ORIGINAL PAPER

Influence of output size of stochastic weather generators on common climate and hydrological statistical indices

Abdullah Alodah[1,2] · Ousmane Seidou[2,3]

Abstract

While Stochastic Weather Generators (SWGs) are used intensively in climate and hydrological applications to simulate hydroclimatic time series and estimate risks and performance measures linked to climate variability, there have been few investigations into how many realizations are required for a robust estimation of these measures. Given the computational cost and time necessary to force climate-sensitive systems with multiple realizations, the estimation of the optimal number of synthetic time series to generate with a particular SWG for a predefined accuracy when estimating a particular risk or performance measure is particularly important. In this paper, the required number of realizations of five SWGs coupled with a SWAT model (the Soil and Water Assessment Tool) needed in order to achieve a predefined Relative Root Mean Square Error is investigated. The statistical indices used are the mean, standard deviation, skewness, and kurtosis of four hydroclimatic variables: precipitation, maximum and minimum temperature, and annual streamflow obtained for each observed and model-generated time series. While the results vary somewhat across SWGs, variables and indicators, they overall show that the marginal improvement decreases dramatically after 25 realizations. The results also indicate that the benefit of generating more than 100 realizations of climate and streamflow data is very minimal. The methodology presented herein can be applied in further investigations of other set of risk indicators, SWGs, hydrological models, and watersheds to minimize the required workload.

Keywords Stochastic weather generators · Stochastic hydrological modeling · Hydrometeorology · Hydrological risk assessment · Climate ensemble · Climate sensitivity · Climate realizations · Hydrological realizations

1 Introduction

Stochastic weather Generators (referred to as SWGs hereafter) are numerical tools employed broadly to simulate the statistical characteristics of observed climate variables and generate random time series that can be used as inputs for climate-sensitive hydrological models (Wheater et al. 2005). The variability in the input translates into variability in the generated hydrological time series. The risk associated with and performance of the modeled water system are assessed by estimating the statistics for the simulated variables. The use of SWG outputs in such studies is convenient, as SWGs can generate long and gap-free synthetic sequences based on historical observations and can be used for water resources planning and management (Vu et al. 2018). A large ensemble of synthetic weather sequences (or realizations) is assumed to represent the internal variability of hydroclimatic variables, consisting mostly of precipitation, maximum temperature, minimum temperature, solar radiation, and relative humidity (Santer et al. 2008) at different spatial and temporal scales (Ailliot et al. 2015).

According to Guenni (1994), SWGs are mainly useful in: (1) extending insufficient or incomplete records that

✉ Abdullah Alodah
a.alodah@qec.edu.sa

1   Department of Civil Engineering, College of Engineering, Qassim University, Buraydah Al Qassim 51431, Saudi Arabia

2   Department of Civil Engineering, Faculty of Engineering, University of Ottawa, 161 Louis Pasteur, Ottawa, ON K1N 6N5, Canada

3   United Nations University Institute for Water, Environment and Health, 204-175 Longwood Rd S, Hamilton, ON L8P 0A1, Canada

constrain the modeling approach (e.g., Fodor et al. 2013; Fatichi et al. 2016), (2) developing datasets for ungauged sites via spatially interpolating model parameters from adjacent areas with sufficient records (e.g., Baffault et al. 1996; Fodor et al. 2013), and, recently, (3) accounting for the uncertainty that arises from natural variability along with anthropogenic forcing in climate-change simulations (e.g., Räisänen and Ruokolainen 2006; Minville et al. 2008; Deser et al. 2012; Thompson et al. 2015). Ailliot et al. (2015) classified SWGs in into four group according to the random number generation process: resampling techniques (e.g., Räisänen and Ruokolainen 2006; Oriani et al. 2014), multivariate autoregressive models applying Box-Jenkins Method (e.g., Box and Jenkins 1976), point process models (e.g., Rodriguez-Iturbe et al. 1987; Onof et al. 2000) and Bayesian hierarchical modeling such as weather type models (e.g., Thompson et al. 2007).

SWGs were introduced initially for hydrological applications requiring long sequences of daily weather data (e.g., Gabriel and Neumann 1962; Todorovic and Woolhiser 1975; Buishand 1977). Since then, SWGs have found wide application in various hydrologic investigations, such as the assessment of anthropogenic climate change impacts (e.g., Zwiers 1996; Eames et al. 2012; Kilsby et al. 2007; Candela et al. 2012), crop yield estimates (e.g., Vesely et al. 2019), ecosystem and food security models (e.g., Stevens and Madani 2016), and in streamflow simulations (e.g., Zhang and Garbrecht 2003; Dubrovský et al. 2004; Alodah and Seidou 2019a) mainly to characterize internal atmospheric variability (or climate noise) (Räisänen and Ruokolainen 2006; Santer et al. 2008; Deser et al. 2012) and particularly under conditions of data scarcity (Breinl et al. 2017). The use of observed climate data in hydrological modeling is always preferable; however, SWGs provide a suitable alternative, as some localized risky events that are not covered fully in the observed set may be overlooked (Räisänen and Ruokolainen 2006; Ivanov et al. 2007; Santer et al. 2008; Vu et al. 2018).

Several authors investigated the abilities of SWGs in representing the statistical properties of observed weather series (e.g., Semenov et al. 1998; Hayhoe 1998, 2000; Qian et al. 2004; Ivanov et al. 2007; Chen et al. 2014; Ailliot et al. 2015; Breinl et al. 2017; Mehan et al. 2017; Gitau et al. 2018; Vesely et al. 2019). Well-known limitations of SWGs include their inability to generate low-frequency variability very well (Soltani and Hoogenboom 2003a) or correctly reproduce the dependence of temperature variables and precipitation amount in wet days on wet/dry spell lengths determinant parameters (Wilby et al. 2004). Alternatively, Panagoulia (2006) showed the great potential of artificial neural network (ANN) models in simulating nonlinear processes of extreme river flows in various climates conditions. The ability of ANN in flow

simulations was also proven even at the level of detailed localized studies taking into consideration the appropriate selection of input variables (Panagoulia et al. 2017).

SWGs are often employed to study the impacts of climatic variability, for instance, in rainfall-runoff simulations (e.g., Dubrovský et al. 2004; Panagoulia 2006), erosion simulations (e.g., Zhang and Garbrecht 2003), simulations of extreme precipitation events (e.g., Furrer and Katz 2008; Semenov 2008), and in climate-change studies (e.g., Kilsby et al. 2007; Kim et al. 2007; Al-Mukhtar et al. 2014; Alodah and Seidou 2019b). Yet, unlike observed weather data which provide only one realization, an unlimited number of weather realizations can be generated (Kim et al. 2018; Vu et al. 2018), and it is very improbable statistically that any two realizations will be identical (i.e., uncorrelated data from a realization to the next one). In general, multiple stochastically-generated time series can provide a broad range of weather possibilities for a detailed sensitivity analysis (Dubrovský et al. 2004; Santer et al. 2008), such as the recently introduced vulnerability-based methods (e.g., bottom-up approaches) for evaluating uncertainty in projected climate change impacts (e.g., Brown et al. 2011; Steinschneider and Brown 2013; Mukundan et al. 2019; Alodah and Seidou 2019b). An ensemble of multiple realizations is recommended in order to characterize the variability in climate data adequately and estimate realistic mean values and variances of meteorological variables (Alodah and Seidou 2019a; Guo et al. 2018; Mehrotra et al. 2006; Anyah and Semazzi 2006; Dubrovský et al. 2004).

Multiple realizations of climate series are increasingly becoming the adopted modeling approach when evaluating the variability of complex climate systems to account for rare occurrences of climate variables (Anyah and Semazzi 2006). Typically, an arbitrary (and commonly limited) number of realizations (ranged from 5 to 1000) is used. Examples of some recent publications utilizing multiple runs of weather generators are presented in Table 1. It is also common to use SWGs to produce a time series that is longer than observed ones (e.g., Kou et al. 2007; Caron et al. 2008; Chen et al. 2012; Eames et al. 2012), although this might lead to biases due to an insufficient sampling of the distribution (Mithen and Black 2011). Therefore, it is recommended that multiple realizations with the same lengths as the training set be used (Dubrovský et al. 2004; Guo et al. 2018). However, the use of multiple realizations requires high-performance computational resources, especially when used in conjunction with a complex impact model. For example, Gitau et al. (2012) analyzed 172 management scenarios and ran a SWAT model 250 times for each of them, for a total of 43,000 runs, using an extremely large computing Condor framework. However, they stated that their work could have taken up to 3.3 years

Table 1 Examples of ensembles of multiple realizations used in the prior hydrological investigations

Stochastic weather generator(s)	Number of realizations	Time series length	Impact models?	References
Met and Roll	30	30-year	SAC-SMA	Dubrovský et al. (2004)
WEGN and WeaGETS	30	300-year	HSAMI	Caron et al. (2008)
MNHMM	100	22-year	Sacramento	Kwon et al. (2011)
WGEN	100	30-year	Several conceptual models	Bastola et al. (2012)
Unique SWG	200	25-year	HEC-HMS and VisualBALAN	Candela et al. (2012)
WXGEN	250	28-year	SWAT	Gitau et al. (2012)
WeaGETS, MulGETS, and KNN	1000	41-year	SWAT	Alodah and Seidou (2019a)
MulGETS	250	30-year	SWAT	Alodah and Seidou (2019b)
WXGEN	10	30-year	Socio-hydrologic model	Han et al. (2019)
Unique Stochastic Model	1000	42-year	Crop simulation models	Hansen and Ines (2005)
WGEN, SIMMETEO, and SWG	100	71-year	Crop simulation models	Apipattanavis et al. (2010)
AR(1) model	20	50-year	Climate only	Alhassoun et al. (1997)
CLIGEN	10	20-year	Climate only	Elliot and Arnold (2001)
HMM, KNN, and Wilks	100	43-year	Climate only	Mehrotra et al. (2006)
Simple resampling method	20	30-year	Climate only	Räisänen and Ruokolainen (2006)
MMLR	50	40-year	Climate only	Jeong et al. (2012)
Improved KNN	50	62-year	Climate only	Steinschneider and Brown (2013)
CLIGEN, LARS-WG, and WeaGETS	50	50-year	Climate only	Mehan et al. (2017)
TripleM	30	30-year	Climate only	Breinl et al. (2017)
AWE-GEN-2d	50	30-year	Climate only	Peleg et al. (2017)
LARS-WG	50	50-year	Climate only	Gitau et al. (2018)
CLIGEN, LARS-WG, and WeaGETS	100	50-year	Climate only	Guo et al. (2018)
AWE-GEN	100	30-year	Climate only	Kim et al. (2018)
Parametric and Non- parametric SWGs	50	17-year	Climate only	Benoit et al. (2019)

to complete via a traditional desktop computer workstation. Thus, given the acknowledged limitations imposed by time and computational expenses, the question related to the required number of realizations to fairly characterize the hydrological space is still open.

This prolonged process, particularly for large watersheds, may be overcome with the help of expensive supercomputers or by identifying a sufficiently representative number of outputs needed to capture the random component of the hydrological model and ultimately reduce the computations. Guo et al. (2018) investigated the numbers of realizations necessary for capturing several statistical characteristics of meteorological variables satisfactorily (i.e., for precipitation and minimum and maximum temperature) generated synthetically by CLIGEN, LARSWG, and WeaGETS. They analyzed increasing

discrete numbers of realizations (1, 25, 50, and 100) and concluded that a weather generator would reproduce essential climate characteristics well by 25 realizations. The current work generally builds on their ideas. However, the statistics considered in their work belong to the *climatic data space* only (precipitation and temperature variables); thus, their findings may not be applicable for hydrological variables, especially due to the non-linearity of the hydrologic response in rainfall-runoff transformations.

Frequently, synthetically generated climate sequences are fed to hydrological models and used after that to examine some risk spaces. This study aims to analyze how the accuracy of the estimates of key statistics evolves with the number of realizations of SWGs. Five SWGs were used to generate ensembles of daily precipitation occurrences and amounts (PCP) and daily maximum (Tmax) and

minimum (Tmin) temperatures coupled with a hydrological model (SWAT) to simulate streamflow. A variety of diagnostic tools were then applied to identify the optimal number of realizations needed for both the climatic and hydrologic variables.

2 Materials and methods

2.1 Study area and available hydro-climatic data

The study area is the South Nation Watershed (SNW), located in Eastern Ontario, Canada. The SNW has a relatively flat area of about 4000 km^2 between 74°22′ to 75°43′ W longitude and 44°40′ to 45°38′ N latitude. The watershed is drained by the South Nation River, which runs northeast for 175 km towards Plantagenet, with a low topographic gradient of only 80 m between its headwaters and the confluence with the Ottawa River. This characteristic maximizes the flood risk and boosts the erosion of riverbanks and agricultural topsoil. The reader is referred to Alodah (2015, 2019) for detailed descriptions of the study area. Climate data were collected for a 41-year period, based on the availability and consistency of the observed data, between 1971 and 2011 at four metrological stations, namely, Russell Station (Climate Identifier (CI): 6107247, Latitude: 45° 15′ 46″N, Longitude: 75° 21′ 34″W, Elevation: 76.2 m), South Mountain Station (CI: 6107955, Latitude: 44° 58′ 00″N, Longitude: 75° 29′ 00″, Elevation: 84.7 m), Morrisburg Station (CI: 6105460, Latitude: 44° 55′ 25″N, Longitude: 75° 11′ 18″W, Elevation: 81.7 m), and St. Albert Station (CI: 6107276, Latitude: 45° 17′ 14″N, Longitude: 75° 03′ 49″W, Elevation: 80 m). In addition, the observed downstream daily discharge data was collected at the Plantagenet Gauging Station (ID: 02LB005, Latitude:45° 31′ 01″ N, Longitude: 74° 58′ 41″ W). There was no missing data in either dataset for the reference period. A detailed description of the observed hydroclimatic data has been presented previously in Alodah and Seidou (2019a).

2.2 Stochastic weather generators

The observed 41-year climate series for maximum air temperature, minimum air temperature, and precipitation from the four metrological stations were fed into five SWGs, namely, the WeaGETS implementing multi-Gamma (referred to as *WG* hereafter) and multi-Exponential (referred to as *WE* hereafter) distributions for wet-day sequences (Chen et al. 2012), MulGETS implementing multi-Gamma (referred to as *MG* hereafter) and multi-Exponential (referred to as *ME* hereafter) distributions for wet-day sequences (Chen et al. 2014), and the k-nearest

neighbor resampling models (Sharif and Burn 2007; Goyal et al. 2013). WeaGETS, a uni-site weather generator from the École de Technologie Supérieure (ÉTS), is a multivariate parametric model that simulates temperature variables conditional to each other based on a normal distribution and using first-order linear auto-regression coupled with constant lag-1 autocorrelation and cross-correlation. It also considers seasonal cycles with the help of Finite Fourier series with two harmonics. The MulGETS, a multi-site weather generator also from ÉTS, is an extension of WeaGETS and has the ability to take into account the spatial attributes of climate data, which is crucial in most hydrological models. For the simulation of a precipitation occurrence, MulGETS uses a two-state (dry or wet) first-order Markov chain with Cholesky factorization, whereas WeaGETS uses a third-order Markov model without parameter smoothing.

A higher-order Markov model is used in WeaGETS since it is recommended for better predicting long dry and wet spells (Bastola et al. 2012; Chen et al. 2012), whereas a first-order Markov chain is the only option in MulGETS. The order determination in the Markov chain models for the simulation of precipitation has been assessed by numerous studies (e.g., Schoof and Pryor 2008; Stowasser 2012). For instance, Schoof and Pryor (2008) examined Markov chains of order 0–3 to replicate monthly precipitation occurrence using the Bayesian information criteria (BIC) and found that the higher-order models performed better in simulating wet spells, while underperformed in the dry spell lengths. The inherent inadequacy of simulating the length of dry and wet spells by the exponential Markov approach proposed by Richardson (1981) was purportedly improved by the serial model- spell length approach (Racsko et al. 1991). Also, temperature variables and precipitation amount in wet days are conditional on wet/dry spell lengths determinant parameters (Wilby et al. 2004). Stowasser (2012), however, indicated that the improvement in producing precipitation statistics when using the theoretically best mixed-order model was minimal in comparison to simpler models.

Both models (WeaGETS and MulGETS) were used twice to simulate the daily wet-day precipitation sequences while implementing two probability distribution functions: a multi-Gamma distribution (a combination of several gamma distributions) and multi-Exponential distribution. The probability distribution functions (PDFs) of the Gamma and Exponential models are:

$$f_{Gamma}(\mathrm{x}) = \frac{(\mathrm{x}/\beta)^{\alpha-1}\exp[-\mathrm{x}/\beta]}{\beta\Gamma(\alpha)} \qquad (1)$$

$$f_{Exp}(\mathrm{x}) = \lambda e^{-\lambda x} \qquad (2)$$

The k-nearest neighbor resampling model (*KNN*) is a daily generator that applies different methodologies based on the nonparametric resampling of an observed climate dataset. Since it is nonparametric, KNN has the advantage of being able to generate unprecedented values in the historical period but within the sampled values. For further information, the reader is referred to a previous paper (Alodah and Seidou 2019a) for a full description of the configurations of the abovementioned stochastic models and their performances.

