



Performance of multiple probability distributions in generating daily precipitation for the simulation of hydrological extremes

Biyang Zhu¹ · Jie Chen¹ · Hua Chen¹

Published online: 19 August 2019

© Springer-Verlag GmbH Germany, part of Springer Nature 2019

Abstract

Stochastic weather generators are statistical models widely used to produce climate time series with similar statistical properties to observed data. They are also used as downscaling tools to generate climate change scenarios for impact studies. Precipitation is one of the main variables simulated by weather generators and is also a key variable for impact studies, especially for hydrology. Precipitation is usually simulated by multiple precipitation models which have been proposed for simulating site-specific precipitation. However, these models' performance in simulating watershed-averaged extreme precipitation, especially in representing hydrological extremes, has not been well-investigated. Accordingly, this study compares the performance of six probability distributions (exponential, gamma, skewed normal, mixed exponential, hybrid exponential/Pareto, and Weibull distributions) and a polynomial curve-fitting method in generating precipitation for simulating hydrological extremes over three basins using a set of extreme indices. The results show that except for the exponential distribution (EXP), all of the methods produce the distribution of observed precipitation at the daily, monthly and annual scales reasonably well for all three river basins. Although the three-parameter hybrid exponential/Pareto distribution (EXPP) overestimates precipitation extremes, other three-parameter models produce extremes accurately. The three-parameter mixed exponential (MEXP) distribution outperforms other models for simulating precipitation extremes. However, with respect to representing hydrological extremes, the MEXP distribution is not the best model. When simulating extreme streamflows with synthetic weather data, the EXP distribution shows the worst performance, while the curve fitting method (PN) performs the best. The inferiority of the EXPP distribution in generating extreme precipitation does not propagate to extreme flow simulations. Meanwhile, the performance of WEB is moderate in terms of representing hydrological extremes. Overall, finding the model that best reproduces precipitation for simulating hydrological extremes is not as clear-cut, since the performance of each model is extreme-indices dependent. Taking all of the indices into account, the MEXP and the PN appear to be superior in representing extreme precipitation and hydrological extremes.

Keywords Precipitation probability distributions · Weather generator · Extreme precipitation · Hydrological modeling

1 Introduction

Long-term precipitation series are usually indispensable for analyzing risk-based hydrological extremes, such as precipitation and flood frequency analysis. However, long periods of observed data are usually not available, especially in northern river basins. To solve this problem, stochastic weather generators have been developed to produce arbitrary lengths of consecutive precipitation data with similar characteristics of the observed data. Even for ungauged regions, weather data can also be generated by interpolating weather generator parameters from adjacent gauged sites (Baffault et al. 1996; Semenov and Brooks

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s00477-019-01720-z>) contains supplementary material, which is available to authorized users.

✉ Jie Chen
jiechen@whu.edu.cn

¹ State Key Laboratory of Water Resources and Hydropower Engineering Science, Wuhan University, Wuhan 430072, People's Republic of China

1999; Wilks 2008). More importantly, weather generators are usually implemented as downscaling tools to generate climate change scenarios for impact studies.

Weather generators usually produce daily precipitation amounts using probability distributions (Woolhiser and Roldán 1982; Harmel et al. 2000; Wan et al. 2005; Jamaludin and Jemain 2007; Liu et al. 2011; Li et al. 2013b, 2014; Chen and Brissette 2014b). They usually include single distributions (e.g. exponential, gamma, Weibull and skewed normal distributions) and mixed distributions (e.g. mixed exponential, hybrid exponential and Pareto distributions). Several studies have evaluated these distributions for their accuracy in generating daily precipitation amounts, especially for extremes at specific sites. For example, Todorovic and Woolhiser (1975) found that the exponential distribution was suitable to produce small and medium precipitation, but not for extremes. While gamma distribution accurately generates low-to-moderate values of precipitation, it underestimates the likelihood of large precipitation (Furrer and Katz 2008). Extreme frequency underestimation was also observed when using the Weibull distribution (Safeeq and Fares 2011). However, various studies have shown that mixed distributions outperform single distributions at simulating extreme precipitation (Li et al. 2012; Chen and Brissette 2014b). For example, Jamaludin and Jemain (2007) compared mixed exponential, mixed gamma, mixed Weibull distributions and their single counterparts and found that the mixed distributions consistently outperform their single counterparts at generating precipitation extremes. More recently, Chen et al. (2015) compared the abilities of gamma, mixed exponential, and hybrid exponential and Pareto distributions in producing entire ranges of precipitation distribution, and found mixed distributions perform better than the gamma distribution for simulating the upper tail of distribution. Furthermore, Chen et al. (2015) also used a polynomial curve fitting approach for capturing extreme precipitation events. A comparison with multiple methods showed the curve fitting method is more accurate than probability distributions at producing the entire range of daily precipitation, especially for extremes.

Precipitation extremes vary drastically in temporal and spatial scales (French et al. 1992; Venugopal et al. 1999), resulting in different performance of each method in generating regional precipitation extremes. For example, Wan et al. (2005) tested exponential, gamma, skewed normal and mixed exponential distributions over 657 stations in Canada, and found that gamma and mixed exponential distributions perform better in winter and warmer months, respectively. Li et al. (2014) found that the skewed normal distribution is appropriate at generating extreme precipitation for the Loess Plateau of China. In addition, the precipitation generator is usually used as a downscaling

tool for assessing climate change impacts on hydrology. For hydrological climate change impact studies, the climate change scenarios generated by a weather generator have been used as inputs to a hydrological model at the watershed scale (Dibike and Coulibaly 2005; Zhang and Liu 2005; Minville et al. 2008; Chen et al. 2010, 2011). When using a lumped hydrological model for impact studies, watershed-averaged precipitation scenarios are required. Even though the performance of precipitation models has been extensively investigated in many studies, they are usually only evaluated for a specific site, rather than at the watershed scale. The performance of precipitation models may be different between station and watershed scales, because the appropriate selection of distribution may be changed when averaging station data to watershed mean. For example, extreme precipitation events averaged in watersheds may not be that extreme as those at stations, particularly for large watersheds. Thus, it is necessary to specifically evaluate the performance of these models for simulating watershed-averaged precipitation, especially for extremes. Moreover, the performance of precipitation models in representing hydrological simulations also needs further investigation.

The performance of precipitation models in representing hydrological extremes has been evaluated in only a few studies. For example, Khalili et al. (2011) compared exponential distribution based single-site and multi-site weather generators for hydrological modelling and found that the exponential distribution performs reasonably well with respect to representing the monthly flow, while underestimating extreme flow in summer and autumn. Li et al. (2013b) investigated the effectiveness of six probability distributions for hydrological modelling and found that the skewed normal distribution performs better than other models at producing runoff volumes, whereas the mixed exponential distribution shows the best performance at producing extremes. Chen et al. (2017) found that the gamma distribution-based model is appropriate at simulating hydrological median and extremes. Even though the above studies attempted to evaluate the performance of precipitation models in representing hydrological extremes, only one or two extreme metrics were used. Studies that specifically investigate the performance of multiple precipitation models in representing hydrological extremes, combine with evaluations using multiple extreme metrics are rare indeed.

