AGU PUBLICATIONS

Water Resources Research

RESEARCH ARTICLE
10.1002/2014WR016827

# Numerical rivers: A synthetic streamflow generator for water resources vulnerability assessments

**Edoardo Borgomeo[1], Christopher L. Farmer[2], and Jim W. Hall[1]**

[1]Environmental Change Institute, School of Geography and the Environment, University of Oxford, Oxford, UK, [2]Mathematical Institute, University of Oxford, Oxford, UK

**Key Points:**
- Nonparametric streamflow generation method based on simulated annealing
- User-specified properties of the time series can be modified
- Enables stress testing against potential climate-induced hydrological changes

**Supporting Information:**
- Supporting Information S1

**Correspondence to:**
E. Borgomeo,
edoardo.borgomeo@ouce.ox.ac.uk

**Citation:**
Borgomeo, E., C. L. Farmer, and J. W. Hall (2015), Numerical rivers: A synthetic streamflow generator for water resources vulnerability assessments, *Water Resour. Res.*, *51*, 5382–5405, doi:10.1002/2014WR016827.

Received 23 DEC 2014
Accepted 11 JUN 2015
Accepted article online 16 JUN 2015
Published online 14 JUL 2015



**Abstract** The vulnerability of water supplies to shortage depends on the complex interplay between streamflow variability and the management and demands of the water system. Assessments of water supply vulnerability to potential changes in streamflow require methods capable of generating a wide range of possible streamflow sequences. This paper presents a method to generate synthetic monthly streamflow sequences that reproduce the statistics of the historical record and that can express climate-induced changes in user-specified streamflow characteristics. The streamflow sequences are numerically simulated through random sampling from a parametric or a nonparametric distribution fitted to the historical data while shuffling the values in the time series until a sequence matching a set of desired temporal properties is generated. The desired properties are specified in an objective function which is optimized using simulated annealing. The properties in the objective function can be manipulated to generate streamflow sequences that exhibit climate-induced changes in streamflow characteristics such as interannual variability or persistence. The method is applied to monthly streamflow data from the Thames River at Kingston (UK) to generate sequences that reproduce historical streamflow statistics at the monthly and annual time scales and to generate perturbed synthetic sequences expressing changes in short-term persistence and interannual variability.

## 1. Introduction

Synthetic hydrology is a tool used to expand the set of plausible streamflow sequences for water resources simulation studies and enable more sampling of hydrological variability [*Matalas*, 1967; *Fiering and Jackson*, 1971; *Hirsch*, 1979]. Applications of synthetic hydrology include reservoir planning [*Jackson*, 1975; *Vogel and Stedinger*, 1988], hydroelectric system operation [*Pereira et al*., 1984], and drought risk characterization [e.g., *Salas et al*., 2005].

The study of synthetic streamflow generation has a long history in hydrology [*Matalas*, 1967; *Rajagopalan et al*., 2010]. Classical synthetic streamflow generation methods include autoregressive lag-1 models [*Thomas and Fiering*, 1962] and autoregressive moving average (ARMA) models [e.g., *Salas and Obeysekera*, 1982; *McLeod et al*., 1977]. These models have been extensively applied in water resources studies; however, their shortcomings have also been recognized [*Sharma et al*., 1997].

Commonly used stochastic streamflow models can only be applied to normally distributed data, thus requiring flows to be transformed to Normality before model fitting [*Prairie et al*., 2006; *Hao and Singh*, 2011]. Furthermore, linear parametric models tend to focus the hydrologist's attention on fitting the model parameters rather than on the statistics of the time series that the synthetic streamflow sequence needs to reproduce. Nonparametric methods such as block bootstrapping [*Vogel and Shallcross*, 1996] and *k*-nearest neighbor resampling [*Lall and Sharma*, 1996] have found many applications in hydrology. These models have the advantage of being data driven and not requiring any data transformation or assumption with respect to the time series' dependence structure. More recently, wavelet-based methods [e.g., *Kwon et al*., 2007; *Keylock*, 2012], entropy theory-based methods [*Hao and Singh*, 2011, 2013; *Srivastav and Simonovic*, 2014], empirical mode decomposition [*Lee and Ouarda*, 2012], and copula theory-based models [e.g., *Lee and Salas*, 2011] for synthetic streamflow generation have been proposed in the literature.

All these approaches were developed and evaluated with the stated aim of reproducing historical observed streamflow characteristics and under the assumption that the future is going to be similar to the past. The

reality of climate change and the need to go beyond methods based on historical data alone mean that water managers require synthetic streamflow generators that have a proven capability to reproduce observed streamflow moment statistics and temporal dependence but that also allow for the generation of streamflow sequences whose properties increasingly depart from the properties of the observed data. Approaches that link downscaled climate model output to hydrological models have been proposed to address this challenge and to generate plausible climate-adjusted streamflow sequences. These studies are based on stochastic weather generators [e.g., *Kilsby et al*., 2007; *Burton et al*., 2010a], regional climate models [e.g., *Prudhomme et al*., 2012; *Addor et al*., 2014], or combinations of these techniques [e.g., *Burton et al*., 2010b] to produce time series of meteorological variables, including rainfall and temperature, that are then used as inputs to hydrological models to project changes in streamflow characteristics [e.g., *Christensen et al*., 2004; *Manning et al*., 2009].

The uncertainty in climate and hydrological models, the limited number of climate projections, and the weakness of climate models at representing important processes such as low-frequency variability [e.g., *Rocheta et al*., 2014] mean that planning decisions and risk assessments based on these approaches may explore a small range of potential future climate change and give limited insight into a water system's vulnerabilities [*Nazemi et al*., 2013; *Brown and Wilby*, 2012]. Furthermore, the uncertainty cascade that characterises the approach driven by hydroclimate projections makes it difficult to attribute vulnerabilities to specific changes in streamflow characteristics (e.g., change in mean, interannual variability, persistence) and to understand to which of these characteristics water resources systems are most vulnerable.

Vulnerability-based, "scenario neutral" or "decision scaling" approaches have been proposed as an alternative to hydroclimatic projection driven assessments and as a way to gain insight into a water system's response to changing hydrological conditions [e.g., *Prudhomme et al*., 2010; *Brown et al*., 2012; *Nazemi et al*., 2013; *Turner et al*., 2014; *Singh et al*., 2014]. These approaches seek to test the response of a given system to a range of user-specified hydroclimatic variables (e.g., change in annual rainfall, change in annual flow volume) and then identify the specific hydroclimatic conditions under which the system is vulnerable.

These vulnerability-based approaches require some kind of "scenario generator" to produce a large number of perturbed hydroclimatic sequences over which to carry out the vulnerability assessment [*Steinschneider and Brown*, 2013] and map out the response of the system to changing streamflow conditions [*Nazemi and Wheater*, 2014]; however, to date, only a few methods for generating streamflows that go beyond the variability of the observed record and that allow users to change low-order statistics been proposed. *Nazemi et al*. [2013] proposed a stochastic flow generation framework based on quantile mapping and copulas to synthesize sequences with altered annual flow volumes and peak arrival times. *Stagge and Moglen* [2013] proposed a stochastic method to generate climate-adjusted daily streamflow data. Other studies have explored the usefulness of paleoreconstructed data to generate sequences that go beyond the variability represented in the historical record [e.g., *Prairie et al*., 2008]. These streamflow generation methods are limited in that they were designed for specific applications and allow for only a few properties of the streamflow sequences to be altered.

This study seeks to develop a synthetic streamflow generation method that can be used as a hydrological scenario generator to assess the vulnerability of water resource systems to changing streamflow characteristics. This method is specifically designed to equip water managers with a streamflow generator that allows for some user-specified streamflow properties to be altered while keeping unchanged some other important properties of the streamflow distribution. In this way, the vulnerability of the water resource system to changes in specific statistics of the streamflow sequence can be assessed.

The proposed synthetic streamflow generation method is based on simulated annealing, a global optimization technique [*Kirkpatrick et al*., 1983] that has found applications in geostatistics [*Farmer*, 1992; *Deutsch and Cockerham*, 1994; *Parks et al*., 2000], operations research [*Eglese*, 1990], water distribution network design [*Cunha and Sousa*, 1999], hydrological model calibration [*Thyer et al*., 1999], and groundwater management problems [*Dougherty and Marryott*, 1991] among others. *Bardossy* [1998] applied simulated annealing to generate daily precipitation time series and recognized its potential for hydrological time series generation. In this paper, we show how simulated annealing can be used to generate synthetic streamflow time sequences that reproduce the properties of the historical record and that represent possible climate-induced changes in user-specified streamflow properties. By allowing the water manager to

define the properties to be included in the algorithm's objective function, this generator provides a tool to obtain streamflow sequences expressing a wide range of potential climate-induced changes.

The paper is structured as follows: In section 2, we explain the proposed method. In section 3, we present the results from the application of the method to the monthly streamflow time series from the River Thames at Kingston, England. In section 4, we discuss our results and conclude.

## 2. Methods

### 2.1. Rationale

The first stochastic model for synthetic streamflow generation was developed by *Sudler* [1927]. Sudler's model was based on the idea of shuffling a deck of cards, where each card corresponded to an annual flow value, and then drawing one card at the time to generate a synthetic flow sequence [*Jackson*, 1975]. The idea of shuffling an observed time series to generate new samples which preserve the statistics of the reference period but have different sequencing is also at the core of bootstrap resampling strategies [e.g., *Vogel and Shallcross*, 1996; *Lall and Sharma*, 1996].

Our method is constructed on this same premise; however, we formulate the streamflow generation problem (i.e., the shuffling in Sudler's model) as a combinatorial optimization problem to be solved with a numerical optimization technique. The combinatorial optimization problem seeks to find a time series which matches a set of statistics (e.g., means, autocorrelation) specified in an objective function, where the objective function computes the difference between a target statistic and the simulated statistic. The target statistic and sequencing specified in the objective function can be taken from the reference period or can be perturbed statistics expressing potential climate change.

To solve the time series optimization problem and minimize the objective function we employ simulated annealing, a heuristic search method first introduced by *Kirkpatrick et al*. [1983]. The suitability of simulated annealing for solving combinatorial optimization problems is well documented in the literature [e.g., *Kirkpatrick et al*., 1983; *van Laarhoven and Aarts*, 1987; *Blum and Roli*, 2003]. Simulated annealing has been shown to work well in applications seeking to reconstruct a field or a configuration, such as a time series, that satisfies particular properties [*Bardossy*, 1998; *Cunha and Sousa*, 1999; *Farmer*, 1992]. Furthermore, compared to other optimization methods, the simulated annealing algorithm is easy to implement, taking up only a few lines of code [*Eglese*, 1990]. The suitability of other optimization algorithms such as genetic algorithms and tabu search for solving our time series optimization problem is not examined in this study and will be explored in future work. For an extended discussion on the relative performance and advantages of simulated annealing compared to other methods the reader is referred to *Simon* [2013].

### 2.2. Simulated Annealing

Simulated annealing is based on the analogy with the physical process of annealing, whereby a material is first melted at high temperature to mobilize its particles and then slowly cooled to force the particles into the low energy state of a highly structured crystalline lattice [*Kirkpatrick et al*., 1983]. If the cooling is done too quickly, the state of minimum energy of the system may not be reached and the material may freeze into an imperfect crystal. In the optimization analogy, the different states of the cooling material are different solutions to the combinatorial optimization problem and the energy of the system is the objective function to be minimized in the optimization.

*Metropolis et al*. [1953] introduced an algorithm to draw samples from a density function simulating a collection of atoms at a given temperature and to ensure that the sampling converges to the required Boltzmann distribution [*Eglese*, 1990]. *Kirkpatrick et al*. [1983] adapted this algorithm to solve combinatorial optimization problems. The algorithm starts with the atoms in a given configuration at a given temperature. In each iteration of the algorithm, an atom is displaced and the change in the energy of the system $\Delta E$ is recorded. If the displacement results in a net decrease in the energy of the system, that is, if $\Delta E < 0$, the displacement is accepted and used as the starting point for the next displacement. If $\Delta E > 0$, the displacement is accepted with a probability [*Kirkpatrick et al*., 1983]:

$$P(\Delta E) = \exp\left(-\frac{\Delta E}{k_b T}\right) \tag{1}$$

where $T$ is the temperature of the system and $k_b$ is a physical constant called the Boltzmann constant. The configuration is accepted if the value of $P(\Delta E)$ is less than a random number drawn from a (0, 1) uniform distribution, otherwise it is rejected and the original configuration is used in the next iteration.

The acceptance probability is proportional to the temperature of the system, thus at higher temperature displacements that increase the energy are more likely to be accepted. This probabilistic approach has the advantage of avoiding getting stuck in suboptimal energy configurations and of allowing for a full exploration of the possible configurations at high temperatures.

In an optimization problem, the energy function is the equivalent to the objective or cost function we are seeking to minimize and the atom configurations are just different configurations of the system (e.g., the streamflow sequence in our case) [*Kirkpatrick et al*., 1983; *Dougherty and Marryott*, 1991]. The temperature does not have any physical meaning, it is just a parameter used to control the probability of accepting system configurations that do not reduce the value of the objective function. In the case of synthetic streamflow generation, the algorithm reorders a randomly generated time series until it "freezes" on a time series that matches a set of desired properties specified in the objective function.

The algorithm starts with a randomly sampled time series. Then it rearranges the elements until an improvement in the objective function is achieved. The rearranged configuration becomes the new configuration of the system and the process is repeated until no further improvements in the objective function can be achieved or a user-specified termination criterion is met.

Five elements are required for the algorithm to be implemented: (1) an initial configuration of the time series that can be obtained by randomly sampling an empirical distribution or a theoretical distribution fitted to the observed streamflow data, (2) a swapping algorithm capable of rearranging the elements in the time series (3), an objective function containing the desired properties that need to be represented in the simulated time series, (4) an annealing schedule of the temperature and number of swaps that need to be performed before the system is cooled, and (5) a termination criterion for the algorithm. The choice of the components to be included in the objective function and of the annealing schedule depends on the time series under consideration and the properties of the time series to be manipulated (e.g., seasonality, means, interannual variability).