2.3 Rainfall-runoff model

The Soil and Water Assessment Tool (SWAT) is a well-known hydrological model that has been used widely for many applications, including the simulation of sediment and nutrient flow but mainly for streamflow simulations (Neitsch et al. 2011). SWAT is a semi-distributed watershed-scale model that relies on hydrologic response units (HRUs) of uniform land and climate characteristics. The SWAT model for this study was first calibrated and validated with the observed climate data using a daily time step and based on the Nash–Sutcliffe efficiency (NSE), the RMSE-observations standard deviation ratio (RSR), and the percent bias (PBIAS). Mehrotra et al. (2006) pointed out that care should be taken when applying NSE alone particularly for its dependence on the size of test samples. Alternatively, more than one metric should be considered (Criss and Winston 2008). The results of the calibration and validation of the model indicate a good fit between the observed and simulated flows (Metric: *Calibration*, *Validation*; NSE: *0.90*, *0.81*; RSR: *0.31*, *0.43*; PBIAS: *−10.0%*, *−8.3%*). The reader is referred to Alodah and Seidou (2019a) for an enhanced description of the SWAT configuration and parameter selection. Next, synthetic climate time series were fed independently into the SWAT model to generate synthetic daily streamflow time series. To examine the hydrological responses to various synthetic climate scenarios, all SWAT parameters were kept unchanged except for the climate input when synthetic climate time series replaced the observed ones, enabling the effect of climate variability on hydrological variables to be tracked.

2.4 Definitions and notations

For additional clarity, the definitions of some terms used herein are given below:

- *Realization* is a random output generated by running a SWG (climate) or the SWAT model with synthetic climate data (streamflow) for a number of years (a 41-year cycle herein), where all realizations are considered equally plausible for a given SWG (the terms "*realization*", "*run*", "iteration" and "scenario" are frequently interchangeable in the prior literature).

- *Cloud* refers to an ensemble of separately-generated realizations (i.e., one thousand herein) of synthetic climate (or streamflow) series accomplished by running a given SWG (coupled with SWAT) 1000 sperate times.

- *Sample* refers to a set of *N realizations* (for example a 10-realization sample), where this set with length *N* is reproduced randomly 10,000 times from the cloud.

To ease comprehension, the following notations are adopted:

- Index *s* goes from 1 to *S* and represent the metrological stations listed below:

  1. Russel
  2. South Mountain
  3. Morrisburg
  4. St. Albert

  where, *S* is the number of stations (4).

- T is the length in years of all climatic and hydrological time series (41-yr series).

- The observed climate and flow time series are denoted as:

  - $PCP_t^{obs,s}$, *t = 1,…,T; s = 1,..,S*, which represents the observed precipitation at time *t* at meteorological station *s*

  - $Tmax_t^{obs,s}$, *t = 1,…,T; s = 1,..,S*, which represents the observed maximum temperature at time *t* at meteorological station *s*

  - $Tmin_t^{obs,s}$, *t = 1,…,T; s = 1,..,S*, which represents the observed minimum temperature at time *t* at meteorological station *s*

  - Q_t^{obs}, *t = 1,…,T,* which represents the observed discharge (OBS Flow) at time *t* at the outlet for the SNW.

- The flow time series at the outlet for the SNW obtained by forcing the SWAT model using observed climate data, called *Simulated Flow using Observed Climate*, is denoted as:

  - $SFOC_t$, *t = 1,…,T.*

The following sections provide a more enhanced description of the main steps involved in the integrated framework: (a) the generation of multiple realizations of climate and streamflow data, and (b) the evaluation criteria used to define the number of realizations needed in hydrological simulations. A schematic illustration of the overall modeling framework is presented in Fig. 1.
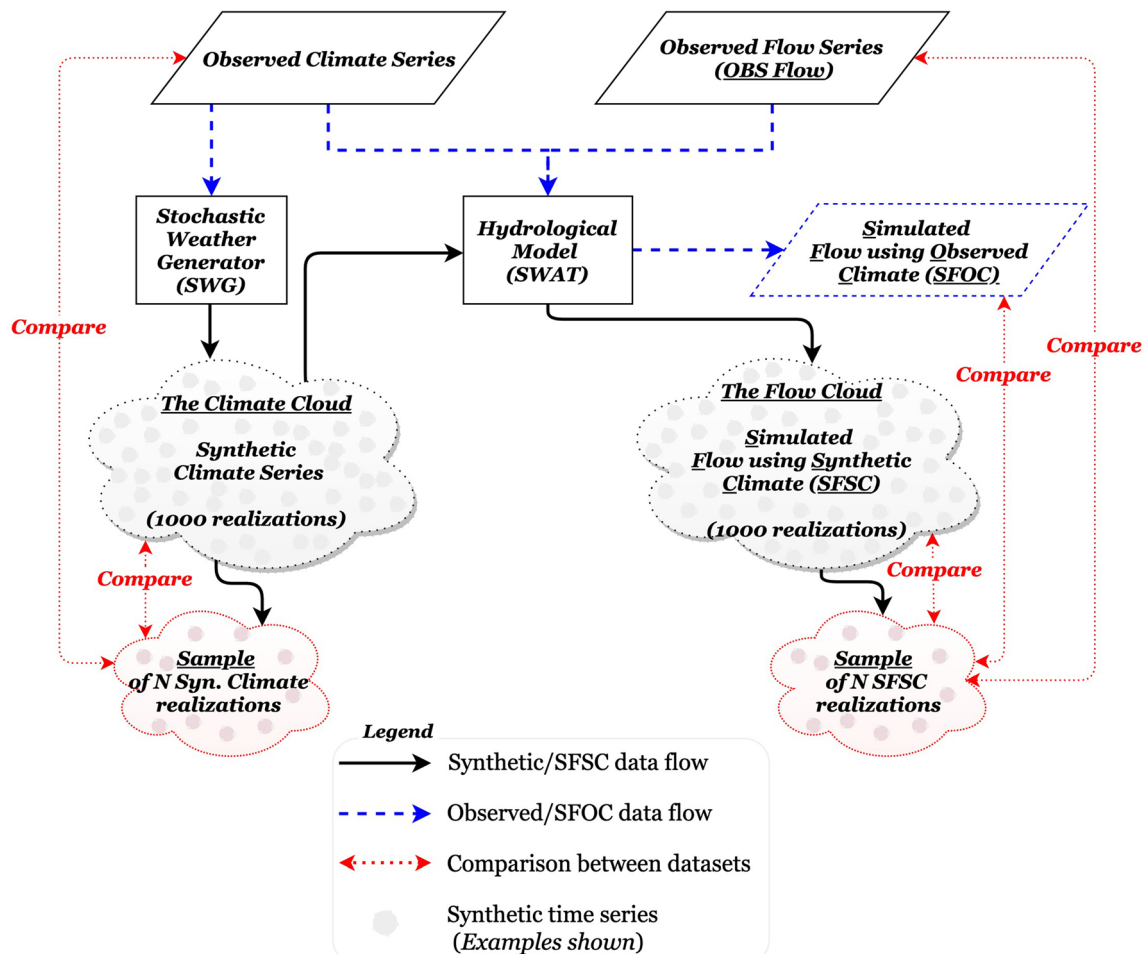
Fig. 1 Schematic representation of the current work, where *N* ranges from 1 to 1000 unique realizations

2.5 Climate and flow cloud generation

In this work, 41 years (1971–2011) of observed climate and streamflow records are used as the reference data from which the deviations are calculated. The synthetic precipitation and temperature time series in this paper are also 41 years long to permit an adequate risk analysis to be conducted (Semenov and Barrow 1997; Elliot and Arnold 2001). Each SWG was run 1000 separate times, resulting in a total of 5000 realizations of weather sequences at a daily time step. Each realization is 41-year long to match the length of the observed climate data, resulting a total of 205,000 synthetic weather years (5 SWGs × 41 years × 1000 realizations). Soltani and Hoogenboom (2003b) found that at least 15 years of historical climate data is required to generate synthetic time series that mimic the observed statistical characteristics. Similarly, the SWAT model was run 5000 separate times, with each run producing a unique 41-year realization of the climate. The choice of 1000 for the number of realizations for each weather generator, despite the excessive computational

demand particularly for the hydrological modeling, was done to form a dense *cloud* of realizations and thus identify a satisfactory number of realizations. The 1000 synthetic time series for precipitation, minimum temperature, and maximum temperature generated using SWGs and representing the climate at station *s* (referred to as *the climate cloud* hereafter) are denoted:

- $PCP_t^{i,s}$, *t = 1,…,T; s = 1,..,S for the precipitation time series*,
- $Tmax_t^{i,s}$, *t = 1,…,T; s = 1,..,S for the maximum temperature time series*,
- $Tmin_t^{i,s}$, *t = 1,…,T; s = 1,..,S for the minimum temperature time series*.

The 1000 streamflow time series obtained by forcing the SWAT model with the synthetic climate time series (referred to as *the flow cloud* hereafter) are each called *Simulated flow using synthetic climate (SFSC)* and denoted as:

- $SFSC_{t,SWG}^i$, *t = 1,…,T; i = 1,…,1000, SWG* ∈ {ME,MG,WE,WG, KNN}.

2.6 Estimation of a statistic *V* using *N* realizations

The following algorithm is used to estimate a statistic *V* using *N* realizations. For *k* between 1 and 10,000:

- Sample without repetition from a subset of size *N* of indices between 1 and 1000, i.e., $\{j_1^k, j_2^k, \ldots, j_N^k\}$.
- A *k*th estimate of the mean value of a statistic *V* is,

$$\mu_k = \frac{1}{N} \sum_{m=1}^{N} \left(\frac{\sum_{t=1}^{T} V_t^{j_m,s}}{T} \right)$$

The more variability in $\{\mu_k\}, k = 1, \ldots, 10000$, the less precise the estimate. The variability in these estimated means can be illustrated using a violin-plot graph. The deviations from Y_{ref} quantify the biases of the estimates.

2.7 Evaluation criteria

Given that a series of samples generated by the SWGs will not be identical, the impact of such variations between the samples is investigated visually using time-series graphs of the simulated sequences, such as sequence plots, running mean plots, and violin and boxplots of the samples.

2.7.1 Visual convergence assessment

An examination using proper graphical techniques can produce a general idea concerning the variable of interest (Ott and Longnecker 2015). Plots of each parameter and the running mean are used to examine the simulation process as the number of realizations increases. A time series plot of the running mean is simple and easy to implement and used to check when a new stochastic generation of flow data is no longer deviating significantly from the mean of previous realizations. The running mean is computed as the mean of all sampled values up to and including the current realization. The plot then shows whether the running mean stabilizes at a realization (randomly ordered) against the mean of all realizations (Smith 2007). These plots will eventually converge to a constant value, which is the mean of all realizations according to the Central Limit Theorem. These visual evaluations should provide general insights, yet they are not sufficient indicators and further statistical analyses must be conducted.

2.7.2 Quantitative assessment

The four key statistics to be estimated from the time series are the mean (μ), standard deviation (σ), and the skewness

(α_3) and kurtosis (α_4) coefficients of the climate or flow variable of interest. For the sake of simplicity, *Y* will be used herein to indicate any of the estimates of above statistics. The statistical measures considered in this paper are the Relative Error (RE), the Relative Root Mean Squared Error ($RMSE_r$) and the Cohen's effect size (*d*). The relative error (RE) refers to the magnitude of the difference between an experimental (sample) value (Y_i) and the known or accepted value (Y_{ref}):

$$RE(\%) = \left(\frac{Y_{ref} - Y_i}{Y_{ref}} \right) \times 100 \tag{3}$$

The root mean squared error (RMSE), also called root-mean-square deviation, is one of the most common metrics used to measure the accuracy of continuous variables via measuring the average magnitude of the error. It is a negatively-oriented score that has a range between 0 to $\propto$, meaning that values closer to 0 are preferable. This metric is particularly useful when a large error cannot be tolerated, as the errors are squared when computing it. The RMSE and relative RMSE ($RMSE_r$) are computed as:

$$RMSE = \sqrt{\frac{1}{KN} \sum_{j=1}^{K} \sum_{i=1}^{N} \left(Y_{ref} - Y_j \right)^2} \tag{4}$$

$$RMSE_r(\%) = \left(\frac{RMSE}{Y_{ref}} \right) \times 100, \tag{5}$$

The improvement in the $RMSE_r$ value obtained by adding one more realization ($RMSE_{r,improvement}$) and the marginal improvement ($RMSE_{r,mar.improvement}$) are defined as:

$$RMSE_{r,improvement} = RMSE_{r,n-1} - RMSE_{r,n} \tag{6}$$

$$RMSE_{r,mar.improvement} = RMSE_{r,n=1} - RMSE_{r,n} \tag{7}$$

where *n* = 2, 3, ..., *N*, and *N* = 1000.

2.7.3 Evaluation of effect size

The Cohen's Effect Size (Cohen's *d*) is a standardized quantitative index that can help in better understanding such large Monte-Carlo-like runs by checking the deviation (or overlap) between two groups in standard deviation units. Cohen's *d* uses the differences in means of the control (or reference) and sample groups and the standard deviation (SD) of the control group (Rosnow and Rosenthal 1996), and determined mathematically as

$$Cohen's\,d = \frac{\left| Mean_{Y_{ref}} - Mean_{Y_i} \right|}{SD_{Y_{ref}}} \tag{8}$$
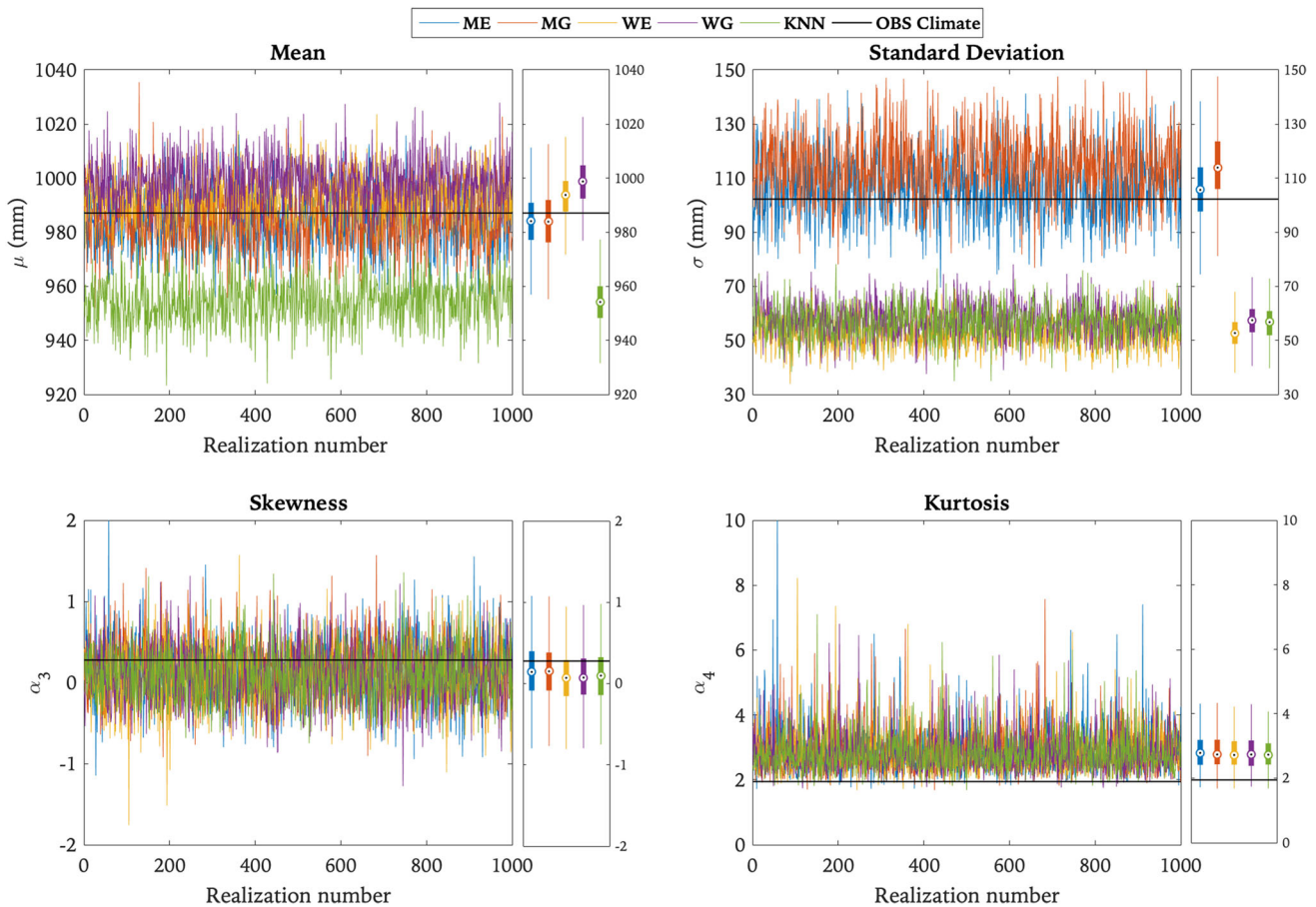
Fig. 2 Plots of the precipitation statistics generated by five SWGs compared to the observed climate values (black line). The side boxes delineate distributions of all realizations with the interquartile range (IQR: q_{25},q_{75}), while the whiskers limits correspond to $q_{25} \pm 1.5$ IQR

The standard deviation of the control group is used following Mehan et al. (2017), assuming variances of the two groups are not similar. This method govern by the control group is also known as Glass's d or Δ (Glass 1976). Large effects mean larger differences in means and lesser overlap between the two distributions. Some rules of thumb were given by Cohen (1988), who stated that the effect size could be interpreted as small ($d < 0.2$), medium ($0.2 < d < 0.8$), and large ($d > 0.8$). Nevertheless, the interpretation of the effect sizes shouldn't rigidly follow Cohen's framework; rather results should be evaluated in the context of prior related literature as suggested by Vacha-Haase and Thompson (2004).