The objectives of this study are to (1) evaluate the performance of multiple precipitation models in generating extreme precipitation at the watershed scale, and (2) evaluate the performance of these models in representing hydrological extreme by using a hydrological model. In addition, the spatial scale effects of precipitation models are investigated by using three river basins with different

drainage areas, as the watershed-averaged precipitation extremes may be affected by watershed size.

2 Data and methods

2.1 Study area and data

Three river basins with different drainage areas located in the Xiangjiang watershed are used in these studies (Fig. 1). It is to insure that they have similar precipitation conditions. The differences in averaged precipitation extremes may be mainly a result from the differences in drainage areas. The streamflow data of the three sub-basins are recorded at the Daxitan, Xiangxiang and Ganxi control stations, which have drainage areas of 3132 km², 6053 km² and 9972 km², respectively. The climate of the Xiangjiang watershed is dominated by a tropical monsoon with maximal precipitation occurring in May and June. The mean annual precipitation and temperature are presented in Table 1 for all three river basins.

Daily precipitation, maximum and minimum temperature (1976–2005) collected from China Meteorological Data Service Center were used to assess the performance of precipitation models. Mean annual precipitation magnitudes vary between 1300 and 1500 mm for all three river basins. The temperature is similar for all three river basins. When evaluating streamflow simulations with synthetic precipitation, observed daily streamflows are required for calibrating the hydrological model. The daily streamflow time series were provided by the Department of Water Resources of Hunan Province for the same period. Streamflow has significant seasonal characteristics, with peak flow series occurring in June during the rainy season (April–September).

2.2 Precipitation generator

Daily precipitation amounts are generated by six parametric distributions and a curve-fitting method. Since precipitation time series are required to drive a hydrological model, a Markov chain based model (Hann et al. 1976; Richardson 1981; Wilks 1989, 1999, 2008; Furrer and Katz

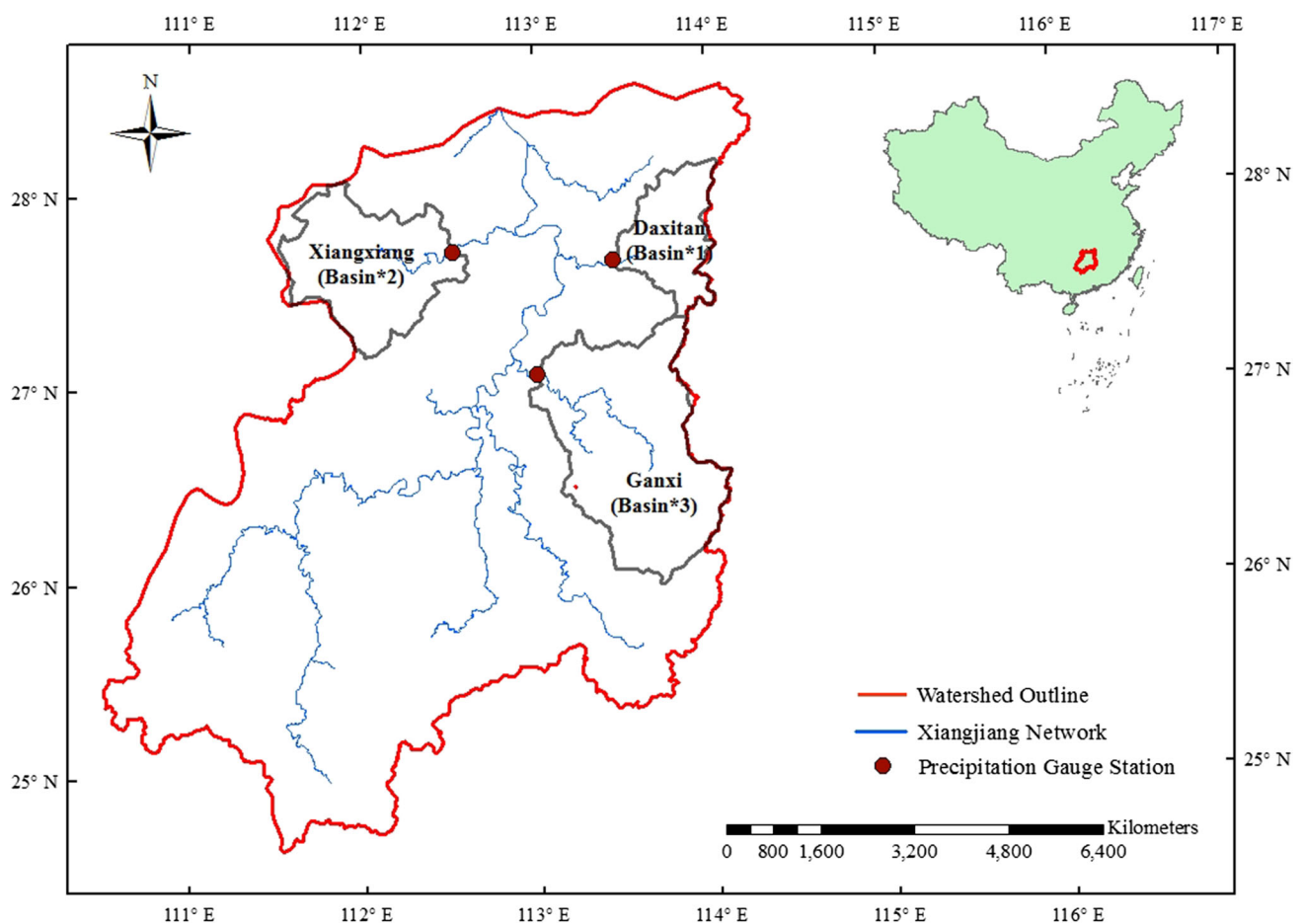


Fig. 1 Location of three river basins in the Xiangjiang watershed and their precipitation gauge control stations

Table 1 General information of the three river basins

River basin	Gauge station	Area (km ²)	Mean annual precipitation (mm)	Mean annual Tmax (°C)	Mean annual Tmin (°C)	Mean annual flow (m ³ /s)
*1	Daxitan	3132	1492	21.23	13.53	88
*2	Xiangxiang	6053	1364	21.33	13.46	263
*3	Ganxi	9972	1559	20.78	13.62	115

2008; Chen et al. 2010; Chen and Brissette 2014a, b) is first applied to generate the sequence of precipitation occurrence for each calendar month. In order to fairly compare all precipitation generators, the same precipitation occurrence is used for all precipitation models. The results may not be exact the same when using various sequences. Thus, it is necessary to run the same order Markov chain seven times to generate different sequences to feed different precipitation models. Previous studies (e.g. Chen et al. 2010; Li et al. 2013a) have shown that the first-order Markov chain performs reasonable well with respect to reproduce the sequence of precipitation occurrence. Some other studies (Chen and Brissette 2014a; Acharya et al. 2017; Vu et al. 2018) also used the second or third-order Markov chains, the advantage, however, is limited.

