Elsevier Editorial System(tm) for Journal of Hydrology
Manuscript Draft

Manuscript Number: HYDROL25555R1

Title: The effect of rain gauge density and distribution on runoff simulation using a lumped hydrological modelling approach

Article Type: Research paper

Keywords: Xiangjiang River basin; Xinanjiang model; Bayesian framework; Rain gauge density; Rain gauge network

Corresponding Author: Mr. Hua Chen, PhD

Corresponding Author's Institution: Wuhan university

First Author: Qiang Zeng, M.sc

Order of Authors: Qiang Zeng, M.sc; Hua Chen, PhD; Chong-Yu Xu, PhD; Meng-Xuan Jie, M.sc; Jie Chen, PhD; Sheng-Lian Guo, PhD; Jie Liu, PhD

Abstract: Most lumped hydrological models use areal average precipitation data as model input. Though weather-radar-based and satellite-based precipitation estimation methods have been proposed in recent years, the rain gauge is still the most widely used precipitation-measuring tool. Optimal selection of rain gauge number and location will improve the accuracy of areal average precipitation estimations with minimum cost. In this study, the impacts of rain gauge density and distribution on lumped hydrological modelling uncertainty with different catchment sizes are analysed. To this end, the performances of a lumped hydrological model, the Xinanjiang model, in a densely gauged river basin, the Xiangjiang River basin, and its sub-basins under different gauge density and distribution are compared. First, seven levels of rain gauge density are defined. For each density level, several samples of different rain gauge distributions are randomly selected. Then, the areal average precipitation of each sample is estimated and used as input to the Xinanjiang model. Finally, the model is calibrated using the shuffled complex evolution (SCE-UA) algorithm, and model uncertainty is evaluated via the Bayesian method. The results show that 1) imperfect precipitation inputs measured by a sparse and irregular rain gauge network will lead to substantial uncertainty in model parameter estimation and flood simulation; 2) the impacts of imperfect precipitation estimates on model efficiency can be reduced to some extent through the adjustment of model parameters; 3) modelling uncertainty is reduced by increasing the rain gauge density or optimizing the rain gauge distribution pattern; and 4) the improvement in lumped model efficiency is no longer significant when the rain gauge density exceeds a certain threshold, but a further increase in rain gauge density will reduce model parameter uncertainty and the width of the runoff confidence interval.

The effect of rain gauge density and distribution on runoff simulation using a lumped hydrological modelling approach

Qiang Zeng[1,2], Hua Chen[1,2*], Chong-Yu Xu[1,3*], Meng-Xuan Jie[1,2], Jie Chen[1,2], Sheng-Lian Guo[1,2], Jie Liu[1,4]

[1.] State Key Laboratory of Water Resources and Hydropower Engineering Science, Wuhan University, Wuhan 430072, China

[2.] Hubei Provincial Collaborative Innovation Center for Water Resources Security, Wuhan University, Wuhan 430072, China

[3.] Department of Geosciences, University of Oslo, P O Box 1047 Blindern, N-0316 Oslo, Norway

[4.] School of Architecture and Civil Engineering, Chengdu University, Chengdu 610106, China

Corresponding to: * chongyu.xu@geo.uio.no; Chua@whu.edu.cn

Abstract

Most lumped hydrological models use areal average precipitation data as model input. Though weather-radar-based and satellite-based precipitation estimation methods have been proposed in recent years, the rain gauge is still the most widely used precipitation-measuring tool. Optimal selection of rain gauge number and location will improve the accuracy of areal average precipitation estimations with minimum cost. In this study, the impacts of rain gauge density and distribution on lumped hydrological modelling uncertainty with different catchment sizes are analysed. To this end, the performances of a lumped hydrological model, the Xinanjiang model, in a densely gauged river basin, the Xiangjiang River basin, and its sub-basins under different gauge density and distribution are compared. First, seven levels of rain gauge density are defined. For each density level, several samples of different rain gauge distributions are randomly selected. Then, the areal average precipitation of each sample is estimated and used as input to the Xinanjiang model. Finally, the model is calibrated using the shuffled complex evolution (SCE-UA) algorithm, and model uncertainty is evaluated via the Bayesian method. The results show that 1) imperfect precipitation inputs measured by a sparse and irregular rain gauge network will lead to substantial uncertainty in model parameter estimation and flood simulation; 2) the impacts of imperfect precipitation estimates on model efficiency can be reduced to some extent through the adjustment of model parameters; 3) modelling uncertainty is reduced by increasing the rain gauge density or optimizing the rain gauge distribution pattern; and 4) the improvement in lumped model efficiency is no longer significant when the rain gauge density exceeds a certain threshold, but a further increase in rain gauge density will reduce model parameter uncertainty and the width of the runoff confidence interval.

Key words: Xiangjiang River Basin, Xinanjiang model, Bayesian framework, rain gauge density, rain gauge network

1. Introduction

Lumped hydrological models are still widely used in flood forecasting and flood risk assessment (e.g., El Alfy, 2016; Jie et al., 2016; Refsgaard et al., 1988; Thiboult & Anctil, 2015; Huang & Hattermann, 2018; Su et al., 2018), water resource assessments (e.g., Kizza et al., 2013; Xu et al., 1996; Koivusalo et al., 2017), and impact studies of climate change on water resources (e.g., Awan et al., 2016; Chen et al., 2007 & 2012; Yan et al., 2016; Guo et al., 2018; Zhuan et al., 2018). However, the performance of lumped hydrological models is substantially influenced by model inputs (Oudin et al., 2006; Chang et al., 2017; Pechlivanidis et al., 2017). Precipitation, as a fundamental process of the hydrological cycle, is the most important forcing input for hydrological modelling and forecasting. Precise estimation of the spatial and temporal characteristics of precipitation is a key factor for accurate runoff simulation. However, precipitation input for a given basin is influenced by many factors, such as the precipitation type (convective, orographic, and frontal), the basin topography and the basin location and land use. The heterogeneous distribution of precipitation input in time and space makes its precise measurement a great challenge (Sattari et al., 2017). There are three commonly used methods for precipitation measurement: the rain gauge-based method, the weather-radar-based method and satellite-based remote sensing. Measurement by rain gauge is direct and of higher quality, but in most regions, rain gauges are too sparsely distributed to represent the spatial variability of precipitation (Villarini et al., 2008). In contrast, weather radar can characterize the spatial variability of precipitation through high spatial resolution measurement but its accuracy is influenced by many factors, including beam shielding, signal attenuation, ground cluster etc. (Germann et al., 2006). As a newly emerging precipitation measurement technology, satellite remote sensing can provide quasi-global precipitation products, but currently, the spatiotemporal resolutions of these products are low (Maggioni et al., 2016). Moreover, the measurements of radar and satellite retrieval products must be calibrated by rain gauge measurement to minimize data biases. Thus, a rain gauge network with a high gauge density and optimum gauge distribution is fundamental for the accurate measurement of precipitation.