2.7.4 The reference values for the key statistics

For any given statistic, several reference statistics can be used to calculate both the RE and the RMSEr. The three reference values for the key *climate* statistics are:

- The statistics calculated from observations ($V_{ref,Y,OBS}$), and

- The average of the statistics calculated from the 1000 realizations in the synthetic climate ($V_{ref,Y,SC}$).

The three reference values for the key *flow* statistics are:

- The estimates of statistics V calculated with observations, denoted by ($V_{ref,OBS}$);
- The estimates of statistics V calculated from the time series simulated via SWAT using the observed climate, denoted by ($V_{ref,SFOC}$); and
- The average of statistics V calculated from the 1000 realizations in the flow cloud, denoted by ($V_{ref,SFSC}$).

3 Results and discussion

The results are presented and discussed in three parts: first, a visual assessment of the synthetically generated climate and flow time series is presented. Second, the effect of the number of SWGs realizations on the accuracy of basic annual climatic indices is assessed. Third, the degree of divergence between the sample and the cloud mean (the
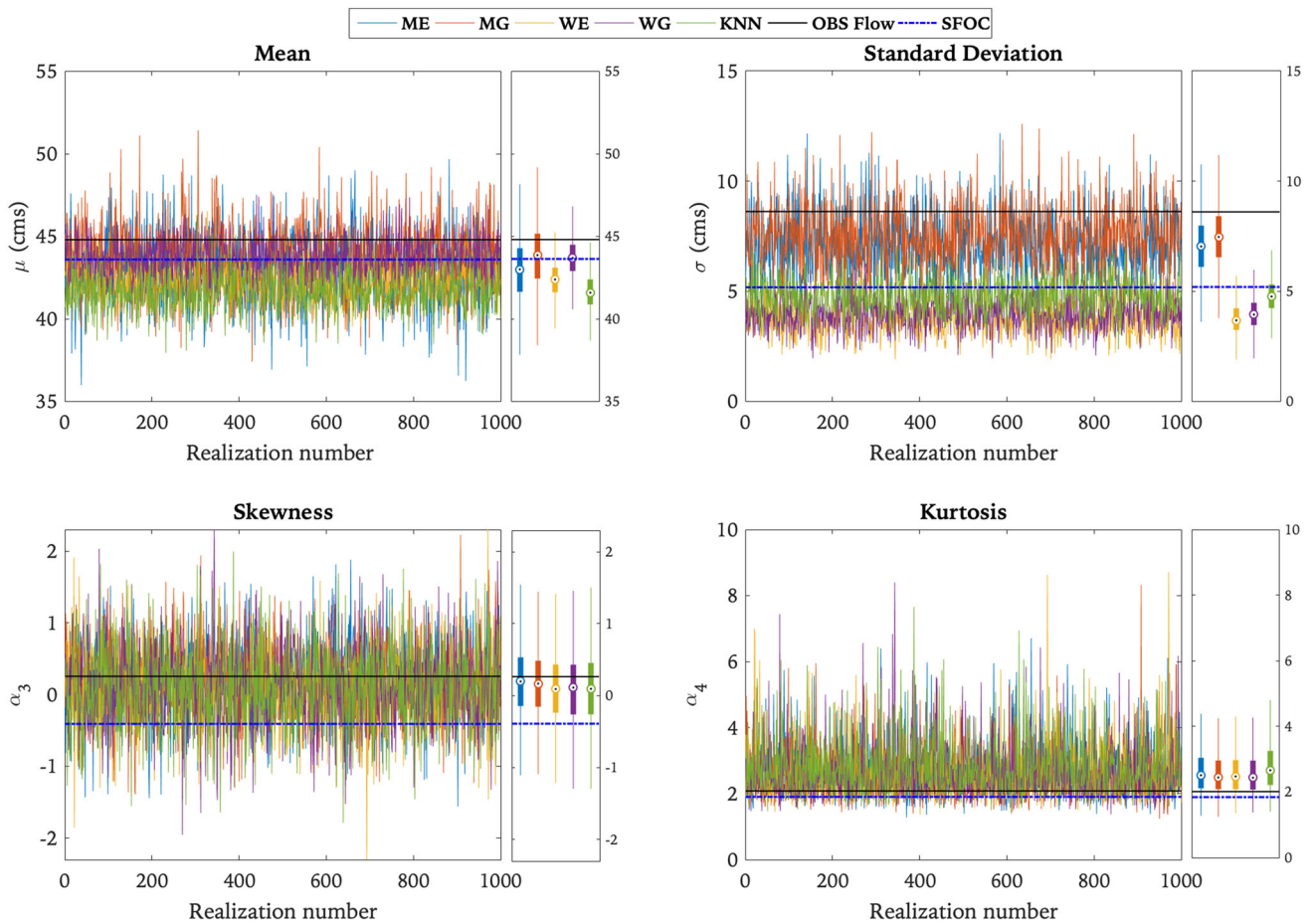
Fig. 3 Plots of mean annual streamflow statistics generated by five SWGs compared to the observed flow and SFOC values (shown by the black and blue dashed lines, respectively). The side boxes delineate distributions of all realizations with the interquartile range (IQR: q_{25}, q_{75}), while the whiskers limits correspond to $q_{25} \pm 1.5$ IQR

control group) is characterized by Cohen's d effect size. Variability is presented via violin and boxplots and graphics of the running mean, the RMSEr, and the RE, where the x-axis in each case represents the number of realizations, which goes from 1 to 1000. The same analysis is performed for each climate and flow variable, performance index, and reference value.

3.1 Visual convergence assessment

Figure 2 shows that the mean annual precipitation estimated by the MulGETS and WeaGETS realizations is reasonably close to the mean of observed values ($\mu_{\text{ref,PCP,OBS}}$), but that the observed values are underestimated by KNN. However, the WeaGETS models (WE and WG) and KNN underestimated the standard deviation
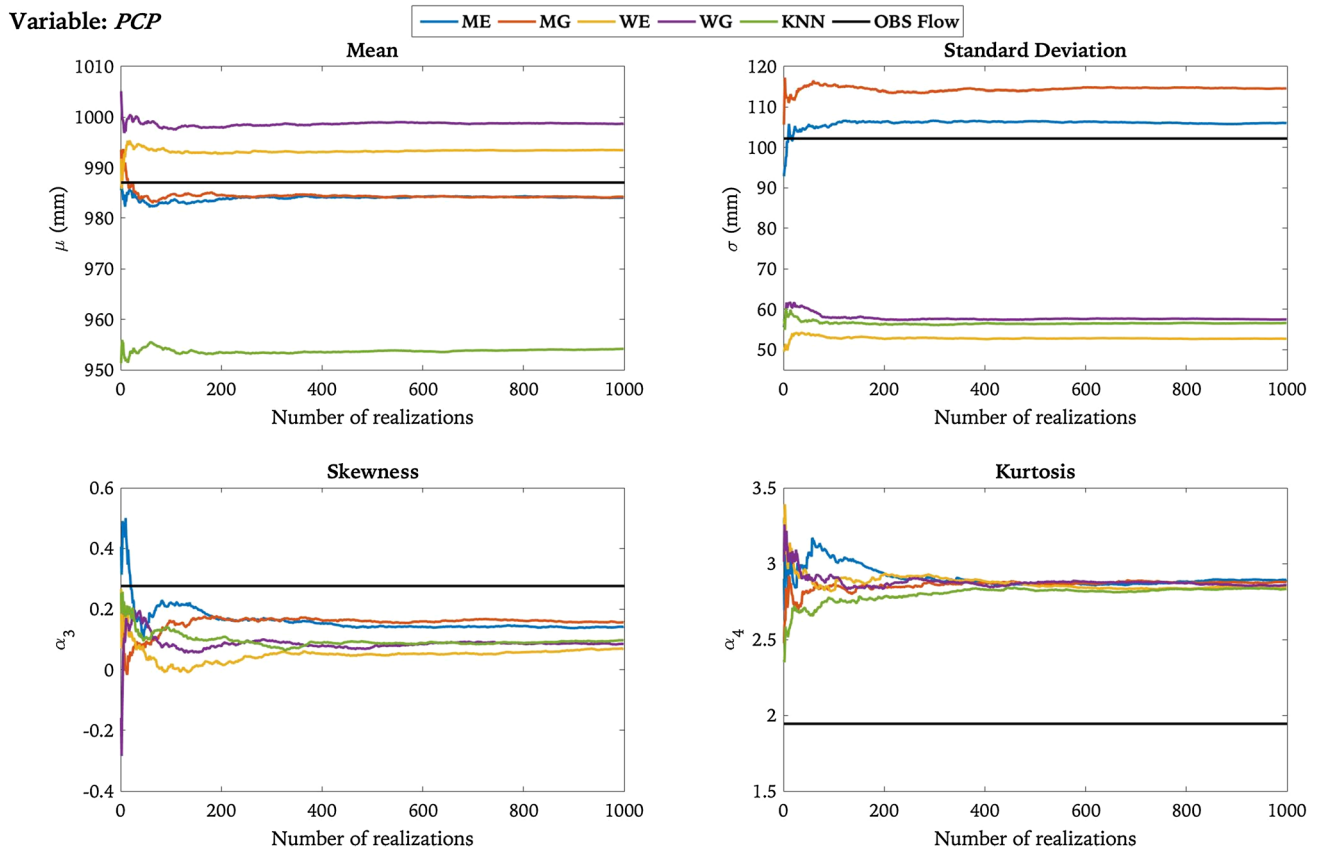
Variable: *PCP*

Fig. 4 Running mean plots for the mean annual precipitation statistics generated by five SWGs in which the order of the realizations is random. The black dashed lines indicate the observed climate values

$\sigma_{\text{ref,PCP,OBS}}$ of the annual precipitation, while both the MulGETS models (ME and MG) were able to capture σ adequately (Fig. 2). The kurtosis coefficients for the synthetic annual precipitation were consistently higher than those for the observed precipitation (Fig. 2). Thus, the results are consistent with the findings of Chen and Brissette (2014), who reported that the kurtosis coefficient of the mean annual precipitation is poorly reproduced by SWGs. The differences among the five models in terms of generating α_3 for the synthetic annual precipitation were not notable.

The interannual σ's of SFSC_{WE} and SFSC_{WG} were underestimated compared to the observed flow and—to a lesser degree—the SFOC (Fig. 3). The interannual variability of SFSC_{KNN} closely matched that of SFOC, while SFSC_{ME} and SFSC_{MG} were between the two reference datasets (mostly underestimated the observed flow but overestimated the SFOC). Interestingly, the SFSCs of all SWGs performed similarly in well reproducing the α_3 of the OBS Flow and overestimating the α_3 of the SFOC. Similar to precipitation results, the poor performance (overestimation) of most outputs of the tested SWGs in replicating the α_4 of the annual streamflow was observed
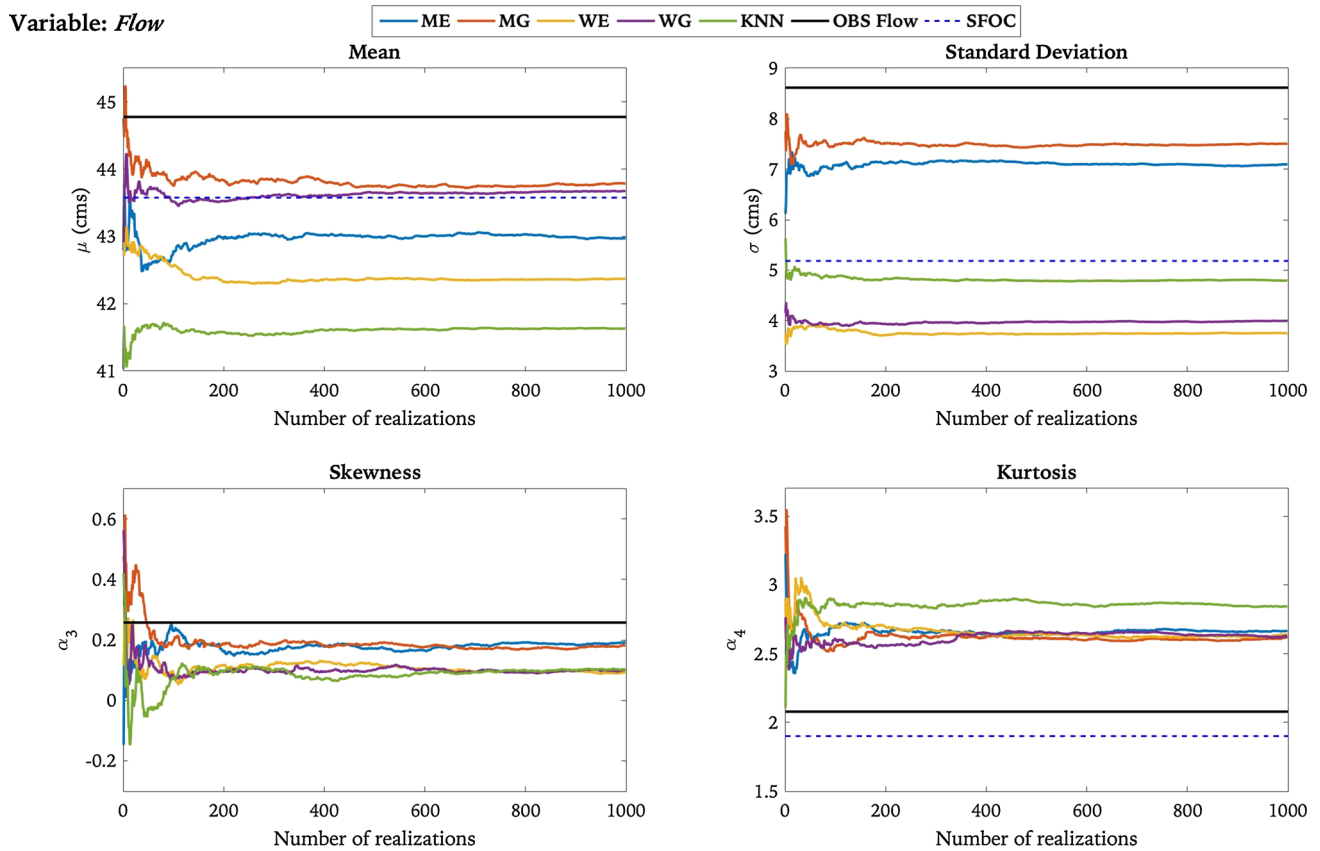
Variable: *Flow*

Fig. 5 Running mean plots for the annual streamflow statistics generated by five SWGs in which the order of the realizations is random. The observed flow (SFOC) values are represented by the black (blue dashed) lines

when compared to both the OBS Flow and SFOC data (Fig. 3). In general, it is fairer to compare the SFSC to the SFOC than to the observed flow, as the first two were both simulated by SWAT and inherited the biases within the model itself. Such figures can help grasp a general idea about the realizations, but further investigation using more sophisticated statistical methods is certainly needed.