The objective function may need to be evaluated several thousand times depending on the annealing schedule selected by the user. The performance of the algorithm can be improved by employing fast calculation procedures [e.g., *Bardossy*, 1998], by compromising a little on the accuracy in the objective function (i.e., accepting solutions whose properties are sufficiently close to the target properties) or by implementing the algorithm on parallel computing facilities.

### 2.3. Constructing Streamflow Time Series With Simulated Annealing

A streamflow time series matching a set of desired properties is generated by iterative improvement. This iterative improvement method follows three major steps. A flowchart of the method is shown in Figure 1.

In the first step, a set $Z = \{z_1, z_2, \ldots, z_k\}$ of properties that we want to be reproduced in the synthetic streamflow sequence is selected. The objective function computes the difference between target $z$ and simulated $z^s$ properties. The objective function is formulated as:

$$O = \frac{1}{O_0} \sum_{k=1}^{K} \left(z_k - z_k^s\right)^2 \tag{2}$$

where $k$ is the number of properties specified in the objective function and $O_o$ is the initial value of the objective function. This normalization is needed to ensure that the objective function starts at 1 [*Deutsch and Cockerham*, 1994]. A weight $w_k$ can be applied to each different component $z_k$ of the objective function. By applying weights, one ensures that each component has the same contribution to the objective function, preventing the components with the largest magnitude from dominating the objective function. Weights can also be used to enforce dimensional consistency and to place more importance on particular components of the objective function.

As noted by *Bardossy* [1998], the objective function should include particular properties $z$ that are considered important. For instance, a synthetic streamflow generation method is typically expected to reproduce

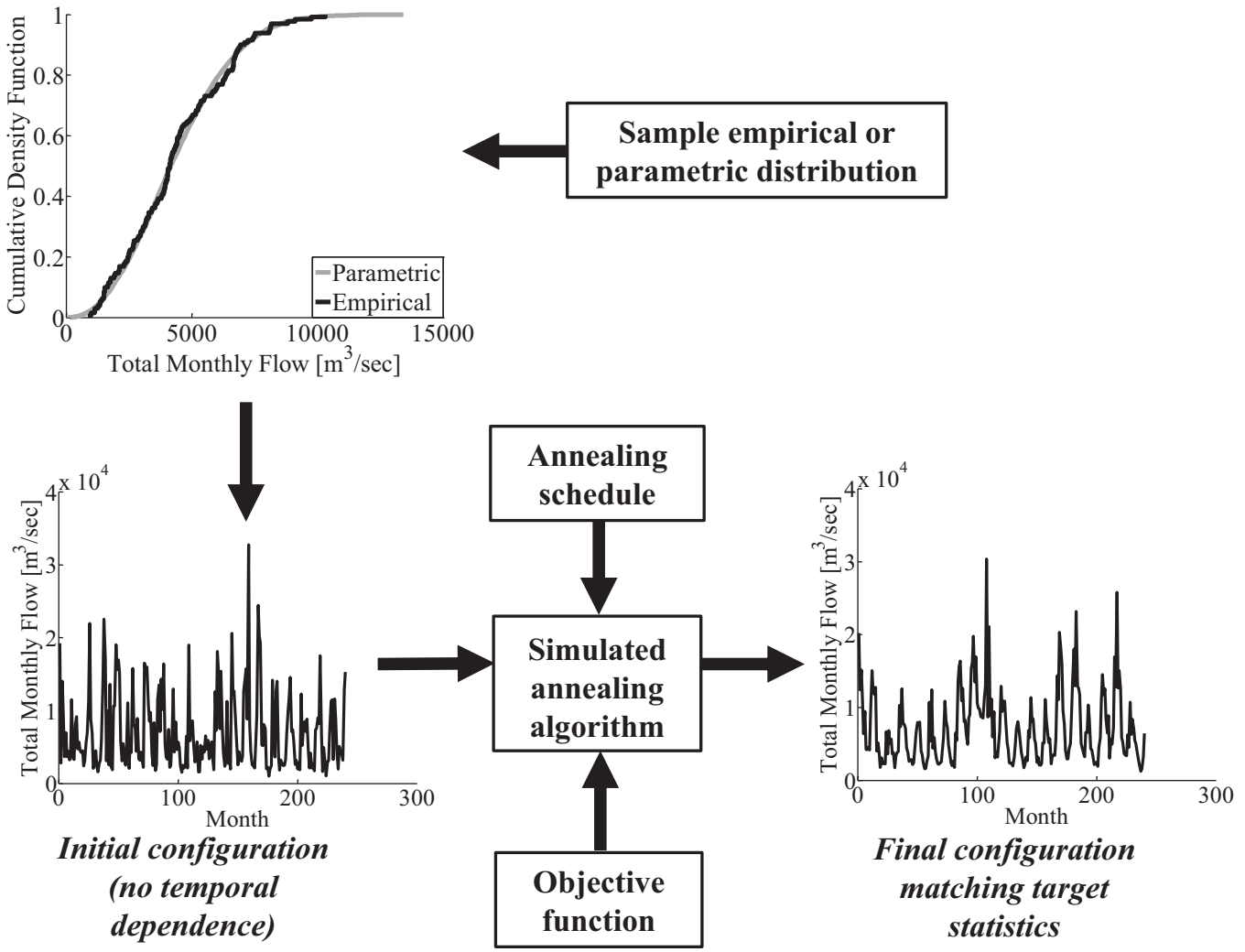


**Figure 1.** Flow chart of the simulated annealing synthetic streamflow generation method.

the mean, standard deviation, and temporal autocorrelation structure of the observed sequence, therefore these properties should be included in the objective function. The flexibility of the method means that several other properties such as autocorrelation of summer months, or Markov Chain transition probabilities can be included in the objective function. Table 1 gives a nonexhaustive list of the statistics one might wish to include in the simulated annealing's objective function.

The target statistical properties in the objective function can be set to the values of the historical sequence or to perturbed values to obtain time series with properties that differ from the historical data and reflect potential climate change induced changes in streamflow characteristics. For instance, in one of the applications proposed in the next section, the standard deviation of the annual totals is adjusted to obtain a streamflow sequence with greater interannual variability. Some statistics may not be independent, meaning that in some applications it may be difficult to perturb some streamflow properties while keeping other properties constant. In these applications, the algorithm can be used to explore trade-offs between each statistics and perturb each statistics as far as theoretically possible as demonstrated in section 3.

The second step consists of generating $n$ streamflow values $x$ for each month $j$ by randomly sampling from the probability distribution of each month $f(x_j)$, where $n$ is the number of years we want to simulate. This distribution $f(x_j)$ for each month can either be a fitted parametric or nonparametric distribution or the empirical distribution based on the observed data, or a combination of the two. For example, a nonparametric distribution might be needed to reproduce the marginal distribution between the 5th and 95th

**Table 1.** Streamflow Properties That Can Be Included in the Objective Function

| Streamflow Properties | Motivation |
|---|---|
| Monthly mean | Adjust monthly flow distribution location and seasonal change |
| Monthly median | Adjust monthly flow distribution location and seasonal change |
| Monthly standard deviation | Adjust variation and spread of monthly flow distribution |
| Skew coefficient | Adjust shape of the flow distribution |
| Interquartile range | Adjust dispersion of monthly flow distribution |
| Quantiles | Adjust flow occurrence at specific values |
| Annual standard deviation | Adjust long-term variability and duration of low flow and high flow periods at interannual time scales |
| Intermonthly autocorrelation at a range of monthly lags | Adjust high-frequency persistence and distribution of prolonged low flows and high flow periods and control the occurrence of rapid changes in river stage |
| Interannual autocorrelation at a range of annual lags | Adjust low-frequency, long-term persistence and correlation at user-specified lags |
| Markov Chain transition probability | Adjust flow event sequencing |

percentiles, while a Generalized Pareto Distribution might be fitted to the tails. The values sampled for each month $x$ are concatenated to form a time series $ts_i$ of length $12n$ of monthly streamflow values. The time series generated at this stage is not expected to preserve autocorrelation statistics, as these are imposed in the next step.

The third step is the reshuffling of the initial time series with the simulated annealing algorithm. The algorithm works as follows:

1. Start with the time series $ts_i$ in a given configuration and an initial temperature $T_0$ (high temperature).
2. Randomly select and swap the values at 2 times $t_1 \neq t_2$ and calculate the change in the objective function $\Delta O$ after the swap.
3. If $\Delta O < 0$, the swap is accepted, as the swap has reduced the objective's function value thus moving the optimization closer to the objective.
4. If $\Delta O \geq 0$, the swap is still accepted with a probability:

$$P(\Delta O) = \exp\left(-\frac{\Delta O}{T}\right) \tag{3}$$

   The probability of accepting a swap decreases as $T$ decreases (so essentially the more the objective function increases, the less likely it is to accept swaps that do not yield any improvements: this is the Metropolis algorithm).
5. After $M$ iterations of steps 2–4, the value of $T$ is lowered and the simulated annealing algorithm is repeated again. The value of $T$ determines how likely we are to accept nonimproving swaps (i.e., swaps where $\Delta O \geq 0$).
6. The algorithm is terminated when the objective function $O = 0$ or when user-specified termination criteria are met.

There is no standard procedure for selecting an annealing schedule (i.e., number of swaps, initial temperature, and number of temperature reductions) [*Kannan et al*., 2005]. In general, the number of swaps at each temperature should be proportional to the number of time steps in the time series and at least equal to the length of the time series to ensure that each value at each time step has the potential to be swapped. *Bardossy* [1998] recommends a number of swaps equal to double the length of the time series in question. The initial temperature should be high enough so that all swaps are accepted at the start of the algorithm [*Eglese*, 1990]. It is important to ensure that the algorithm does not spend too much time at high temperatures where all swaps are accepted because this may result in wasted computing time.

### 2.4. Model Verification and Validation

The performance of the simulated annealing algorithm can be evaluated using verification and validation statistics. As explained by *Stedinger and Taylor* [1982], model verification is required to show that the streamflow generation model correctly reproduces the flow statistics it is designed to reproduce, in this case the statistics included in the objective function. In general, a synthetic streamflow generation method

should preserve moment statistics, such as the monthly means and standard deviations, and also the time dependence structure (intermonthly and interannual autocorrelation) of the observed flows [*Hao and Singh*, 2012] and it is recommended that these properties are included in the objective function.

Model verification can be carried out by comparing the observed and simulated statistics with dot and line plots or box plots and by verifying that the objective function converges to a minimum as the simulated annealing's temperature is reduced. The number of accepted swaps can be plotted against the temperature reductions to evaluate the algorithm's speed of convergence and to confirm that the probability of accepting nonimproving swaps decreases with temperature.

Model validation is needed to demonstrate that the streamflow model faithfully reproduces other flow characteristics which are not explicitly included in the model formulation (i.e., in the simulated annealing's objective function) but which are still important for water resources management applications. Specific statistics not included in the objective function (e.g., specific quantiles, transition probabilities) should be compared using dot and line plots or box plots as in model verification. Of particular importance to water management studies is the ability of any synthetic streamflow generation method to correctly represent the drought characteristics of the observed series. Drought statistics, including average and maximum drought length and deficits, of the observed and simulated flows should be compared with box plots. Quantile-quantile plots, marginal probability density function plots, and flow duration curves should also be used for model validation to gain a more comprehensive view of the impacts of the simulated annealing algorithm on the whole streamflow distribution.

## 3. Application

The monthly naturalized flows observed for the River Thames at Kingston, in the south of England, from 1883 to 2012 inclusive, were used to demonstrate and test the proposed method. The Thames River at Kingston drains an area of about 10,000 km$^2$ and is characterized by a temperate climate, with streamflow minima normally occurring during the summer months. The Thames River Basin has been classified as seriously water stressed by the *Environment Agency* [2008] and climate change impact assessments in the area suggest a reduction in water availability [*Diaz-Nieto and Wilby*, 2005; *Manning et al*., 2009], underscoring the need to understand the vulnerability of water resources in the basin to changing streamflow characteristics.

The proposed method is applied to generate: (i) monthly streamflow sequences with the same characteristics as the observed record and (ii) monthly streamflow sequences with some properties perturbed to represent potential climate change impacts on streamflow characteristics. This latter application demonstrates how the proposed method can be used to perturb specific properties of the time series while keeping other properties unaltered and generate streamflow sequences for use in vulnerability assessments of water supply systems of the type described by *Nazemi and Wheater* [2014].