Having found the precipitation occurrence, precipitation amounts on wet days are modeled using six probability distributions and a curve-fitting method. The six parametric distributions include one-parameter exponential distribution (EXP), two-parameter gamma (GAM) and Weibull (WEB) distributions and three-parameter skewed normal (SN), mixed exponential (MEXP) and hybrid exponential/Pareto (EXPP) distributions. The six precipitation models can also be classified into four single distributions (EXP, GAM, WEB and SN) and two mixed distributions (MEXP and EXPP). In addition, a second-order polynomial-based curve fitting method (PN) is also used for comparison (Chen et al. 2015). More details of these methods are presented in the supplementary material. The parameters of the probability distributions and the curve-fitting method are estimated by the maximum likelihood method for every two-week period over 30 years, except for the SN distribution where the method of moments is applied for parameter estimation. In order to obtain the true expectancy of a weather generator, the length of 100 years is generate using each method. Short time series could result in biases due to the random nature of the stochastic process. Previous studies (e.g. Richardson 1981; Furrer and Katz 2008) have shown that 100 years is sufficient to represent the true expectancy of a weather generator.

It is worth noting that random number control is required to avoid the generation of outliers when using SN and EXPP distributions (Chen and Brissette 2014b). As for

SN, random number control is particularly necessary because negative values exist when the absolute skew coefficient amount is greater than 4.5 (Meyer 2011; Li et al. 2013b). Unreasonable extreme precipitation events may be generated by the EXPP distribution because the GP distribution in EXPP produces a positive tail index, leading to the overestimation of extreme precipitation (Li et al. 2013b; Chen and Brissette 2014b). According to previous studies, an appropriate solution is to insure synthetic precipitation amounts lower than the probable maximum precipitation (PMP) (Chen and Brissette 2014b; Chen et al. 2015).

Temperatures are required to drive a hydrological model to simulate streamflow time series. Maximum and minimum temperatures are generated by a WeaGETS (Chen et al. 2012) weather generator, which produces daily temperature time series using a first-order linear autoregressive model (Richardson 1981).

2.3 Hydrological modeling

The Xin'anjiang (XAJ) model is used for hydrological simulations in this study (Zhao 1992). XAJ is a conceptual-lumped hydrological model widely applied in humid and semi-humid regions of China for hydrological modeling and forecasts (Li et al. 2009, 2012; Shi et al. 2011). The XAJ model consists of three sub-models: a three-layer evapotranspiration sub-model, a runoff generation sub-model and a runoff routing sub-model (Li et al. 2009; Zhao 1992; Lü et al. 2013). The runoff is separated into three water sources: surface flow, interflow and subsurface flow. The model includes 16 free parameters with four parameters that account for evapotranspiration, three for runoff-yield, four for separation and five for confluence.

The XAJ model requires daily precipitation and evaporation as inputs. Since the gauged evaporation is not available, it is calculated by Hargreaves and Samani's scheme (1985) using observed maximum and minimum temperature. The gauged daily streamflow time series is needed for model calibration and validation. As For observed flow, a 20-year observed streamflow time series (1976–1995) is used for model calibration and a 10-year time series (1996–2005) is used for model validation.

Model parameters are optimized using the Shuffled Complex Evolution Method (Duan et al. 1993) with the Nash–Sutcliffe coefficient (NSE) as an objective function. The NSEs are 0.90, 0.85 and 0.86 for calibration and 0.87, 0.89 and 0.87 for validation over the Daxitan, Xiangxiang and Ganxi river basins, respectively. Overall, the XAJ model performs reasonably well for all three river basins, with NSEs being greater than 0.85.

2.4 Statistical analysis

The precipitation models are first evaluated with respect to reproducing the mean, standard deviation (Std), interquartile range (IQR) and skewness coefficients (CS) of daily, monthly and annual precipitation. The mean relative errors and mean absolute relative errors are calculated for these statistics. The precipitation models' performance in representing extreme precipitation is evaluated using three extreme indices: annual maximum daily precipitation (Rx1day), and total annual precipitation resulting from events exceeding the 95th (Rx95TOT) and the 99th percentile (Rx99TOT). The Kolmogorov–Smirnov (KS) test is used to test the similarity of distributions between observed and synthetic data at a significant level of $P = 0.05$. Using synthetic precipitation, daily and monthly streamflows are simulated by the XAJ model and compared to observed streamflow. To avoid errors resulting from the hydrological model, streamflow generated using observed weather data are used to represent observed flow (Li et al. 2013b). Five extreme hydrological indices are also used to evaluate the synthetic precipitation in representing streamflow. These indices are the mean duration of high pulses (HPD), the number of high pulses (HPC), the annual maximum 1-day flow (Qx1day), the annual maximum 3-day flow (Qx3day) and the annual maximum 7-day flow (Qx7day). Among them, the HPD and HPC are two high pulse-based indices. A high pulse is defined as flow series in a year exceeding the 75th percentile of annual flow.

3 Results

3.1 Precipitation amount generation

3.1.1 Annual and monthly precipitation characteristics

Table 2 shows the relative error of the mean and standard deviation (Std) of the monthly and annual precipitation simulated by all seven precipitation models. All models produce the mean of annual and monthly precipitation very well, with absolute relative errors (AREs) consistently less than 6%. Among the seven precipitation models, the PN method produces the mean of the monthly (and yearly) streamflows with the largest error. Although little

differences exist among the models in term of generating mean values, there are significant differences in generating the Stds of monthly and annual precipitation. The difference between the Stds of observed and generated annual precipitation over basin *3 is larger than that over the other two river basins. The performances of the SN and PN models are superior in producing the Std for all river basins, especially for basin *2. Furthermore, the KS test shows that the EXP- and PN-generated precipitation time series are significantly different from observed time series for basin *1 and basin *2. In general, the three-parameter distributions perform better than the one- and two-parameter distributions for all three river basins.

3.1.2 Daily precipitation characteristics

Figure 2 presents four statistical metrics (mean, Std, IQR and CS) of observed and generated daily precipitation. All seven models perform similarly with respect to reproducing the mean values of daily precipitation for all three river basins. The EXP underestimates the Std by more than 20%, whereas the other models simulate the Std of daily precipitation with relative errors (REs) ranging between -10 and 10% for all three river basins. In terms of the IQR, among all the models, the EXPP performs the worst with a mean relative error (MRE) of IQR ranging between 13.13 and 22.53% for all three river basins. The GAM, SN and MEXP models generate CS of daily precipitation accurately, with AREs less than 3%. WEB and EXPP appear to be worse than the other models, especially for basin *1 with REs of 25.05% and 93.06% , respectively. The over-estimation of CS may be because some models generate too many high values, especially when using the EXPP.