The annual precipitation and streamflow statistics are plotted as a function of the number of realizations in Figs. 4 and 5. The running mean plots show the mean of previous realizations up to and including each iteration displayed on the x-axis. Such figures show how the running mean highly fluctuates at the beginning of the sequence, making it difficult to construct robust confidence intervals. The statistics for the outputs of the five weather generators,

however, do not differ much after 100 realizations. That is, almost all parameter estimates appear to stabilize around 100 realizations. Biases caused by the stochastic generation of the cloud are clearly outweighed eventually by the increased number of realizations, as stated in Räisänen and Ruokolainen (2006), as it is the case for any Monte Carlo experiment (Cunha et al. 2014). That is, the approximation or the performance gains can be improved by increasing the number of realizations to achieve a certain level of precision. We are not presenting graphs for temperature parameters due to lack of space, but similar patterns were observed.
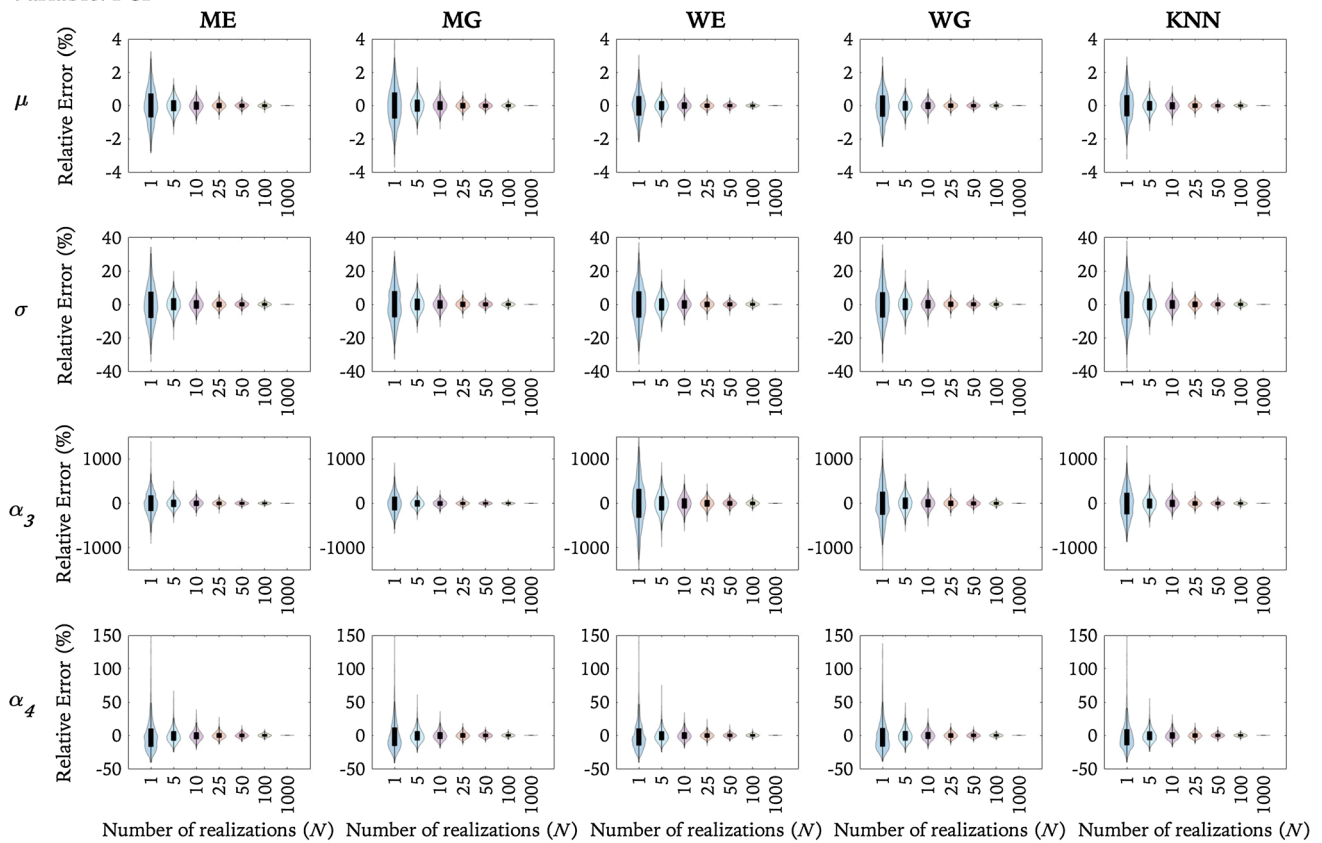
Variable: *PCP*

Fig. 6 Violin plots of the relative errors (%) of the main *annual* precipitation statistics for the *N*-realization samples used to estimate these statistics from the cloud; an *N-realization* sample is derived from 10,000 different randomly selected SFSC sets

3.2 Variations in the spread, RMSEr's, and REs for key statistics as a function of the number of realizations

3.2.1 Climate space

As explained in the methodology section, the spread of the estimates was visualized using violin and box plots. Violin plots are accompanied with black boxplots (25th and 75th percentiles representing interquartile ranges, and 1.5 times the IQR whiskers). These plots were generated using the functions by Bastian Bechtold available on the GitHub repository (Violin plots for Matlab https://github.com/ bastibe/Violinplot-Matlab). As expected, the variability in each of the indicators decreases as the number of realizations increases (Figs. 6, 7, and 8). The use of a single realization is not recommended due to the high error expected, particularly for applications that depend heavily on higher moments, such as an assessment of extremes. For instance, the precision when estimating the α_3 of the annual precipitation using one realization can be off by more than 500%. Once the number of realizations increases, the expected error decreases dramatically. This decrease in the expected error is particularly clear for higher moments at 25 realizations and higher. Moreover, the use of more than 100 realizations seems very unnecessary.
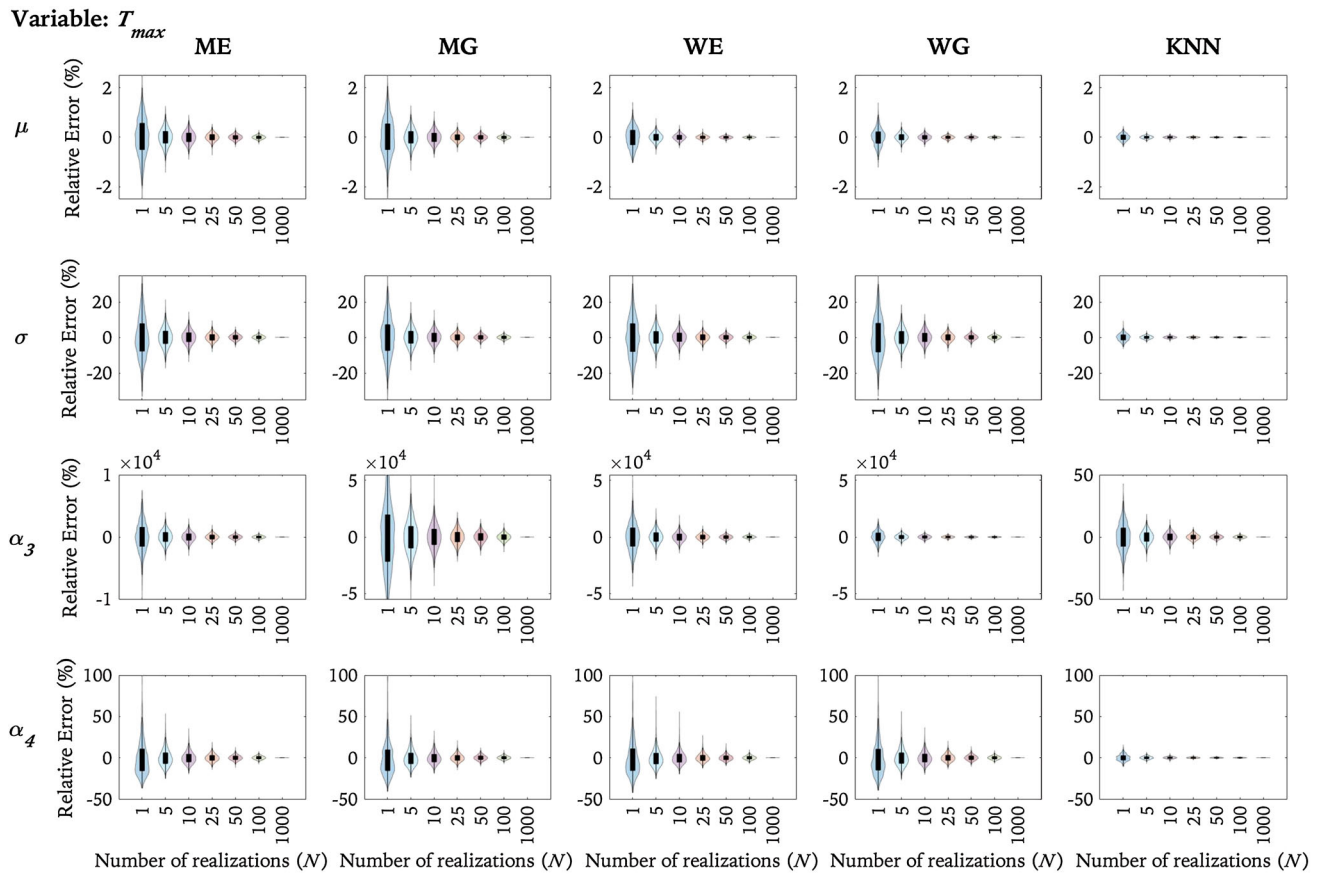
Fig. 7 Violin plots of the relative errors (%) of the main annual maximum temperature statistics for the *N*-realization samples used to estimate these statistics from the cloud; an *N-realization* sample is derived from 10,000 different randomly selected SFSC sets

The marginal improvements in the RMSEr's of PCP, Tmax, and Tmin as the number of realizations increases are shown respectively in Figs. 9, 10, and 11, where the synthetic climate using *N* realizations (relative to using *N-1* realizations) is compared to the two reference datasets: the climate cloud (synthetic climate) and the observed climate. Tables 2, 3, and 4 present a similar comparison for the three climate variables but relative to the results of just a single realization. These results are consistent with the previous findings suggesting that after 100 realizations, the marginal improvement in the RMSEr becomes insignificant (e.g., less than a 0.21 (1.09) % improvement across SWGs in μ_{Tmax} (σ_{Tmax}) when adding 900 realizations).

Also, 25 realizations appear to be reasonably adequate, particularly for the first two moments (e.g., less than a 0.46 (2.34) % improvement across SWGs in μ_{Tmax} (σ_{Tmax}) when adding 975 realizations). The results are very similar for the temperature variables, whereas precipitation indicators require even fewer realizations.

3.2.2 Hydrological space

For the streamflow data, Fig. 12 presents the REs of the key annual streamflow statistics, including the mean, standard deviation, skewness, and kurtosis. The variability of each RE as a function of different numbers of
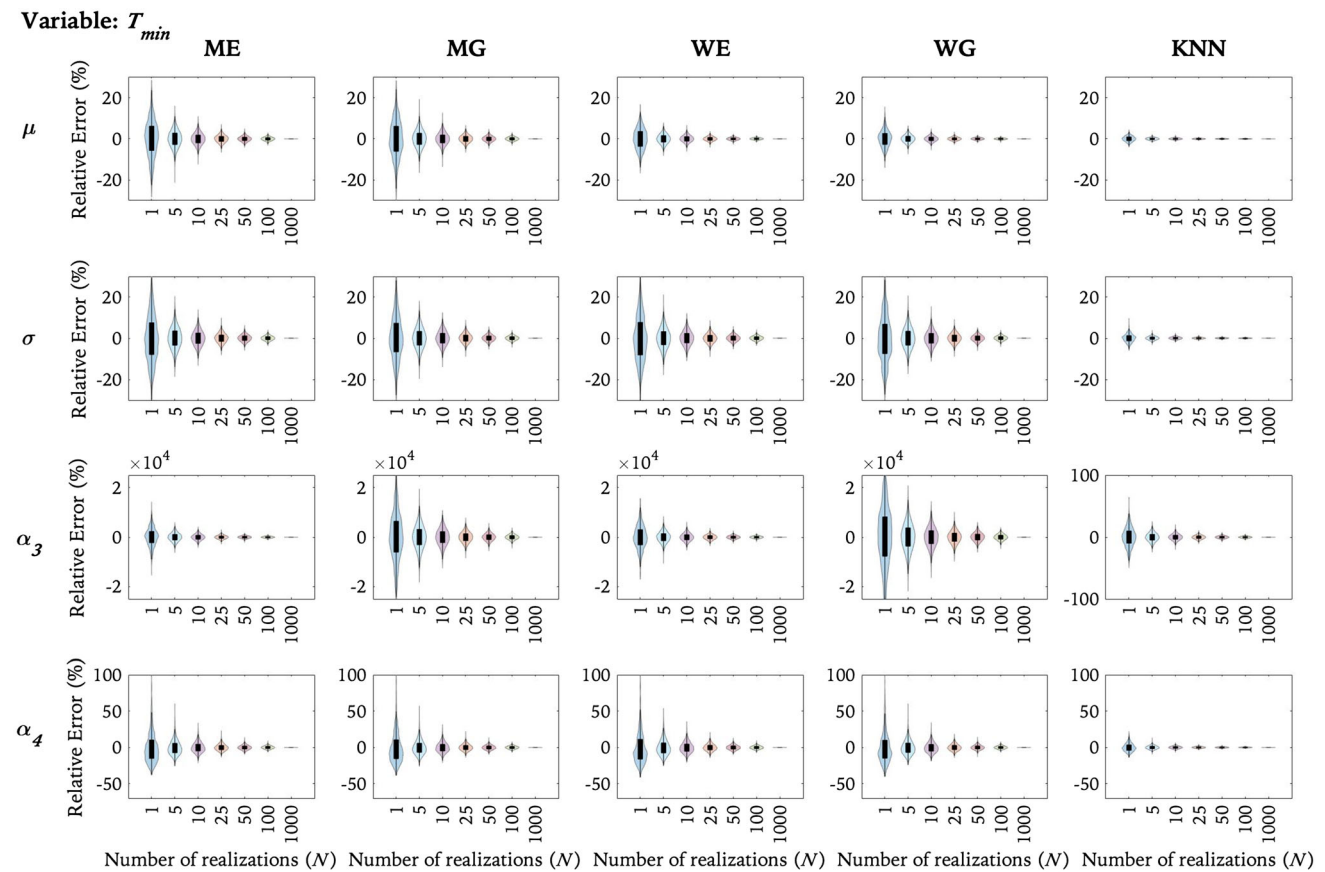
Variable: T_{min}

Fig. 8 Violin plots of the relative errors (%) of the main annual minimum temperature statistics for the *N*-realization samples used to estimate these statistics from the cloud; an *N-realization* sample is derived from 10,000 different randomly selected SFSC sets

realizations (1, 5, 10, 25, 50, 100, and 1000) is represented using violin plots, each of which is based on ten thousand *N*-realization samples randomly taken from the cloud containing all SFSC time series. Figure 12 strongly suggests that a sole realization is not sufficient for representing SWGs in hydrological modeling. Similar to the situation for the climate variables, 100 realizations seem adequate, with very low relative errors across different statistics.

The marginal improvements in the RMSEr's of the streamflow statistics are plotted in Fig. 13, and Table 5 lists these improvements as functions of the number of

realizations. In Fig. 13, the estimates are compared to the three reference values (SFSC, SFOC, and OBS Flow). The estimate is compared to the estimate obtained using a single realization (Table 5). The results are consistent with the previous findings, which suggest that after 100 realizations, the marginal improvement in the RMSEr becomes insignificant (e.g., less than a 0.55% improvement for all three reference datasets and across all SWGs for μ when compared to the μ calculated from 1000 realizations). Also, 25 realizations appear to be reasonably adequate, particularly for the first two moments (e.g., less than a 1.78%
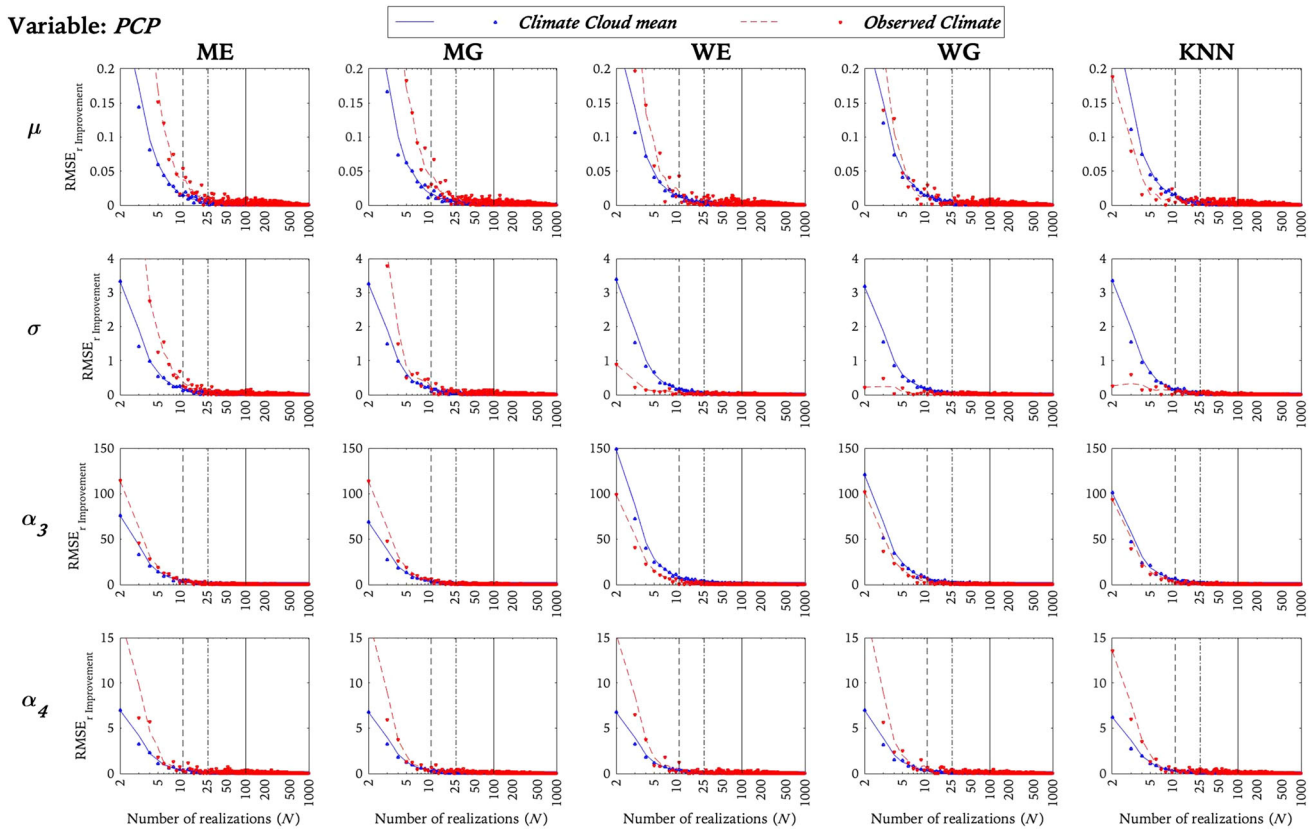
Fig. 9 Improvement in the *RMSEr's* of the main annual precipitation statistics for the *N*-realization samples generated by the five SWGs. The *RMSEr's* is calculated using either the observations (red line) or the cloud mean (blue line) as reference. The *N*-realization samples are derived from 10,000 different randomly selected SFSC sets. Scattered markers represent actual results for which the lines are slightly smoothed by moving averages with spans of 3. Vertical black dashed, dash-dotted, and solid lines represent 10, 25, and 100 realizations, respectively

improvement for all three reference datasets and across the SWGs for μ when compared to the μ calculated from 1000 realizations).