### 3.1. Historical Streamflow

In this application, 100 flow sequences, each 130 years long, were generated to assess the ability of the proposed method to reproduce the streamflow statistics of the observed sequence. The initial time series $ts_i$ was generated by randomly sampling from the empirical distributions of the monthly flows observed for the Thames at Kingston from 1883 to 2012. The empirical distribution was preferred over parametric or nonparametric distributions to avoid having to make assumptions with respect to distribution parameters and density estimation methods for nonparametric distributions. To generate streamflow sequences that reproduce the historical flow data, three components were included in the objective function:

1. Monthly means.

2. Monthly standard deviations.

3. Intermonthly up to lag-8 autocorrelation.

The objective function was formulated as follows:

$$O=\frac{1}{O_0}\left[\sum_{c=1}^{8}\left(\rho_c-\rho_c^s\right)^2+\sum_{j=1}^{12}\left(\bar{m}_j-\bar{m}_j^s\right)^2+\sum_{j=1}^{12}\left(s_j-s_j^s\right)^2\right] \tag{4}$$

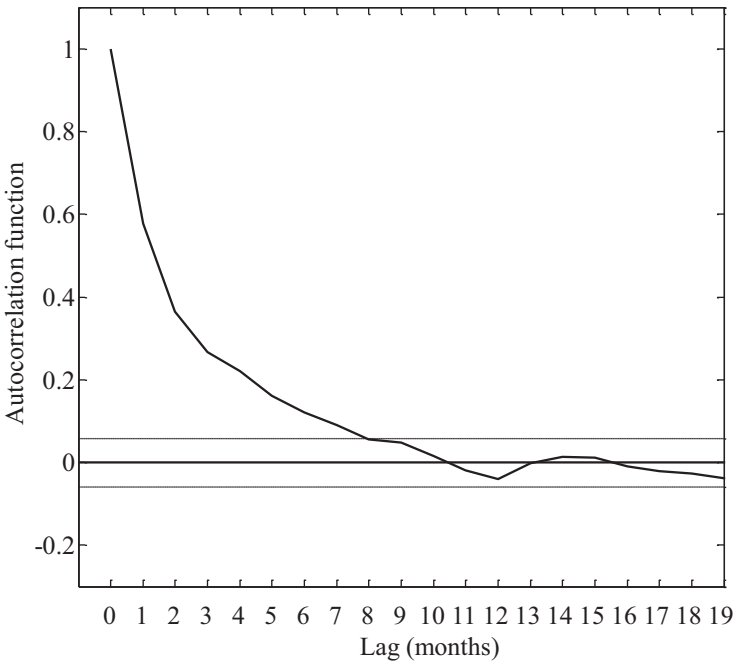


**Figure 2.** Autocorrelation function of the deseasonalized monthly streamflow totals observed for the Thames at Kingston (1883–2012).

where $\rho_c$ and $\rho_c^s$ are the intermonthly autocorrelation coefficients of the observed and of the simulated sequences, respectively, $\bar{m}_j$ and $\bar{m}_j^s$ are the means of the monthly flow totals for month $j$ for the observed and simulated sequences, respectively, and $s_j$ and $s_j^s$ are the standard deviations of monthly flow totals for month $j$ for the observed and simulated sequences, respectively. The monthly means and standard deviations were selected as components of the objective function to preserve the seasonality signal in the simulated time series. The intermonthly autocorrelation was selected because the observed deseasonalized monthly flow totals show significant intermonthly autocorrelation at the 5% significance level up to lag-8, as shown in Figure 2. We note that a smaller or larger number of lags may be included in the objective function depending on the data at hand and the hydrologist's preferences. These streamflow statistics are broadly considered as the most important flow properties that should be preserved by any synthetic hydrology method [*Sharma et al.*, 1997; *Hao and Singh*, 2012].

To demonstrate the effectiveness of our method at reproducing historical streamflow characteristics, we compare our results with streamflow sequences generated with a linear lag-1 autoregressive process traditionally used in hydrology [*Loucks and van Beek*, 2005]. A monthly AR(1) model with seasonally varying coefficients was fitted to the logarithmically transformed observed monthly flows. In this model, the flow $X$ for month $j$ is calculated as follows [*Thomas and Fiering*, 1962]:

$$X_j = \bar{X}_j + \rho_{j,\,j-1}\frac{\sigma_j}{\sigma_{j-1}}\left(X_{j-1} - \bar{X}_{j-1}\right) + \sigma_j\sqrt{1-\rho_{j,\,j-1}^2}\,\varepsilon_t \tag{5}$$

where $\bar{X}_j$ is the mean streamflow for month $j$, $\rho_{j,\,j-1}$ is the lag-1 autocorrelation coefficient between months $j$ and $j-1$, $\sigma_j$ is the standard deviation of the flow in month $j$, and $\varepsilon_t$ is the noise term with mean 0 and variance 1.

An example of the monthly streamflow series generated with the method is given in Figure 3. The initial temperature $T_0$ was set to 100 and the number of temperature reductions was set to 20. For each temperature, we carried out 3200 iterations. With this annealing schedule, it took approximately 4 h of continuous computing on a 3.4 GHz processor to generate 100 synthetic streamflow sequences. Figure 4 shows that the value of the objective function approaches zero after 20 temperature reductions. Similarly, Figure 5 shows the number of accepted swaps against temperature reductions. The number of accepted swaps rapidly decreases with decreasing temperature, as expected given that at lower temperatures there is a lower

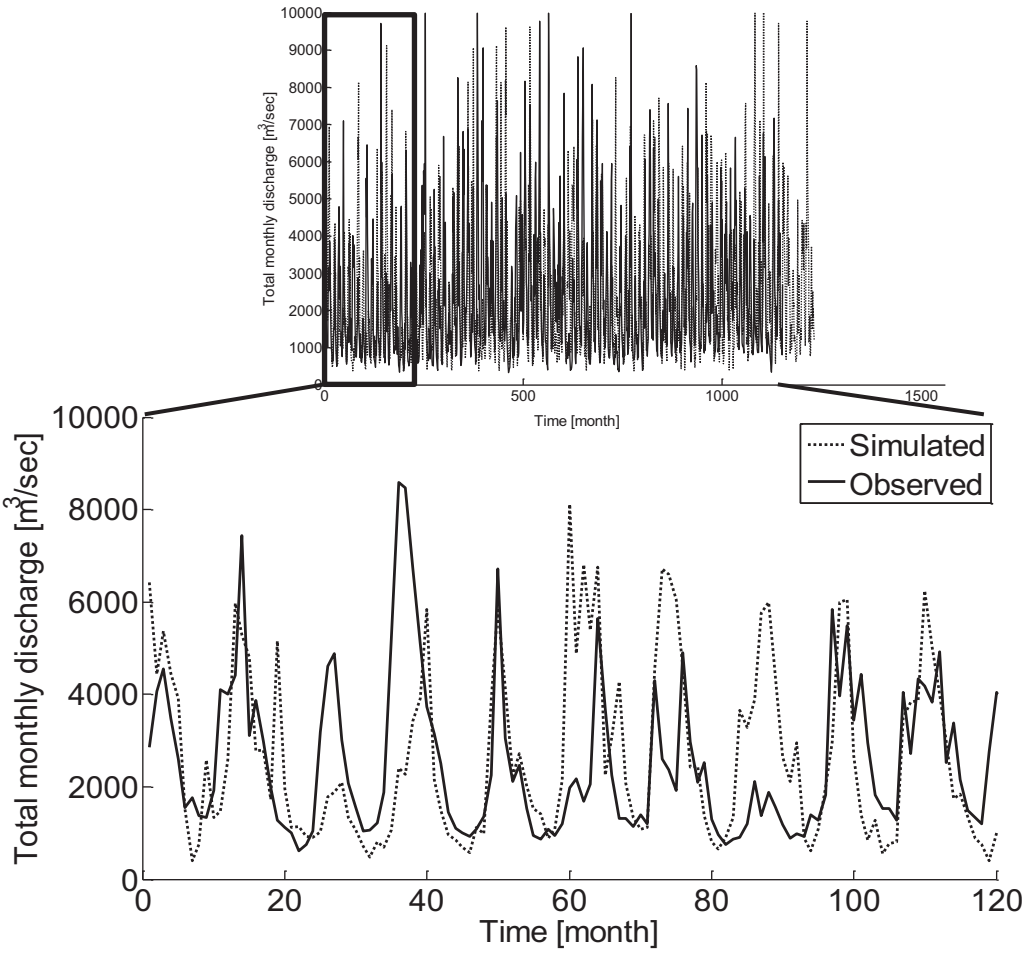


**Figure 3.** Simulated and observed total monthly flow for the Thames at Kingston.

probability of accepting nonimproving swaps and suggesting that the time series is "freezing" on a structure that matches the selected properties.

To assess the performance of the algorithm, we also examine the trade-offs between the three components in the objective function in equation (4). Analyzing the trade-offs between the three components of the objective function can play a useful role in assessing algorithm performance and in determining whether or not multiobjective search should be accommodated into the simulated annealing schedule. The trade-offs between the three different objectives in equation (4) (matching the target mean, standard deviation, and autocorrelation structure at the monthly time scale) are shown in Figure 6. Figure 6 shows the value of the three objectives at each iteration for three representative temperature values. As the temperature is reduced, all three objective functions converge to zero, suggesting that our formulation of the optimization problem in equation (4) is capable of meeting all three objectives (i.e., generating a time series matching the target statistics). These results demonstrate that our method is able to reproduce any desired statistics provided it is consistent with the other statistics. Future work should explore the potential for multiobjective search algorithms to solve this time series generation problem.

To verify the streamflow generation method and assess whether or not it is capable of reproducing the statistics included in the objective function, we plotted the monthly means and standard deviations of the simulated series against the first two moments of the observed series, as shown in Figure 7. All simulated series had monthly means and standard deviations within 5% of the observed value (solid black line in Figure 7).

The simulated series also successfully reproduce the intermonthly autocorrelation up to lag-8. To display the efficacy of the simulated annealing algorithm, Figure 8 compares the observed (black line) and

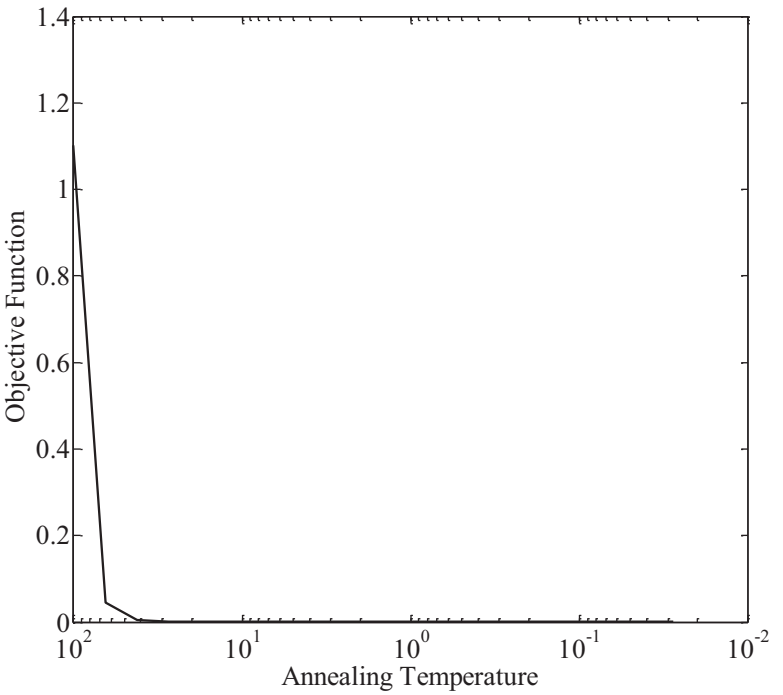


**Figure 4.** Value of the objective function against temperature reductions for the simulated historical sequence.

simulated intermonthly autocorrelation function (boxes) of the deseasonalized monthly flows. The intermonthly autocorrelation for the trial time series bootstrapped at the start of the algorithm is also shown in Figure 8 (circles). Figure 8 shows how the simulated annealing is capable of imposing an intermonthly autocorrelation structure starting from a time series with no autocorrelation structure. The method generates streamflow sequences that match the desired properties included in the objective function, thus showing

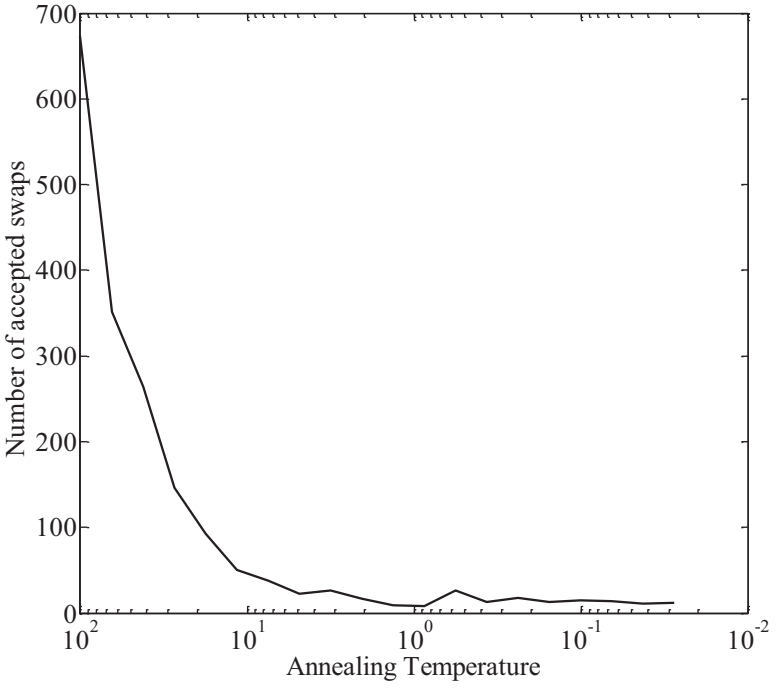


**Figure 5.** Number of accepted swaps during the simulation of the historical sequence against the annealing temperature.