Several studies have discussed the influence of rain gauge density and gauge distribution on the accuracy of precipitation estimation and hydrological modelling (Bárdossy and Das, 2008; Bras and Rodríguez-Iturbe, 1976; Girons et al., 2015; Krstanovic and Singh, 1992; Morrissey et al., 1995). Moulin et al. (2009) observed that uncertainty in the mean areal rainfall estimation is a key factor that leads to rainfall-runoff modelling error. St-Hilaire et al. (2003) used the areal average precipitation estimated from rain gauge networks with different densities as the input data for a lumped hydrological model to study the impact of rain gauge density on runoff simulation results. They found that the accuracy of runoff simulation was significantly improved when a dense rain gauge network was used, especially for the peak discharge simulation. Xu et al. (2013) further showed that the accuracy of rainfall estimation and hydrological model performance increased gradually with the increase of rain gauge density up to a certain threshold, while good performance of a lumped model was also observed with fewer rain gauges when an optimum rain gauge distribution pattern was used.

Recently, considerable attention has been focused on the uncertainties in hydrological modelling, including input uncertainty, parameter uncertainty, model structural uncertainty and output uncertainty. Many uncertainty estimation frameworks have been developed and tested in the literature (Beven and Binley, 1992; Camacho et al., 2015; Engeland et al., 2005; Vrugt et al., 2003). With these uncertainty estimation frameworks, the research on model parameters is no longer restricted to the calibrated optimum parameter set and now includes the posterior distributions of model parameters considering the parameter equifinality. In addition, the model output is not simply a single simulated runoff series but, a confidence interval describing model uncertainty. As precipitation is the main forcing data input in hydrological models, errors embedded in it will introduce considerable uncertainty in the model parameter estimation and the runoff simulation (Younger et al., 2009). Kavetski et al. (2006a) developed a Bayesian total error analysis framework (BATEA) to transparently analyse the input uncertainty separately from other uncertainties by introducing latent variables to characterize the rainfall error. Similarly, Ajami et al. (2007) used an error model with several hyper-parameters to simulate the rainfall error and developed a framework, the Integrated Bayesian Uncertainty Estimator (IBUNE), to explicitly account for the main hydrological modelling uncertainties. Li et al. (2011) perturbed the observed precipitation using normally distributed multipliers to simulate the systematic error and random error in precipitation data. They used the perturbed precipitation data as the input to a lumped hydrological model to study the sensitivity of two uncertainty estimation methods, the GLUE and the Bayesian method, to precipitation errors. Because rainfall sampling uncertainty is not well understood, for computational convenience, the precipitation error in the aforementioned studies was considered by a simple multiplicative or additive error model with model parameters calibrated or determined by experience. However, these error models are rarely validated because the "true precipitation" is difficult to acquire.

For a certain rain gauge network, the errors in precipitation data are predominately of two types: (1) point measurement error and (2) spatial interpolating error (McMillan et al., 2012). For the first type, precipitation estimates are influenced by the gauge type, wind effects and evaporation, etc. For the second type, no matter which interpolation method is used, the accuracy of interpolation is affected by rain gauge density and gauge distribution. Some key papers related to the impacts of imperfect precipitation inputs on hydrological modelling are listed in Table 1. From the deterministic perspective, some studies focus on the evaluation of the impact of precipitation errors on model parameter calibration and runoff simulation accuracy. Some of these studies directly used the precipitation data from randomly or specifically selected rain gauge combinations under different gauge densities (Anctil et al., 2006; Andréassian et al., 2001; Bárdossy and Das, 2008; Dong et al., 2005; St-Hilaire et al., 2003; Xu et al., 2013), whereas others used precipitation error models to simulate the two aforementioned types of error in precipitation data (Oudin et al., 2006; Xu et al., 2006). With respect to uncertainty, some studies emphasize the assessment of precipitation uncertainty in lumped hydrological modelling through the Bayesian inference method (i.e., Ajami et al., 2007; Kavetski et al., 2006b), which can explicitly or implicitly evaluate the uncertainty from diverse sources. However, to our knowledge, no paper has evaluated the impact of rain gauge density and gauge distribution on model parameter posterior distribution and the confidence interval of runoff simulation. Investigating these impacts on lumped hydrological modelling with respect to certainty and uncertainty is important for designing rain gauge networks and improving the structure and performance of hydrological models.

< Table 1 here please >

Thus, the objectives of this study are: (i) to assess the influence of rain gauge density and gauge distribution on lumped hydrological model performance; (ii) to investigate the effects of rain gauge density on modelling uncertainty; and (iii) to evaluate the effects of gauge distribution pattern on model parameter inference and model uncertainty using Bayesian framework.

The structure of this paper is as follows. Section 2 introduces the study area and data used in this study. Then, details about the lumped model, the model calibration method and the model uncertainty estimation method are given in Section 3. In Section 4, the results corresponding to the three objectives of the study are demonstrated and analysed. These results are discussed and compared to those of other studies in Section 5. Finally, major conclusions are drawn and recommendations related to the design of rain gauge networks are given in Section 6.

2. Study area and data

The Xiangjiang River basin in central-south China was selected as the study area (Fig. 1). This basin covers an area of approximately 94,660 km^2 with a total river length of 856 km. The basin elevation ranges from 2100 m above sea level on the southern boundary to 330 m a.s.l. on the northern river plain. This basin is controlled by the Mongolia high pressure system in winter and dominated by a southeast monsoon in summer (Xu et al., 2013). The monsoon climate and undulating terrain lead to a heterogeneous distribution of precipitation in both time and space. Nearly two-thirds of the 1450 mm annual mean precipitation occur in the rainy season from April to September. The climate of this region is generally warm and humid, and the monthly mean air temperature ranges from 4°C (February) to 30°C (July). Four sub-basins—Xiangxiang, Ganxi, Hengyang and Xiangtan, with sizes of 6053, 9972, 52,150 and 81,638 km^2, respectively—were studied.

< Figure 1 here please >

The study area has been densely instrumented in recent decades. There are 188 rain gauges evenly distributed in the basin, ensuring the achievement of the upscaling strategy in research on the impacts of rain gauge density on lumped modelling performance. Daily precipitation data are available from these rain gauges from 1990 to 2005. Corresponding pan evaporation data were acquired from 11 evaporation gauges in this basin. The study area was divided into four sub-basins depending on the available discharge stations (Fig. 1). Information from the different sub-basins is summarized in Table 2. These hydro-meteorological data are quality controlled by the Hydrology and Water Resources Bureau of Hunan Province, China. They have been used in many other studies (e.g., Li et al., 2015; Xu et al., 2015a, b) for various research purposes.

< Table 2 here please >

For each of the 4 basins, seven rain gauge density levels were defined, each corresponding to a percentage of the available rain gauge. Samples of rain gauge combinations were randomly selected from the complete gauge network for each density level in each basin. The maximum sampling number was set to 1000 for density levels with possible combinations exceeding 1000. However, when the possible combinations were less than 1000, all possible combinations were enumerated. For example, when 9 rain gauges are sampled from the 188 rain gauges of Xiangtan Basin, there are $188!/(179!\times 9!) \approx 10^{14}$ possible combinations. As it is impossible to consider all these combinations, we randomly selected 1000 combinations. However, when 2 gauges were sampled from the 19 rain gauges of Xiangxiang Basin, there were only $19!/(17!\times 2!) = 171$ possible combinations. Therefore, all the combinations were sampled for further analysis. The rain gauge number of each density level and the sampling number are listed in Table 3. After the sampling procedure, the areal average precipitation of each sample was derived by the Thiessen method. This method was chosen because it is commonly used in lumped hydrological modelling.