3.3 Impact of the number of realizations on Cohen's d effect size

Cohen's *d* values of precipitation, maximum temperature, minimum temperature, and streamflow statistics are presented in Figs. 14, 15, 16, and 17, respectively. Such figures clearly indicate a very large effect size, as expected, when a single realization is used for all variables, statistics, and SWGs. 5, 10 and 25 realizations are not sufficiently enough but the effect size diverges decreasingly from zero as the number of realizations is increasing. Moreover, the upper quartile of the 10,000 different randomly selected sample sets across variables, SWGs, and statistics show that in more than 75%, the effect size deemed to be small ($d < 0.2$) after 50 realizations. Further, figures demonstrate that 100 realizations are exhibiting even a smaller effect size of the randomly selected sets as Cohen's *d* values are always below 0.5 (the horizontal red dash-dotted line) indicating small to medium effect size. Similar interpretive
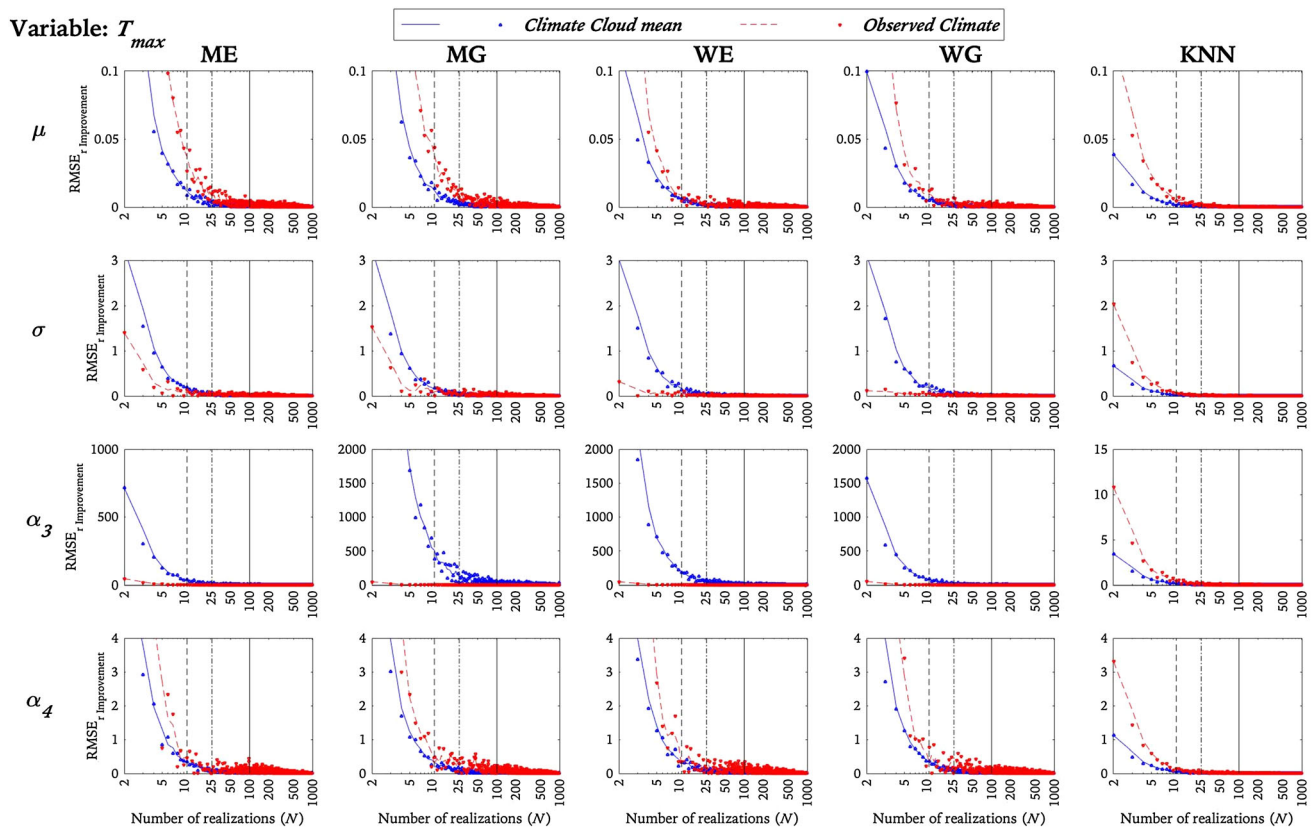
Fig. 10 Improvement in the *RMSEr's* of the main annual maximum temperature statistics for the *N*-realization samples generated by the five SWGs. The *RMSEr* is calculated using either the observations (red line) or the cloud mean (blue line) as reference. The *N*-realization samples are derived from 10,000 different randomly selected SFSC sets. Scattered markers represent actual results for which the lines are slightly smoothed by moving averages with spans of 3. Vertical black dashed, dash-dotted, and solid lines represent 10, 25, and 100 realizations, respectively

framework was also followed in the prior related literature (cf. Mehan et al. 2017; Guo et al. 2018).

3.4 Discussion

The main finding of this work is that, while a larger number of realizations may provide a better representation of climate variability, a limited problem-dependant number of realizations can provide robust estimates of key risk statistics. In this particular application, the marginal improvements in the RMSEr's of all statistics (climatic and hydrological variables) are not substantial after 25

realizations, particularly for the first two moments (i.e., μ and σ) and to a lesser extent for higher moments (i.e., α_3 and α_4). Cohen's *d*, which characterizes the degree of divergence between the sample and the cloud mean (the control group), was used to compare the effect sizes as the number of realizations is increasing. Like any Monte Carlo simulation relying on repeated random sampling, the marginal value of a realization decreases as the number of realizations increase. The findings of these metrics suggest that for this particular SWAT model and this particular set of SWGs, going beyond 100 realizations is redundant with a higher computational cost, as the improvement afterward
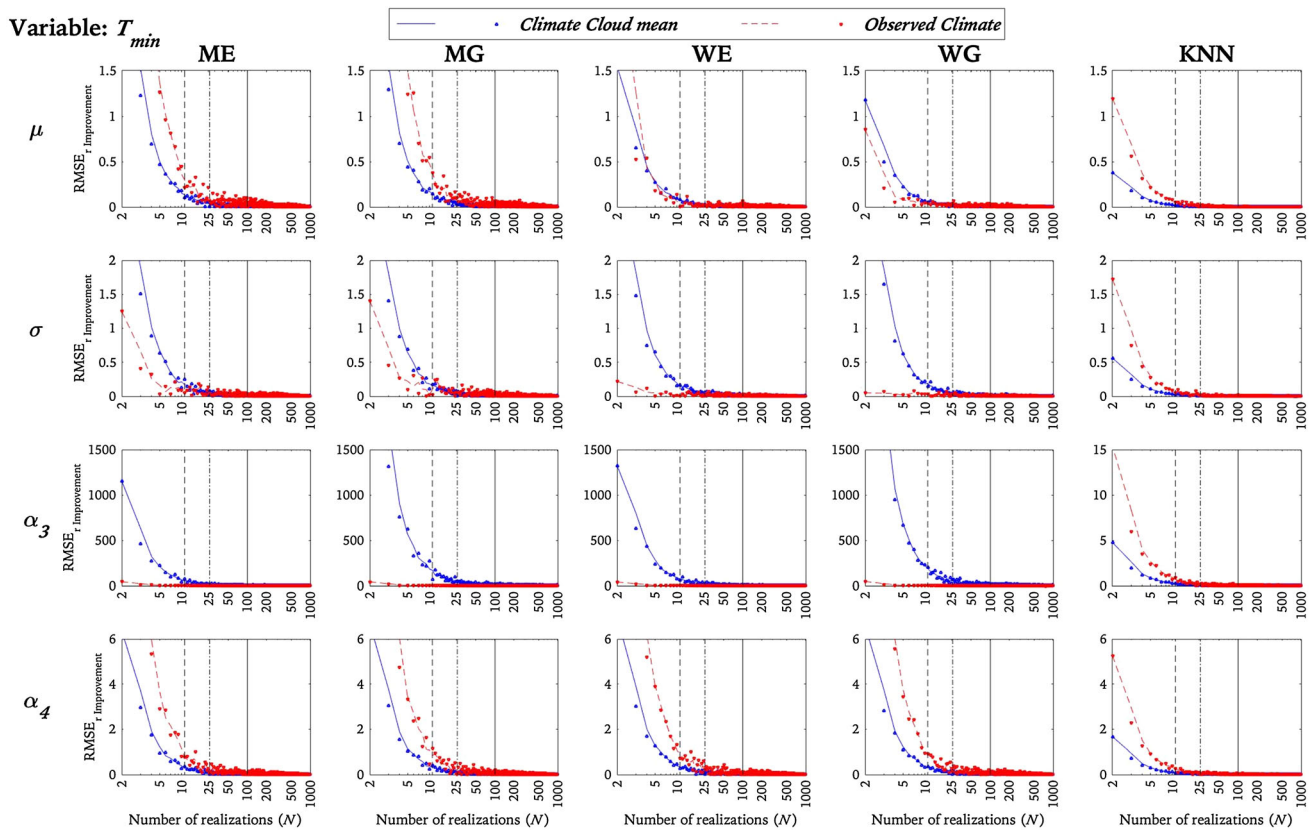
Fig. 11 Improvement in the *RMSEr's* of the main annual minimum temperature statistics for the *N*-realization samples generated by five SWGs versus the counterparts. The *RMSEr* is calculated using either the observations (red line) or the cloud mean (blue line) as reference. The *N*-realization samples are derived from 10,000 different randomly selected SFSC sets. Scattered markers represent actual results for which lines are slightly smoothed by moving averages with spans of 3. Vertical black dashed, dash-dotted, and solid lines represent 10, 25, and 100 realizations, respectively

is very minimal even for higher moments. Such results will surely serve to better account for risk in quantitative analysis and decision making in a variety of water and environmental related problems with minimal computational and time requirements.

An interesting finding is that there are systematic biases contained within the weather generators that lead to the SFSC and SFOC to be different from the observed flow. However, increasing the number of realizations cannot reduce these biases. That is, repeated runs of a given SWG that tends to misestimate a particular variable will not be useful in obtaining a correct characterization of the observed variable. A few ways to decrease such biases include improving the SWGs, selecting a SWG with minimal bias, and/or improving the rainfall-runoff model. Alternatively, one can consider generating a large dataset of realizations and then select a number of realizations that better represents the observed set, as suggested by Gitau et al. (2018). However, the latter approach still presents a challenge, as it can be computationally expensive and time-consuming. The simplest of these solutions is to use the methodology presented herein to select the number of realizations that leads to the feasibly lowest RMSEr or RE for the problem at hand (i.e., when the relative

Table 2 Marginal improvements in *RMSEr* ($RMSE_{r,mar.improvement}$) of two precipitation reference datasets obtained by using *N* realizations relative to a single realization

SWG	N	Synthetic Climate (PCP)				Observed Climate (PCP)			
		μ	σ	α_3	α_4	μ	σ	α_3	α_4
ME	5	0.58	6.27	142.48	13.57	1.69	18.23	205.77	31.08
	10	0.71	7.83	176.22	16.87	2.02	22.43	244.79	34.62
	25	0.84	9.12	206.82	19.84	2.31	25.01	274.30	37.75
	50	0.90	9.79	222.73	21.36	2.41	25.98	286.04	38.50
	100	0.95	10.27	233.46	22.37	2.46	26.54	291.13	38.88
	1000	1.05	11.36	258.24	24.77	2.51	27.06	295.82	39.30
MG	5	0.63	6.28	126.99	13.11	1.86	12.79	205.97	28.18
	10	0.78	7.78	156.70	16.12	2.25	15.16	244.57	31.63
	25	0.92	8.99	182.07	18.86	2.55	15.96	273.10	34.15
	50	0.99	9.68	195.74	20.26	2.67	16.40	284.33	35.16
	100	1.04	10.16	205.83	21.24	2.74	16.61	290.43	35.34
	1000	1.15	11.20	226.93	23.47	2.80	16.75	294.98	35.72
WE	5	0.47	6.41	285.88	12.95	0.99	0.93	175.79	26.54
	10	0.58	7.79	356.21	15.78	1.14	0.93	207.24	29.59
	25	0.68	9.11	417.09	18.59	1.25	1.08	226.50	32.37
	50	0.73	9.80	448.82	19.97	1.29	1.10	232.03	33.13
	100	0.76	10.27	472.18	20.91	1.31	1.13	236.06	33.69
	1000	0.84	11.34	521.60	23.08	1.32	1.13	239.04	33.95
WG	5	0.49	6.13	226.69	13.12	0.73	0.93	177.44	29.27
	10	0.62	7.60	283.24	15.99	0.83	1.00	207.32	32.51
	25	0.72	8.91	330.35	18.76	0.89	1.08	227.91	35.48
	50	0.78	9.58	356.13	20.20	0.91	1.12	235.86	35.97
	100	0.82	10.05	373.48	21.16	0.93	1.11	240.65	36.19
	1000	0.90	11.11	412.88	23.38	0.94	1.14	243.46	36.65
KNN	5	0.52	6.51	191.85	11.95	0.31	0.84	163.73	24.35
	10	0.64	7.93	234.66	14.67	0.35	1.14	196.25	27.17
	25	0.74	9.28	275.80	17.17	0.36	1.33	215.33	29.15
	50	0.80	9.98	296.28	18.49	0.37	1.30	220.19	30.11
	100	0.84	10.47	311.13	19.43	0.37	1.30	223.98	30.65
	1000	0.92	11.56	343.50	21.46	0.37	1.33	227.40	30.77

improvement becomes very small). Alternative methods for assessing SWGs include statistical tests of significance, such as the *t* test and F-test (e.g., Min et al. 2011; Chen and Brissette 2014); χ^2 goodness-of-fit test (e.g., Semenov et al. 1998); nonparametric tests, such as the Wilcoxon rank sum test, the Kolmogorov–Smirnov (K–S) test, and Mann–Whitney test (e.g., Zhang and Garbrecht 2003; Qian et al. 2004; Chen et al. 2010); the RMSEs of various statistics of interest (e.g., Mehrotra et al. 2006); and employing distance techniques, such as the Mahalanobis distance between statistics derived from observed and simulated time series (e.g., Alodah and Seidou 2019a).