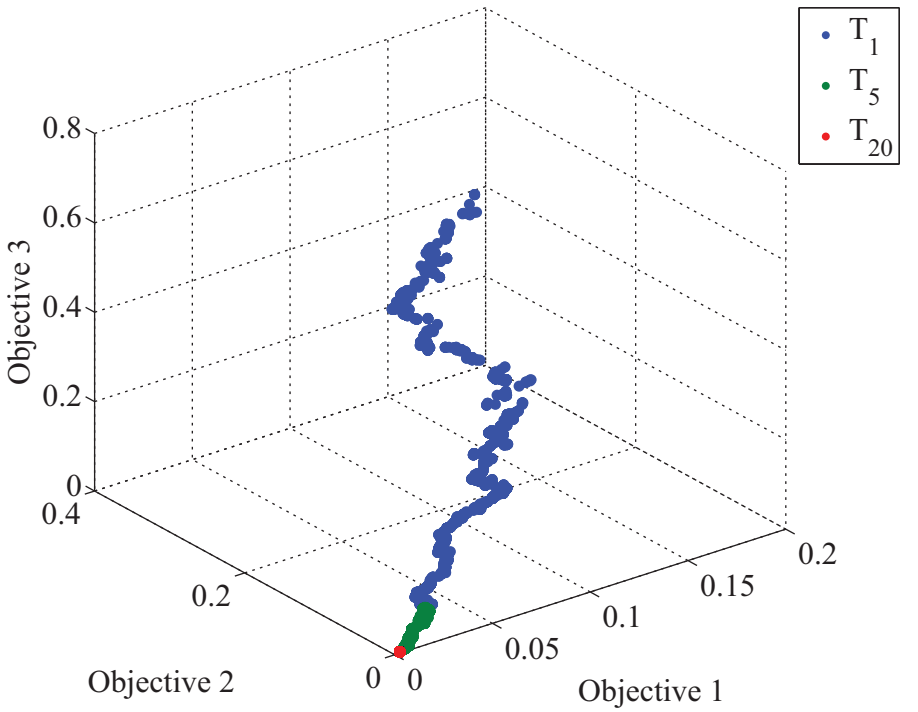


**Figure 6.** Trade-off curve for objective 1 (matching monthly means), objective 2 (matching monthly standard deviation), and objective 3 (matching autocorrelation structure) for 3200 iterations and for three different annealing temperatures.

that it faithfully reproduces the statistics it was designed to reproduce (i.e., fulfilling the model verification requirements).

A second set of diagnostics was examined to assess whether the synthetic series preserved some of the properties of the observed series which were not included in the objective function (model validation). This is to demonstrate that the simulated annealing algorithm reproduces the properties specified in the objective function but at the same time preserves additional important properties of the flow distribution. Figure 9 shows the maximum, minimum, Q95, and skewness box plots for the 100 synthetic monthly sequences and for the observed data (solid black line). The maximum and minimum flows are well preserved in the

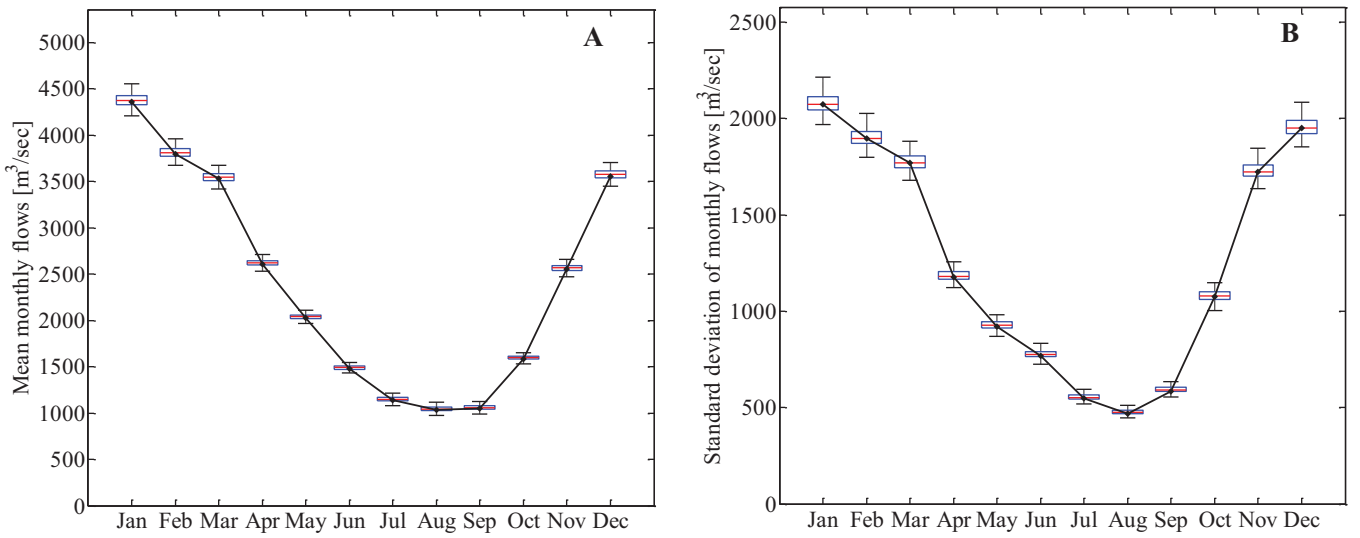


**Figure 7.** (a) Monthly means and (b) standard deviations for 100 simulated historical monthly streamflow sequences each 130 years long. The black line shows the same statistics for the observed data.

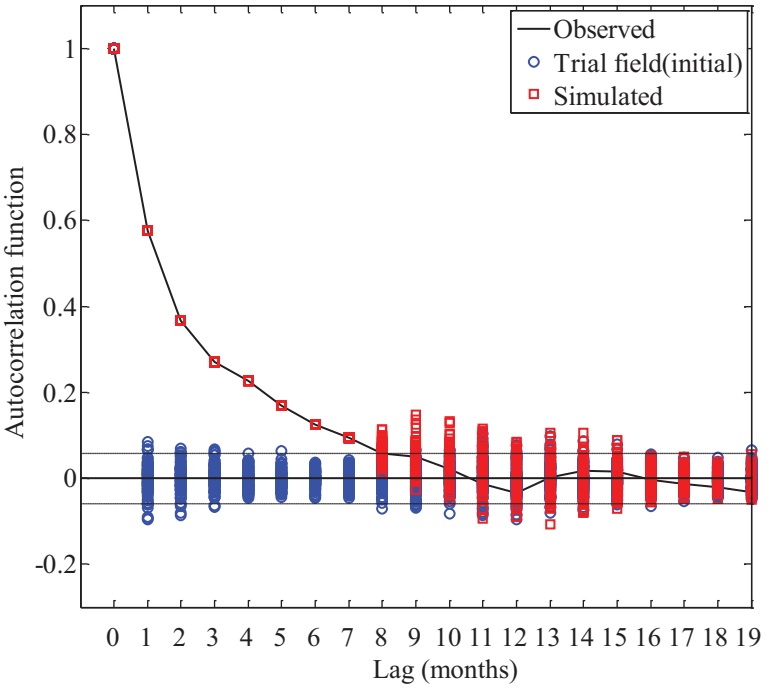


**Figure 8.** Autocorrelation function for 100 simulated deseasonalized monthly streamflow sequences at the start (circles) and end (squares) of the simulated annealing algorithm.

synthetic sequence (Figures 9a and 9b) and so are the Q95 flows (Figure 9c). The synthetic streamflow sequences show a slight underestimation of the skewness for the summer months (Figure 9d), but still match the skewness of the observed monthly flows better than the AR(1) process (see supporting information). To achieve a better match to the observed data, the monthly skewness coefficients can be included in the objective function.

The mean, standard deviation, and lag-1 correlation of the annual streamflow totals are shown in Figure 10. Although these properties were not included in the objective function, the synthetic streamflow sequences show a good match with the same statistics for the observed sequence (black horizontal line in Figure 10), suggesting that our method fulfills the requirements for model validation. The proposed model is capable of preserving important properties at the monthly and annual temporal scales without requiring any temporal disaggregation procedure. Figure 10 also shows box plots of the means, standard deviations, and interannual lag-1 autocorrelation coefficients for 100 streamflow series generated with the AR(1) process. The sequences generated with the AR(1) process show a greater spread around the mean than the sequences generated with simulated annealing and a systematic underestimation of the interannual variability and of the lag-1 interannual correlation coefficient of the observed data (black line).

To compare our method with the traditional AR(1) process, we also estimated the marginal probability density functions for the monthly flows generated with the two methods. Figure 11 shows box plots of the marginal probability densities for 100 sequences of monthly flows simulated with the synthetic streamflow generator and the AR(1) process for February and July. The synthetic streamflow generator shows a better match to the historical marginal probability densities (solid line in Figure 11) and better captures the shape of the density function. QQ plots of the simulated and historical data also show that our method preserves the underlying distributions of the monthly flows and that the shuffling algorithm does not result in unrealistic monthly flow distributions (see supporting information).

### 3.2. Climate Perturbed Streamflows

In this section, we apply the method to generate perturbed streamflow sequences. We demonstrate how by including particular streamflow properties in the objective function, we are able to control the outcomes of the adjustment, obtaining the changes desired in some streamflow characteristics while keeping some

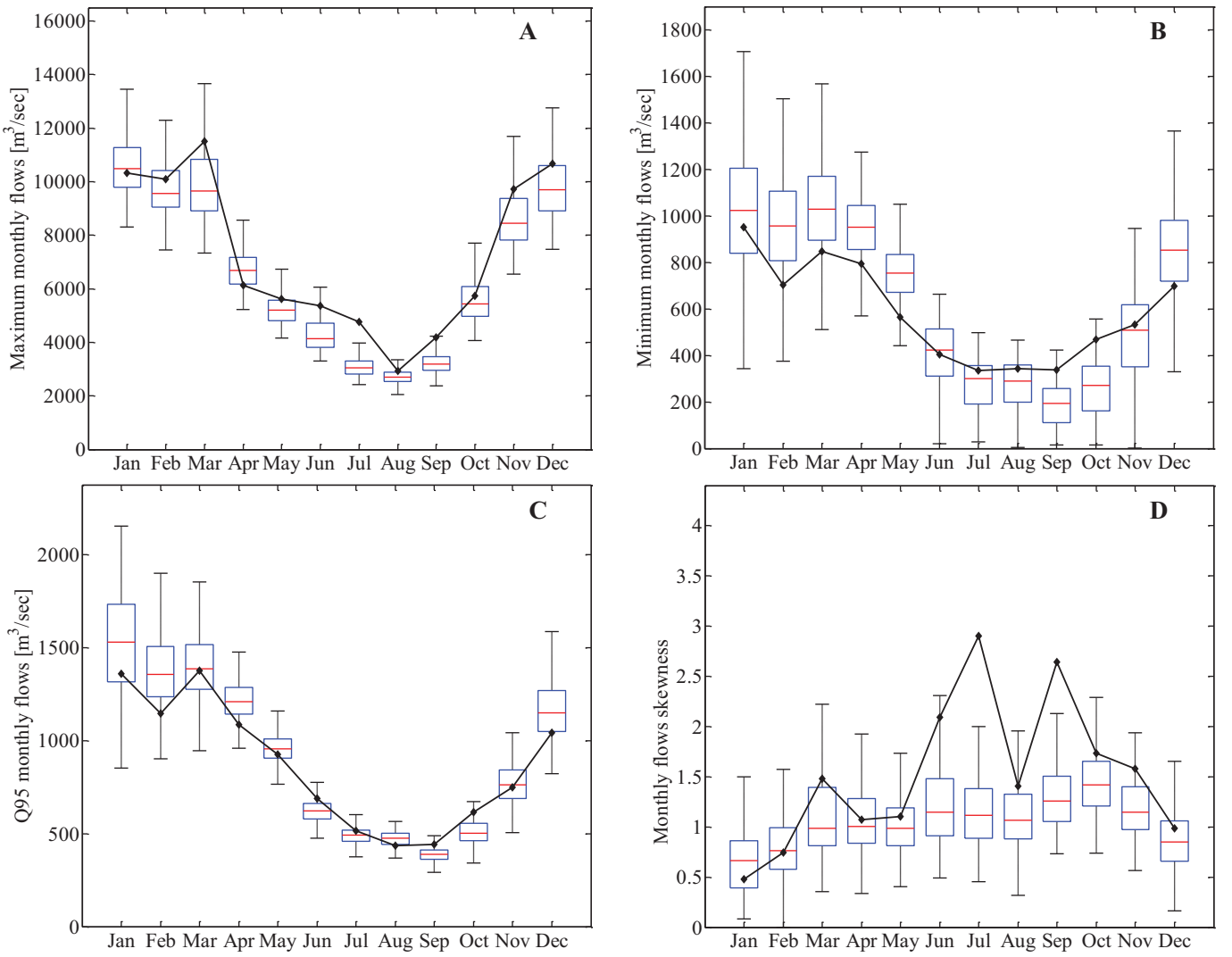


**Figure 9.** Monthly (a) maximum, (b) minimum, (c) Q95, and (d) skewness statistics for 100 simulated historical monthly streamflow sequences each 130 years long. The black lines show the same statistics for the observed data.

other statistics unaltered. This is particularly important when testing water resource system sensitivities to climate-induced changes in streamflow characteristics, because it allows for the simulation of specific changes at both annual and monthly time scales in the streamflow time series.

Several properties of the time series can be adjusted to simulate changes in the hydroclimatological regime. For instance, the strength of the interannual autocorrelation or the degree of intraannual variability can be perturbed to simulate the effects of a changing climate on streamflow characteristics. An increase in the strength of interannual autocorrelation and variability could result in longer periods of below (or above) average streamflows, which could increase the likelihood of multiyear droughts or floods. Changes in high-frequency persistence and intraannual variability could result in changing patterns of streamflow occurrences, with most streamflow occurring in a few months of the year.

