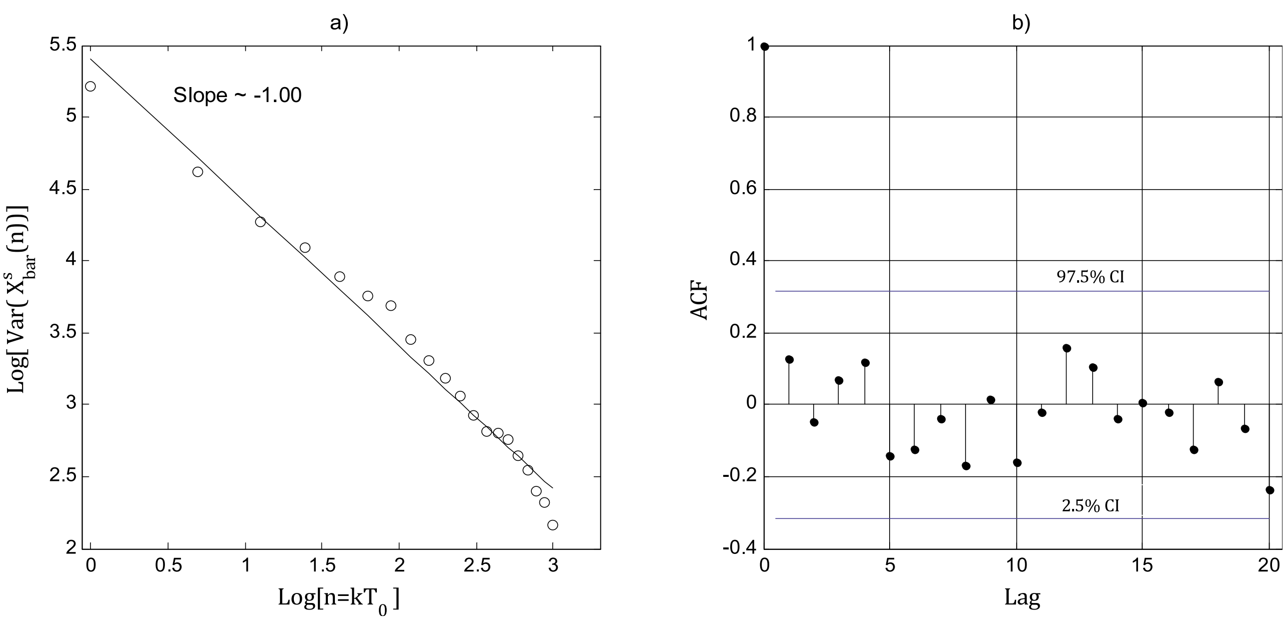

Figure 4. (a) Illustrates the variance-time plot of the observed streamflow in cubic meters per second at a log-scale. The annual time scales are selected as an integer multiplier of the streamflow annual period (years). The variance of the sample mean approximately decays proportional to n^{-1}, indicating that the streamflow time series is a weak dependent periodic process. (b) The sample autocorrelation function (ACF) of the annually averaged streamflow for the entire period of the observation. The autocorrelation dies off quickly and becomes negligible after the first lag, indicating short memory at the annual scale.

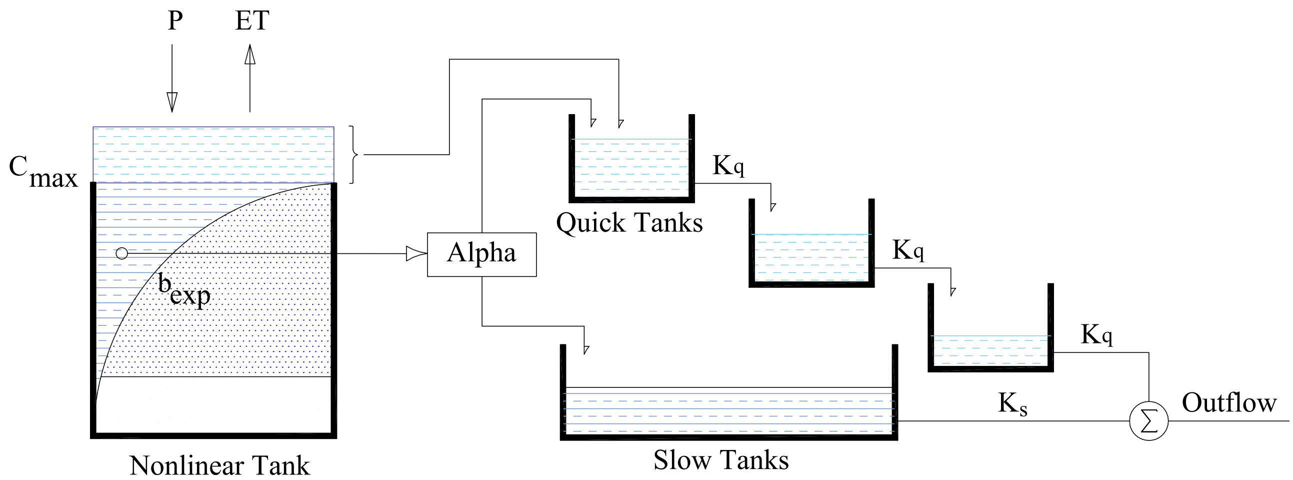

Figure 5. The conceptual structure of the five-parameter HyMod model.

4.3. Optimization Setup and Calibration Experiments

4.3.1. Optimization Method and Criterion

[31] To implement the proposed block bootstrap parameter estimation scheme, we used the SCE-UA algorithm [*Duan et al.*, 1992] to minimize the root-mean-square error (RMSE) of differences between observed and simulated streamflow:

$$\theta_{\mathbf{opt}} = \underset{\theta_{\min} \le \theta \le \theta_{\max}}{ArgMin} \sqrt{\frac{\sum_{n=1}^{N_0} \left[X_n^{os} - X_n^{es}(\theta)\right]^2}{N_0}}, \tag{6}$$

where $\theta = \{\theta_1, \theta_2, \ldots, \theta_{n_h}\}$ denotes the feasible parameter space of the n_h hydrologic parameters, X_n^{os} is the observed streamflow, $X_n^{es}(\theta)$ is the estimated streamflow that is a function of the parameters and model structure, and N_0 signifies the total number of observations that have been used in the calibration process. We then implicitly computed the conditional expected value $\theta_{\mathbf{opt}} = \mathbb{E}[\theta | X_N^{OS}]$ of the parameters given a finite length of observations, such that $\mathbb{E}|x_n^{os} - x_n^{es}(\theta_{\mathbf{opt}})|^2 \le \mathbb{E}|x_n^{os} - x_n^{es}(\theta)|^2$ for all sets of feasible real valued θ.