One limitation of the present work is that the results are specific to a particular hydrological model on a particular

Table 3 Marginal improvements in *RMSEr* (*RMSE$_{r,mar.improvement}$*) of two maximum temperature reference datasets obtained by using *N* realizations relative to a single realization

SWG	N	Synthetic climate (Tmax)				Observed climate (Tmax)			
		μ	σ	α_3	α_4	μ	σ	α_3	α_4
ME	5	0.42	6.42	1351.87	12.17	1.34	2.11	74.70	29.94
	10	0.53	7.90	1657.07	15.19	1.67	2.46	83.22	35.15
	25	0.62	9.23	1944.42	17.84	1.95	2.73	90.18	38.34
	50	0.67	9.95	2089.46	19.21	2.09	2.85	92.18	39.41
	100	0.70	10.48	2190.06	20.17	2.18	2.94	92.34	40.07
	1000	0.77	11.57	2418.07	22.30	2.32	2.94	93.52	40.46
MG	5	0.43	6.20	18012.49	12.88	1.36	2.28	70.78	33.29
	10	0.54	7.56	22297.20	15.77	1.69	2.75	82.49	37.84
	25	0.63	8.94	26099.98	18.56	1.98	3.02	87.79	41.25
	50	0.67	9.59	27944.27	19.86	2.12	3.02	90.34	42.44
	100	0.70	10.05	29327.05	20.82	2.23	3.05	91.42	42.77
	1000	0.78	11.12	32349.77	22.96	2.44	3.08	92.27	43.36
WE	5	0.22	5.92	7365.25	13.16	0.49	0.43	72.32	35.22
	10	0.27	7.43	9071.86	16.28	0.57	0.51	85.17	40.58
	25	0.32	8.67	10584.92	19.14	0.62	0.47	90.27	43.46
	50	0.34	9.33	11364.65	20.67	0.64	0.48	92.55	45.00
	100	0.36	9.82	11947.48	21.70	0.65	0.49	93.55	45.71
	1000	0.40	10.87	13177.97	24.02	0.66	0.49	94.41	46.19
WG	5	0.19	6.19	2852.45	13.17	0.47	0.35	77.10	35.97
	10	0.24	7.72	3521.50	16.13	0.55	0.36	88.91	40.25
	25	0.28	9.01	4121.78	18.84	0.61	0.44	94.49	43.68
	50	0.30	9.67	4424.45	20.22	0.63	0.43	96.81	44.85
	100	0.31	10.16	4651.55	21.20	0.64	0.41	97.93	45.32
	1000	0.35	11.23	5136.53	23.36	0.65	0.43	99.02	45.86
KNN	5	0.07	1.18	6.41	2.15	0.23	3.45	19.76	6.16
	10	0.09	1.47	7.87	2.67	0.29	4.17	24.01	7.42
	25	0.11	1.71	9.19	3.13	0.33	4.66	27.24	8.36
	50	0.11	1.83	9.90	3.37	0.36	4.84	28.69	8.71
	100	0.12	1.93	10.39	3.54	0.38	4.95	29.54	8.89
	1000	0.13	2.12	11.48	3.92	0.41	5.04	30.36	9.08

watershed and to particular SWGs. However, the methodology can be applied to any case in which multiple weather generators are compared, and where there is a strong incentive to limit the number of simulations, for instance to save time and computational resources. The current paper required 5,000 runs (scenarios) of the SWAT model, and the computation time required to complete these scenarios was almost a month on a typical desktop computer workstation (Intel Core i7-4790 Processor @ 3.60 GHz (8 CPUs), 16 GB (2 × 8 GB) RAM, 1 TB disk), exclusive of

Table 4 Marginal improvements in *RMSEr* ($RMSE_{r,mar.improvement}$) of two minimum temperature reference datasets obtained by using *N* realizations relative a single realization

SWG	N	Synthetic climate (Tmin)				Observed climate (Tmin)			
		μ	σ	α_3	α_4	μ	σ	α_3	α_4
ME	5	5.03	6.31	2108.35	12.04	14.53	2.02	69.78	38.38
	10	6.26	7.76	2585.52	14.99	17.84	2.40	77.83	47.41
	25	7.31	9.15	3034.60	17.70	20.25	2.65	84.66	54.45
	50	7.84	9.85	3241.87	18.94	21.10	2.62	86.21	57.09
	100	8.23	10.35	3408.78	19.90	21.65	2.70	86.95	58.78
	1000	9.08	11.44	3762.85	21.99	22.12	2.70	88.16	60.46
MG	5	5.11	6.11	5881.90	12.33	15.25	2.03	67.72	39.32
	10	6.35	7.48	7281.93	15.14	18.78	2.54	77.99	47.48
	25	7.41	8.85	8495.20	17.83	21.39	2.78	81.76	54.88
	50	7.95	9.51	9083.50	19.09	22.41	2.81	83.99	57.90
	100	8.36	9.97	9537.18	20.04	23.06	2.84	85.11	59.57
	1000	9.24	11.03	10523.34	22.11	23.72	2.84	86.01	61.27
WE	5	2.87	6.10	2627.83	12.91	3.95	0.40	71.97	40.98
	10	3.54	7.54	3268.96	16.08	4.48	0.39	80.42	50.31
	25	4.12	8.83	3840.86	18.87	4.71	0.43	86.16	57.61
	50	4.45	9.46	4126.73	20.20	4.83	0.43	88.24	60.29
	100	4.66	9.96	4342.01	21.20	4.91	0.45	89.38	61.95
	1000	5.13	11.00	4799.23	23.46	4.92	0.44	90.39	63.59
WG	5	2.25	6.23	6879.67	12.30	1.20	0.16	70.64	39.40
	10	2.77	7.68	8491.17	15.24	1.38	0.20	80.76	48.17
	25	3.25	8.99	9922.20	17.88	1.43	0.24	85.73	55.20
	50	3.50	9.67	10678.55	19.16	1.45	0.23	88.08	57.95
	100	3.67	10.15	11216.60	20.11	1.47	0.21	89.14	59.54
	1000	4.06	11.22	12401.88	22.17	1.48	0.23	90.22	61.25
KNN	5	0.72	1.07	8.59	3.11	2.29	3.18	27.18	9.70
	10	0.90	1.32	10.71	3.87	2.85	3.86	33.68	11.96
	25	1.05	1.54	12.51	4.52	3.31	4.40	38.80	13.80
	50	1.12	1.67	13.42	4.86	3.55	4.61	41.10	14.55
	100	1.18	1.75	14.10	5.09	3.73	4.71	42.52	15.08
	1000	1.31	1.93	15.57	5.63	4.05	4.83	43.96	15.58

the subsequent time spent in the post-processing of the outputs. The time involved could even be higher for larger watersheds or a longer simulation period. Thus, the use of a small but adequate representative number of realizations, as determined herein, can significantly minimize the computational challenge and reduce the simulation time without losing much information (e.g., it would take less than a day for 25 realizations on a 3.60 GHz Intel Core i7
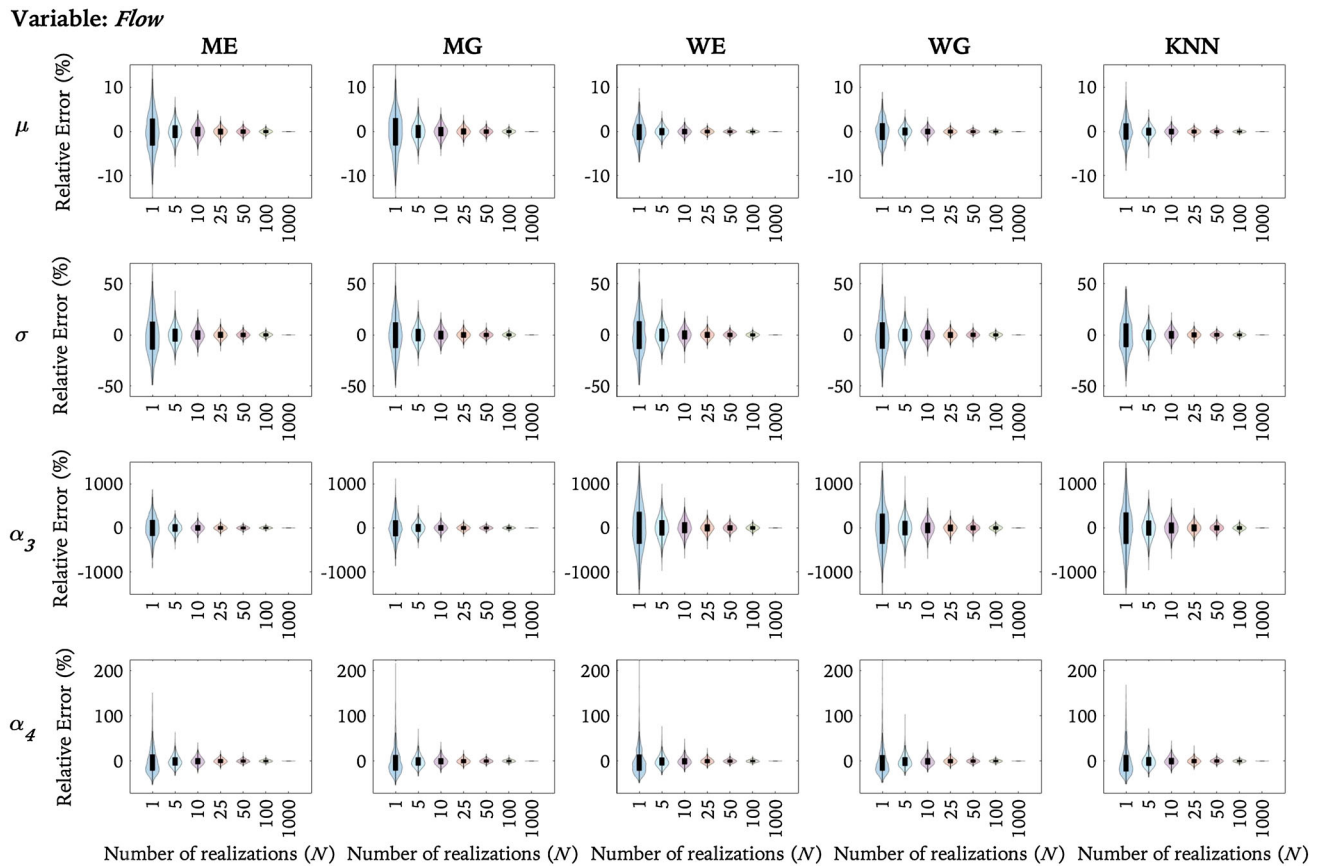
Variable: *Flow*

Fig. 12 Violin plots of the relative errors (%) of the main *annual* streamflow statistics for the *N*-realization samples used to estimate these statistics from the cloud; an *N-realization* sample is derived from 10,000 different randomly selected SFSC sets

CPU with a 16 GB RAM machine). Furthermore, the methodology presented in this paper has the advantage of making a straightforward link between the number of realizations and common statistical indicators and is more likely to appeal to practitioners.

Indeed, it can be argued that the results depend on the SWG, the hydrologic model, and the risk parameter used. High quantiles of flow, and hydrological parameters, such as sediments, would perform differently. Therefore, the results of this work can be further extended to include multiple hydrological models and more such risk parameters. However, the vast majority of risk statistics derived from environmental models are related to the four first moments of the time series that were examined in this paper. We also used a relatively complex hydrological model that is used worldwide, suggesting that the results of this study would be a reasonably informed guess in most practical cases where the modeler does not want to run an
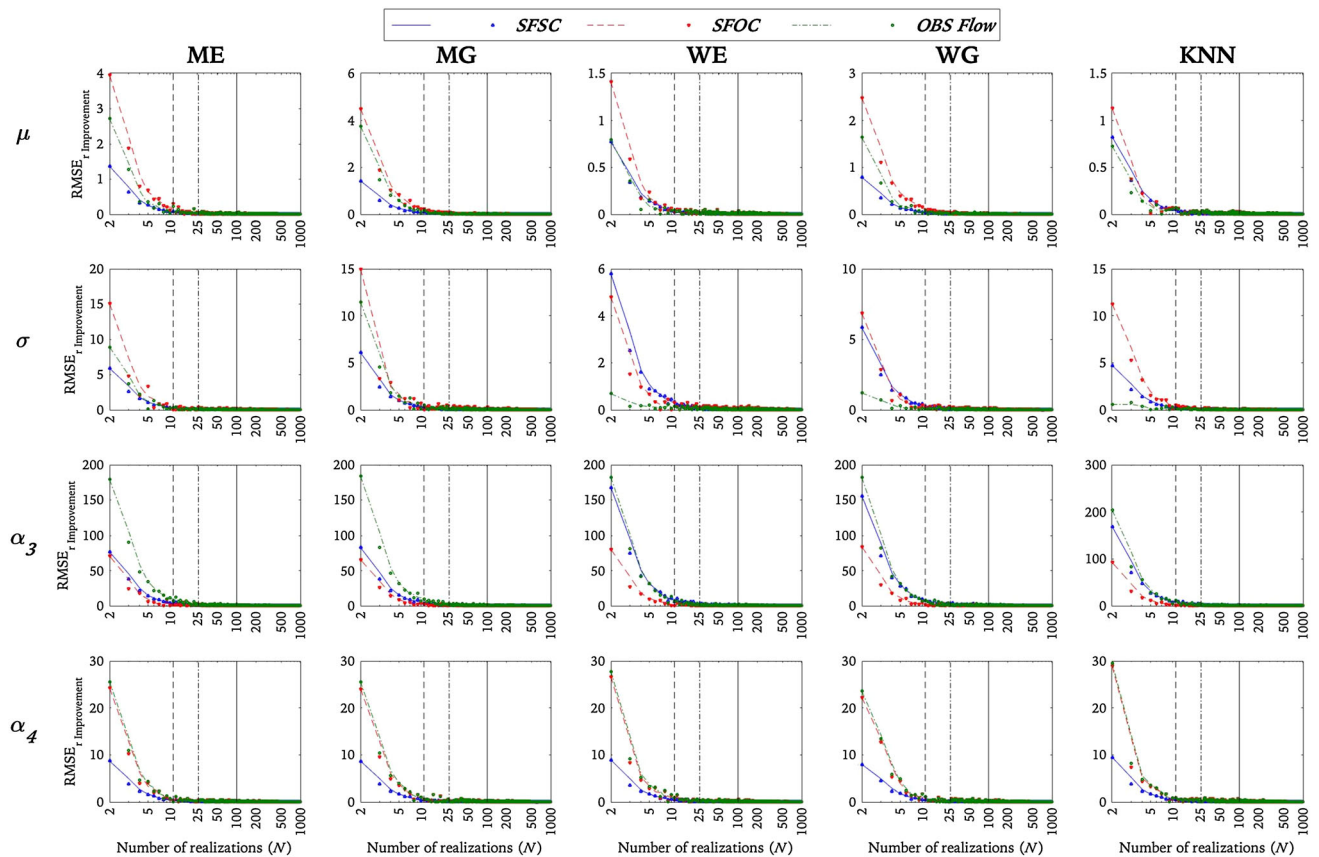
Fig. 13 Improvement in the *RMSEr's* of the main annual streamflow statistics of the *N*-realization samples generated by five SWGs. The *RMSEr* is calculated using either the observed flow (green line), the simulated flow using a synthetic climate (SFSC, purple line), simulated flow using the observed climate (SFSC, blue line). The *N-realization* SFSC sample is derived from 10,000 different randomly selected SFSC sets. Scattered markers represent actual results for which lines are slightly smoothed by moving averages with spans of 3. Vertical black dashed, dash-dotted, and solid lines represent 10, 25, and 100 realizations, respectively

experiment to determine the optimal number of realisations. Our findings are comforted by Guo et al. (2018), who found that the optimal number of realizations is 25 by using a different set of SWGs and risk parameters on a different watershed.

4 Conclusions

In summary, five SWGs coupled with a SWAT model are used to generate multiple time series for four hydroclimatic variables at four climatic stations and one hydrometric station on the South Nation Watershed located in Ontario, Canada. The investigated variables, including precipitation,

Table 5 Marginal improvements in *RMSEr* ($RMSE_{r,mar.improvement}$) of three reference datasets obtained by adding *N* realizations relative to a single realization