Three examples of climate perturbed streamflows are provided here. The same simulated annealing schedule used to generate the synthetic sequences in section 3.1 (20 temperature reductions, 3200 iterations at each temperature) was applied to generate the perturbed sequences. In the first example (referred to as SA 1), we generate 100 sequences of monthly flows, each 130 years long, with increased high-frequency persistence but with unchanged long-term variability. The increase in high-frequency persistence is represented by applying a perturbation factor $p = 1.5$ to the observed intermonthly lag-1 autocorrelation

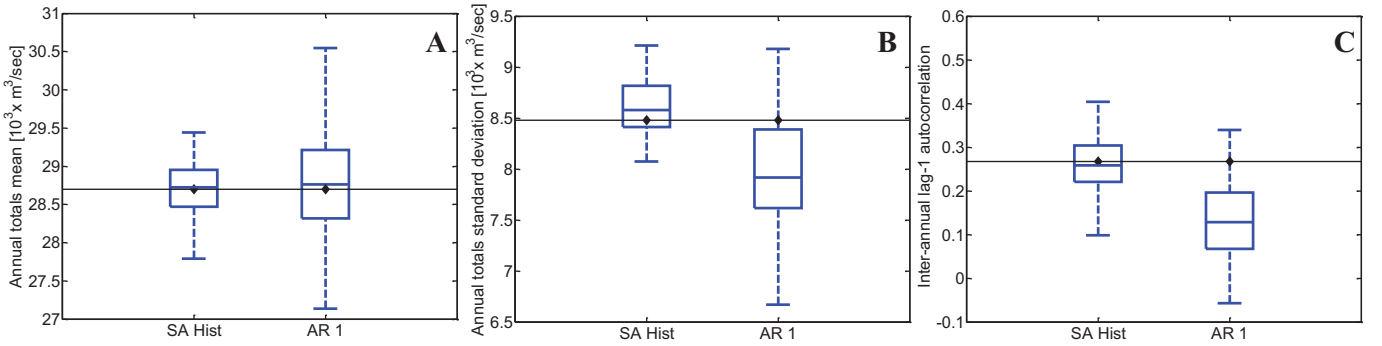


**Figure 10.** Box-plots of the (a) mean, (b) standard deviation, and (c) interannual lag-1 autocorrelation statistics for 100 simulated sequences using simulated annealing (SA Hist) and an AR 1 process. The horizontal black lines show the same statistics for the observed data.

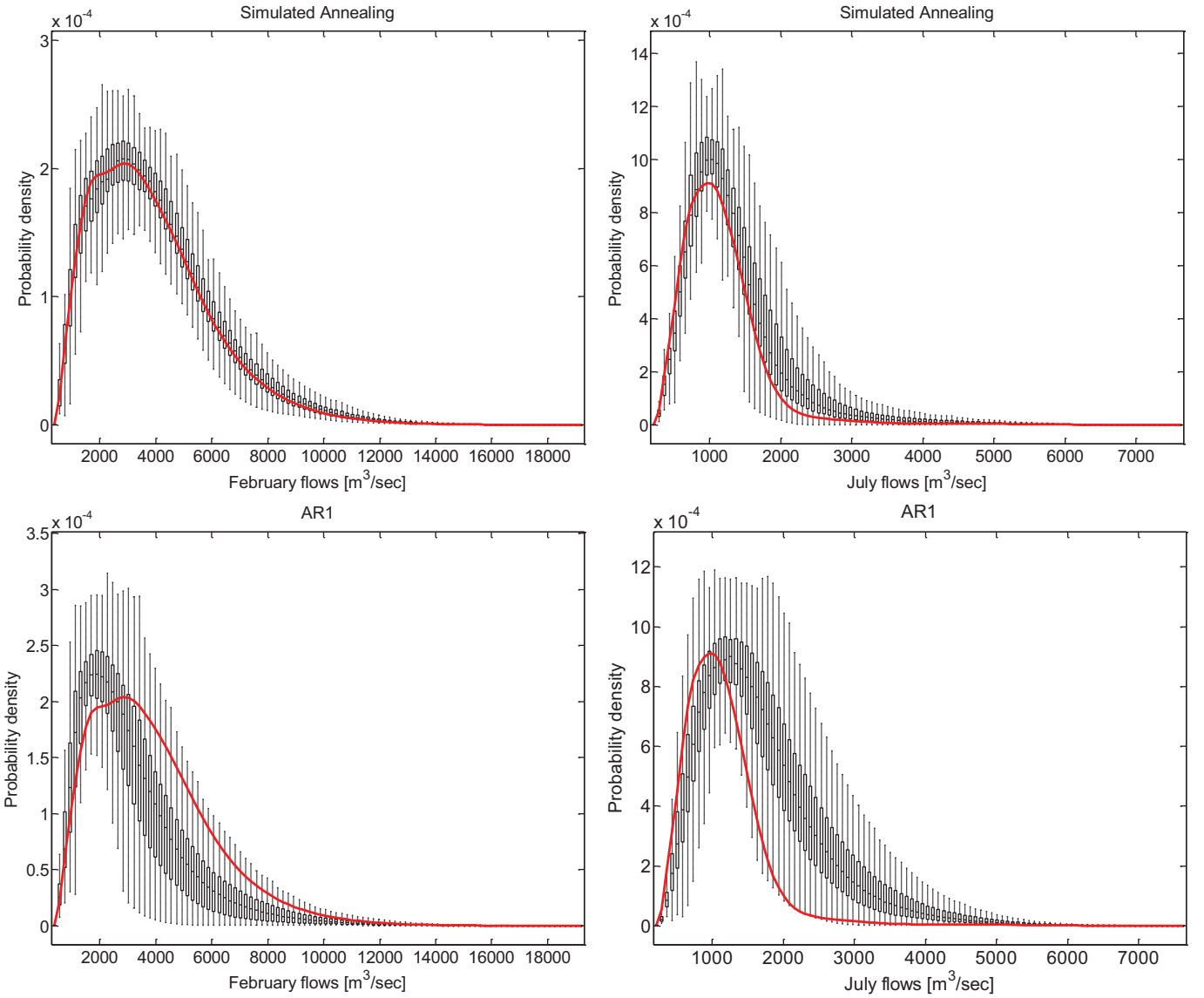


**Figure 11.** Marginal probability density functions estimated by kernel density estimation for July and February flow. (top) The solid line indicates the historical record and the box-plots show the estimates obtained for 100 realizations of the simulated annealing algorithm and (bottom) of the AR 1 process.

coefficient, representing a 50% increase in the correlation between consecutive months in the time series. The long-term variability is kept unaltered by setting the target annual standard deviation to the historical period value.

To achieve this desired change, we include the following properties in the objective function: (i) monthly means, (ii) monthly standard deviations, (iii) perturbed intermonthly lag-1 autocorrelation, (iv) annual standard deviation, and (v) interannual lag-1 autocorrelation. The objective function is formulated as follows:

$$O=\frac{1}{O_0}\left[(d-d^s)^2+\ (\sigma-\sigma^s)^2+(p\rho_1-\rho_1^s)^2+\sum_{c=2}^{8}(\rho_c-\rho_c^s)^2+\sum_{j=1}^{12}\left(\bar{m}_j-\bar{m}_j^s\right)^2+\ \sum_{j=1}^{12}\left(s_j-s_j^s\right)^2\right] \tag{6}$$

where $p$ is the perturbation factor, $d$ and $d^s$ are the interannual lag-1 autocorrelation coefficients of the observed and simulated sequences, respectively, $\sigma$ and $\sigma^s$ are the standard deviations of the annual totals of the observed and simulated sequences, respectively. For this first example, we compare our results with a flow series generated with the AR(1) process described above with the same perturbation factor $p$ imposed on the autocorrelation coefficient.

In the second example (referred to as SA 2), we generate 100 sequences of monthly flows, each 130 years long, with an increase in the magnitude of interannual variability but with unchanged persistence structure at both the monthly and annual levels. The increase in interannual variability is represented by applying a $p = 1.3$ perturbation factor to the standard deviation of the annual totals of the observed streamflow sequence. The persistence structure is preserved by including the observed intermonthly and interannual autocorrelation coefficients in the objective function. The objective function includes (i) monthly means, (ii) monthly standard deviations, (iii) intermonthly autocorrelation, (iv) interannual lag-1 autocorrelation, and (v) perturbed standard deviation of the annual totals. The objective function is:

$$O=\frac{1}{O_0}\left[(d-d^s)^2+\ (p\sigma-\sigma^s)^2+\ \sum_{c=1}^{8}(\rho_c-\rho_c^s)^2+\sum_{j=1}^{12}\left(\bar{m}_j-\bar{m}_j^s\right)^2+\sum_{j=1}^{12}\left(s_j-s_j^s\right)^2\right] \tag{7}$$

In the third example (referred to as SA 3), we generate 100 sequences of monthly flows, each 130 years long, with a 15% ($p = 0.85$) decrease in the magnitude of monthly means of the summer months but with unchanged persistence and variability structures at the monthly and annual levels. The objective function includes (i) monthly means for the autumn, winter, and spring months (ii) perturbed monthly means for the summer months (iii) monthly standard deviations, (iii) intermonthly autocorrelation, and (v) interannual lag-1 autocorrelation. The objective function is:

$$O=\frac{1}{O_0}\left[(d-d^s)^2+\ \sum_{c=1}^{8}(\rho_c-\rho_c^s)^2+\sum_{j=6}^{8}\left(p\bar{m}_j-\bar{m}_j^s\right)^2+\sum_{\substack{j=1\\ j\neq 6,7,8}}^{12}\left(\bar{m}_j-\bar{m}_j^s\right)^2+\sum_{j=1}^{12}\left(s_j-s_j^s\right)^2\right] \tag{8}$$

The results of the applications were analyzed by plotting the first two moments and the autocorrelation structure of the monthly and annual flows and also the Q95 of the monthly flows. Figure 12 shows the mean monthly flows generated for each of the three different perturbations and also the mean monthly flows generated by perturbing the intermonthly lag-1 autocorrelation coefficient in the AR(1) process. Perturbing the lag-1 intermonthly autocorrelation coefficient in the AR (1) and SA 1 examples does not result in any major departures from the monthly means of the observed sequence (Figures 12a and 12b). The SA 2 sequences also preserve the monthly means as specified in the objective function (Figure 12c). Figure 12d shows that the perturbation of the summer mean monthly flows was successful, resulting in sequences with summer mean monthly flows 15% lower than the observed values.

The standard deviations of the monthly flows are shown in Figure 13. The sequences generated with the perturbed AR (1) process (Figure 13a) show a greater spread around the observed standard deviation than the sequences generated with the simulated annealing algorithm. In general, all sequences generated with the simulated annealing algorithm display a close match with the standard deviation of the observed as specified in the objective function.

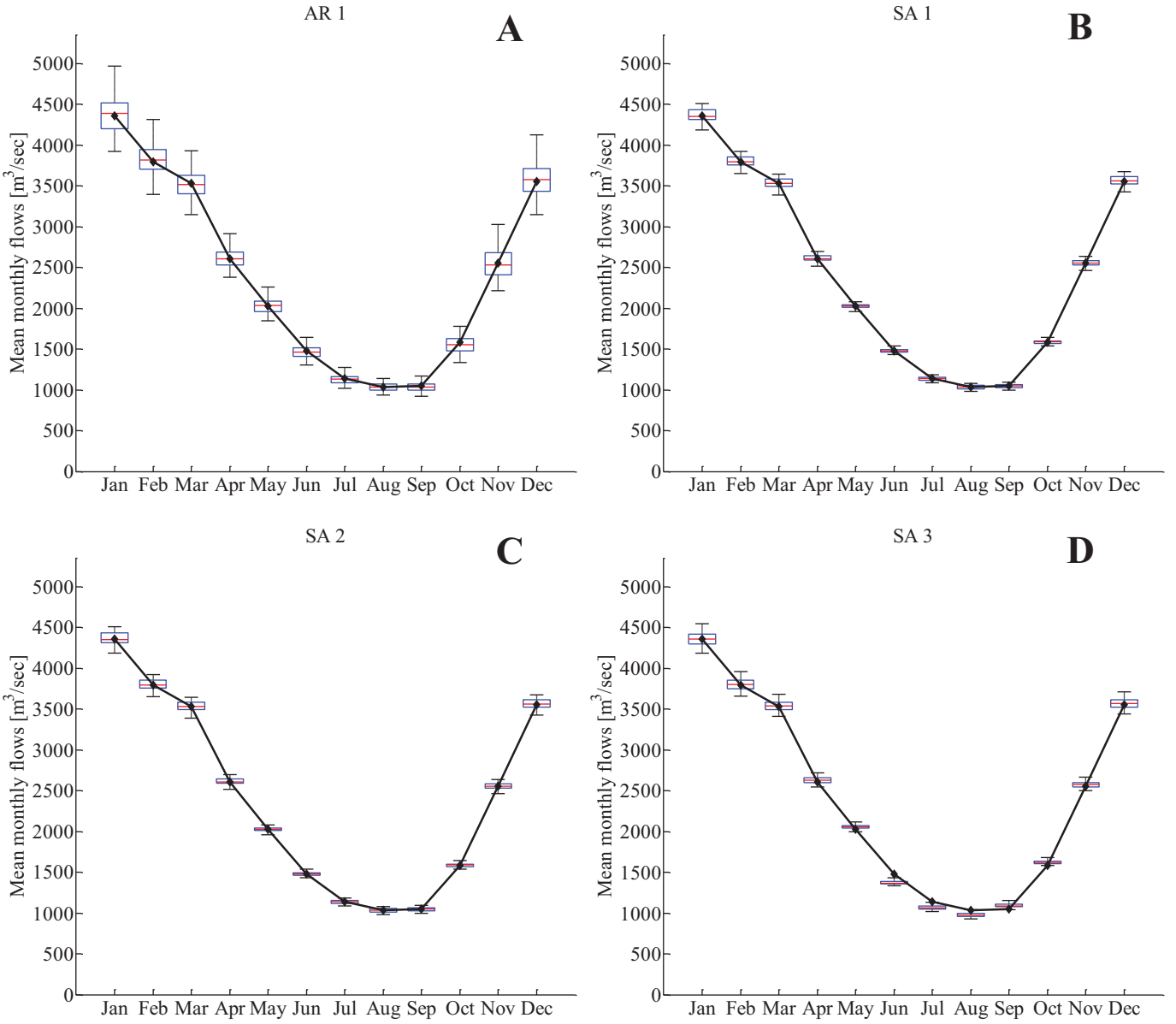


**Figure 12.** Box-plots of mean monthly flows for 100 realizations generated by perturbing (a) the intermonthly lag-1 autocorrelation coefficient in the AR 1 process, (b) the intermonthly lag-1 autocorrelation coefficient in the SA, (c) the annual standard deviation, and (d) the mean monthly flows of the summer months. The black lines show the same statistics for the observed data.