< Table 3 here please >

3. Methods

3.1 Xinanjiang model

The Xinanjiang model, which is a deterministic lumped hydrological model, is used in this study. This model was developed by Zhao and his colleagues (Zhao et al., 1980). "Its main feature is the concept of runoff formation on repletion of storage, which means that runoff is not generated until the soil moisture content reaches field capacity" (Zhao et al., 1995). This concept fits for the runoff generation mechanism of humid and semi-humid regions. The Xinanjiang model has been widely and successfully applied in southern China for flood forecasting (Jie et al., 2016; Zhao, 1992), design flood estimation (Zeng et al., 2016) and water resources assessment (Zhang et al., 2009). The flowchart of this model is shown in Fig. 2. Symbols in the solid boxes of Fig. 2 are the model inputs, outputs and state variables. The model inputs are the basin average precipitation (P) and pan evaporation (Epan) measured with an "E601"-type evaporation pan with a surface area of 3000 cm^2. Model outputs include the actual evapotranspiration (E) and the simulated runoff at the outlet of the study area (Q_{sim}). Symbols outside the solid boxes of Fig. 2 are the parameters of this model, and their explanations are listed in Table 4.

< Figure 2 here please >

< Table 4 here please >

The Xinanjiang model used in this study (where nearly all the precipitation falls in the form of rainfall and thus a snow routine does not need to be considered) consists of the following four major routines: (i) evapotranspiration calculation; (ii) runoff production; (iii) runoff separation; and (iv) flow routing. In this model, the study area is represented by a stack of soil layers including the upper layer, the lower layer and the deep layer with water storage capacity represented by parameters UM, LM and WM-UM-LM, respectively. Potential evapotranspiration is commonly calculated by the multiplication of a pan coefficient KE with the measured pan evaporation (McVicar et al., 2007; Xu et al., 2006) in China and other countries where pan evaporation data are more available than other meteorological data needed to calculate PET. In this study, KE is considered one of the key model parameters that directly influence the achievement of water balance. Water moisture in the soil is evaporated layer by layer from top to bottom. When the precipitation input is greater than the potential evapotranspiration, runoff is generated where the soil water content reaches the water storage capacity. The excess water is first stored in a free water reservoir with areal mean storage capacity of parameter SM. The storage capacity of the free water reservoir is heterogeneously distributed in the catchment, and this uneven special distribution is described by parameter EX. Thus, surface runoff, RS, generates in those areas where the free water content reaches its storage capacity. The rest of the water in the reservoir is separated into interflow, RI, and ground flow, RG, with the ratio of parameter KI and parameter KG, respectively. Surface runoff is further routed to the catchment outlet through a unit hydrograph with parameters N and NK. Interflow and ground flow are routed to a catchment outlet through a linear reservoir with parameters CI and CG, respectively. The runoff simulation at the catchment outlet is the sum of the three routing results.

3.2 Model calibration method

A global automatic optimization algorithm, the shuffled complex evolution (SCE-UA; see Duan et al. 1992) algorithm, is applied in this study for model calibration. The objective of model calibration is to minimize the mean square error of modelling runoff.

$$MSE = \frac{\sum \left(Q_{obs}^{t} - Q_{sim}^{t}\right)^2}{N} \tag{1}$$

where Q_{obs}^{t} and Q_{sim}^{t} are the daily observed and simulated runoffs, respectively, at time t, and N is the length of daily runoff data used for model calibration.

Two commonly used indices, the Nash-Sutcliffe efficiency (NS) and the relative volume error (RE), are used to evaluate the performance of the Xinanjiang model. The functions of the two indices are expressed as follows:

$$NS = 1 - \frac{\sum \left(Q_{obs}^{t} - Q_{sim}^{t}\right)^2}{\sum \left(Q_{obs}^{t} - \overline{Q_{obs}}\right)^2} \tag{2}$$

$$RE = \frac{\sum \left(Q_{sim}^{t} - Q_{obs}^{t}\right)}{\sum Q_{obs}^{t}} \tag{3}$$

where $\overline{Q_{obs}}$ is the mean of the daily observed runoff series. NS represents the ratio between the residual variance and the observed data variance. A “perfect” model fit is found when the value of NS equals one. RE determines how well the water balance is maintained. A positive RE indicates overestimation of total water volume and vice versa.

To minimize the initial condition influence on model performance, data from 1990 are used for warming up the model, while data from 1991 to 2005 are used for model calibration. The areal average precipitation samples of each gauge density level are used separately as the inputs to the Xinanjiang model. For each model input, the 15 model parameters are calibrated via the SCE-UA algorithm. To achieve the first objective of this study, model calibration results are analysed and compared to investigate the impact of differences in rain gauge density and gauge distribution on hydrological modelling. The precipitation sample producing the maximum model NS efficiency is chosen to be the “best precipitation sample” of the corresponding rain gauge density level for further analysis.

3.3 The Bayesian framework

To achieve the last two objectives of this study, the Bayesian framework is used to investigate the modelling uncertainty under different rain gauge density and gauge distribution. In the philosophy of the Bayesian method, unknown model parameters are treated as random quantities rather than a specified parameter set (Bernardo and Smith, 1994). Parameter distributions are inferred through the Bayesian formula, which considers the information in the observed data and prior experience together. The Bayesian formula is expressed as follows:

$$p(\theta,\varphi \mid \eta) = \frac{p(\eta \mid \theta,\varphi) \times p(\theta,\varphi)}{\int p(\eta \mid \theta,\varphi) \times p(\theta,\varphi) d\theta d\varphi} \quad (4)$$

where θ represents the hydrological model parameters, φ represents the statistical error model parameters, and η represents the transformed runoff observation. The priori distribution of model parameters $p(\theta,\varphi)$ is modified to the posterior distribution $p(\theta,\varphi\,|\,\eta)$ through the information of the observed data, which is represented by the likelihood function $p(\eta\,|\,\theta,\varphi)$. The denominator of the right-hand side of Eq. 4 is the occurrence probability of the observed data. For a given observed data set, the denominator is a normalization constant for projecting the numerator into probability space.

3.3.1 Likelihood function

In the Bayesian formula, the likelihood function is a probability density function for the observed data conditioned on the model parameters, which is equal to the conditional probability of simulation errors. However, the simulation errors might be related to the hydrological processes and highly correlated in time (Engeland and Gottschalk, 2002). Their characteristics are difficult to precisely depict by any known statistically correct likelihood function (Gupta et al., 1998), but their chief components can be described by statistical models. The simulation error at time t is calculated with the following formula:

$$\delta_t = T(Q_{sim,t}) - T(Q_{obs,t}) \qquad (5)$$

where $Q_{obs,t}$ and $Q_{sim,t}$ represent the daily observed runoff and the simulated runoff at time t, respectively, and T is a transformation function for obtaining normally distributed homoscedastic simulation errors. There are two commonly used transformation functions in the literature, including the Box-Cox transformation (Li et al., 2011; Yang et al., 2007) and the normal quantile transformation (Krzysztofowicz, 1997; Li et al., 2010). In the present study, the logarithmic transformation, which belongs to a subset of the Box-Cox transformation, is used for transformation efficiency and computational convenience. This transformation has been used in many other studies (e.g., Beven and Freer, 2001; Engeland and Gottschalk, 2002; Thiemann et al., 2001). Furthermore, the AR(1) error model (Eq. 6) is used in this study to characterize the autocorrelation property of the simulation errors.

$$\delta_t - \mu = \alpha(\delta_{t-1} - \mu) + \varepsilon_t \tag{6}$$

Here, α and μ are the autoregressive coefficient and the arithmetic mean of δ, respectively, and ε_t represents the residuals that are assumed to fit for an independent normal distribution with zero mean and constant variance σ^2. As the systematic error can be avoided by the adjustment of model parameter KE, μ is set to zero. Parameters of the error model, α and σ^2, are inferred simultaneously with the parameters of the hydrological model.