| SWG	N	Streamflow (SFOC)				Streamflow (SFSC)				Streamflow (OBS Flow)			
		μ	σ	α_3	α_4	μ	σ	α_3	α_4	μ	σ	α_3	α_4
ME	5	7.30	25.31	119.37	42.59	2.58	11.26	150.92	16.40	4.70	14.91	351.84	45.39
	10	8.75	27.55	134.23	49.14	3.18	13.90	185.64	20.42	5.38	17.73	430.93	52.83
	25	9.96	29.10	146.16	52.81	3.75	16.26	217.89	23.73	5.94	19.40	498.71	57.12
	50	10.35	30.31	149.08	53.81	4.01	17.49	233.76	25.55	6.07	19.58	529.46	58.41
	100	10.61	30.27	151.01	54.69	4.22	18.36	245.21	26.78	6.17	20.04	547.96	59.39
	1000	10.79	30.67	152.53	55.25	4.66	20.26	271.34	29.57	6.22	20.19	569.54	60.06
MG	5	8.25	22.32	114.69	42.13	2.63	10.90	157.65	16.49	6.62	19.32	344.81	45.42
	10	10.13	25.09	131.59	48.60	3.23	13.44	194.66	20.31	7.72	22.30	421.20	52.85
	25	11.78	26.54	139.32	52.87	3.80	15.67	226.59	23.86	8.56	24.30	484.25	57.87
	50	12.53	27.00	143.07	54.30	4.07	16.78	244.01	25.71	8.85	24.97	513.25	59.59
	100	13.01	27.43	144.61	55.02	4.27	17.59	255.97	26.95	8.99	25.21	529.91	60.44
	1000	13.56	27.65	146.05	55.63	4.72	19.42	282.57	29.77	9.13	25.55	547.54	61.19
WE	5	2.41	7.97	133.42	42.32	1.44	10.84	316.82	16.37	1.35	1.23	336.25	45.35
	10	2.78	9.13	153.59	49.63	1.79	13.54	391.76	20.56	1.54	1.43	403.62	53.63
	25	2.96	9.93	166.72	54.08	2.09	15.78	457.25	24.08	1.61	1.63	451.17	58.76
	50	3.03	9.98	171.59	55.88	2.25	16.99	492.42	25.89	1.65	1.57	470.14	60.79
	100	3.08	10.12	173.30	56.29	2.35	17.81	517.44	27.17	1.68	1.60	481.07	61.37
	1000	3.11	10.23	175.37	56.84	2.61	19.67	570.76	30.07	1.70	1.63	490.04	62.06
WG	5	4.66	11.46	139.17	44.84	1.48	10.82	295.18	16.53	2.73	2.12	337.98	47.66
	10	5.84	12.56	162.17	51.57	1.86	13.47	366.79	20.80	3.16	2.09	406.53	55.57
	25	6.78	13.67	173.43	55.76	2.18	15.70	427.75	24.50	3.48	2.36	457.92	60.60
	50	7.24	13.98	178.00	57.76	2.34	16.81	460.04	26.43	3.57	2.41	477.91	62.86
	100	7.54	14.08	180.45	58.16	2.45	17.67	483.34	27.69	3.61	2.40	488.35	63.44
	1000	7.92	14.18	182.23	59.04	2.72	19.53	533.48	30.73	3.65	2.41	498.54	64.46
KNN	5	1.73	21.30	152.63	43.95	1.56	9.19	314.20	17.24	1.06	1.72	371.78	46.04
	10	2.09	24.88	175.37	50.85	1.92	11.23	384.80	21.32	1.32	1.97	443.76	53.51
	25	2.16	27.81	186.78	55.40	2.24	13.25	449.57	24.99	1.34	2.30	503.86	58.45
	50	2.24	28.82	192.81	56.13	2.41	14.24	484.70	26.75	1.40	2.40	526.45	59.39
	100	2.24	29.25	196.46	56.78	2.53	14.96	508.68	28.11	1.39	2.39	536.67	60.13
	1000	2.27	29.69	198.02	57.60	2.79	16.53	560.89	31.08	1.41	2.42	548.85	61.02
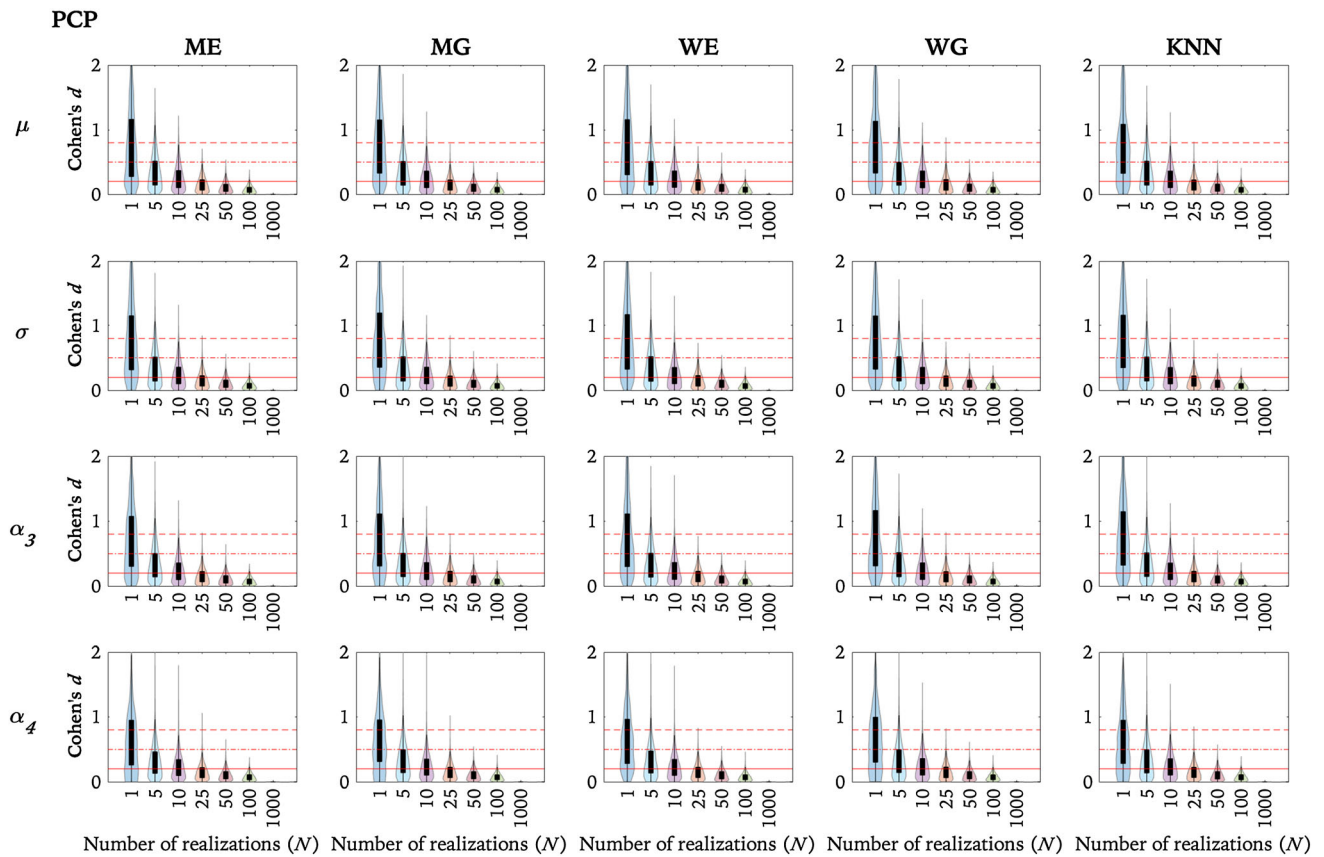
Fig. 14 Range of Cohen's *d* results of precipitation statistics where an *N-realization* sample is derived from 10,000 different randomly selected sets from the five SWGs. Horizontal red dashed, dash-dotted, and solid lines represent 0.8, 0.5, and 0.2 Cohen's *d* values, respectively

maximum and minimum air temperature and streamflow, are used to determine the optimal output size of the cloud representing SWGs. Four risk and performance indicators, namely, the mean, standard deviation, skewness, and kurtosis of these variables were estimated to assess the level of agreement between synthetic time series and observations. The number of realizations required to reach a predefined Relative Root Mean Square Error is then investigated to ultimately conduct computationally inexpensive impact studies. Using the two error metrics, namely, *RE* and *RMSE*, and the effect-size metric (Cohen's *d*), it was shown that when the number of realizations is high, the considered five weather generators perform somewhat similarly in terms of reproducing the risk and performance indicators. Overall, the results indicate that there is no very major benefit from generating more than 25 realizations in hydrological modeling. Applications requiring more precision (e.g., analysis of hydro-climatic extreme events)
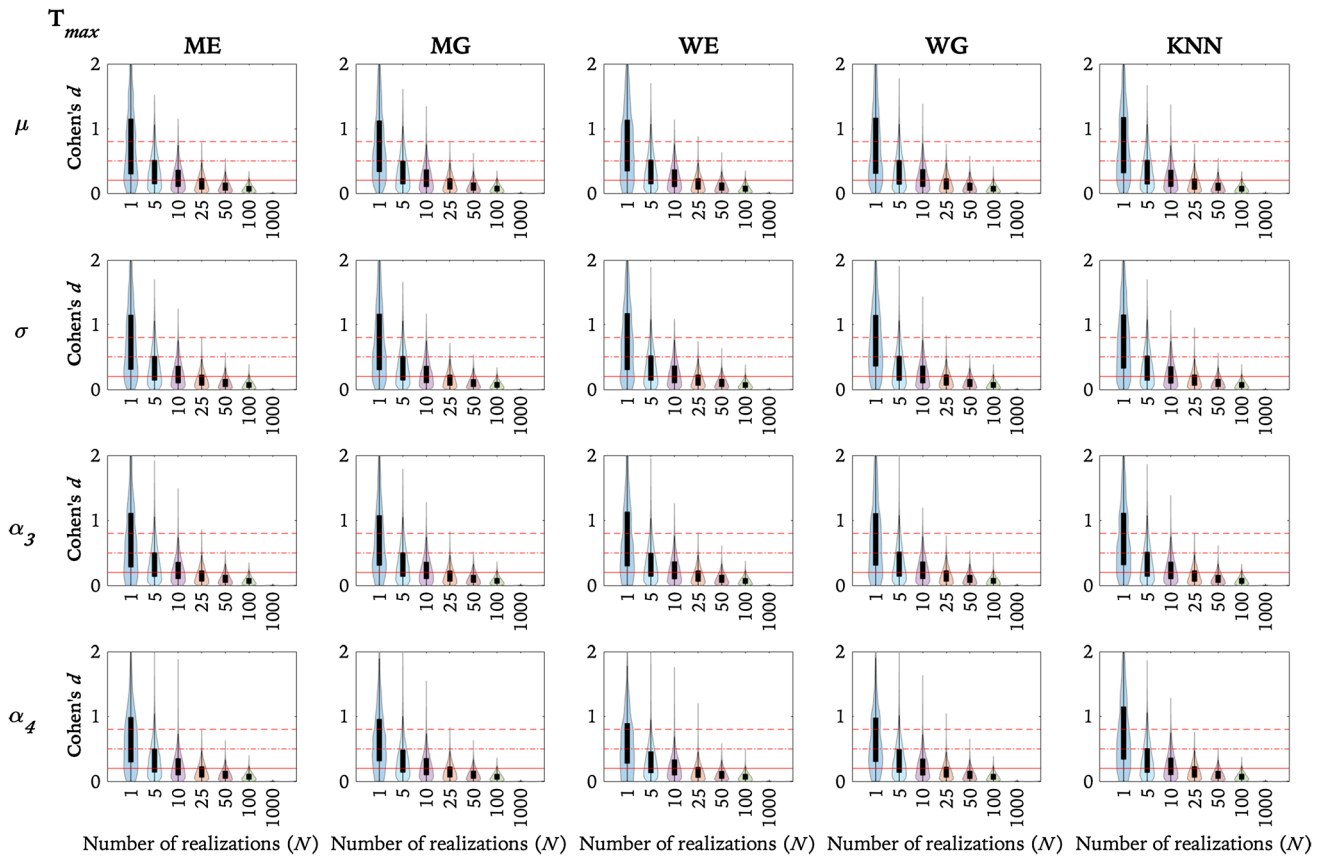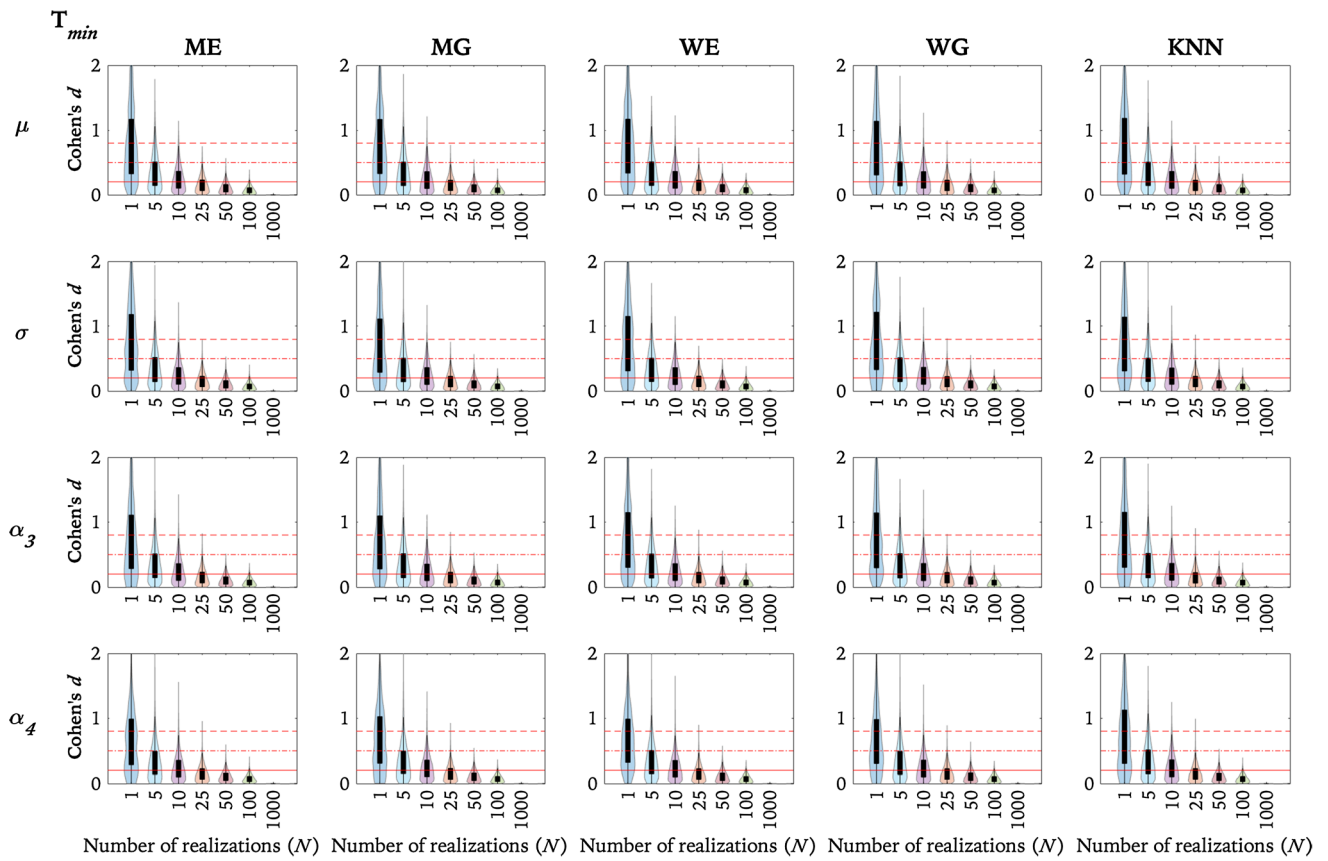
Fig. 15 Range of Cohen's d results of maximum temperature statistics where an N-realization sample is derived from 10,000 different randomly selected sets from the five SWGs. Horizontal red dashed, dash-dotted, and solid lines represent 0.8, 0.5, and 0.2 Cohen's d values, respectively
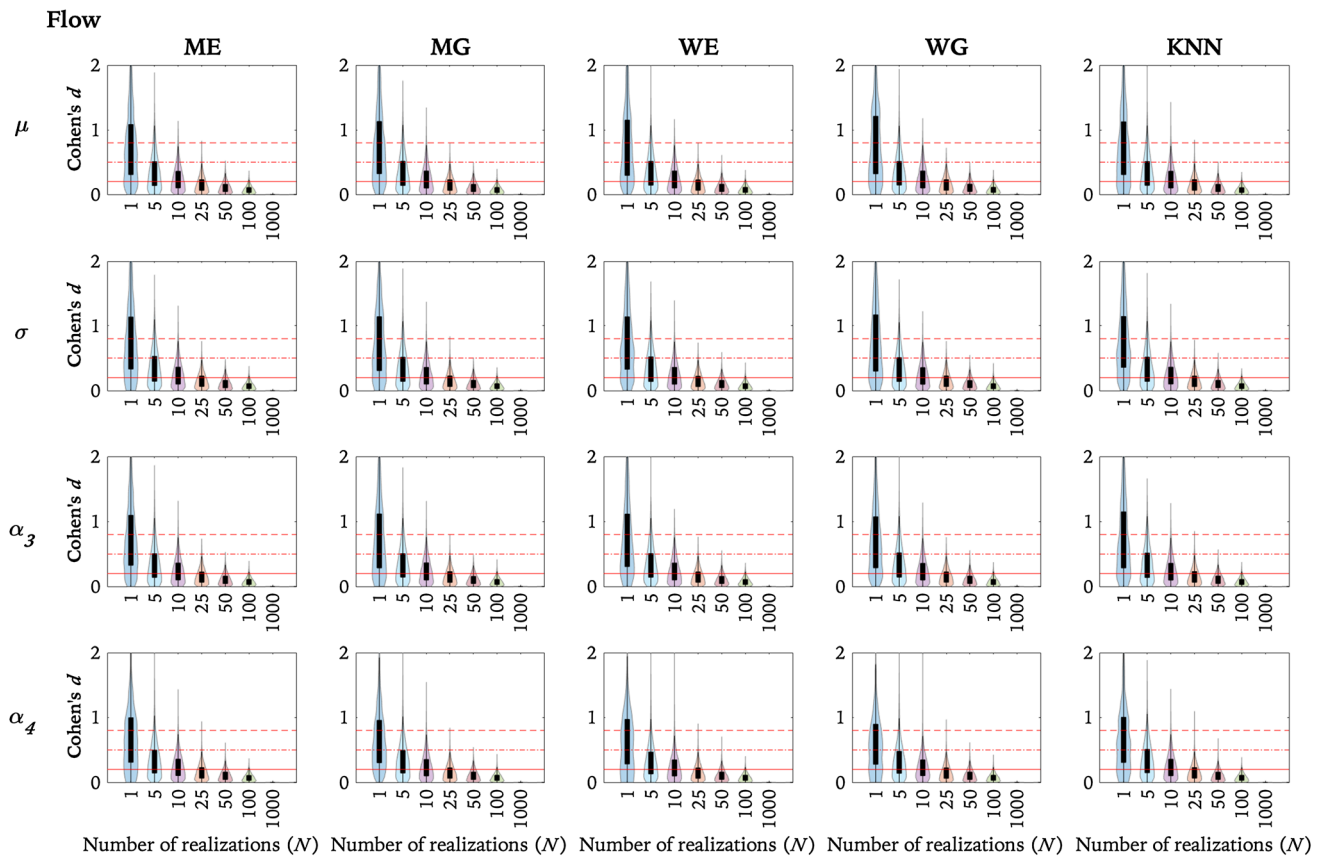
Fig. 16 Range of Cohen's *d* results of minimum temperature statistics where an *N-realization* sample is derived from 10,000 different randomly selected sets from the five SWGs. Horizontal red dashed, dash-dotted, and solid lines represent 0.8, 0.5, and 0.2 Cohen's *d* values, respectively

Flow

Fig. 17 Range of Cohen's d results of streamflow statistics where an *N-realization* sample is derived from 10,000 different randomly selected sets from the five SWGs. Horizontal red dashed, dash-dotted, and solid lines represent 0.8, 0.5, and 0.2 Cohen's d values, respectively

may use 100 realizations, as the results obtained from 100 realizations are not notably different from those obtained using 1000 realizations. Adopting a smaller, but carefully chosen, number of realizations can significantly reduce the workload on analysts and therefore benefit a larger audience in risk assessment studies, particularly when high-performance machines are not easily accessible.