The autocorrelation functions for the monthly sequences are shown in Figure 14. The autocorrelation functions for the AR(1) and SA 1 examples where the intermonthly lag-1 autocorrelation coefficient was perturbed are shown in Figures 14a and 14b. The autocorrelation function of the AR(1) process shows that the perturbation at lag-1 also influences subsequent lags, whereas the autocorrelation function for the sequences generated with our method only shows the desired perturbation at lag-1 (Figure 14b). The SA 2 sequences (Figure 14c) preserve the intermonthly autocorrelation structure up to lag-8, as specified in the objective function. At lags greater than 8, the SA 2 synthetic sequences show a greater autocorrelation than the observed, suggesting that increasing the interannual variability (i.e., increasing the annual standard deviation) also increases the long-term memory of the sequence. The flexibility of the proposed method means that the extent to which the intermonthly autocorrelation changes as a result of the perturbation can be controlled by including lags greater than 8 in the objective function.

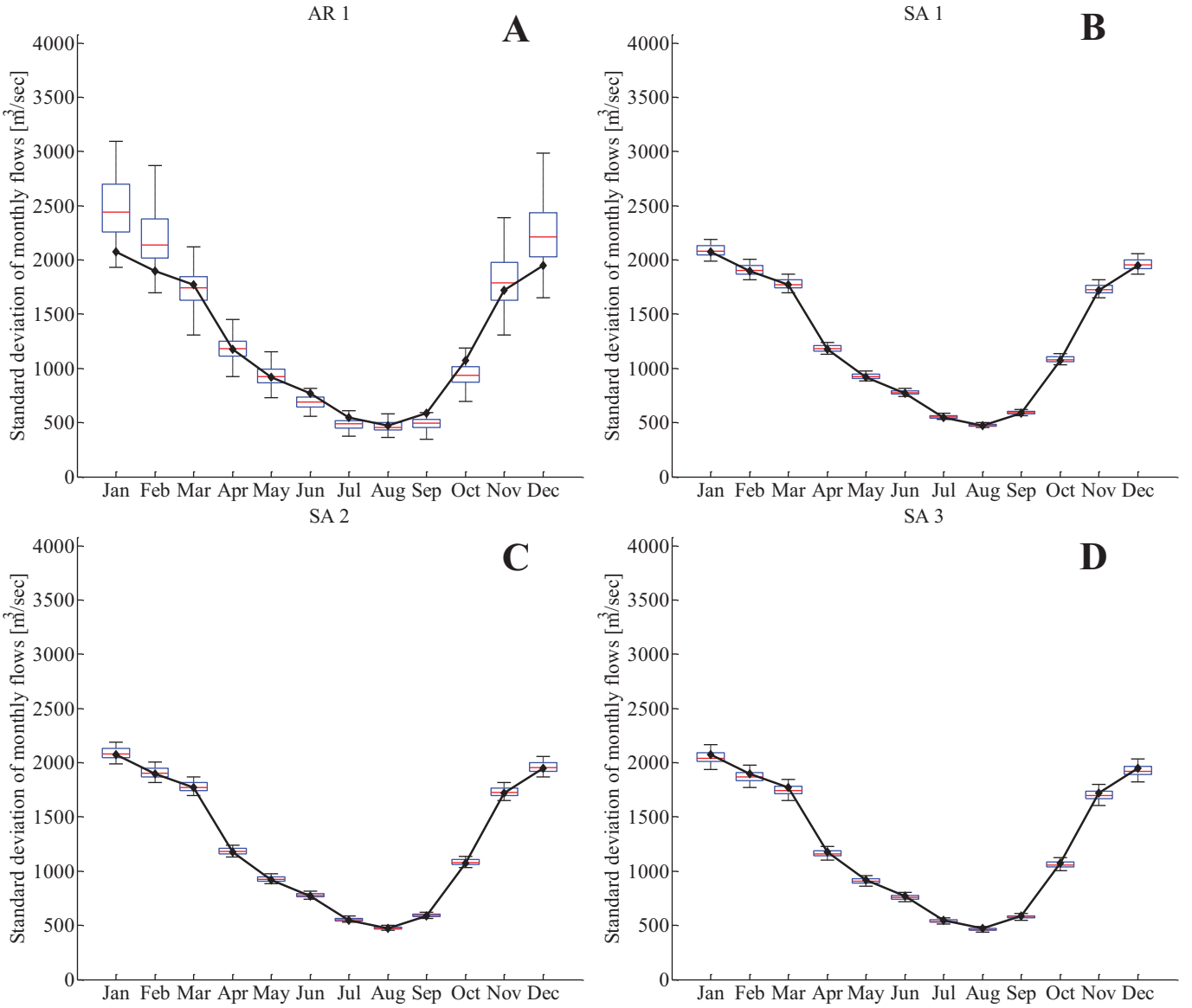


**Figure 13.** Box-plots of the standard deviation of monthly flows for 100 realizations generated by perturbing (a) the intermonthly lag-1 autocorrelation coefficient in the AR 1 process, (b) the intermonthly lag-1 autocorrelation coefficient in the SA, (c) the annual standard deviation, and (d) the mean monthly flows of the summer months. The black lines show the same statistics for the observed data.

The monthly Q95 of the perturbed sequences are shown in Figure 15. The monthly Q95 were not included in the objective function and this explains the spread around the observed value (black line) in Figures 15b–15d. Figure 15d shows that perturbing the summer monthly means does not result in a significant departure in the summer monthly Q95. Even though the monthly Q95 were not included in the objective function, the perturbed sequences still show a good match with the observed data, providing further validation for our method.

Figure 16 shows the annual means, annual standard deviations, and interannual lag-1 autocorrelation coefficients calculated for the examples. The black horizontal lines in Figure 16 represent the statistics for the observed series.

At the annual time scale the flow series generated with the simulated annealing method show a closer match to the observed annual mean (Figure 16a) than the sequences generated with the AR (1) process, even though the annual means were not included in the objective function.

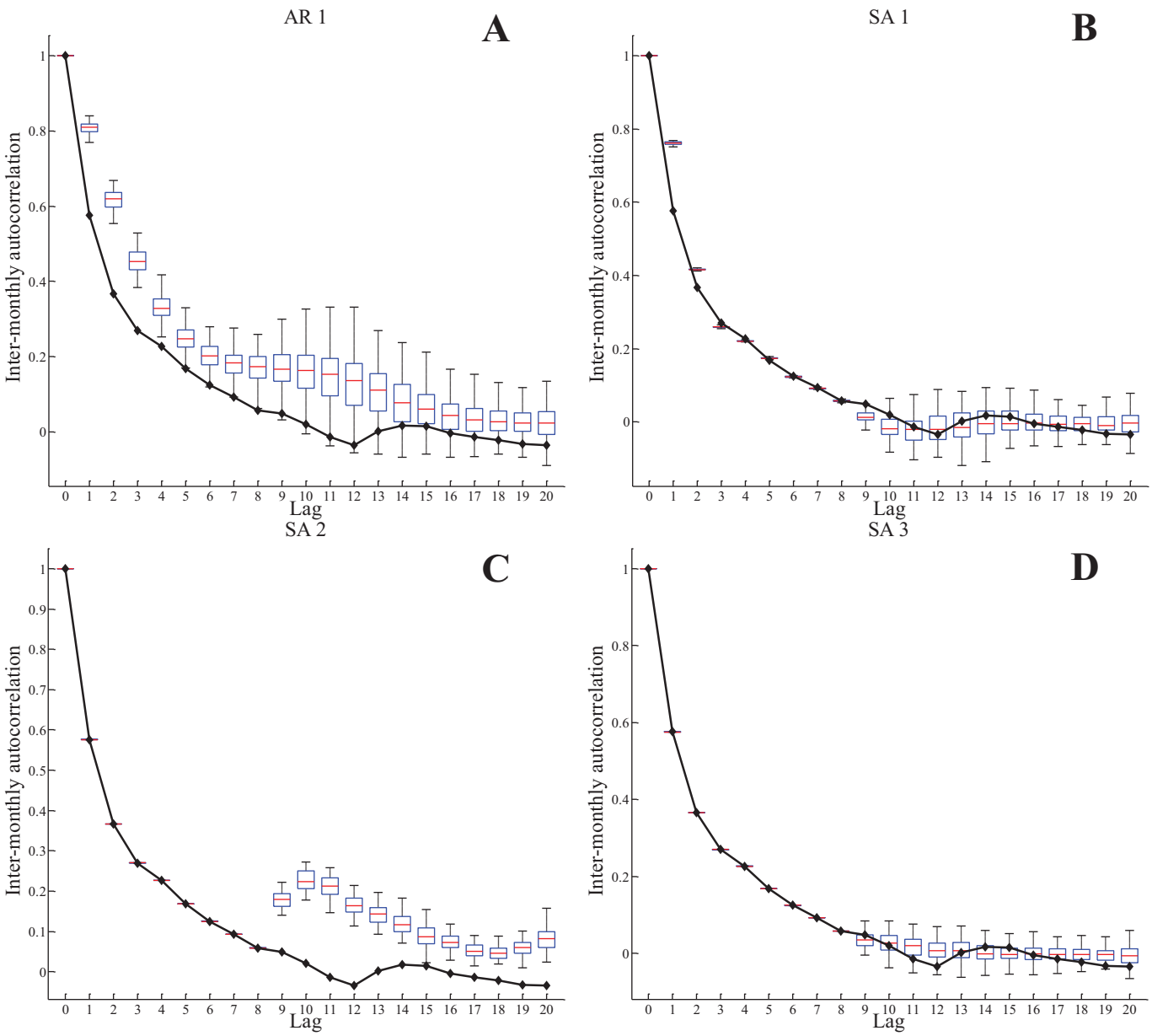


**Figure 14.** Box-plots of the monthly autocorrelation function for 100 realizations generated by perturbing (a) the intermonthly lag-1 autocorrelation coefficient in the AR 1 process, (b) the intermonthly lag-1 autocorrelation coefficient in the SA, (c) the annual standard deviation, and (d) the mean monthly flows of the summer months. The black lines show the same statistics for the observed data.

In the AR(1) process, when the intermonthly autocorrelation coefficient is increased, other important statistics of the flow series change as a result, particularly at the annual time scale. The standard deviation of the annual flows increases (Figure 16b) and so does the spread of annual mean values (Figure 16a). The proposed streamflow generator, on the other hand, is capable of preserving the desired interannual levels of variability (i.e., the same standard deviation of historical annual flows) and persistence (i.e., the same interannual autocorrelation) while also increasing the short-term persistence. Compared to traditional AR processes, this method has the advantage of allowing for the perturbation of some key properties of the streamflow time series without altering some other properties of the time series. This overcomes the known limitation encountered when trying to preserve statistics at both the monthly and annual time scales using traditional streamflow simulation models.

Figure 16 shows the results from the second example (SA 2) where the standard deviation of the annual totals was perturbed but the autocorrelation coefficients at both monthly and annual time scales were kept

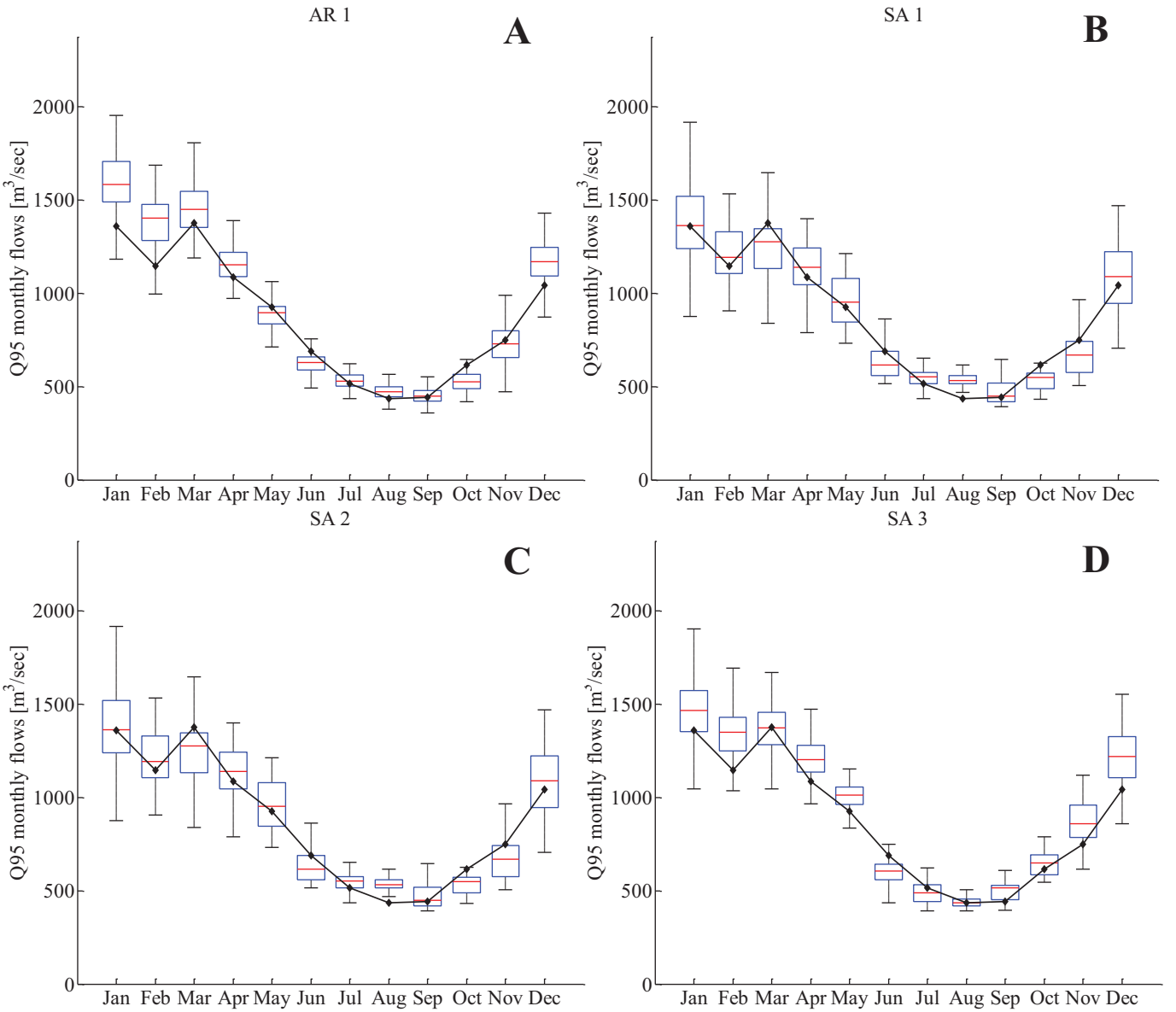


**Figure 15.** Box-plots of the monthly Q95 for 100 realizations generated by perturbing (a) the intermonthly lag-1 autocorrelation coefficient in the AR 1 process, (b) the intermonthly lag-1 autocorrelation coefficient in the SA, (c) the annual standard deviation, and (d) the mean monthly flows of the summer months. The black lines show the same statistics for the observed data.

constant. A similar perturbation would be difficult to perform using traditional autoregressive processes because with these methods statistics at the annual and monthly time scales cannot be controlled at the same time.