4.3.2. Dependence of Optimum Parameters on Length of Calibration Period

[32] As mentioned previously, while the model parameters may be considered to be time-invariant properties of the hydrologic system, their calibrated estimates are functionally dependent on the length and properties of the calibration period (the data sample) and various sources of uncertainty. Figure 6 illustrates the functional dependence of the optimal parameter estimates on length of calibration period, with some of them appearing to tend asymptotically toward a limit with increasing amounts of data. These results were obtained by running the SCE-UA algorithm with number of complexes equal to $2n_h + 1 = 11$, and with default stopping criteria as proposed by *Duan et al.* [1992]; for each case, 10 independent optimization runs were performed to ensure that the global minima were reached with very high probability. Note that, unlike the result reported by *Yapo et al.* [1996] for the Sacramento model, the parameter of the HyMod model do not seem to stabilize after only 11 years of data but continue to depend on the data set, even when more than 20 years of calibration data are used.

[33] Considering the optimal parameters to be the conditional expectation, $\theta_{\mathbf{opt}} = \mathbb{E}[\theta | X_N^{os}]$ of a finite random quantity (see section 4.3.1), the law of large numbers indicates that the optimal parameters will tend toward their population values with increasing calibration sample size. However, two questions arise: (1) Is the observation time series large enough to ensure that the estimated parameters are close to the population value? (2) What confidence limits can be placed around the estimated parameters? The answers to these questions were explored using the Block Bootstrap method, as discussed next.

5. Block Bootstrap Calibration, Assumptions, and Results

[34] The MBB approach was used to generate sample estimates of the hydrologic model parameters and to quantify the associated uncertainties, using the first 15 years of data (water years 1948 to 1963) for block bootstrap sampling and model calibration. To account for the periodic structure of the observed time series, nonoverlapping annual blocks of data were used to construct the pseudo time series as follows. The notation $\boldsymbol{X} = \{B_1, \ldots, B_j, \ldots, B_k\}$ is used to indicate the k years of data used for model calibration, where each $B_j = \{x_1, x_2, \ldots, x_{365}\}$, indicates the daily observations of a non-leap year. Since each set X and B_j contains coupled observations regarding multiple hydrologic fluxes, the coupled observations of streamflow, precipitation, and evapotranspiration are represented by $\boldsymbol{X} = \{X^s; X^p; X^e\}$. To construct each bootstrap pseudo time series, *iid* samples are drawn uniformly with replacement from the set $j \in \{1,2,\ldots, k\}$ and then the associated annual block of observations B_j is

Table 1. Description of the HyMod Parameters With Their Feasible Ranges

Parameters	Description	Minimum	Maximum
$C_{\max}$	Maximum storage capacity of watershed	1	1000
$B_{\exp}$	Spatial variability of the soil moisture capacity	0	2
Alpha	Distribution factor	0	1
K_S	Slow-flow tank residence time	0.0002	0.1
K_q	Quick-flow tanks residence time	0.1	0.99

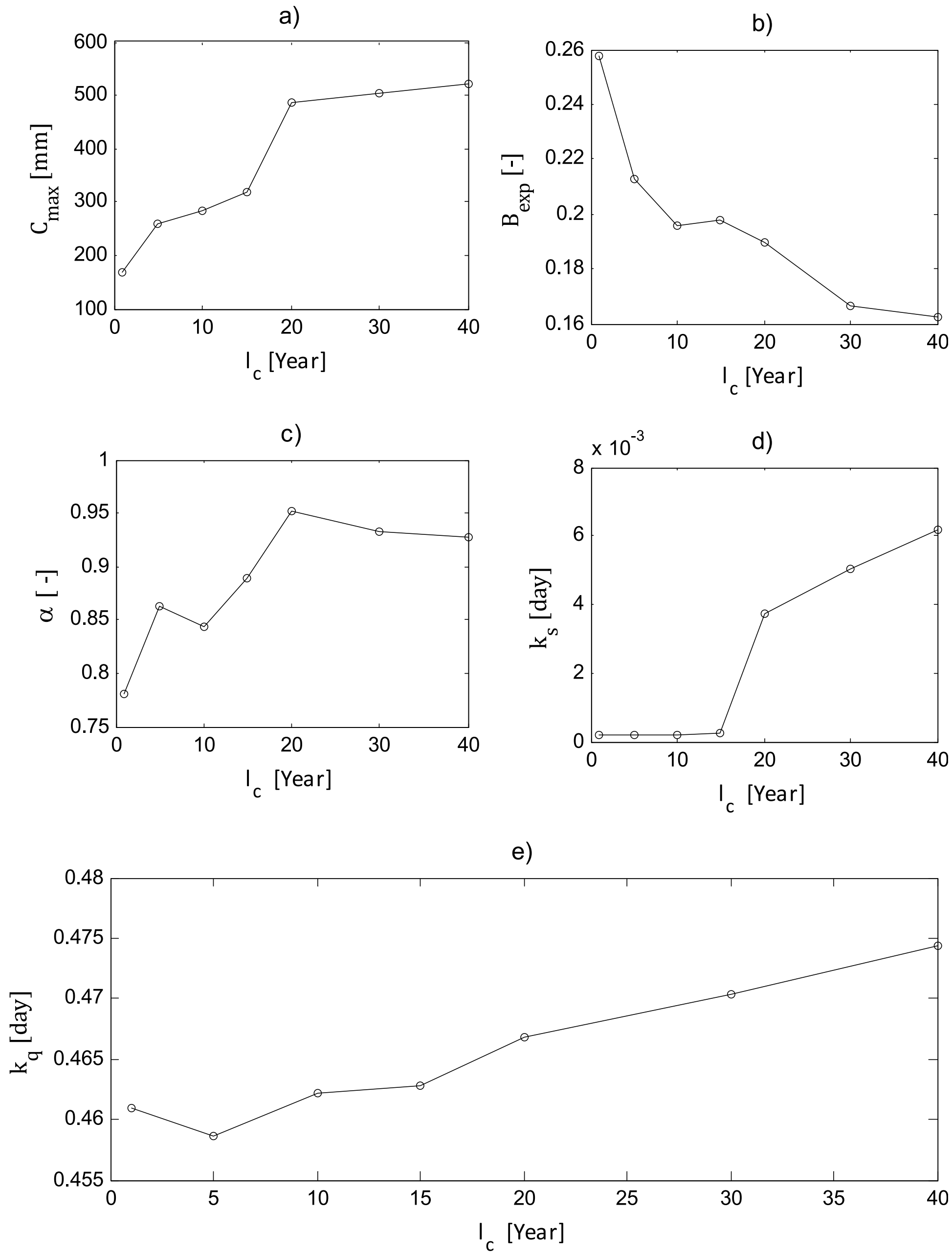

Figure 6. Dependence of the HyMod model optimal parameter estimates as a function of calibration length in years. The computed estimates are clearly not independent of the calibration length.