3.2 Extreme precipitation events

Figure 3 presents three extreme indices (Rx1day, Rx95pTOT and Rx99pTOT) of observed daily precipitation and synthetic daily precipitation generated by seven precipitation models. Generally, the EXP performs the worst in simulating extreme indicators, with REs of medium values of -30.56% , -24.75% and -21.52% for Rx1day, Rx95pTOT and Rx99pTOT, respectively. The WEB and MEXP models perform better than the others, with AREs of medians being less than 19% for the three extreme indices. When using the EXPP, the medium values of Rx1day are overestimated by more than 10% for all river basins, whereas those of Rx95TOT are accurately simulated with RE being less than 5%. Generally, none of the models generate the inter-annual variation of extreme indices accurately for all three river basins, even though SN and MEXP perform better than the other models for both Rx1day and Rx99TOT. The EXPP and PN models

Table 2 Mean and standard deviation (Std) of annual and monthly observed precipitation and precipitation generated by seven models (EXP, GAM, WEB, SN, MEXP, EXPP and PN) for three river basins

Time scale	River basin	Source	Observed precipitation	Relative error (%)						
				EXP	GAM	WEB	SN	MEXP	EXPP	PN
Yearly	Basin *1	Mean	1492	1	2.1	− 0.2	0.1	0.6	− 0.7	2.6
		Std	186	− 22	10.8	4.3	− 4.3	− 10.8	2.7	− 2.7
		KS test#	−	0	0	0	0	0	0	0
	Basin *2	Mean	1364	− 1	− 0.3	− 0.4	1.7	− 1.7	− 2.3	5.6
		Std	185	− 12.9	− 9.7	− 5.9	2.7	− 4.8	12.4	2.7
		KS test#	−	0	0	0	0	0	0	1
	Basin *3	Mean	1559	1.2	1.2	0.5	2.3	0.9	0.8	4.4
		Std	277	− 45.5	− 48	− 39.4	− 33.9	− 36.1	− 35.4	− 39
		KS test#	−	0	0	0	0	0	0	0
Monthly	Basin *1	Mean	124.32	− 0.5	0.7	− 1.6	− 1.4	− 0.8	− 2.1	1.1
		Std	84.55	− 16.4	− 4.4	− 7.9	− 10.8	− 10.1	− 9.5	− 5.3
		KS test#	−	1	0	0	0	0	0	0
	Basin *2	Mean	113.64	− 1	− 0.3	− 0.5	1.7	− 1.7	− 2.3	5.7
		Std	79.3	− 12.6	− 11.2	− 7.1	− 3.8	− 10	− 5.2	− 1.8
		KS test#	−	1	0	0	0	0	0	1
	Basin *3	Mean	129.95	− 0.4	− 0.4	− 1.1	0.7	− 0.7	− 0.8	2.8
		Std	85.4	− 18.5	− 15.7	− 11.5	− 12	− 11.9	− 9	− 8.7
		KS test#	−	0	0	0	0	0	0	0

The relative error is presented for generated precipitation

KS test: The result is 0 if the observed and synthetic data are from the same distribution, and is 1 otherwise

overestimate the inter-annual variation of Rx1day with MREs of 43.20% and 73.74%, respectively, across all three river basins.

Table 3 presents the KS test results for three precipitation extreme indices for all seven precipitation models and three river basins. The KS test rejects the null hypothesis that the generated and observed data are from the same distribution if the value is equal to one in Table 3. The EXP model appears to be the worst with most times of rejection. Additionally, MEXP and WEB perform the best when the KS test accepts the null hypothesis for all indices and basins. In general, the performances of the three-parameter models are similar to those of the two-parameter distributions but superior to that of the one-parameter distribution.

3.3 Monthly flow modeling

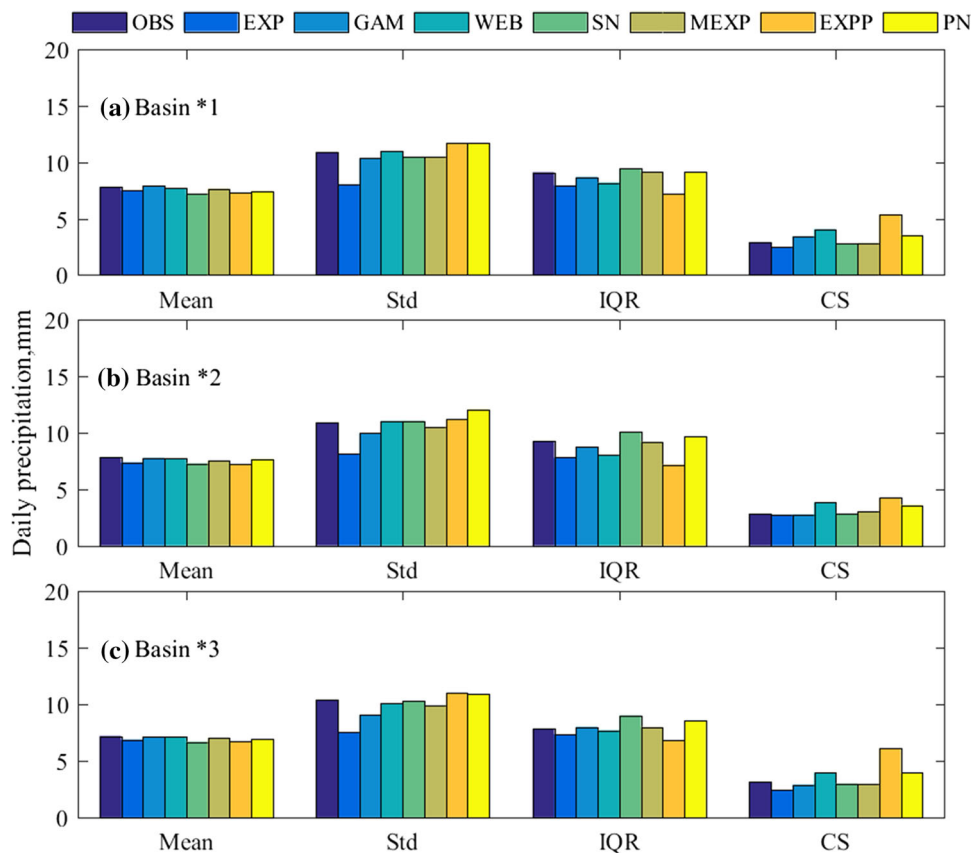
Figure 4 shows the mean and standard deviation (Std) of monthly streamflows generated using synthetic and observed climate data series. All seven models produce monthly mean streamflows reasonably well for all three river basins, with MREs ranging between − 17.14 and 5.45%. EXP and SN appear to be worse than other models, with mean absolute relative errors (MAREs) ranging

between 13.13 and 17.14% for the three river basins. Although PN shows the best performance in terms of producing mean monthly flow, it poorly simulates the July mean flow for basin *2 with a RE of 47.84%. In terms of generating the Std of monthly flow, the MAREs of basin *2 are obviously larger than those of the other two river basins. In particular, EXP appears to be the worst model with the largest MARE of standard deviation: 52.12% for basin *2. The performances of the three-parameter MEXP and SN models are better than that of the two-parameter GAM, but worse than that of the WEB model for all three basins. The PN model performs the best at simulating the Std of monthly streamflows, especially for basin *2.