To validate the streamflow generation method further and explore the effects of perturbing some statistics on the drought properties of the streamflow sequence, we compute four basic drought statistics: the average drought length, the maximum drought length, the average drought deficit and the maximum drought deficit. A drought event is defined using a threshold method [*Hisdal et al.*, 2004], where the threshold is arbitrarily set to the Q75 of the annual observed flows. A drought event starts when the simulated flows fall below this threshold and terminates when flows rise above it. The drought deficit is calculated as the cumulative deviation between the streamflow and the threshold, whereas the drought duration is the number of years spent below the threshold. The average length and deficit values are calculated for each 130 year long simulated time series, for each of the 100 synthetic simulations. The maximum length and deficit are calculated similarly, and correspond to the longest and most severe drought event in each simulation.

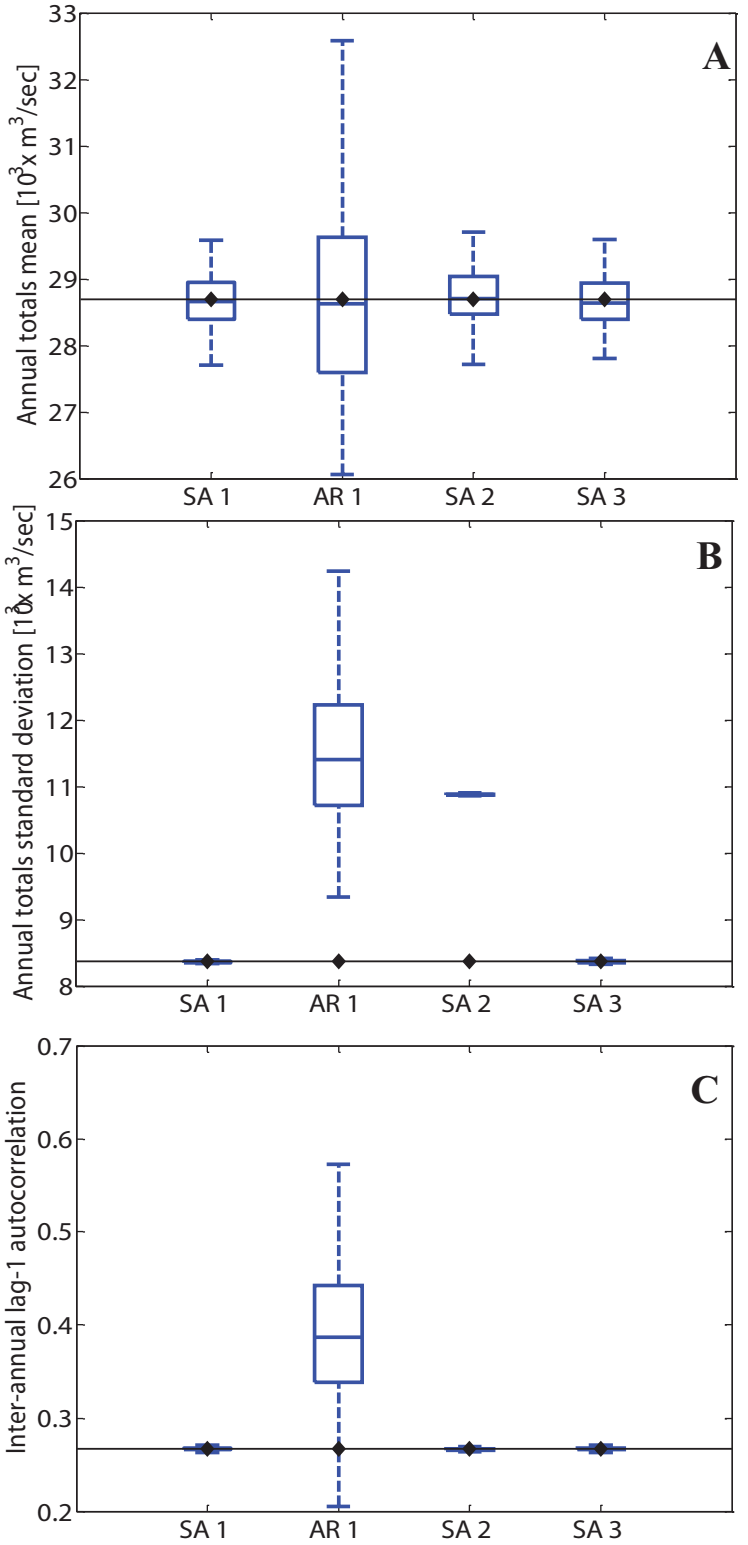


**Figure 16.** Box-plots of the (a) mean, (b) standard deviation, and (c) interannual lag-1 autocorrelation coefficient for annual flows generated by perturbing the intermonthly persistence (SA1 and AR1), the interannual variability (SA 2) and the mean monthly flows of summer months (SA 3). The black lines show the same statistics for the observed data.

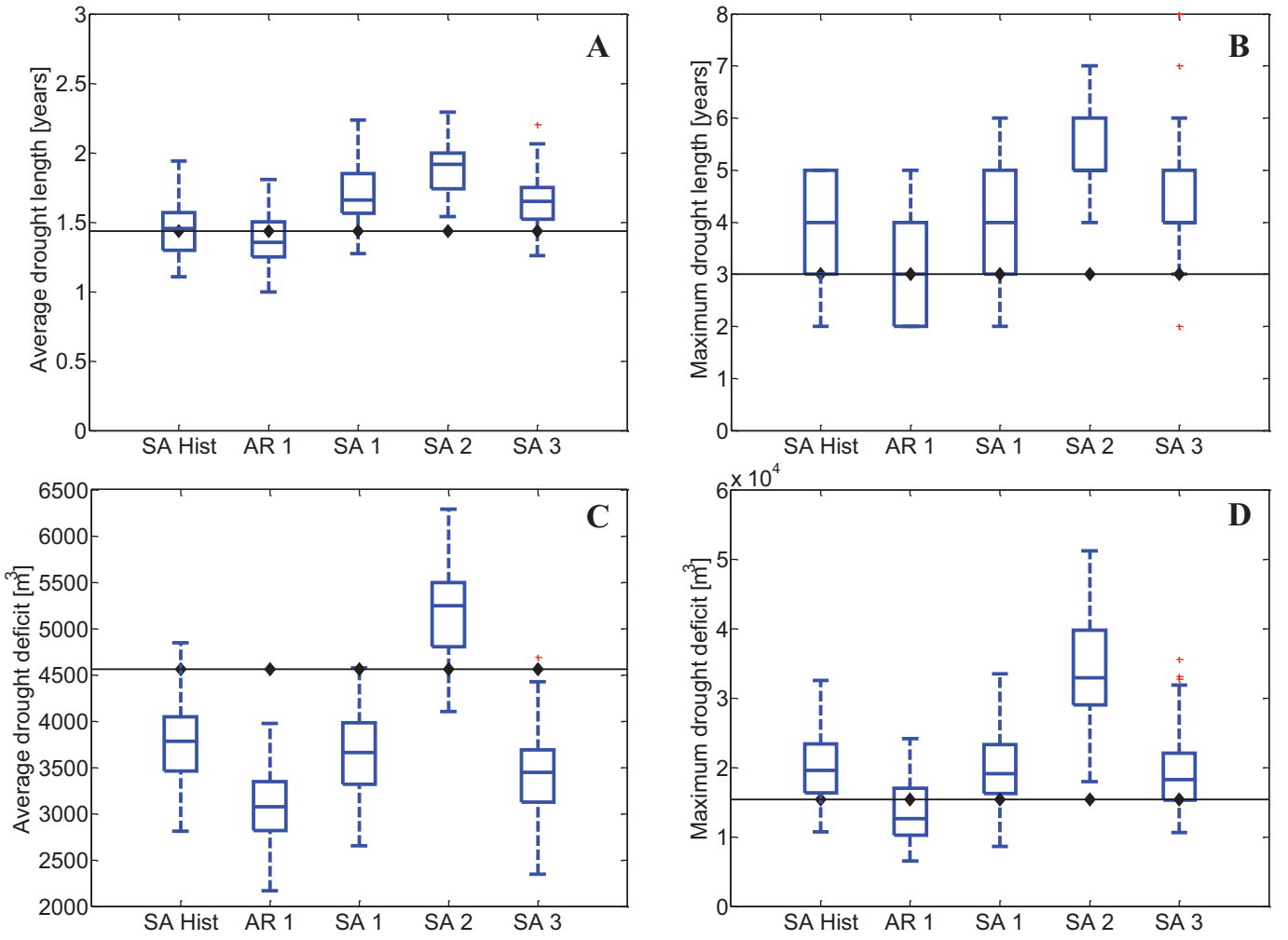


**Figure 17.** Box-plots of (a) average and (b) maximum drought length and (c) average and (d) maximum deficit for the annual totals of the 100 realizations for each application. The horizontal black lines show the same statistics for the observed data.

Figures 17a and 17b show the average drought length and maximum drought length calculated for the sequences generated with the unperturbed AR(1) model and simulated annealing model (SA Hist) and for the perturbed sequences SA 1, SA 2, and SA 3. The AR(1) shows a slight underestimation of the observed average drought length (solid black line), while the synthetic sequences generated with simulated annealing show a good match (SA Hist). Altering the intermonthly autocorrelation lag-1 coefficient (SA 1) and the interannual variability (SA 2) leads to sequences with greater average drought lengths (Figure 17a) and also longer droughts in absolute terms (Figure 17b). Similarly, reducing the summer monthly means results in an increase in the average drought length in the sequences (Figure 17a).

Figures 17c and 17d show the average and maximum drought deficits, respectively. The historical model SA Hist shows a slight underestimation of the average drought deficit as does the AR(1) process. Perturbing the intermonthly lag-1 autocorrelation coefficient (SA 1) and the summer monthly means (SA 3) result in longer drought events on average (Figure 17a) with smaller deficits (Figure 17c). On the other hand, altering the interannual variability (SA 2 example) increases both the drought duration and deficit characteristics of the sequence, providing water managers with a tool to test their systems' vulnerability to multiyear droughts.

## 4. Summary and Conclusions

We have presented a synthetic streamflow generation method for use in water resources vulnerability assessments. The method is nonparametric and data driven, using a simulated annealing algorithm to impose desired statistical properties on the synthetic streamflow sequence. The application of the method

to the River Thames using monthly observed data shows that it is capable of reproducing relevant streamflow statistics. It was shown that the method can be used to generate synthetic streamflow sequences matching both statistics of the observed data and perturbed statistics representing possible climate change induced changes in streamflow characteristics.

This method offers an alternative to propagating synthetic weather sequences (from climate models or weather generators) through rainfall-runoff models. It provides flexible inputs to bottom-up and vulnerability-based assessments to test water resource system sensitivity to changes in specific streamflow characteristics [e.g., *Nazemi and Wheater*, 2014]. Compared to weather generator and hydrological model based approaches for generating climate-perturbed streamflow sequences [e.g., *Manning et al.*, 2009], this method allows for greater flexibility in controlling which specific properties of the streamflow distribution are adjusted to simulate the effects of climate change. Furthermore, the sequences generated with this method are not conditional on any hydrological model parameterization or structure. As our method seeks to reproduce (and perturb) observed flows directly, it does not require translation from naturalized to observed flows. This approach thus allows water resources managers to focus on the impacts of changing streamflow characteristics rather than on the different sources of uncertainty associated with hydrological models.

The method has several advantages over more traditional parametric synthetic streamflow generation techniques. First, our approach avoids a priori selection of a stochastic model and, as with other nonparametric approaches [e.g., *Sharma et al.*, 1997], it reproduces the temporal dependence structure of the streamflow time series without requiring any assumptions about distributions and linearity. Our method shifts the hydrologist's attention from fitting model parameters to the statistical properties that are considered important and that need to be included in the objective function. Second, we demonstrated how our approach allows one to manipulate specific streamflow properties while keeping some other properties fixed. Third, our method allows one to preserve seasonal and annual statistics without the need to resort to disaggregation procedures.

Further research is needed to extend the method to multisite synthetic streamflow generation, where spatial autocorrelation statistics between different sites could be inserted in the objective function to generate spatially coherent streamflow sequences. Future work will also explore the role of covariates (e.g., the North Atlantic Oscillation) in controlling streamflow characteristics, the multiobjective trade-offs between different properties specified in the objective function and approaches to reduce the time complexity of the proposed method.