sampled from $\boldsymbol{X}$. The sampling is repeated k times with replacement, and the B^*_j values are concatenated to form the $\boldsymbol{X}^*_b = \{B^*_1, B^*_2, \ldots, B^*_k\}$. By repeating the procedure several times, $b = 1, \ldots, 100$ pseudo time series are generated and the hydrologic model is then calibrated using each pseudo time series, to arrive at bootstrapped estimates of the optimal parameter sets. These estimates are used for density estimation and confidence interval analysis of the hydrologic parameter estimators (e.g., mean, median, and/or quantiles).

5.1. Confidence Interval Analysis

[35] A major advantage of bootstrapping over analytical methods is its ability to provide nonparametric sample distributions for complex estimators when analytical forms do not exist. In particular, the sampling distribution obtained via bootstrapping enables estimation of the confidence intervals on a population estimate and thereby facilitates hypothesis testing. Given the bootstrap sampling distribution $\hat{\boldsymbol{F}}$ of the estimator, the $(1 - 2\alpha)$ 100% percentile confidence limits (CI) are provided by the interval $[\hat{\boldsymbol{F}}^{-1}(\alpha), \hat{\boldsymbol{F}}^{-1}(1 - \alpha)]$, where $\hat{\boldsymbol{F}}^{-1}(\alpha)$ denotes the α quantile of the cumulative distribution. To cope with the fact that the population and empirical bootstrap sample distribution may differ markedly in their location and shape, we follow the Bias Corrected and Accelerated (BCA) CI analysis method introduced by *Efron* [1987] and *Efron and Tibshirani* [1993]. This approach assumes that there exists a nonpivotal transformation of an estimator that makes it normally distributed

Table 2. Bootstrap CI Estimation of the Hydrologic Parameters Statistics

(1)		(2)			(3)			(4)		
Parameters		2.5% Quantile			20% Trimmed Mean			97.5% Quantile		
C_{max}	$t(\theta^*)$[a]		274.9394			375.1943			522.9268	
	BCA-CI[b]	253.1991		287.6108	360.80		389.40	495.4718		555.8929
B_{exp}	$t(\theta^*)$		0.1609			0.2184			0.2942	
	BCA-CI	0.1498		0.1655	0.2102		0.2265	0.2852		0.3110
Alpha	$t(\theta^*)$		0.8087			0.9156			0.9999	
	BCA-CI	0.7809		0.8269	0.9023		0.9286	0.9907		1.00
K_S	$t(\theta^*)$		0.0002			0.0014			0.0147	
	BCA-CI	0.0002		0.0002	0.0008		0.0023	0.0098		0.0190
K_q	$t(\theta^*)$		0.4548			0.4682			0.4762	
	BCA-CI	0.4516		0.4589	0.4672		0.4693	0.4752		0.4810

[a]The variable $t(\theta^*)$ represents the bootstrap estimator of the sampled optimum hydrologic parameters.
[b]BCA-CI represents the computed confidence bound based on the bootstrap bias corrected and accelerated method.