3.4 Extreme hydrological events

The observed and simulated mean duration of high pulses (HPD) and the number of high pulses (HPC) are presented in Fig. 5. Generally, the HPD is reasonably represented by all seven precipitation models for basin *3, with median values of REs ranging between − 1.14% (WEB) and 12.84% (EXPP). However, HPD is not well-represented for basin *1 and basin *2, with REs ranging from 19.67 to 30.32%, and from − 37.37 to − 15.71%, respectively. More specifically, the GAM and EXPP models perform

Fig. 2 Mean, standard deviation (Std), interquartile range (IQR) and skewness coefficients (CS) of observed and generated daily precipitation from seven models (EXP, GAM, WEB, SN, MEXP, EXPP and PN) for **a** basin *1, **b** basin *2, **c** basin *3



worse than the others at producing the HPD for basin *2 and basin *3, respectively. In addition, the inter-annual variation of simulated HPD using generated precipitation series is larger than that of the observed HPD. WEB appears to be superior at producing the median HPD, and SN is superior at simulating the inter-annual variation of the HPD. Furthermore, most of the models underestimate the mean value of HPC for all three river basins, except the GAM, SN and PN models for basin *2. The HPC is well-reproduced by all of the models for basin *2 and basin *3, with AREs less than 16%, but for basin *1, the AREs are more than 22%. The EXP model is arguably the worst at representing the HPC of annual discharge, especially for basin *1.

The distributions of the Qx1day, Qx3day and Qx7day series are shown in Fig. 6 as box plots. Generally, the EXP model cannot adequately produce Qx1day, Qx3day and Qx7day series for basin *2 and basin *3 with its RE of medians ranging between -51.31 and -26.81% , even though it performs better for basin *1 with RE less than 14%. WEB, EXPP and PN represent Qx1day similarly and accurately for all three basins, especially for basin *1 with AREs less than 4%. The MEXP performs better than the single EXP, GAM and SN distributions, but worse than the single WEB distribution. Additionally, the MEXP model

performs the worst at representing the inter-annual variation of Qx1day, Qx3day and Qx7day, with REs ranging between -58.14 and -23.46% . In general, PN is the best model at representing the medium and inter-annual variation of Qx1day, Qx3day and Qx7day.

4 Discussion

This study assessed the performance of seven precipitation models in generating precipitation time series for simulating extreme flow. For simulating precipitation distribution at the daily, monthly and annual scales, all the models perform similarly and reasonably well for all three river basins, with the exception of EXP. Compared to models with two or one parameter(s), the three-parameter models have advantages in terms of generating precipitation extremes, with the exception of EXPP. This is as expected, as complex models usually include a component/parameter that specifically considers the upper tail of daily precipitation distribution (Chen et al. 2015). A relatively poor performance was observed for the EXPP model because it generates a few unreasonable high values. The main problem with the EXPP was identified in a few studies (e.g. Li et al. 2013b; Chen et al. 2015). Although a PMP-based

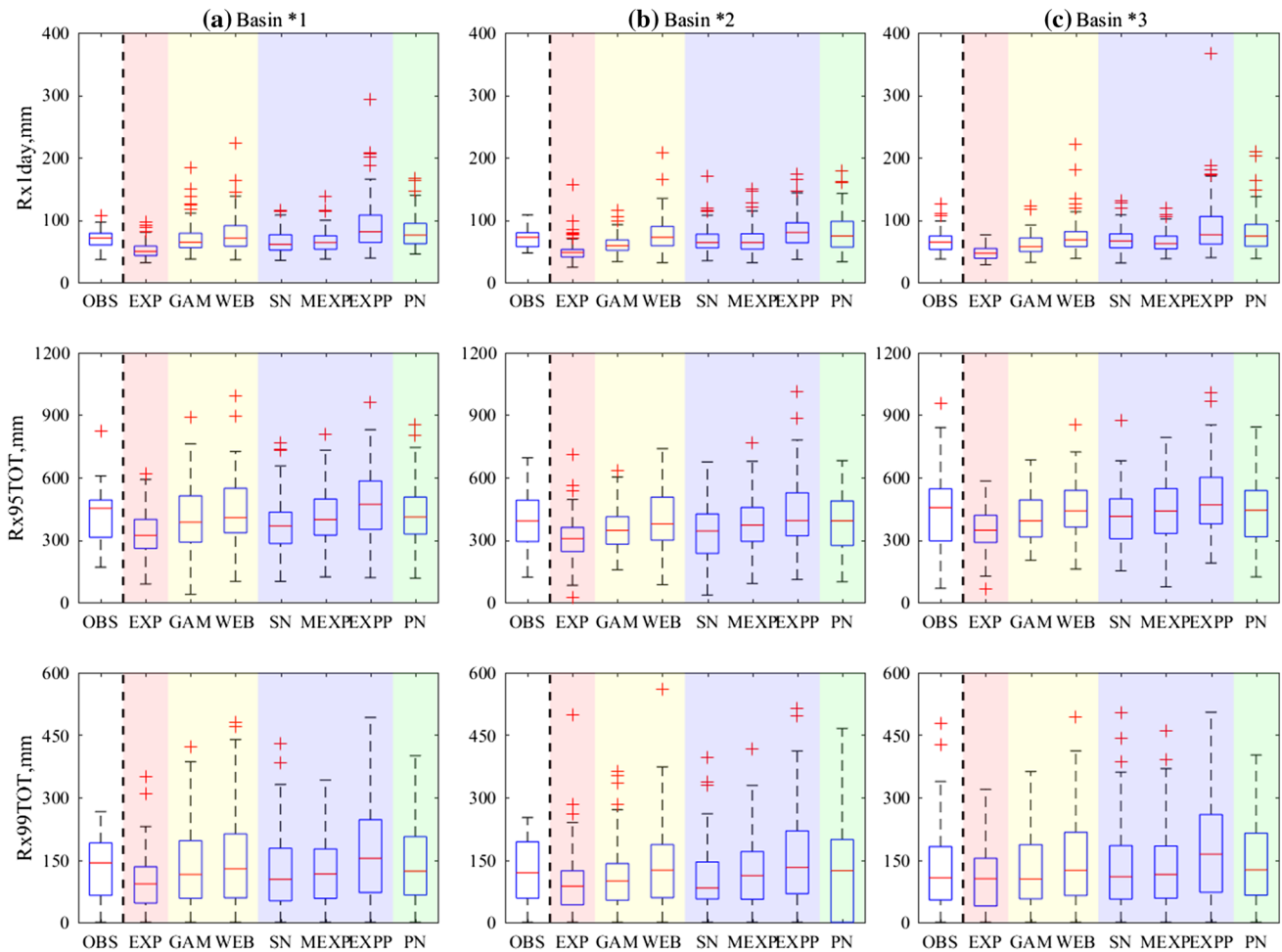


Fig. 3 Boxplots of Rx1day, R95pTOT and R99pTOT extreme indices of observed and generated daily precipitation for three river basins. Colored red represents one-parameter distribution, yellow for two-

parameter distributions, blue for three-parameter distributions and green for the curve-fitting method