Based on the AR(1) error model assumption, the likelihood function is expressed as follows:

$$p(\delta \mid \theta, \phi) = (1-\alpha^2)^{1/2}(2\pi\sigma^2)^{-N/2} \exp\left[-\frac{1-\alpha^2}{2\sigma^2}\delta_1^2 + \sum_{i=2}^{N} -\frac{(\delta_t - \alpha\delta_{t-1})^2}{2\sigma^2}\right] \tag{7}$$

where N represents the number of days and the other notations are as defined above.

3.3.2 Prior distribution

To reduce the number of model runs, only the parameters sensitive to the precipitation input are inferred in this study, while the other parameters, which characterize the catchment properties, are kept at their optimum value calibrated by the SCE-UA algorithm using the areal average precipitation estimated by the complete rain gauge network as model input. These free parameters include the evaporation ratio KE, which influences the water balance; the interflow and ground flow separation parameters, KI and KG, respectively; the recession coefficients of interflow and ground flow, CI and

CG, respectively; and the surface flow routing parameter NK. In the absence of information about the free model parameters, uniform prior distribution of model parameters is used. For the error model parameters, the prior distribution of autoregressive coefficient α is considered a uniform distribution. For the standard deviation coefficient σ of AR(1) residuals, Jeffreys' non-informative prior is used, and its prior density is proportional to σ^{-1}. The prior distributions of eight parameters, including six hydrological model parameters and two error model parameters, are listed in Table 5.

< Table 5 here please >

3.3.3 Adaptive MCMC algorithm

An adaptive Markov chain Monte Carlo method, called the single-component adaptive Metropolis (SCAM) algorithm (Haario et al., 2005), is adopted to sample from the posterior parameter distribution. This algorithm can be considered a single-component Metropolis-Hastings algorithm with the component's proposal distribution adapted during the sampling process. In the implementation of this algorithm, the abovementioned eight parameters are updated one by one, and after the updating of all parameters, one iteration is completed. The proposal distribution of each parameter is a normal distribution centred on the present parameter value with a variance adapted iteration by iteration. The algorithm is conducted in the following steps:

1. Let $x_{i,t}$ be the ith parameter of the tth iteration, where i~[1,8];
2. Sample one candidate point $y_{i,t}$ from the proposal distribution of the ith parameter, $y_{i,t} \sim N(x_{i,t}, v_{i,t})$ normal distribution centred on current point;
3. Accept the candidate point with the probability calculated as follows:

$$a(y_{i,t}, x_{i,t+1}) = \min[1, \frac{\pi(x_{1,t+1},..., x_{i-1,t+1}, y_{i,t},...x_{8,t})}{\pi(x_{1,t+1},..., x_{i-1,t+1}, x_{i,t},...x_{8,t})}] \quad (8)$$

$$x_{i,t+1} = \begin{cases} y_{i,t} \text{ with probability } a(y_{i,t}, x_{i,t+1}) \\ x_{i,t} \text{ with probability } 1 - a(y_{i,t}, x_{i,t+1}) \end{cases} \quad (9)$$

where π represents the posterior distribution of the model parameter set;

4. Update the variance of the proposed distribution of the i^th parameter (Eq. 10);

$$v_{i,t+1} = \begin{cases} v_{i,0}, & t \le t_0 \\ sVar(x_{0,i},..., x_{i,t}) + s\upsilon, & t > t_0 \end{cases} \quad (10)$$

where t_0 denotes the number of iterations below which the variance of the proposal distribution for the i^th parameter is a constant $v_{i,0}$. After t_0 iterations, the variance of the proposal distribution is updated to the sampling variance multiplied by a scaling factor, s, which equals 2.4 (Gelman et al., 1996) in this study. The υ used in Eq. 10 is a small constant for preventing the variance of the proposal distribution from shrinking to zero.

To avoid the convergence of the Markov chain into a local optimum region, four independent Markov chains with different initial states randomly chosen from the parameter space are adopted in this study. The widely used potential scale reduction score $\sqrt{R}$ is used to check the convergence of the Markov chains. The detailed calculation steps of $\sqrt{R}$ were illustrated by (Gelman and Rubin, 1992).

3.3.4 Confidence intervals for runoff simulation

To study the impacts of rain gauge density on hydrological modelling uncertainty, the "best precipitation sample" with the largest NS value for each rain gauge density level is used as a model input. Two hundred thousand parameter sets of each of the four Markov chains are sampled after convergence. Thus, a total of 800,000 parameter sets are sampled from their posterior distributions. Then, the parameter posterior distribution and runoff confidence interval under different rain gauge densities are derived and compared. These parameter posterior distributions and runoff confidence intervals are further used as a model uncertainty baseline for evaluating the model uncertainty due to the difference of rain gauge distribution. To analyse the influence of rain gauge distribution on hydrological modelling uncertainty, all the 1000 precipitation samples, derived from randomly sampled gauge combinations for a given rain gauge density level, are separately used as model input. Eight hundred parameter sets for each precipitation sample are generated after convergence. Then, these parameter sets are combined to represent the parameter uncertainty caused by the difference of rain gauge distribution for a given rain gauge density. In addition, a total of 800,000 parameter sets are sampled out for the 1000 precipitation samples of a certain rain gauge density level.

The 95% confidence intervals for runoff simulation due to parameter uncertainty are estimated from the modelling runoffs with parameter sets sampled above. The 95% confidence intervals for runoff simulation considering the parameter uncertainty and model uncertainty are derived from the modelling runoffs adding the model residuals that are characterized by an AR(1) model. Indices used to measure the derived 95% confidence intervals of runoff simulations are the average relative interval length (ARIL) (Jin et al., 2010) and the percentage of observations that are contained in the intervals (CI95) (Li et al., 2009). These two indices are calculated as follows:

$$CI95 = \frac{N_{in}}{N} \times 100\% \qquad (11)$$

$$ARIL = \frac{1}{N} \sum \frac{Limit_{Upper,t} - Limit_{Lower,t}}{Q_{obs,t}} \qquad (12)$$

where N_{in} is the number of observations contained in the 95% confidence interval; N represents the number of days; $Limit_{Upper,t}$ and $Limit_{Lower,t}$ are the upper and lower boundaries of the 95% confidence interval, respectively; and other notations are as defined above.