**Acknowledgments**

The research described in this paper was supported by the Oxford Martin Programme on Resource Stewardship. Edoardo Borgomeo is funded by the Engineering and Physical Sciences Research Council, the Environment Agency (science project SC120053), and Thames Water. This work has also been supported by the Natural Environment Research Council (Consortium on Risk in the Environment: Diagnostics, Integration, Benchmarking, Learning and Elicitation (CREDIBLE) grant NE/J017302/1). We would like to thank Francesca Pianosi, Glenn Watts, Chris Lambert, and Keith Colquhoun for comments on an earlier version of this manuscript and Michael Simpson, Raghav Pant, and Mohammad Mortazavi for useful suggestions. We thank the associate editor and three reviewers for their comments that helped to improve the article. The streamflow data for the Thames at Kingston used in this study can be obtained from the National River Flow Archive (http://www.ceh.ac.uk/data/nrfa/data/station.html?39001).

## References

Addor, N., O. Rössler, N. Köplin, M. Huss, R. Weingartner, and J. Seibert (2014), Robust changes and sources of uncertainty in the projected hydrological regimes of Swiss catchments, *Water Resour. Res.*, *50*, 7541–7562, doi:10.1002/2014WR015549.

Bardossy, A. (1998), Generating precipitation time series using simulated annealing, *Water Resour. Res.*, *34*(7), 1737–1744.

Blum, C., and A. Roli (2003), Metaheuristics in combinatorial optimization: Overview and conceptual comparison, *ACM Comput. Surv.*, *35*(3), 268–308.

Brown, C., and R. L. Wilby (2012), An alternate approach to assessing climate risks, *Eos Trans. AGU*, *93*(41), 401.

Brown, C., Y. Ghile, M. Laverty, and K. Li (2012), Decision scaling: Linking bottom-up vulnerability analysis with climate projections in the water sector, *Water Resour. Res.*, *48*, W09537, doi:10.1029/2011WR011212.

Burton, A., H. J. Fowler, C. G. Kilsby, and P. E. O'Connell (2010a), A stochastic model for the spatial-temporal simulation of nonhomogeneous rainfall occurrence and amounts, *Water Resour. Res.*, *46*, W11501, doi:10.1029/2009WR008884.

Burton, A., H. J. Fowler, S. Blenkinsop, and C. G. Kilsby (2010b), Downscaling transient climate change using a Neyman-Scott Rectangular Pulses stochastic rainfall model, *J. Hydrol.*, *381*(1–2), 18–32.

Christensen, N. S., A. W. Wood, N. Voisin, D. P. Lettenmaier, and R. N. Palmer (2004), The effects of climate change on the hydrology and water resources of the Colorado river basin, *Clim. Change*, *62*, 337–363.

Cunha, M., and J. Sousa (1999), Water distribution network design optimization: Simulated annealing approach, *J. Water Resour. Plann. Manage.*, *125*(4), 215–221.

Deutsch, C., and P. Cockerham (1994), Practical considerations in the application of simulated annealing to stochastic simulation, *Math. Geol.*, *26*(1), 67–82.

Diaz-Nieto, J., and R. Wilby (2005), A comparison of statistical downscaling and climate change factor methods: Impacts on low flows in the River Thames, United Kingdom, *Clim. Change*, *69*(2–3), 245–268, doi:10.1007/s10584-005-1157-6.

Dougherty, D. E., and R. A. Marryott (1991), Optimal groundwater management. 1: Simulated annealing, *Water Resour. Res.*, *27*(10), 2493–2508.

Eglese, R. W. (1990), Simulated annealing: A tool for operational research, *Eur. J. Oper. Res.*, *46*, 271–281.

Environment Agency (2008), *Water Resources in England and Wales—Current State and Future Pressures*, Bristol, U. K.

Farmer, C. L. (1992), *Numerical rocks, in* The Mathematics of Oil Recovery *(Based on the Proceedings of a conference on the Mathematics of Oil Recovery held at Robinson College, Cambridge, July 1989)*, edited by P. R. King, pp. 437–447, Oxford Univ. Press, Oxford.

Fiering, M. B., and B. B. Jackson (1971), Synthetic Streamflows, Water Resour. Monogr. Ser., vol. 1, 89 pp., AGU, Washington, D. C.

Hao, Z., and V. P. Singh (2011), Single-site monthly streamflow simulation using entropy theory, *Water Resour. Res.*, *47*, W09528, doi:10.1029/2010WR010208.

Hao, Z. and V. P. Singh (2012), Entropy-copula method for single-site monthly streamflow simulation, *Water Resour. Res.*, *48*, W06604, doi: 10.1029/2011WR011419.
Hao, Z., and V. P. Singh (2013), Modeling multisite streamflow dependence with maximum entropy copula, *Water Resour. Res.*, *49*, 7139–7143, doi:10.1002/wrcr.20523.
Hirsch, R. M. (1979), Synthetic hydrology and water supply reliability, *Water Resour. Res.*, *15*(6), 1603–1615.
Hisdal, H., L. M. Tallaksen, B. Clausen, E. Peters, and A. Gustard (2004), Hydrological drought characteristics, in Hydrological Drought—Processes and Estimation Methods for Streamflow and Groundwater, Dev. in Water Sci., vol. 48, edited by L. M. Tallaksen and H. A. J. van Lanen, pp. 139–198, Elsevier Sci., Netherlands.
Jackson, B. B. (1975), The use of streamflow models in planning, *Water Resour. Res.*, *11*(1), 54–63.
Kannan, S., S. Mary Raja Slochanal, and N. Padhy (2005), Application and comparison of metaheuristic techniques to generation expansion planning problem, *IEEE Trans. Power Syst.*, *20*(1), 466–475.
Keylock, C. J. (2012), A resampling method for generating synthetic hydrological time series with preservation of cross-correlative structure and higher-order properties, *Water Resour. Res.*, *48*, W12521, doi:10.1029/2012WR011923.
Kilsby, C. G., P. D. Jones, A. Burton, A. C. Ford, H. J. Fowler, C. Harpham, P. James, A. Smith, and R. L. Wilby (2007), A daily weather generator for use in climate change studies, *Environ. Modell. Software*, *22*(12), 1705–1719, doi:10.1016/j.envsoft.2007.02.005.
Kirkpatrick, S., C. Gelatt Jr., and M. Vecchi (1983), Optimization by simulated annealing, *Science*, *220*(4598), 671–680.
Kwon, H.-H., U. Lall, and A. F. Khalil (2007), Stochastic simulation model for nonstationary time series using an autoregressive wavelet decomposition: Applications to rainfall and temperature, *Water Resour. Res.*, *43*, W05407, doi:10.1029/2006WR005258.
Lall, U., and A. Sharma (1996), A nearest neighbor bootstrap for time series resampling, *Water Resour. Res.*, *32*(3), 679–693.
Lee, T., and J. Salas (2011), Copula-based stochastic simulation of hydrological data applied to Nile River flows, *Hydrol. Res.*, *42*(4), 318–330.
Lee, T., and T. Ouarda (2012), Stochastic simulation of nonstationary oscillation hydroclimatic processes using empirical mode decomposition, *Water Resour. Res.*, *48*, W02514, doi:10.1029/2011WR010660.
Loucks, D. P., and E. van Beek (2005), Water Resources System Planning and Management: An Introduction to Methods, Models and Applications, U. N. Educ., Sci. and Cult. Organ., Paris.
Manning, L. J., J. W. Hall, H. J. Fowler, C. G. Kilsby, and C. Tebaldi (2009), Using probabilistic climate change information from a multimodel ensemble for water resources assessment, *Water Resour. Res.*, *45*, W11411, doi:10.1029/2007WR006674.
Matalas, N. C. (1967), Mathematical assessment of synthetic hydrology, *Water Resour. Res.*, *3*(4), 937–945.
McLeod, A. I., K. W. Hipel, and W. C. Lennox (1977), Advances in Box-Jenkins modelling. 2: Application, *Water Resour. Res.*, *13*(3), 577–586.
Metropolis, N., A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller (1953), Equation of state calculations by fast computing machines, *J. Chem. Phys.*, *21*(6), 1087–1092.
Nazemi, A., and H. S. Wheater (2014), Assessing the vulnerability of water supply to changing streamflow conditions, *Eos Trans. AGU*, *95*(32), 288.
Nazemi, A., H. S. Wheater, K. P. Chun, and A. Elshorbagy (2013), A stochastic reconstruction framework for analysis of water resource system vulnerability to climate-induced changes in river flow regime, *Water Resour. Res.*, *49*, 291–305, doi:10.1029/2012WR012755.
Parks, K. P., L. R. Bentley, and A. S. Crowe (2000), Capturing geological realism in stochastic simulations of rock systems with Markov statistics and simulated annealing, *J. Sediment. Res.*, *70*, 803–813.
Pereira, M. V. F., G. C. Oliveira, C. C. G. Costa, and J. Kelman (1984), Stochastic streamflow models for hydroelectric systems, *Water Resour. Res.*, *20*(3), 379–390.
Prairie, J., K. Nowak, B. Rajagopalan, U. Lall, and T. Fulp (2008), A stochastic nonparametric approach for streamflow generation combining observational and paleoreconstructed data, *Water Resour. Res.*, *44*, W06423, doi:10.1029/2007WR006684.
Prairie, J. R., B. Rajagopalan, T. J. Fulp, and E. A. Zagona (2006), Modified K-NN model for stochastic streamflow simulation, *J. Hydrol. Eng.*, *11*(4), 371–378.
Prudhomme, C., R. L. Wilby, S. Crooks, A. L. Kay, and N. S. Reynard (2010), Scenario-neutral approach to climate change impact studies: Application to flood risk, *J. Hydrol.*, *390*, 198–209.
Prudhomme, C., A. Young, G. Watts, T. Haxton, S. Crooks, J. Williamson, H. Davies, S. Dadson, and S. Allen (2012), The drying up of Britain? A national estimate of changes in seasonal river flows from 11 Regional Climate Model simulations, *Hydrol. Processes*, *26*, 1115–1118, doi: 10.1002/hyp.8434.
Rajagopalan, B., J. Salas, and U. Lall (2010), *Stochastic methods for modeling precipitation and streamflow, in* Advances in Data-Based Approaches for Hydrologic Modeling and Forecasting, edited by B. Sivakumar and R. Berndtsson, pp. 17–52, World Sci., Singapore.
Rocheta, E., M. Sugiyanto, F. Johnson, J. Evans, and A. Sharma (2014), How well do general circulation models represent low-frequency rainfall variability?, *Water Resour. Res.*, *50*, 2108–2123, doi:10.1002/2012WR013085.
Salas, J. D., and J. T. B. Obeysekera (1982), ARMA Model identification of hydrologic time series, *Water Resour. Res.*, *18*(4), 1011–1021, doi: 10.1029/WR018i004p01011.
Salas, J. D., C. Fu., A. Cancelliere, D. Dustin, D. Bode, A. Pineda, and E. Vincent (2005), Characterizing the severity and risk of drought in the Poudre river, Colorado, *J. Water Resour. Plann. Manage.*, *131*(5), 383–393.
Sharma, A., D. G. Tarboton, and U. Lall (1997), Streamflow simulation: A nonparametric approach, *Water Resour. Res.*, *33*(2), 291–308.
Simon, D. (2013), *Evolutionary Optimization Algorithms*, John Wiley, Hoboken, N. J.
Singh, R., T. Wagener, R. Crane, M. Mann, and L. Ning (2014), A vulnerability driven approach to identify adverse climate and land use combinations for critical hydrologic indicator thresholds: Application to a watershed in Pennsylvania, USA, *Water Resour. Res.*, *50*, 3409–3427, doi:10.1002/2013WR014988.
Srivastav, R. K., and S. P. Simonovic (2014), An analytical procedure for multi-site, multi-season streamflow generation using maximum entropy boostrapping, *Environ. Modell. Software*, *59*, 58–75.
Stagge, J. H., and G. E. Moglen (2013), A nonparametric stochastic method for generating daily climate-adjusted streamflows, *Water Resour. Res.*, *49*, 6179–6193, doi:10.1002/wrcr.20448.
Stedinger, J. R., and M. R. Taylor (1982), Synthetic stream flow generation. 1: Model verification and validation, *Water Resour. Res.*, *18*(4), 909–918.
Steinschneider, S., and C. Brown (2013), A semiparametric multivariate, multisite weather generator with low-frequency variability for use in climate risk assessments, *Water Resour. Res.*, *49*, 7205–7220, doi:10.1002/wrcr.20528.
Sudler, C. E. (1927), Storage required for the regulation of stream flow, *Trans. Am. Soc. Civ. Eng.*, *91*, 622–660.
Thomas, H. A., and M. B. Fiering (1962), Mathematical synthesis of stream flow sequences for the analysis of river basin simulation, in *Design of Water Resources Systems*, edited by A. Maas et al., Harvard Univ. Press, Cambridge, Mass.

Thyer, M., G. Kuczera, and B. C. Bates (1999), Probabilistic optimization for conceptual rainfall-runoff models: A comparison of the shuffled complex evolution and simulated annealing algorithms, *Water Resour. Res.*, *35*(3), 767–773.

Turner, S. W. D., D. Marlow, M. Ekström, B. G. Rhodes, U. Kularathna, and P. J. Jeffrey (2014), Linking climate projections to performance: A yield-based decision scaling assessment of a large urban water resources system, *Water Resour. Res.*, *50*, 3553–3567, doi:10.1002/2013WR015156.

van Laarhoven, P. J. M., and E. H. L. Aarts (1987), *Simulated Annealing: Theory and Applications*, D. Reidel, Amsterdam, Netherlands.

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

Vogel, R. M., and J. R. Stedinger (1988), The value of stochastic streamflow models in over year reservoir design application, *Water Resour. Res.*, *24*(9), 1483–1490.