Table 3 KS test results for extreme precipitation indices with a significance level of $P = 0.05$

River basin	Index	KS test						
		EXP	GAM	WEB	SN	MEX	PEXP	PN
Daxitan (*1)	Rx1day	1	0	0	0	0	1	0
	Rx95TOT	1	0	0	1	0	0	0
	Rx99TOT	1	0	0	0	0	0	0
Xiangxiang (*2)	Rx1day	1	1	0	0	0	1	1
	Rx95TOT	1	0	0	0	0	0	0
	Rx99TOT	1	0	0	0	0	0	0
Ganxi (*3)	Rx1day	1	0	0	0	0	1	1
	Rx95TOT	1	0	0	0	0	0	0
	Rx99TOT	0	0	0	0	0	0	0
Number of passing	1/9	8/9	9/9	8/9	9/9	6/9	7/9	

KS test: The value is 0 if the observed and synthetic data comes from the same distribution and 1 otherwise

method was used to control the random numbers, the values generated by EXPP still vary in a very large range. This may be explained if the threshold value of PMP was not

appropriate, so the use of this distribution to simulate daily precipitation should be interpreted with caution. The WEB method performs better than the other single distributions

Fig. 4 Mean and standard deviation (Std) of observed and generated monthly streamflows for three river basins

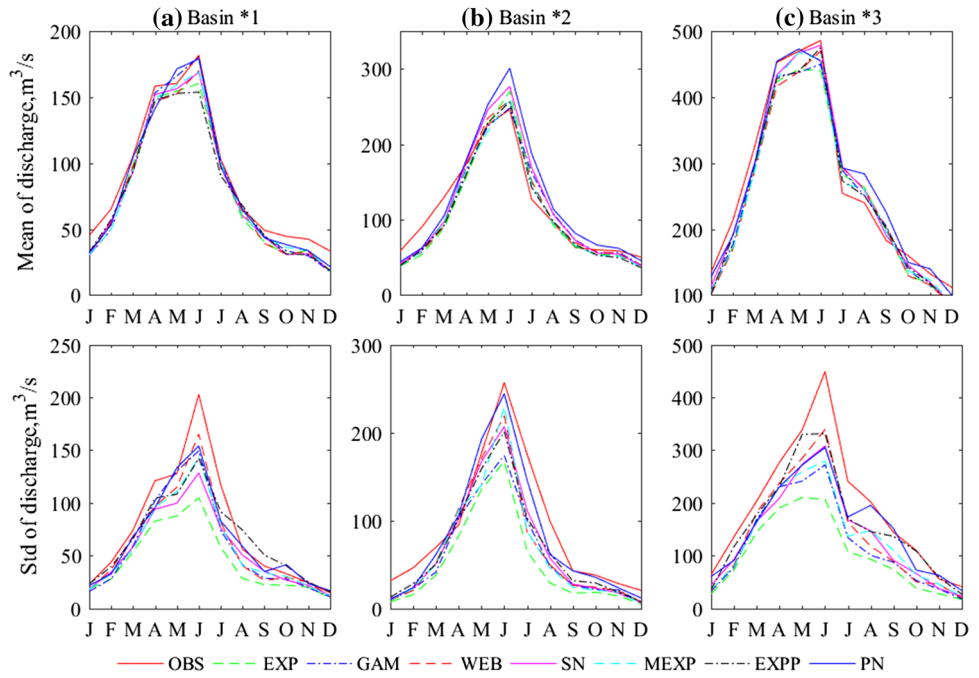
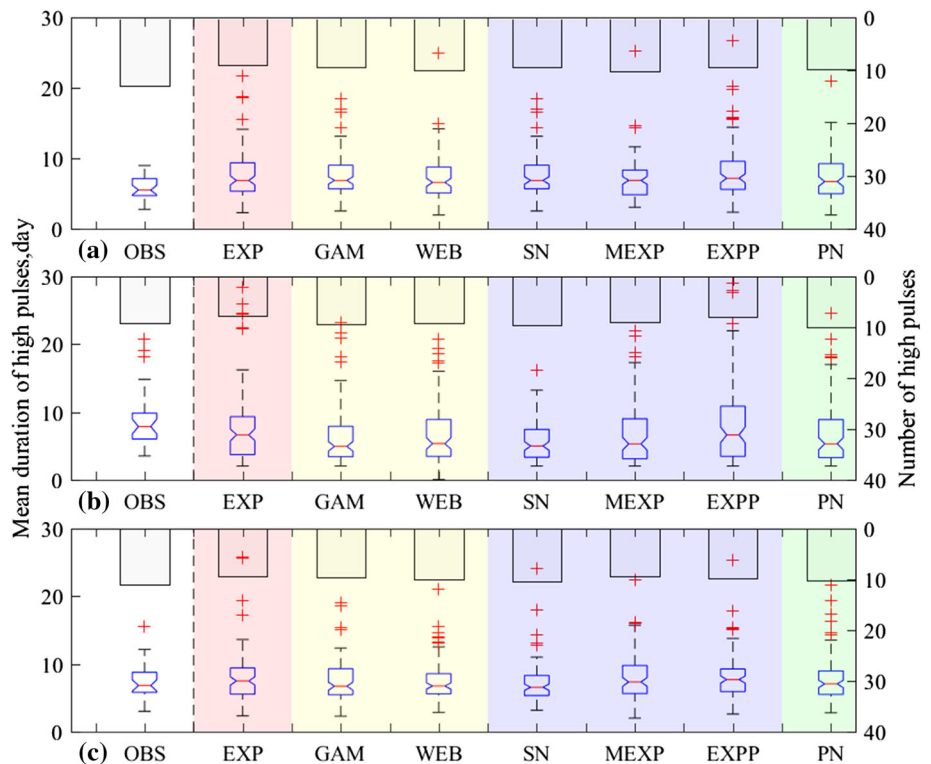


Fig. 5 The mean duration (boxplots) and the number (inverted bar) of observed high pulses and high pulses simulated with the precipitation generated by seven models for **a** basin *1, **b** basin *2 and **c** basin *3



with respect to reproducing extreme precipitation. Unlike this study, results from Safeeq and Fares (2011) show that the WEB distribution fails to reproduce precipitation extremes with the underestimation of extreme events. Overall, this study shows the MEXP distribution to be the best model at generating daily precipitation, which is consistent with most previous studies (Woolhiser and

Roldán 1982; Wilks 1999; Wan et al. 2005). The combination of two exponential distributions contributes to its good performance.