4. Results

In the following three sections, the model performances under different rain gauge density and distribution patterns for the four sub-basins are first analysed. Then, the influence of rain gauge density and gauge distribution on model parameter posterior distribution and on total modelling uncertainty is investigated.

4.1 Model performance under different rain gauge density and distributions

The Nash-Sutcliffe efficiency, NS, and the relative volume errors, RE, for the runoff simulations with the precipitation samples as model inputs for each rain gauge density level were calculated. The boxplots of the two indices under different rain gauge densities for each sub-basin are shown in Fig. 3. For the four sub-basins, model performance is improved with the increase of rain gauge density, while the improvement is no longer significant when the rain gauge number exceeds a threshold. For the smallest basin, Xiangxiang, approximately 10 rain gauges (approximately 605 km^2 per gauge) are required to achieve stable model performance. For the largest basin, Xiangtan, 38 rain gauges (approximately 2148 km^2 per gauge) are necessary. This phenomenon indicates that a denser rain gauge network is required for a smaller basin to achieve stable model efficiency. Fig. 3 also shows that a higher rain gauge density leads to robust model performance with a better value of indices and a smaller variability of indices. When rain gauge density is lower than the threshold, the impacts of rain gauge distribution on model performance is more obvious. For illustrative purposes, Fig. 4 and Fig. 5 show the rain gauge distributions with maximum, median and minimum NS values for the Ganxi Basin and Hengyang Basin, respectively. For both basins, it is seen that when the gauge density is low, relatively evenly distributed rain gauges (Fig. 4(a), (d) and Fig. 5(a), (d)) will give better model simulation efficiency. If there are some gauges evenly distributed in the study area already, adding more gauges in the upstream mountainous region (see Fig. 4(g), (j) and Fig. 5(g), (j)) is beneficial for the improvement of model efficiency. However, if rain gauges are all concentrated in parts of the study area (Fig. 4(l) and Fig. 5(i)), then it is difficult to achieve good model efficiency even with a large rain gauge number.

< Figure 3 here please >

< Figure 4 here please >

< Figure 5 here please >

To demonstrate the impacts of rain gauge density on model parameter estimation, the standard deviations (STDs) of the 1000 calibrated parameter sets corresponding to the 1000 precipitation samples for each rain gauge density level were calculated. For illustrative purposes, Hengyang Basin and Xiangtan Basin are considered here. The STD value of each parameter was normalized by subtracting the minimum value and then dividing by the difference between the maximum value and the minimum value. The normalized STD values of each parameter for different rain gauge densities are shown in Fig. 6. In general, the normalized STD values for almost all model parameters (except for EX and IMP in Xiangtan Basin) display a descending trend with the increase of rain gauge density, which means the difference of rain gauge distribution will have less influence on model parameter estimation when more rain gauges are used for areal average precipitation estimation. The abnormal trends of parameters EX and IMP may be caused by the equifinality problem in model calibration. In the Xinanjiang model, EX describes the spatial distribution of free water storage capacity. Its value is influenced by the areal mean free water storage capacity SM. These parameters together determine the amount of surface water generation for a precipitation event. Fig. 6(b) shows that although the normalized STD value of EX increases when the rain gauge number changes from 19 to 38, the STDs of both parameters generally show a decreasing trend with the increase in rain gauge number. For parameter IMP, which represents the ratio of impermeable area to basin area, the STD ranges from 0.0023 to 0.0027 when different rain gauge densities are used. Therefore, this parameter is not sensitive to the rain gauge density, and thus an abnormal behaviour of the normalized STD of this parameter is found in Fig. 6 (b).

< Figure 6 here please >

The impact of rain gauge density on runoff simulation is investigated by comparing the flow duration curves (FDCs) of the observed runoff and those of the simulated runoffs. The results of the Xiangtan Basin are shown in Fig. 7 as a demonstration. Fig. 7 shows that a wider 95% confidence interval of simulated FDC is found when fewer rain gauges are used for model calibration, especially for an extreme flood with low exceedance probability. This phenomenon indicates that (1) the impact of rain gauge density on runoff simulation is greater for high floods and (2) this impact is reduced with the increase in the number of rain gauges used for model calibration.

< Figure 7 here please >

The above results reveal that rain gauge density and gauge distribution have considerable impacts on the model parameter estimation as well as the runoff simulation. Although the improvement of model performance is no longer significant when the rain gauge density exceeds a threshold (Fig. 3), reduced uncertainties in model parameter estimation (Fig. 6) and high flood simulation (Fig. 7) are observed with further increases in rain gauge density.

4.2 Influence of rain gauge density on modelling uncertainty

This section investigates how much the model uncertainty varies with the gauge density for the optimum rain gauge distribution pattern obtained under each rain gauge density. The “best precipitation sample” of each rain gauge density derived in Section 3.2 is first used as the model input. The model parameters’ posterior distributions are then investigated using the Bayesian method. The 95% confidence intervals of simulated runoffs are finally derived to analyse the model uncertainty under different rain gauge densities. Boxplots of 800,000 model parameter sets sampled from their posterior distribution with the SCAM algorithm are shown in Fig. 8. It can be seen that the posterior distribution of parameter KE, which is related to water balance, varies significantly when different rain gauge densities are used for areal average precipitation estimation. The insufficiency of the rain gauge number and the shortcoming of the areal average precipitation estimation method cause significant variation of KE when different rain gauge densities are used. For the other parameters, difference of posterior distributions are also significant when the rain gauge density is low. However, when the rain gauge density is greater than a certain threshold, this difference is no longer obvious and model parameters converge to a stable posterior distribution. These thresholds are different for different parameters. For the interflow separation parameter KI and recession parameter CI, 56 rain gauges are needed to obtain a stable posterior distribution. For the ground water separation parameter KG and the surface runoff routing parameter NK, 94 rain gauges are required. The ground water recession parameter CG seems less sensitive to the rain gauge density, as only 19 rain gauges are needed to derive a stable CG posterior distribution. Generally, from the variation of model parameters with rain gauge density, it seems that the impact of rain gauge density is different on different model parameters, and a greater impact is found on model parameters related to the water balance calculation.

< Figure 8 here please >

The 95% confidence intervals of simulated runoffs under different rain gauge density levels are derived, and then the interval measures are calculated and listed in Table 6. There is no significant difference in the interval length and containing ratio of the 95% confidence interval due to parameter uncertainty under different rain gauge density levels. However, when both the parameter uncertainty and the model uncertainty are considered, the interval length decreases with the increase of rain gauge number, which means that model uncertainty can be reduced by increasing the rain gauge density. However, this improvement is no longer significant when the number of rain gauges exceeds 38. This threshold is also applicable for the model performance index, the NS value. For illustrative purposes, the 95% confidence intervals due to parameter uncertainty and those due to both parameter uncertainty and model uncertainty for floods with a peak value exceedance probability (PEP) of 5%, 50% and 95% are shown in Fig. 9. When the same rain gauge number is used, the interval length due to parameter uncertainty is much shorter than that due to parameter uncertainty and model uncertainty. Thus, parameter uncertainty is less significant than model uncertainty in this case. Under different rain gauge density levels, the 95% confidence intervals seem to be narrower when more rain gauges are used, especially for high flood runoffs (Fig. 9 (c), (f)).