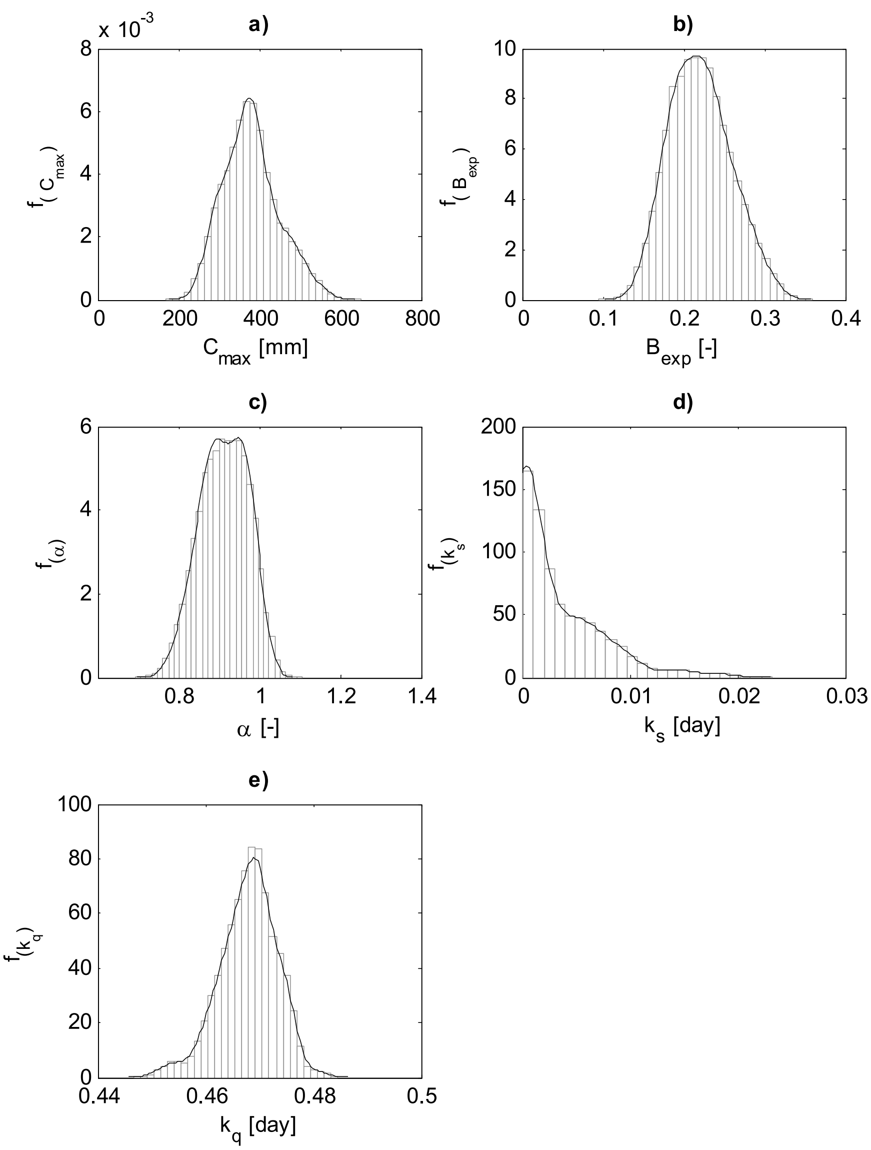

Figure 7. The smoothed bootstrap density estimation using a Gaussian kernel. One thousand smoothed bootstrap replicates are generated using equations (3) and (4) for construction of the densities.

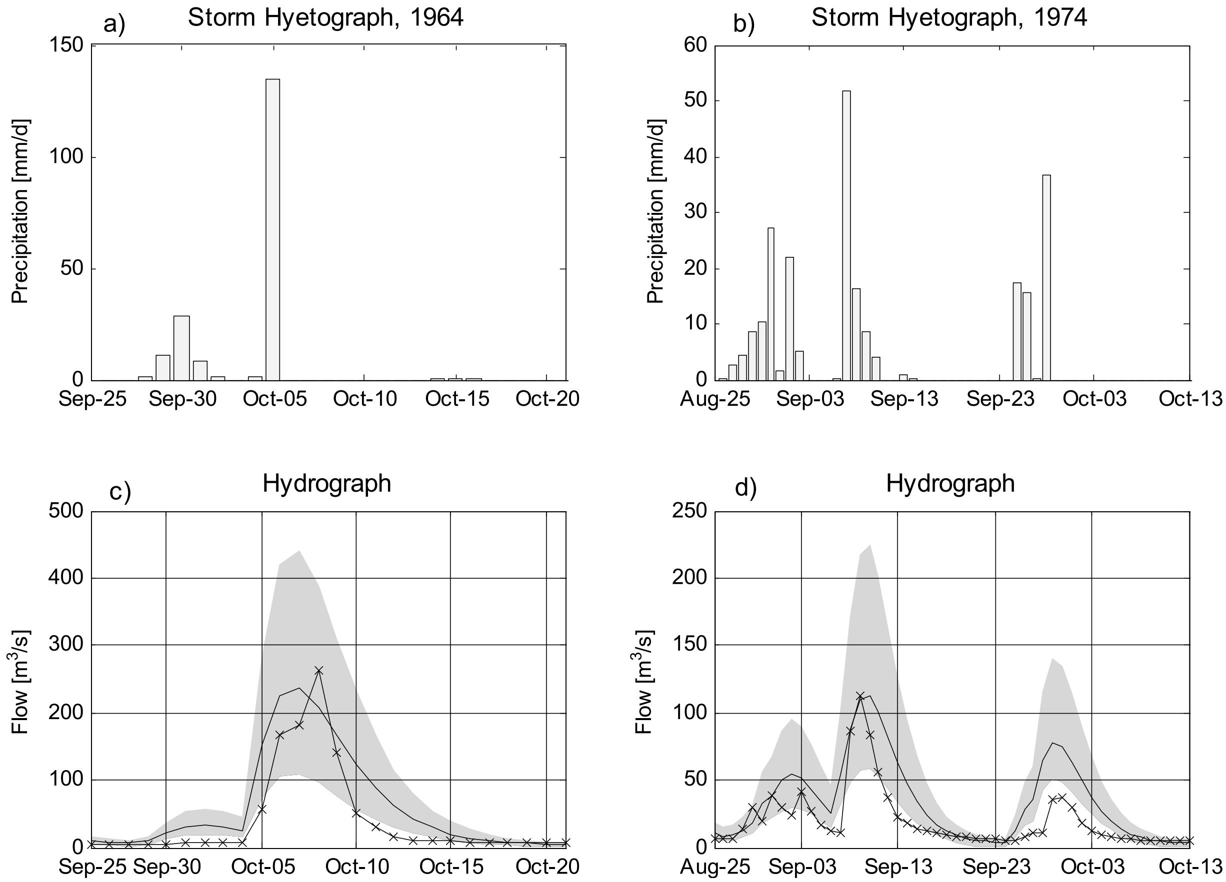

Figure 8. (a, b) Hyetograph of the selected storms. (c, d) Characterization of streamflow uncertainty. The dashed lines indicate the observations and the solid line indicates the mean of the ensembles. The shaded areas show the 95% confidence intervals of forecast delineated by the streamflow replicates.

and requires nonparametric estimation of a bias and an acceleration coefficient from the data. Accordingly, the formalism of the BCA confidence bound estimation can be expressed as

$$L_b = \hat{\boldsymbol{F}}^{-1}\left[\boldsymbol{\Phi}\left(z_0 - \frac{z_{[1-\alpha]} - z_0}{1 + a\left(z_{[1-\alpha]} - z_0\right)}\right)\right]$$
$$U_b = \hat{\boldsymbol{F}}^{-1}\left[\boldsymbol{\Phi}\left(z_0 + \frac{z_{[0]} - z_{[\alpha]}}{1 + a\left(z_{[0]} - z_{[\alpha]}\right)}\right)\right], \quad (7)$$

where L_p and U_p denote the lower and upper confidence bounds, respectively; z_0 represents the bias coefficient; $\Phi(\bullet)$ is the standard normal CDF; a is the acceleration constant; and $z_{[\alpha]}$ is the α percentile of a normal distribution. For lucid explanations of the method and computation of the coefficients, please see *Efron and Tibshirani* [1993] and *Lunneborg* [2000].

[36] Table 2 shows the results of the BCA confidence bound analysis of the mean and the quantiles of the sampled parameter estimates. Here, we follow the recommendation of *Lunneborg* [2000] and *Wilcox* [2005] and present 20% trimmed estimators (computed from the 80% high probability density region) because these make the analysis more robust in the presence of outliers (that might arise, e.g., due to the small but finite probability that the SCE-UA optimizer does not converge to the global optimum). Given that the bootstrap parameter estimates are sampled in a least-square sense, the trimmed mean can be used as the best-unbiased parameter estimate for streamflow forecasting. Now, comparing Figure 6 with the confidence bounds estimated in columns 2 and 4 of Table 2, we can see that the parameter estimates obtained by the conventional approach without bootstrapping (section 4.3.2) using the entire 40-year data period lie within the 95% quantiles estimated using the shorter 15-year period of calibration. This indicates that the confidence intervals estimated via bootstrapping on a 15-year period appear to be reasonable.