The same seven models were further compared to assess their performance in generating daily precipitation time series for simulating streamflow time series, especially for hydrological extremes. In general, the precipitation time

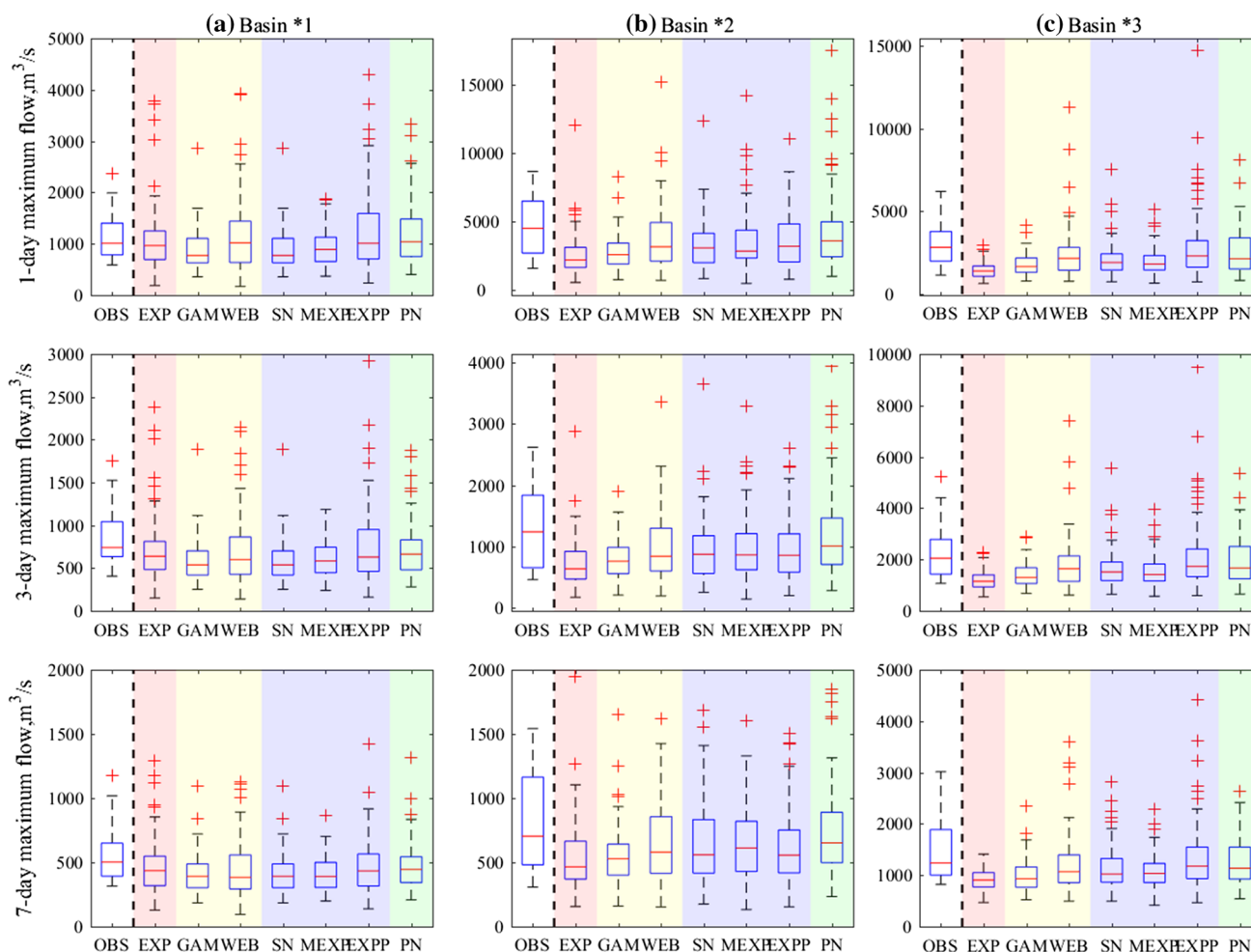


Fig. 6 Boxplots of annual maximum 1-day (Qx1day), 3-day (Qx3day), 7-day (Qx7day) observed and simulated flows for three river basins

series generated by all seven models can reasonably represent the monthly mean discharge. However, with the exception of PN, they perform poorly at representing the variation of monthly flow. In addition, the performance of each model at generating extreme precipitation did not fully translate to generate extreme flow events. For example, while the MEXP is the best model at reproducing extreme precipitation indices, it underestimates the extreme flow indices. The EXPP and PN models somewhat overestimate extreme precipitation indices, while they reasonably represent the extreme flow indices. This may be because EXPP and PN generate a large amount of events with low precipitation. The best model for reproducing precipitation to simulate hydrological extremes may be dependent upon the chosen index. In this study, WEB appears to be the best model at reproducing high pulses, while PN performs the best at generating annual mean 1-day, 3-day and 7-day maximum flows. In general, taking into account all of the chosen criteria, it appears that the PN model is the best choice for all three river basins.

Generally, the performance of these models in generating daily precipitation at the river basin scale was similar to that at the station scale. In terms of generating precipitation extremes, EXP was always the worst model for the underestimation of extremes for each station and each basin. Combining the results of this study and those of previous similar studies (Chen et al. 2012; Chen and Brissette 2014b; Acharya et al. 2017), the MEXP appears to be superior in reproducing extreme precipitation at river basin and multiple station scales. In terms of spatial scale effects, the drainage area of watersheds does not appear to have an impact on the performance of precipitation models in generating extreme precipitation. However, this study only evaluates the performance of different precipitation models in generating watershed-averaged precipitation for streamflow simulations over monsoon-characterized watersheds in China. For watersheds with different climatic and hydrological characteristics, the performance of various precipitation models may not be the same. For example, a previous study (Li et al. 2013b) has shown that the

MEXP distribution performs the best at simulating streamflow for watersheds characterized by snowfall in the winter. Even for watersheds with similar climate, various watershed characteristics may also show different sensitivities to precipitation inputs. Moreover, the performance of precipitation models may also be dependent on evaluation metrics. There is no single model that performs better than other models for all evaluation metrics. All these could be avenues for future studies.

5 Conclusion

The study investigates the performances of seven precipitation models in generating daily precipitation for simulating hydrological time series, especially extremes over three sub-basins in the Xiangjiang watershed. The following conclusions are drawn:

1. All seven models produce mean precipitation reasonably well. However, the three-parameter models generally perform better than the two-parameter and one-parameter distributions for simulating precipitation variance, with the exception of EXPP. Especially, the one-parameter EXP performs the worst at simulating the entire range of daily precipitation distribution.
2. MEXP is the best model at simulating extreme precipitation. In addition to MEXP, WEB and SN perform well at generating the medium value and the inter-annual variation for precipitation extremes, respectively. EXPP model performs more poorly than the others at simulating annual maximum precipitation at the basin scale.
3. Generally, the three-parameter models perform better than the one- and two-parameter models at representing hydrological extremes. In particular, PN shows the best performance with respect to simulating hydrological extremes among all precipitation models, even though WEB model also shows reasonable performance in representing high pulse.
4. The results of this study show that the most appropriate model cannot be determined when using different extreme indices. Overall, MEXP and PN outperform the other models in representing extreme precipitation and hydrological extremes.

Acknowledgements This work was partially supported by the National Natural Science Foundation of China (Grant No. 51779176, 51539009), the Overseas Expertise Introduction Project for Discipline Innovation (111 Project) funded by the Ministry of Education and State Administration of Foreign Experts Affairs P.R. China (Grant No. B18037), and the Thousand Youth Talents Plan from the Organization Department of the CCP Central Committee (Wuhan University, China). The authors are grateful to the China

Meteorological Data Sharing Service System and the Department of Water Resources of Hunan Province for providing meteorological and hydrological data.