< Table 6 here please >

< Figure 9 here please >

4.3 Influence of rain gauge distribution on modelling uncertainty

Parameters sampled from their posterior distributions with the 1000 random precipitation samples as separate model inputs are merged together to represent the impacts of rain gauge distribution on model parameter estimation. Fig. 10 shows the boxplots of these 800,000 sampled parameter sets. When the rain gauge density is low, a difference in rain gauge distribution will lead to biased parameter estimation to compensate for the errors in areal average precipitation estimation. With the increase of rain gauge density, this parameter bias is reduced as the medians of parameter samples approach a constant value. A reduced box width in Fig. 10 is also found for each parameter with the increase of rain gauge density. Thus, the parameter estimation uncertainty caused by the difference of rain gauge distribution can be reduced by the increased rain gauge density. However, when more than 38 rain gauges are used, this improvement is no longer significant except for parameter KE.

< Figure 10 here please >

The STDs of parameter samples are shown in Table 7. The STDs of almost all parameters are reduced with the increase in rain gauge density. For certain model parameters, the STD of parameter samples with random precipitation inputs (1000 random precipitation samples) is larger than that of parameter samples with fixed precipitation inputs (“best precipitation sample”). The difference between the two STDs represents the parameter uncertainty induced by the difference of rain gauge distribution. Fig. 11 demonstrates the ratio of parameter STD between random precipitation input and fixed precipitation input under different rain gauge densities. It seems that (1) the larger ratios between the two STDs indicate greater uncertainty in parameter estimation, which is caused by the rain gauge distribution, especially when the rain gauge number is less than 38 for the Xiangtan Basin; (2) for parameter KE, no matter how many gauges are used, obvious parameter uncertainty caused by gauge distribution still exists; and (3) the large ratio of the STDs of KE indicates the severe impact of rain gauge distribution on model water balance.

< Table 7 here please >

< Figure 11 here please >

The 95% confidence intervals of runoff simulations for flows with peak exceedance probability (PEP) of 5%, 50%, and 95% using randomly selected precipitation samples as model inputs are shown in Fig. 12. This figure shows that under the same rain gauge density, the confidence interval due to model parameter uncertainty is smaller than that due to model uncertainty. It also demonstrates a significant reduction in the interval length when more gauges are available, especially for the confidence interval due to parameter uncertainty. Comparing Fig. 12 and Fig. 9, the parameter uncertainty induced confidence interval is larger when the uncertainty caused by the rain gauge distribution is considered. For the entire simulated runoff series, the indices of the 95% confidence intervals are listed in Table 8. For the confidence interval due to model parameter uncertainty, Table 8 shows that the average relative interval length is reduced when more rain gauges are available. The same conclusion applies to the confidence interval due to both model uncertainty and parameter uncertainty. However, it seems the improvement of total model uncertainty is not significant when the rain gauge number is larger than 94. Comparing Table 8 and Table 6, the confidence interval length due to both parameter uncertainty and model uncertainty is wider if the impact of rain gauge distribution is considered, especially when the number of rain gauges is less than 94. This increase of interval length seems to be mainly caused by the parameter uncertainty, as the interval due to parameter uncertainty increased most significantly with decreased rain gauge density.

< Table 8 here please >

< Figure 12 here please >

5. Discussion

5.1 Model performance under different rain gauge density and distributions

Section 4.1 revealed the impacts of rain gauge density and gauge distribution on lumped hydrological modelling performance. It is worth noting that precipitation interpolation will also cause errors in daily areal average precipitation estimation, which, however, is not the main focus of this paper. The impacts of diverse gauge distributions on lumped hydrological modelling are comparable as long as the same areal average precipitation derivation method is used. Many precipitation interpolation methods are available in the literature, such as the Thiessen polygons, Kriging, thin smooth plate splines, etc. (Ruelland et al., 2008), and some studies suggest using factors such as topography indexes or terrain elevation as covariates in precipitation interpolation (Diodato, 2005). However, it seems that the relationships between precipitation and these factors are less clear on a daily time scale than on longer time scales (Johansson and Chen, 2003). Moreover, when the gauge density is low, it is impossible to implement some complicated interpolation methods. For example, the Kriging method requires relatively high-density gauge network data to derive the semi-variogram (Ruelland et al., 2008). Thus, we chose the Thiessen polygons method because it is the most widely used method in lumped hydrological modelling and it requires much less computation time than other sophisticated precipitation interpolation methods such as Kriging or thin smooth plate splines (Ruelland et al., 2008).

In accordance with other studies (Anctil et al., 2006; Bárdossy and Das, 2008; Dong et al., 2005; Xu et al. 2013), our results show that there is a threshold of rain gauge density above which the improvement in model efficiency indices is no longer significant. We also found that for lumped hydrological modelling in a smaller basin, a denser rain gauge network is required to achieve good model efficiency. One interesting finding of our research is that even though no significant improvements in model efficiency indices are found when the rain gauge density is greater than a given threshold, a further increase in rain gauge density will reduce the variability of model parameters and increase the accuracy of the peak flow simulation.

Similar to the findings of Andréassian et al. (2001) and Xu et al. (2013), we found that only a few gauges can lead to very good model performance if these gauges are properly distributed. Thus, we can achieve good model efficiency through the optimization of rain gauge network. Dong et al. (2005) and Xu et al. (2013) revealed that the geographical location of rain gauges will impact the model simulation results. They also suggested that orographic precipitation should be considered in designing the spatial configuration of rain gauges. Through the comparison of rain gauge distributions that lead to different model efficiencies at several gauge density levels (Fig. 4 and Fig. 5), it seems that when gauge density is low, an evenly distributed network of rain gauges is beneficial for lumped hydrological modelling. When more gauges are available, additional gauges should be installed in mountainous regions, where orographic rain is more likely. This is only a qualitative suggestion of rain gauge network design. To quantitatively determine the number and location of rain gauges, Anctil et al. (2006) used the Genetic algorithm to optimize the rain gauge network. However, the method they used is only suitable for discarding redundant rain gauges from a dense gauge network.

To expand a rain gauge network, methods based on entropy or Kriging theory (Chen et al. 2008) are recommended.

5.2 Influence of rain gauge density on modelling uncertainty

Though some rain gauge networks will lead to good mean model efficiency, model uncertainty under these gauge networks should be analysed regarding the well-known parameter equifinality problem. Kavetski et al. (2006b) and Ajami et al. (2007) investigated the model input uncertainty in hydrological modelling through Bayesian theory. They found that precipitation errors will affect the confidence interval of simulated runoff considerably. However, they focused on the development and testing of uncertainty estimation frameworks, and no specific attention was paid to evaluating the impacts of rain gauge density and distribution on modelling uncertainty. In Section 4.2, we tested how much model uncertainty varies with the rain gauge density if the "best precipitation sample" of each density level is used as a model input. We found that the parameter equifinality-induced modelling uncertainties are similar for each gauge density level. In addition, the parameter uncertainty is relatively smaller than the model uncertainty, which is consistent with the results of other uncertainty evaluation papers (Engeland et al., 2005; Li et al., 2011). However, the model parameter posterior distribution varies considerably under different rain gauge densities, while no significant difference in the modelling efficiency is found, especially when the rain gauge number exceeds 38 for the Xiangtan Basin, indicating that precipitation errors can be compensated by adjusting model parameters.