[37] The results presented in Table 2 and Figure 7 suggest that the estimated parameters exhibit varying degrees of uncertainty. For example, the error function appears to be relatively flat with respect to parameter $C_{\max}$, resulting in a relatively wide range of variability. Meanwhile the cost function is highly sensitive to parameter k_q resulting in a smaller range of uncertainty. Furthermore, the distributions of some of the parameters (e.g., k_s) are clearly skewed.

5.2. Parameter Uncertainty and Ensemble Streamflow Simulation

[38] The process of generating sample estimates that span the parameter space is generally limited by computational

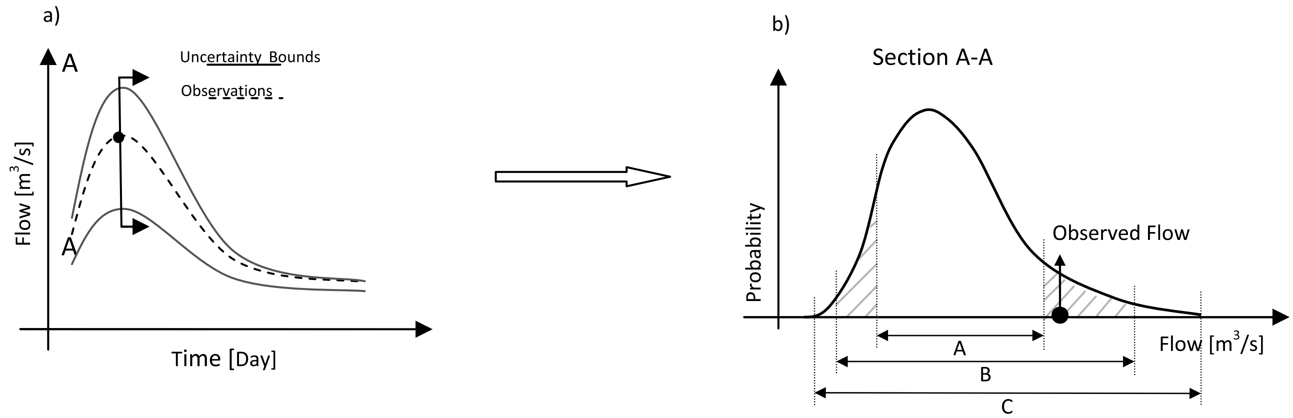

Figure 9. (a) Schematic of streamflow uncertainty bounds and location of the observed data. (b) Probability regions are defined using quantile levels $q_\alpha = \hat{\boldsymbol{F}}^{-1}(\alpha)$ associated with different probabilities of nonexceedance α. Regions A, B, and C refer to quantile levels [$q_{0.25}$, $q_{0.75}$], [$q_{0.025}$, $q_{0.975}$], and [Max, Min], respectively.

cost. If only small sample sizes can be generated, the discrete nature of the sampled parameter distributions may result in poor mapping of the statistical distribution of parameter uncertainty into the streamflow forecasts. Here, we can employ the method of smoothed bootstrapping to generate a more useful estimate of the parameter sampling density. The sampled hydrologic parameter sets are perturbed with an optimal Gaussian kernel function, using a plug-in method to estimate the kernel bandwidth [*Bowman and Azzalini*, 1997], resulting in the smoothed density estimates for the model parameters shown in Figure 7.

[39] A small complication that can arise in using the smoothed density estimates to approximate the probability densities of streamflow forecasts is that the estimated distribution may extend beyond the independently specified feasible range (Table 1) for some of the model parameters. Therefore, we treat all such samples (that violate the feasible ranges) as ill-posed (inadmissible) and do not include them in the generation of sample streamflow trajectories. Figure 8 shows ensemble one-day ahead forecast simulations for two storm events in 1964 and 1974 that were not used for model calibration; we see that the uncertainty bounds provide qualitatively realistic assessments of streamflow prediction uncertainty and generally encompass the observations except during periods of low flow. Causes for this may include numerous factors not considered in this study, including the relative simplicity of the hydrologic model used, errors in the precipitation data, the use of lumped rather than distributed modeling of the watershed, and the choice of the RMSE criterion which tends to give less weight to the simulation of low-flow events.

[40] To provide a quantitative assessment of the streamflow uncertainty intervals obtained by our approach, we compute a Quantitative Forecast Probability Measure (QFPM) to evaluate how well the distribution of ensemble trajectories encompasses the observations. The assessment is presented for five periods consisting of cumulatively consecutive water years 1964–1968 of the evaluation period. For each period, the QFPM is computed by calculating the average fraction of ensemble members falling within each of three different quantile levels $q_\alpha = \hat{\boldsymbol{F}}^{-1}(\alpha)$ (see Figure 9 and Table 3). For quantile levels A, B, and C, the theoretical average fraction of ensembles should be close to 0.5, 0.95, and 1.0, respectively (though slightly smaller numbers should be expected because of the finite sample size). As can be seen, the QFPMs are somewhat smaller than the theoretically expected values; this is perhaps not surprising since (as mentioned earlier) we have not accounted for all sources of uncertainty, most notably model structural uncertainty. What is important, however, is that the QFPM scores remain consistent with increasing length of evaluation period, indicating a statistical consistency in our approach to treating the data. For complementary information on how the quality of ensemble forecast can be quantitatively assessed, please refer to *Borga* [2002], *Hossain and Anagnostou* [2005], and *Moradkhani et al.* [2006], among others.