References

Ailliot P, Allard D, Monbet V, Naveau P (2015) Stochastic weather generators: an overview of weather type models. J de la Soc Française de Stat 156(1):101–113

Alhassoun S, Sendil U, Al-Othman AA, Negm AM (1997) Stochastic generation of annual and monthly evaporation in Saudi Arabia. Can Water Resour J 22(2):141–154

Al-Mukhtar M, Dunger V, Merkel B (2014) Evaluation of the climate generator model CLIGEN for rainfall data simulation in Bautzen catchment area, Germany. Hydrol Res 45(4–5):615–630

Alodah A (2015). Development of Climate Change Scenarios for the South Nation Watershed. M.A.Sc. Thesis, Civil Engineering, Université d'Ottawa/University of Ottawa. http://dx.doi.org/10.20381/ruor-2801

Alodah A (2019). Stochastic assessment of climate-induced risk for water resources systems in a bottom-up framework. Doctoral dissertation. Civil Engineering, Université d'Ottawa/University of Ottawa. http://dx.doi.org/10.20381/ruor-24004

Alodah A, Seidou O (2019a) The adequacy of stochastically generated climate time series for water resources systems risk and performance assessment. Stoch Environ Res Risk Assess 33(1):253–269. https://doi.org/10.1007/s00477-018-1613-2

Alodah A, Seidou O (2019b) Assessment of climate change impacts on extreme high and low flows: an improved bottom-up approach. Water 11(6):1236. https://doi.org/10.3390/w11061236

Anyah RO, Semazzi FHM (2006) Climate variability over the Greater Horn of Africa based on NCAR AGCM ensemble. Theoret Appl Climatol 86(1–4):39–62

Apipattanavis S, Bert F, Podestá G, Rajagopalan B (2010) Linking weather generators and crop models for assessment of climate forecast outcomes. Agric For Meteorol 150(2):166–174

Baffault C, Nearing MA, Nicks AD (1996) Impact of CLIGEN parameters on WEPP-predicted average annual soil loss. Trans ASAE 39(2):447–457

Bastola S, Murphy C, Fealy R (2012) Generating probabilistic estimates of hydrological response for Irish catchments using a weather generator and probabilistic climate change scenarios. Hydrol Process 26(15):2307–2321

Benoit L, Vrac M, Mariethoz G (2019) Accounting for rain type non-stationarity in sub-daily stochastic weather generators. Hydrol. Earth Syst. Sci. Discuss. https://doi.org/10.5194/hess-2019-562

Box G, Jenkins G (1976) Time series analysis, forecasting and control, 2nd edn. Holden-Day, San Francisco

Breinl K, Di Baldassarre G, Girons Lopez M, Hagenlocher M, Vico G, Rutgersson A (2017) Can weather generation capture precipitation patterns across different climates, spatial scales and under data scarcity? Sci Rep 7(1):5449

Brown C, Werick W, Leger W, Fay D (2011) A decision-analytic approach to managing climate risks: application to the Upper Great Lakes 1. JAWRA J Am Water Resour Assoc 47(3):524–534

Buishand TA (1977) Stochastic modelling of daily rainfall sequences (Doctoral dissertation, Veenman)

Candela L, Tamoh K, Olivares G, Gomez M (2012) Modelling impacts of climate change on water resources in ungauged and data-scarce watersheds. Application to the Siurana catchment (NE Spain). Sci Total Environ 440:253–260

Caron A, Leconte R, Brissette F (2008) An improved stochastic weather generator for hydrological impact studies. Can Water Resour J 33:233–256

Chen J, Brissette FP (2014) Comparison of five stochastic weather generators in simulating daily precipitation and temperature for the loess plateau of China. Int J Climatol 34(10):3089–3105

Chen J, Brissette FP, Leconte R (2010) A daily stochastic weather generator for preserving low-frequency of climate variability. J Hydrol 388(3–4):480–490

Chen J, Brissette FP, Leconte R, Caron A (2012) A versatile weather generator for daily precipitation and temperature. Trans ASABE 55(3):895–906

Chen JF, Brissette X, Zhang J (2014) A multi-site stochastic weather generator for daily precipitation and temperature. Trans ASABE 2014:1375–1391. https://doi.org/10.13031/trans.57.10685

Cohen J (1988) Statistical power analysis for the behavioral sciences, 2nd edn. Lawrence Earlbaum Associates, Hillsdale

Criss RE, Winston WE (2008) Do Nash values have value? Discussion and alternate proposals. Hydrol Process 22(14):2723–2725. https://doi.org/10.1002/hyp.7072

Cunha A Jr, Nasser R, Sampaio R, Lopes H, Breitman K (2014) Uncertainty quantification through the Monte Carlo method in a cloud computing setting. Comput Phys Commun 185(5):1355–1363

Deser C, Phillips A, Bourdette V, Teng H (2012) Uncertainty in climate change projections: the role of internal variability. Clim Dyn 38(3):527–546. https://doi.org/10.1007/s00382-010-0977-x

Dubrovský M, Buchtele J, Žalud Z (2004) High-frequency and low-frequency variability in stochastic daily weather generator and its effect on agricultural and hydrologic modelling. Clim Change 63(1–2):145–179

Eames M, Kershaw T, Coley D (2012) A comparison of future weather created from morphed observed weather and created by a weather generator. Build Environ 56:252–264

Elliot W, Arnold C (2001) Validation of the weather generator CLIGEN with precipitation data from Uganda. Trans ASAE 44:53–58

Fatichi S, Ivanov VY, Paschalis A, Peleg N, Molnar P, Rimkus S, Kim J, Burlando P, Caporali E (2016) Uncertainty partition challenges the predictability of vital details of climate change. Earth's Future 4(5):240–251. https://doi.org/10.1002/2015EF000336

Fodor N, Dobi I, Mika J, Szeidl L (2013) Applications of the MVWG multivariable stochastic weather generator. Sci World J. https://doi.org/10.1155/2013/571367

Furrer EM, Katz RW (2008) Improving the simulation of extreme precipitation events by stochastic weather generators. Water Resour Res. https://doi.org/10.1029/2008WR007316

Gabriel KR, Neumann J (1962) A Markov chain model for daily rainfall occurrence at Tel Aviv. Q J R Meteorol Soc 88:90–95

Gitau MW, Chiang LC, Sayeed M, Chaubey I (2012) Watershed modeling using large-scale distributed computing in Condor and the Soil and Water Assessment Tool model. Simulation 88(3):365–380

Gitau MW, Mehan S, Guo T (2018) Weather generator effectiveness in capturing climate extremes. Environ Process 5(1):153–165

Glass GV (1976) Primary, secondary, and meta-analysis of research. Educ Res 5(10):3–8

Goyal MK, Burn DH, Ojha CSP (2013) Precipitation simulation based on k-nearest neighbor approach using gamma kernel. J Hydrol Eng 18:481–487

Guenni L (1994) Spatial interpolation of the parameters of stochastic weather models. Climate Change, Uncertainty and Decision Making, Institute for Risk Research, University of Waterloo, Ontario and IGBP-BAHC, Berlin, 61–79

Guo T, Mehan S, Gitau MW, Wang Q, Kuczek T, Flanagan DC (2018) Impact of number of realizations on the suitability of simulated weather data for hydrologic and environmental applications. Stoch Environ Res Risk Assess 32(8):2405–2421

Han B, Benner SG, Flores AN (2019) Including Variability across Climate Change Projections in Assessing Impacts on Water Resources in an Intensively Managed Landscape. Water 11(2):286

Hansen JW, Ines AV (2005) Stochastic disaggregation of monthly rainfall data for crop simulation studies. Agric For Meteorol 131:233–246

Hayhoe HN (1998) Relationship between weather variables in observed and WXGEN generated data series. Agric For Meteorol 90:203–214

Hayhoe HN (2000) Improvements of stochastic weather data generators for diverse climates. Clim Res 14(2):75–87. https://doi.org/10.3354/cr014075

Ivanov VY, Bras RL, Curtis DC (2007) A weather generator for hydrological, ecological, and agricultural applications. Water Resour Res 43(W10):406. https://doi.org/10.1029/2006WR005,364

Jeong DI, St-Hilaire A, Ouarda TB, Gachon P (2012) Multisite statistical downscaling model for daily precipitation combined by multivariate multiple linear regression and stochastic weather generator. Clim Change 114(3–4):567–591

Kilsby CG, Jones PD, Burton A, Ford AC, Fowler HJ, Harpham C, Wilby RL (2007) A daily weather generator for use in climate change studies. Environ Model Softw 22(12):1705–1719

Kim BS, Kim HS, Seoh BH, Kim NW (2007) Impact of climate change on water resources in Yongdam Dam Basin, Korea. Stoch Environ Res Risk Assess 21:355

Kim J, Tanveer ME, Bae DH (2018) Quantifying climate internal variability using an hourly ensemble generator over South Korea. Stoch Environ Res Risk Assess 32(11):3037–3051

Kou X, Ge J, Wang Y, Zhang C (2007) Validation of the weather generator CLIGEN with daily precipitation data from the Loess Plateau, China. J Hydrol 347(3–4):347–357

Kwon HH, Sivakumar B, Moon YI, Kim BS (2011) Assessment of change in design flood frequency under climate change using a multivariate downscaling model and a precipitation-runoff model. Stoch Environ Res Risk Assess 25(4):567–581

Mehan S, Guo T, Gitau M, Flanagan DC (2017) Comparative study of different stochastic weather generators for long-term climate data simulation. Climate 5(2):26

Mehrotra R, Srikanthan R, Sharma A (2006) A comparison of three stochastic multi-site precipitation occurrence generators. J Hydrol 331(1–2):280–292

Min YM, Kryjov VN, An KH, Hameed SN, Sohn SJ, Lee WJ, Oh JH (2011) Evaluation of the weather generator CLIGEN with daily precipitation characteristics in Korea. Asia-Pacific J Atmos Sci 47(3):255–263

Minville M, Brissette F, Leconte R (2008) Uncertainty of the impact of climate change on the hydrology of a nordic watershed. J Hydrol 358:70–83

Mithen S, Black E (2011) Water, life and civilisation: climate, environment and society in the Jordan Valley. Cambridge University Press, Cambridge

Mukundan R, Acharya N, Gelda RK, Frei A, Owens EM (2019) Modeling streamflow sensitivity to climate change in New York City water supply streams using a stochastic weather generator. J Hydrol: Reg Stud 21:147–158

Neitsch SL, Arnold JG, Kiniry JR, Williams JR (2011) Soil and water assessment tool theoretical documentation version 2009. Texas Water Resources Institute

Onof C, Chandler RE, Kakou A, Northrop P, Wheater HS, Isham V (2000) Rainfall modelling using Poisson-cluster processes: a review of developments. Stoch Environ Res Risk Assess 14(6):384–411

Oriani F, Straubhaar J, Renard P, Mariethoz G (2014) Simulation of rainfall time-series from different climatic regions using the direct sampling technique. Hydrol Earth Syst Sci Discuss 11:3213–3247

Ott RL, Longnecker MT (2015) An introduction to statistical methods and data analysis. Thomson Learning, Inc. Duxbury, Pacific Grove, CA, Nelson Education

Panagoulia D (2006) Artificial neural networks and high and low flows in various climate regimes. Hydrol Sci J 51(4):563–587

Panagoulia D, Tsekouras GJ, Kousiouris G (2017) A multi-stage methodology for selecting input variables in ANN forecasting of river flows. Glob Nest J 19:49–57

Peleg N, Fatichi S, Paschalis A, Molnar P, Burlando P (2017) An advanced stochastic weather generator for simulating 2-D high-resolution climate variables. J Adv Model Earth Syst 9(3):1595–1627

Qian B, Gameda S, Hayhoe H, De Jong R, Bootsma A (2004) Comparison of LARS-WG and AAFC-WG stochastic weather generators for diverse Canadian climates. Clim Res 26(3):175–191

Racsko P, Szeidl L, Semenov M (1991) A serial approach to local stochastic weather models. Ecol Model 57(1–2):27–41

Räisänen J, Ruokolainen L (2006) Probabilistic forecasts of near-term climate change based on a resampling ensemble technique. Tellus A: Dyn Meteorol Oceanogr 58(4):461–472

Richardson CW (1981) Stochastic simulation of daily precipitation, temperature, and solar radiation. Water Resour Res 17(1):182–190

Rodriguez-Iturbe I, Cox DR, Isham V (1987) Some models for rainfall based on stochastic point processes. Proc R Soc Lond A. Math Phys Sci 410(1839):269–288

Rosnow RL, Rosenthal R (1996) Computing contrasts, effect sizes, and counternulls on other people's published data: general procedures for research consumers. Pyschol Methods 1:331–340

Santer BD, Thorne PW, Haimberger L, Taylor KE, Wigley TML, Lanzante JR, Karl TR (2008) Consistency of modelled and observed temperature trends in the tropical troposphere. Int J Climatol: J R Meteorol Soc 28(13):1703–1722

Schoof JT, Pryor SC (2008) On the proper order of markov chain model for daily precipitation occurrence in the contiguous

United States. J Appl Meteorol Climatol 47(9):2477–2486. https://doi.org/10.1175/2008JAMC1840.1

Semenov MA (2008) Simulation of extreme weather events by a stochastic weather generator. Clim Res 35(3):203–212

Semenov MA, Barrow EM (1997) Use of a stochastic weather generator in the development of climate change scenarios. Clim Change 35:397–414

Semenov MA, Brooks RJ, Barrow EM, Richardson CW (1998) Comparison of the WGEN and LARS-WG stochastic weather generators for diverse climates. Clim Res 10(2):95–107

Sharif M, Burn DH (2007) Improved K-nearest neighbor weather generating model. J Hydrol Eng 12(1):42–51

Smith BJ (2007) BOA: an R package for MCMC output convergence assessment and posterior inference. J Stat Softw 21(11):1–37

Soltani A, Hoogenboom G (2003a) A statistical comparison of the stochastic weather generators WGEN and SIMMETEO. Clim Res 24(3):215–230

Soltani A, Hoogenboom G (2003b) Minimum data requirements for parameter estimation of stochastic weather generators. Clim Res 25(2):109–119. https://doi.org/10.3354/cr025109

Steinschneider S, Brown C (2013) A semiparametric multivariate, multisite weather generator with low-frequency variability for use in climate risk assessments. Water Resour Res 49(11):7205–7220

Stevens T, Madani K (2016) Future climate impacts on maize farming and food security in Malawi. Sci Rep 6:36241

Stowasser Markus (2012) Modelling rain risk: a multi-order markov chain model approach. J Risk Finance 13(1):45–60

Thompson C, Thomson P, Zheng X (2007) Fitting a multisite rainfall model to New Zealand data. J Hydrol 340:25–39

Thompson DWJ, Barnes EA, Deser C, Foust WE, Phillips AS (2015) Quantifying the role of internal climate variability in future climate trends. J Clim 28(16):6443–6456. https://doi.org/10.1175/JCLI-D-14-00830.1

Todorovic P, Woolhiser DA (1975) A stochastic model of n-day precipitation. J Appl Meteorol 14:17–24

Vacha-Haase T, Thompson B (2004) How to estimate and interpret various effect sizes. J Couns Psychol 51:473–481. https://doi.org/10.1037/0022-0167.51.4.473

Vesely FM, Paleari L, Movedi E, Bellocchi G, Confalonieri R (2019) Quantifying uncertainty due to stochastic weather generators in climate change impact studies. Sci Rep 9:1–8

Vu TM, Mishra AK, Konapala G, Liu D (2018) Evaluation of multiple stochastic rainfall generators in diverse climatic regions. Stoch Environ Res Risk Assess 32(5):1337–1353

Wheater HS, Chandler RE, Onof CJ, Isham VS, Bellone E, Yang C, Segond ML (2005) Spatial-temporal rainfall modelling for flood risk estimation. Stoch Environ Res Risk Assess 19(6):403–416

Wilby RL, Charles SP, Zorita E, Timbal B, Whetton P, Mearns LO (2004) Guidelines for use of climate scenarios developed from statistical downscaling methods. Supporting material of the Intergovernmental Panel on Climate Change, available from the DDC of IPCC TGCIA, 27

Zhang X, Garbrecht JD (2003) Evaluation of CLIGEN precipitation parameters and their implication on WEPP runoff and erosion prediction. Trans ASAE 46:311

Zwiers F (1996) Interannual variability and predictability in an ensemble of AMIP climate simulations conducted with CCC GCM2. Clim Dyn 12:825–847