5.3 Influence of rain gauge distribution on modelling uncertainty

When the impacts of rain gauge distribution are considered in Section 4.3, it seems that the parameter uncertainty under each gauge density level is much greater than the parameter uncertainty when the "best precipitation sample" is used as a model input. This phenomenon indicates that precipitation errors due to the rain gauge distribution will substantially affect model parameter estimation. Similar to the findings of Andréassian et al. (2001), this study reveals that the responses of model parameters to precipitation errors are different. For parameters related to the water balance calculation, more rain gauges are required to obtain stable parameter posterior distributions. With the increase in rain gauge number, both model uncertainty and parameter uncertainty induced by differences in gauge distribution reduce gradually. However, there is still a threshold (94 for the Xiangtan Basin) above which the reduction of model uncertainty is no longer obvious. The reason for this may be that the increased rain gauge number can only reduce the model input error, whereas the uncertainty due to the model conceptualization and model parameter equifinality still exists.

6. Conclusions

This study analysed the influence of rain gauge density and gauge distribution on the performance of a lumped hydrological model with different catchment sizes both in a deterministic sense and in terms of model uncertainty. The model inputs were the areal average precipitation samples estimated from randomly selected rain gauge combinations for different density levels. The performance of the Xinanjiang model was compared using different model inputs. The modelling uncertainty was further investigated using a Bayesian framework.

The results revealed that with the increase in rain gauge number, the model performance indices of the four sub-basins increase significantly when the rain gauge number is less than a certain threshold. Above this threshold, a further increase in the number of rain gauges will provide less improvement for these indices, although decreases of uncertainty in model parameter estimation and flood simulation continue.

The study shows that required rain gauge density for achieving adequate modelling results depends on the basin size. In this study area, for medium-size sub-basins with a drainage area of thousands of square kilometres, 10 to 15 evenly distributed rain gauges are necessary to obtain a stable model performance. For large sub-basins with a drainage area larger than 50,000 square kilometres, 30 to 50 evenly distributed rain gauges are required. Though the threshold number of rain gauges decreases with the decrease in basin size, the corresponding gauge density increases. Thus, a smaller basin requires a denser rain gauge network. However, below the threshold, there are still some rain gauge networks that lead to good model performance, indicating the potential to improve model efficiency through rain gauge network optimization. Furthermore, the threshold number of rain gauges is not only related to the basin size but also to the spatiotemporal characteristics of the precipitation. For different regions, this threshold may be different. The numerical relationship between the required rain gauge number and basin size needs to be investigated in further studies in diverse climatic regions.

The comparison of model parameter posterior distribution and modelling uncertainty between different rain gauge density levels indicates that when the best rain gauge network of each density level is used, the total modelling uncertainty is reduced with the increase in rain gauge number, especially for the high flood runoffs. However, the posterior distributions of some parameters are quite different for different gauge density levels. To enable a robust parameter estimation, different rain gauge numbers are required for different model parameters. A higher rain gauge density is necessary for those parameters that are directly involved in calculating the water balance.

This study also found that when differences in rain gauge distribution are considered, a lower rain gauge density leads to biased parameter estimation and large parameter uncertainty. Through the increased rain gauge density, this parameter uncertainty can be reduced, but no significant improvement of model performance is found when the gauge density level is greater than a certain threshold due to the uncertainty embedded in the model conceptualization.

The aforementioned conclusions may be suitable for other basins when lumped hydrological models are used to characterize the relationships between precipitation and runoff, but the threshold of the rain gauge density will be different. Moreover, the behaviour of distributed hydrological models with the number of rain gauges and their spatial distributions will be different. Additional studies in other basins using different hydrological models are needed to reach a general conclusion and guidance for the design of rain gauge networks.

Acknowledgements

The study was financially supported by the National Key Research and Development Program （2017YFA0603702） and the National Natural Science Fund of China (51539009; 51339004). The authors are grateful to the anonymous reviewers, the editor and the associate editor for their thorough review as well as their valuable comments and suggestions, which not only greatly improved the quality of the paper but also benefitted our research in general.

References

Ajami, N. K., Duan, Q., Soroosh, S., 2007. An integrated hydrologic bayesian multimodel combination framework: confronting input, parameter, and model structural uncertainty in hydrologic prediction. Water Resour. Res. 43, W01403, doi:10.1029/2005WR004745.

Anctil, F., Lauzon, N., Andréassian, V., Oudin, L., Perrin, C., 2006. Improvement of rainfall-runoff forecasts through mean areal rainfall optimization. J. hydrol. 328(3), 717-725.

Andréassian, V., Perrin, C., Michel, C., Usart-Sanchez, I., Lavabre, J., 2001. Impact of imperfect rainfall knowledge on the efficiency and the parameters of watershed models. J. Hydrol. 250(1), 206-223.

Awan, U.K., Liaqat, U.W., Choi, M., Ismaeel, A., 2016. A SWAT modeling approach to assess the impact of climate change on consumptive water use in Lower Chenab Canal area of Indus basin. Hydrol. Res. 47(5), 1025-1037, DOI: 10.2166/nh.2016.102**.**

Bárdossy, A., Das, T., 2008. Influence of rainfall observation network on model calibration and application. Hydrol. Earth Syst. Sc., 12(1), 77-89.

Bernardo, J. M., Smith A. F. M., 1994. Bayesian theory. Wiley, Chichester.

Beven, K., Binley A., 1992. The future of distributed models: Model calibration and uncertainty prediction. Hydrol. Process. 6(3), 279–298.

Beven, K., Freer, J., 2001. Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. J. Hydrol. 249(1), 11-29.

Bras, R. L., Rodríguez-Iturbe, I., 1976. Network design for the estimation of areal mean of rainfall events. Water Resour. Res. 12(6), 1185-1195.

Camacho, R. A., Martin, J. L., McAnally, W., Díaz- Ramirez, J., Rodriguez, H., Sucsy, P., Zhang, S. 2015. A comparison of Bayesian methods for uncertainty analysis in hydraulic and hydrodynamic modeling. JAWRA J. Am. Water Resour. As. 51(5), 1372-1393.

Chang, C. H., Wu, S. J., Hsu, C. T., Shen, J. C., Lien, H. C. 2017. An evaluation framework for identifying the optimal raingauge network based on spatiotemporal variation in quantitative precipitation estimation. Hydrol. Res. 48(1), 77-98.

Chen, H., Guo, S., Xu, C. Y., Singh, V. P., 2007. Historical temporal trends of hydro-climatic variables and runoff response to climate variability and their relevance in water resource management in the hanjiang basin. J. of Hydrol. 344(3-4), 171-184.

Chen, H., Xu, C. Y., Guo, S., 2012. Comparison and evaluation of multiple GCMs, statistical downscaling and hydrological models in the study of climate change impacts on runoff. J. Hydrol. 434–435, 36-45.

Chen, Y. C., Wei, C., Yeh, H. C., 2008. Rainfall network design using kriging and entropy. Hydrol. Process. 22(3), 340.

Diodato, N., 2005. The influence of topographic co-variables on the spatial variability of precipitation over small regions of complex terrain. Int. J. Climatol. 25(3), 351-363.