[41] To illustrate the effectiveness of the proposed method under conditions where the model and data are free of errors, Figure 10 shows the results of a synthetic study conducted using the same model and data. The HyMod model was run with a selected set of "true" parameter values θ^s = {250, 0.40, 0.84, 0.005, 0.45} for the water year 1948–1963, and an ensemble simulation was generated for an independent evaluation period 1963–1964 using the bootstrapping methodology proposed here. The results clearly show that the 95% confidence intervals suitably bracket the synthetically generated "true" discharge observations (solid circles). Analysis of the results showed that the density of the flow uncertainty is highly skewed and peaks very close to the

Table 3. Quantitative Forecast Probability Measures for Five Overlapping Evaluation Periods

Probability Regions	QFPM of Evaluation Period Water-Years				
	1964	1964–1965	1964–1966	1964–1967	1964–1968
A (0.5)[a]	0.41	0.36	0.38	0.35	0.34
B (0.95)	0.76	0.75	0.75	0.72	0.68
C (0.99)	0.93	0.90	0.91	0.90	0.90

[a]Shows different probability bands. See Figure 9 for explanations. The numbers in parentheses indicate the theoretical optimal value. QFPM, quantitative forecast probability measures.

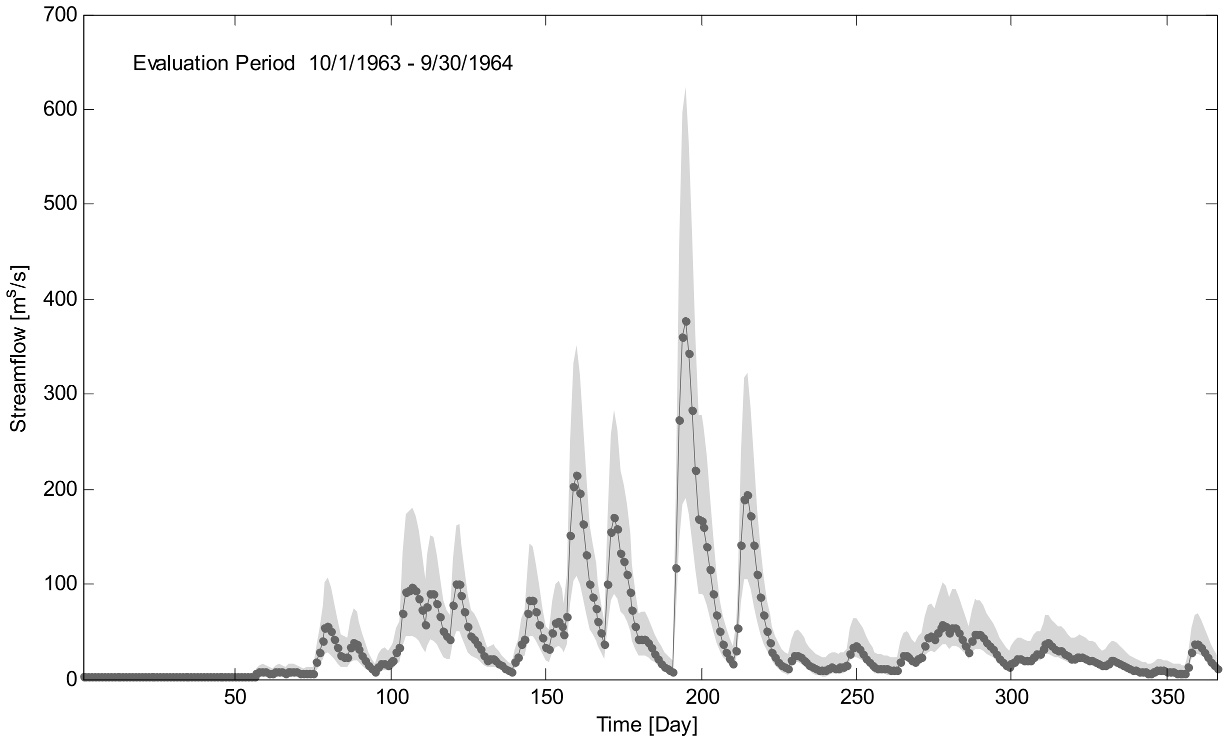

Figure 10. Projected synthetic characterization of the bootstrap uncertainty estimation into the streamflow space for the water year 1963–1964. In the absence of structural model error and data error, the 95% confidence bound of the simulated ensembles encompass the synthetic observations (solid circles) reasonably.

lower uncertainty bound during low flows, due mainly to the density of the slow tank residence time parameter K_s being positively skewed.

6. Computational Considerations

[42] This method has asymptotic convergence properties meaning that by increasing the number of simulated pseudo time series, estimation error will be reduced. Implementation of the method can impose significant computational costs, in particular, when the hydrologic model has a complex structure and precise estimates are desired. All reported simulations were run by use of a Linux desktop machine with a double quad core Intel-Xeon 3.2 GHz processor with no parallelization in the programming. Construction of each pseudo time series and estimation of the associated optimal parameter estimates took 300–500 s. Future parallel implementations will help to reduce the run times and allow the methods to be applied in the case of more complex hydrologic models.

7. Summary and Conclusion

[43] Methods of statistical inference can aid in the application of hydrologic modeling within a probabilistic context, helping to improve short- and long-term predictability of river flows, and bringing a proper consideration of uncertainty into the decision-making process. This study has explored the use of the block bootstrap approach for generating robust estimates of the parameters of hydrologic models by accounting for the fact that the calibration data constitute a sample of finite size drawn from an unknown underlying distribution. Because of this, the estimated parameters of any hydrologic model will have functional dependence on the size, hydrologic properties and error characteristics of the available data, and this dependence can be characterized in a probabilistic sense that accounts for periodic stationarity and short-range dependence. The approach was illustrated using a simple case study showing that a consistent characterization of parameter uncertainty can be achieved using limited length periods of data.

[44] To properly apply the block bootstrap methods to account for other sources of uncertainty will require several extensions. Sources of unbiased observation error that can be expressed via a measurement error model can easily be incorporated. Perhaps more important, the use of MBB to account for the stochastic properties of model structural error should be explored. Other issues to be examined include the sensitivity of the results to block size and bandwidth estimation method (during smoothed bootstrapping). While methods exist to analytically derive explicit expression for the optimal block length for population parameters such as the mean and variance [e.g., *Hall et al.*, 1995], methods for finding the optimal block length in the context of hydrologic parameter estimation need to be determined.