References

- Acharya N, Frei A, Chen J, Decristofaro L, Owens EM (2017) Evaluating stochastic precipitation generators for climate change impact studies of New York City's primary water supply. *J Hydrometeorol* 18:879–896
- Baffault C, Nearing MA, Nicks AD (1996) Impact of CLIGEN parameters on WEPP-predicted average annual soil loss. *Trans ASAE* 39:447–457
- Chen J, Brissette FP (2014a) 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 (2014b) Stochastic generation of daily precipitation amounts: review and evaluation of different models. *Clim Res* 59(3):189–206
- Chen J, Brissette FP, Leconte R (2010) A daily stochastic weather generator for preserving low-frequency of climate variability. *J Hydrol* 388(3):480–490
- Chen J, Brissette FP, Leconte R (2011) Uncertainty of downscaling method in quantifying the impact of climate change on hydrology. *J Hydrol (Amsterdam)* 401(3–4):190–202
- Chen J, Brissette FP, Leconte R (2012) WeaGETS—a matlab-based daily scale weather generator for generating precipitation and temperature. *Proc Environ Sci* 13:2222–2235
- Chen J, Brissette FP, Zielinski PA (2015) Constraining frequency distributions with the probable maximum precipitation for the stochastic generation of realistic extreme events. *J Extreme Events* 2(2):1550009
- Chen J, Chen H, Guo S (2017) Multi-site precipitation downscaling using a stochastic weather generator. *Clim Dyn* 50(22):1–18
- Dibike YB, Coulibaly P (2005) Hydrologic impact of climate change in the Saguenay watershed: comparison of downscaling methods and hydrologic models. *J Hydrol* 307(1–4):163
- Duan QY, Gupta VK, Sorooshian S (1993) Shuffled complex evolution approach for effective and efficient global minimization. *J Optim Theor Appl* 76(3):501–521
- French MN, Krajewski WF, Cuykendall RR (1992) Rainfall forecasting in space and time using a neural network. *J Hydrol* 137(1–4):1–31
- Furrer EM, Katz RW (2008) Improving the simulation of extreme precipitation events by stochastic weather generators. *Water Resour Res* 44(4):317–322
- Haan CT, Allen DM, Street JO (1976) A Markov chain model of daily rainfall. *Water Resour Res* 2(3):443–449
- Hargreaves GH, Samani ZA (1985) Reference crop evapotranspiration from temperature. *Appl Eng Agric* 1(2):96–99
- Harmel RD, Richardson CW, King KW (2000) Hydrologic response of a small watershed model to generated precipitation. *Trans ASAE* 43(43):1483–1488
- Jamaludin S, Jemain AA (2007) Fitting daily rainfall amount in peninsular malaysia using several types of exponential distributions. *J Appl Sci Res* 3(10):1027–1036
- Khalili M, Brissette F, Leconte R (2011) Effectiveness of multi-site weather generator for hydrological modeling 1. *Jawra J Am Water Resour Assoc* 47(2):303–314
- Li HX, Zhang YQ, Chiew FHS, Xu SG (2009) Predicting runoff in ungauged catchments by using xianjiang model with modis leaf area index. *J Hydrol* 370(1):155–162

- Li C, Singh VP, Mishra AK (2012) Simulation of the entire range of daily precipitation using a hybrid probability distribution. *Water Resour Res* 48(3):W3521
- Li C, Singh VP, Mishra AK (2013a) A bivariate mixed distribution with a heavy-tailed component and its application to single-site daily rainfall simulation. *Water Resour Res* 49(2):767–789
- Li Z, Brissette F, Chen J (2013b) Finding the most appropriate precipitation probability distribution for stochastic weather generation and hydrological modelling in nordic watersheds. *Hydrol Process* 27(25):3718–3729
- Li Z, Brissette F, Chen J (2014) Assessing the applicability of six precipitation probability distribution models on the loess plateau of china. *Int J Climatol* 34(2):462–471
- Liu Y, Zhang WC, Shao YH, Zhang KX (2011) A comparison of four precipitation distribution models used in daily stochastic models. *Adv Atmos Sci* 28(4):809–820
- Lü H, Hou T, Horton R, Zhu Y, Chen X, Jia YW, Wang W, Fu XL (2013) The streamflow estimation using the xinjiang rainfall runoff model and dual state-parameter estimation method. *J Hydrol* 480(4):102–114
- Meyer C (2011) General description of the CLIGEN model and its history. USDA-ARS National Soil Erosion Laboratory, West Lafayette
- Minville M, Brissette F, Leconte R (2008) Uncertainty of the impact of climate change on the hydrology of a nordic watershed. *J Hydrol (Amsterdam)* 358(1–2):70–83
- Richardson CW (1981) Stochastic simulation of daily precipitation, temperature, and solar radiation. *Water Resour Res* 17:182–190
- Safeeq M, Fares A (2011) Accuracy evaluation of climgen weather generator and daily to hourly disaggregation methods in tropical conditions. *Theor Appl Climatol* 106(3–4):321–341
- Semenov M, Brooks R (1999) Spatial interpolation of the Lars-wg stochastic weather generator in great Britain. *J Jpn Soc Bus Ethics Stud* 11(2):205–216
- Shi P, Chen C, Srinivasan R, Zhang XS, Cai T, Fang XQ, Qu SM, Chen X, Li QF (2011) Evaluating the swat model for hydrological modeling in the xixian watershed and a comparison with the xaj model. *Water Resour Manage* 25(10):2595–2612
- Todorovic P, Woolhiser DA (1975) A stochastic model of n-day precipitation. *J Appl Meteorol* 14(1):17–24
- Venugopal V, Foufoula-Georgiou E, Sapozhnikov V (1999) A space-time downscaling model for rainfall. *J Geophys Res Atmos* 104(D16):19705–19721
- 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
- Wan H, Zhang X, Barrow E (2005) Stochastic modelling of daily precipitation for canada. *Atmosphere* 43(1):23–32
- Wilks DS (1989) Conditioning stochastic daily precipitation models on total monthly precipitation. *Water Resour Res* 25:1429–1439
- Wilks DS (1999) Interannual variability and extreme-value characteristics of several stochastic daily precipitation models. *Agric For Meteorol* 93(3):153–169
- Wilks DS (2008) High-resolution spatial interpolation of weather generator parameters using local weighted regressions. *Agric For Meteorol* 148(1):111–120
- Woolhiser DA, Roldán J (1982) Stochastic daily precipitation models: 2. A comparison of distributions of amounts. *Water Resour Res* 18(5):1461–1468
- Zhang XC, Liu WZ (2005) Simulating potential response of hydrology, soil erosion, and crop productivity to climate change in Changwu tableland region on the loess plateau of china. *Agric For Meteorol* 131(3–4):142
- Zhao RJ (1992) The xinjiang model applied in china. *J Hydrol* 135(1–4):371–381

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.