Appendix

[45] The bootstrap approach to uncertainty estimation of the sample mean of independent data is illustrated in this appendix. Let $\theta(\boldsymbol{F})$ be the mean of an unobservable popu-

lation distribution $\boldsymbol{F}$, from which a sample observation $\boldsymbol{X} = \{x_1, x_2, \ldots, x_n\}$ is available:

$$\mu(\boldsymbol{F}) = \int x d\boldsymbol{F}(x). \tag{A1}$$

Assuming that the variance of the X is finite, define the pivotal statistic:

$$\boldsymbol{R}_n(\boldsymbol{X}, \boldsymbol{F}) = \sqrt{n}[\bar{x}_n - \mu(\boldsymbol{F})] \tag{A2}$$

where $\bar{x}_n$ is the sample mean and,

$$\mu(\hat{\boldsymbol{F}}) = \int x d\hat{\boldsymbol{F}}(x) = \frac{1}{n}\sum_{i=1}^{n} x_i = \bar{x}_n \tag{A3}$$

$$\sigma^2(\hat{\boldsymbol{F}}) = \frac{1}{n}\sum_{i=1}^{n} (x_i - \bar{x}_n)^2. \tag{A4}$$

As a consequence of the strong law of large numbers, $\mu(\hat{\boldsymbol{F}}) \rightarrow \mu(\boldsymbol{F})$ and $\sigma^2(\hat{\boldsymbol{F}}) \rightarrow \sigma^2(\boldsymbol{F})$ as n tends to the population size (i.e., the mean and variance converge to the population values almost surely). On the basis of the *Glivenko-Cantelli* [1933] theorem it can be shown that the empirical distribution of $\hat{\boldsymbol{F}}$ converges to $\boldsymbol{F}$ in the sense that $\sup_x |\hat{\boldsymbol{F}}(x) - \boldsymbol{F}(x)| \rightarrow 0$ almost surely.

[46] Now consider the bootstrap samples $\boldsymbol{X}^* = \{x_1^*, x_2^*, \ldots, x_n^*\}$ from $\hat{\boldsymbol{F}}$ and define the bootstrap mean:

$$\bar{x}_n^b = \frac{1}{n}\sum_{i=1}^{n} x_i^*. \tag{A5}$$

On the basis of the weak law of large numbers, it can be shown that the distribution of the $\boldsymbol{R}^*_n(\boldsymbol{X}, \boldsymbol{F}) = \sqrt{n}[\bar{x}_n^b - \mu(\hat{\boldsymbol{F}})]$ converges weakly to the distribution of $\boldsymbol{R}_n(\boldsymbol{X}, \boldsymbol{F})$ as $b \rightarrow \infty$, which we know to be a zero mean standard normal distribution with variance $\sigma^2(\boldsymbol{F})$.

[47] In the case of the sample mean distribution, the closed form of the distribution exists and can be used for confidence interval estimation. However, bootstrap resampling can be used for confidence bound estimation of *any* functional of $\boldsymbol{F}$ (e.g., $\mu(\boldsymbol{F}) = g[\int \psi(x) d\boldsymbol{F}(x)]$) provided that $\boldsymbol{F}$ is a Fréchet differentiable function [*Lahiri*, 2003].

[48] **Acknowledgments.** M.E. thanks Efi Foufoula for her support and valuable comments during the research. The many suggestions and comments provided by three anonymous reviewers have also greatly improved the paper.

References

Beran, J. (1994), *Statistics for Long-Memory Processes*, *Monogr. Stat. Appl. Probab.*, vol. 61, 315 pp., Chapman and Hall, New York.

Beven, K. J. (1993), Prophecy, reality and uncertainty in distributed hydrological modeling, *Adv. Water Resour.*, *16*, 41–51.

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

Blasone, R. S., J. A. Vrugt, H. Madsen, D. Rosbjerg, B. A. Robinson, and G. A. Zyvoloski (2008), Generalized likelihood uncertainty estimation (GLUE) using adaptive Markov chain Monte Carlo sampling, *J. Adv. Water Resour.*, *31*, 630–648.

Borga, M. (2002), Accuracy of radar rainfall estimates for streamflow simulation, *J. Hydrol.*, *267*, 26–39.

Bowman, A. W., and A. Azzalini (1997), *Applied Smoothing Techniques for Data Analysis, the Kernel Approach with S-Plus Illustrations*, 191 pp., Oxford Univ. Press, New York.

Boyle, D. P., H. V. Gupta, and S. Sorooshian (2000), Toward improved calibration of hydrologic models: Combining the strengths of manual and automatic methods, *Water Resour. Res.*, *36*(12), 3663–3674.

Burnash, R. J. C. (1995), The NWS River Forecast System—catchment modeling, in *Computer Models of Watershed Hydrology*, edited by V. P. Singh, pp. 311–366, Water Resour. Publ., Highlands Ranch, Colo.

Carlstein, E. (1986), The use of subseries values for estimating the variance of a general statistic from a stationary sequence, *Ann. Stat.*, *14*, 1171–1179.

Duan, Q., S. Sorooshian, and V. Gupta (1992), Effective and efficient global optimization for conceptual rainfall-runoff models, *Water Resour. Res.*, *28*, 1015–1031.

Efron, B. (1979), Bootstrap methods: Another look at jackknife, *Ann. Stat.*, *7*, 1–26.

Efron, B. (1982), *The Jackknife, the Bootstrap and the Other Resampling Plans*, Society for Industrial and Applied Mathematics, Philadelphia, PA.

Efron, B. (1987), Better bootstrap confidence intervals, *J. Am. Stat. Assoc.*, *82*(397), 171–200.

Efron, B., and R. Tibshirani (1993), *An Introduction to the Bootstrap*, 430 pp., Chapman and Hall/CRC, Boca Raton, Fla.

Gupta, H. V., S. Sorooshian, and P. O. Yapo (1998), Towards improved calibration of hydrologic models: Multiple and non-commensurable measures of information, *Water Resour. Res.*, *34*(4), 751–763.

Hall, P., J. L. Horowitz, and B. Y. Jing (1996), On blocking rules for bootstrap with dependent data, *Biometrika*, *82*(3), 561–574.

Horowitz, J. L. (2003), Bootstrap methods for Markov processes, *Econometrica*, *71*(4), 1049–1082.

Hossain, F., and E. N. Anagnostou (2005), Numerical investigation of the impact of uncertainties in satellite rainfall estimation and land surface model parameters on simulation of soil moisture, *Adv. Water Resour.*, *28*, 1336–1350.

Kavetski, D., G. Kuczera, and S. W. Franks (2006), Bayesian analysis of input uncertainty in hydrological modeling: 2. Application, *Water Resour. Res.*, *42*, W03408, doi:10.1029/2005WR004376.

Künsch, H. R. (1989), The jackknife and the bootstrap for general stationary observations, *Ann. Stat.*, *17*(3), 1217–1241.

Lahiri, S. N. (1995), On the asymptotic behavior of the moving block bootstrap normalized sums of heavy-tailed random variables, *Ann. Stat.*, *23*(4), 1331–1349.

Lahiri, S. N. (1999), Theoretical comparison of block bootstrap methods, *Ann. Stat.*, *27*(1), 386–404.

Lahiri, S. N. (2003), *Resampling Methods for Dependent Data*, 371 pp., Springer, New York.

Liu, R. Y., and K. Singh (1992), Moving blocks jackknife and bootstrap capture weak dependence, in *Exploring the Limits of Bootstrap*, edited by R. Lepage and L. Billards, pp. 225–248, John Wiley, New York.

Lunneborg, C. E. (2000), *Data Analysis by Resampling Concepts and Applications*, 561 pp., Duxbury Press, Pacific Grove, Calif.

Moore, R. J. (1985), The probability-distributed principle and runoff production at point and basin scales, *J. Hydrol. Sci.*, *30*(2), 273–297.

Moradkhani, H., K. L. Hsu, H. Gupta, and S. Sorooshian (2005a), Uncertainty assessment of hydrologic model states and parameters: Sequential data assimilation using the particle filter, *Water Resour. Res.*, *41*(5), W05012, doi:10.1029/2004WR003604.

Moradkhani, H., S. Sorooshian, H. V. Gupta, and P. R. Houser (2005b), Dual state–parameter estimation of hydrological models using ensemble Kalman filter, *Adv. Water Resour.*, *28*(2), 135–147.

Moradkhani, H., K. Hsu, Y. Hong, and S. Sorooshian (2006), Investigating the impact of remotely sensed precipitation and hydrologic model uncertainties on the ensemble streamflow forecasting, *Geophys. Res. Lett.*, *33*, L12401, doi:10.1029/2006GL026855.

Nelder, J. A., and R. Mead (1965), A simplex method for function minimization, *Comput. J.*, *7*(4), 308–313.

Politis, D. (2001), Resampling time series with seasonal components, *Frontiers in Data Mining and Bioinformatics: Proceedings of the 33rd Symposium on the Interface of Computing Science and Statistics*, Orange County, California, June 13–17.

Politis, D. N., and J. P. Romano (1994), The stationary bootstrap, *J. Am. Stat. Assoc.*, *89*, 1303–1313.

Politis, D. N., J. P. Romano, and T. L. Lai (1992), Bootstrap confidence bands for spectra and cross-spectra, *IEEE Trans. Signal Process.*, *40*(5), 1206–1215.

Rosenblatt, M. (1965), A central limit theorem and a strong mixing condition, *Proc. Natl. Acad. Sci. U. S. A.*, *42*, 43–47.

Sharma, A., D. G. Tarboton, and U. Lall (1997), Streamflow simulation: A nonparametric approach, *Water Resour. Res.*, *33*(2), 291–308.

Silverman, B. W., and G. A. Young (1987), The bootstrap: To smooth or not to smooth? *Biometrika*, *74*(3), 469–479.

Sorooshian, S., and J. A. Dracup (1980), Stochastic parameter estimation procedures for hydrologic rainfall-runoff models: Correlated and heteroscedastic error cases, *Water Resour Res.*, *16*(2), 430–442.

Sorooshian, S., Q. Duan, and V. K. Gupta (1993), Calibration of rainfall-runoff models: Application of global optimization to the soil moisture accounting model, *Water Resour. Res.*, *29*(4), 1185–1194.

Srinivas, V. V., and K. Srinivasan (2005), Hybrid moving block bootstrap for stochastic simulation of multi-site multi-season streamflows, *J. Hydrol.*, *302*, 307–330.

Srinivas, V. V., and K. Srinivasan (2006), Hybrid matched-block bootstrap for stochastic simulation of multiseason streamflows, *J. Hydrol.*, *329*, 1–15.

Synowiecki, R. (2007), Consistency and application of moving block bootstrap for nonstationary time series with periodic and almost periodic structure, *Bernoulli*, *13*(4), 1151–1178.

Thiemann, M., M. Trosset, H. Gupta and S. Sorooshian (2001), Bayesian recursive parameter estimation for hydrologic models, *Water Resour. Res.*, *37*(10), 2521–2535.

Tolson, B. A., and C. A. Shoemaker (2007), Dynamically dimensioned search algorithm for computationally efficient watershed model calibration, *Water Resour. Res.*, *43*, W01413, doi:10.1029/2005WR004723.

Tukey, J. W. (1977a), *Exploratory Data Analysis*, Addison-Wesley, Reading, Mass.

U.S. Army Corps of Engineers (1975), Post disaster report on April 1974 floods in Mississippi in the Mobile District, Mobile Dist., Mobile, Ala.

Vogel, R. M., and A. L. Shallcross (1996), The moving blocks bootstrap versus parametric time series models, *Water Resour. Res.*, *32*(6), 1875–1882.

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.

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

Wilcox, R. R. (2005), *Introduction to Robust Estimation and Hypothesis Testing*, 2nd ed., 516 pp., Elsevier Acad., San Diego, Calif.

Yapo, P. O., H. V. Gupta, and S. Sorooshian (1996), Automated calibration of conceptual rainfall-runoff models sensitivity to calibration data, *J. Hydrol.*, *181*, 23–48.

Young, G. A. (1990), Alternative smoothed bootstraps, *J. R. Stat. Soc., Ser. B.*, *52*(3), 477–484.

Zoubir, A. M., and D. R. Iskander (2004), *Bootstrap Techniques for Signal Processing*, 217 pp., Cambridge Univ. Press, Cambridge, U.K.

M. Ebtehaj, Saint Anthony Falls Laboratory, National Center for Earth Surface Dynamics, Department of Civil Engineering, University of Minnesota, 2 3rd Ave. SE, Minneapolis, MN, 55414, USA. (ebteh001@umn.edu)

H. V. Gupta, Department of Hydrology and Water Resources, University of Arizona, 1133 E James E. Rogers Way, Rm 122, PO Box 210011, Tucson, AZ 85721, USA.

H. Moradkhani, Department of Civil and Environmental Engineering, Maseeh College of Engineering and Computer Science, Portland State University, PO Box 751, Portland, OR 97207-0751, USA.