Dong, X., Dohmen-Janssen, C. M., Booij, M. J., 2005. Appropriate Spatial Sampling of Rainfall or Flow Simulation/Echantillonnage Spatial de la Pluie Approprié pour la Simulation D'écoulements. Hydrolog. Sci. J. 50(2).

Duan, Q., Soroosh, S., Vijai, G., 1992. Effective and efficient global optimization for conceptual rainfall-runoff models. Water Resour. Res. 28(4), 1015-1031.

El Alfy, M., 2016. Assessing the impact of arid area urbanization on flash floods using GIS, remote sensing, and HEC-HMS rainfall–runoff modeling. Hydrol. Res. 47(6), 1142-1160, DOI: 10.2166/nh.2016.133

Engeland, K., Gottschalk, L., 2002. Bayesian estimation of parameters in a regional hydrological model. Hydrol. Earth Syst. Sc. 6(5), 883-898.

Engeland, K., Xu, C. Y., Gottschalk, L., 2005. Assessing uncertainties in a conceptual water balance model using Bayesian methodology/Estimation bayésienne des incertitudes au sein d'une modélisation conceptuelle de bilan hydrologique. Hydrolog. Sci. J. 50(1), 45-63.

Gelman, A., Roberts, G., Gilks, W., 1996. Efficient metropolis jumping rules, in: J. M. Bernardo, J. O. Berger, A. P. Dawid and A. F. M. Smith (Eds.), Bayesian statistics 5. Oxford Univ. Press, New York, 599-608.

Gelman, A., Rubin D. B., 1992. Inference from iterative simulation using multiple sequences. Stat. Sci. 4, 457-472.

Germann, U., Galli, G., Boscacci, M., Bolliger, M., 2006. Radar precipitation measurement in a mountainous region. Q. J. Roy. Meteor. Soc. 132(618), 1669-1692.

Girons Lopez, M., Wennerström, H., Nordén, L. Å., Seibert, J., 2015. Location and Density of Rain Gauges for the Estimation of Spatial Varying Precipitation. Geogr. Ann. A. 97(1), 167-179.

Guo, Y. X., Fang, G. H., Wen, X., Lei, X. H., Yuan, Y., Fu, X. Y., 2018. Hydrological responses and adaptive potential of cascaded reservoirs under climate change in Yuan River Basin. Hydrol. Res. DOI: 10.2166/nh.2018.165

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

Haario, H., Saksman, E., Tamminen, J., 2005. Componentwise adaptation for high dimensional MCMC. Computation. Stat. 20(2), 265-273.

Huang, S. C., Hattermann, F. F., 2018. Coupling a global hydrodynamic algorithm and a regional hydrological model for large-scale flood inundation simulations. Hydrol. Res. DOI: 10.2166/nh.2017.061

Jie, M. X., Chen, H., Xu, C. Y., Zeng, Q., Tao, X. E., 2016. A comparative study of different objective functions to improve the flood forecasting accuracy. Hydrol. Res. 47(4), 718-735.

Jin, X., Xu, C.Y., Zhang, Q., Singh, V.P., 2010. Parameter and modeling uncertainty simulated by GLUE and a formal Bayesian method for a conceptual hydrological model. J. Hydrol. 383(3), 147-155.

Johansson, B., Chen, D., 2003. The influence of wind and topography on precipitation distribution in Sweden: Statistical analysis and modelling. Int. J. Climatol. 23(12), 1523-1535.

Kavetski, D., Kuczera, G., Franks, S. W., 2006a. Bayesian analysis of input uncertainty in hydrological modeling: 1. Theory. Water Resour. Res. 42(3), 446-455.

Kavetski, D., Kuczera, G., Franks, S. W., 2006b. Bayesian analysis of input uncertainty in hydrological modeling: 2. Application. Water Resour. Res. 42(3), 216-224.

Kizza, M., Guerrero, J.L., Rodhe, A., Xu, C.Y., Ntale, H.K., 2013. Modelling catchment inflows into Lake Victoria: Regionalisation of the parameters of a conceptual water balance model. Hydrol. Res. 44, 789-808.

Koivusalo, H., Turunen,M., Salo, H., Haahti, K., Nousiainen, R., Warsta, L., 2017. Analysis of water balance and runoff generation in high latitude agricultural fields during mild and cold winters. Hydrol. Res. 48(4), 957-968.

Krstanovic, P. F., Singh V. P., 1992. Evaluation of rainfall networks using entropy: I. Theoretical development. Water Resour. Manag. 6(4), 279-293.

Krzysztofowicz, R., 1997. Transformation and normalization of variates with specified distributions. J. Hydrol. 197(1), 286–292.

Li, H., Beldring, S, Xu, C. Y., 2015. Stability of model performance and parameter values on two catchments facing changes in climatic conditions. Hydrolog. Sci. J. 60(7-8), 1317-1330, DOI:10.1080/02626667.2014.978333.

Li, L., Xia, J., Xu, C. Y., Singh, V. P., 2010. Evaluation of the subjective factors of the glue method and comparison with the formal bayesian method in uncertainty assessment of hydrological models. J. Hydrol. 390(3), 210-221.

Li, L., Xia, J., Xu, C.Y., Chu, J.T. Wang, R., 2009. Analyse the sources of equifinality in hydrological model using glue methodology, in Hydroinformatics in hydrology, hydrogeology and water resources, Proceedings of Symposium JS.4, Joint Convention of the International Association of Hydrological Sciences (IAHS) and the International Association of Hydrogeologists (IAH), Hyderabad, India, 6-12 September 2009, 130-138.

Li, L., Xu, C. Y., Xia, J., Engeland, K., Reggiani, P., 2011. Uncertainty estimates by Bayesian method with likelihood of AR (1) plus Normal model and AR (1) plus Multi-Normal model in different time-scales hydrological models. J. Hydrol. 406(1), 54-65.

Maggioni, V., Meyers, P. C., Robinson, M. D., 2016. A review of merged high-resolution satellite precipitation product accuracy during the tropical rainfall measuring mission (TRMM) era. J. Hydrometeorol. 17(4), 160209115331003.

McMillan, H., Krueger, T., Freer, J., 2012. Benchmarking observational uncertainties for hydrology: rainfall, river discharge and water quality. Hydrol. Process. 26(26), 4078-4111.

McVicar, T. R., Van Niel, T. G., Li, L., Hutchinson, M. F., Mu, X., Liu, Z., 2007. Spatially distributing monthly reference evapotranspiration and pan evaporation considering topographic influences. J. Hydrol. 338(3), 196-220.

Morrissey, M. L., Maliekal, J. A., Greene, J. S., Wang, J., 1995. The uncertainty of simple spatial averages using rain gauge networks. Water Resour. Res. 31(8), 2011-2017.

Moulin, L., Gaume, E., Obled, C., 2009. Uncertainties on mean areal precipitation: assessment and impact on streamflow simulations. Hydrol. Earth Syst. Sc. 13(2), 99-114.

Oudin, L., Perrin, C., Mathevet, T., Andréassian, V., Michel, C., 2006. Impact of biased and randomly corrupted inputs on the efficiency and the parameters of watershed models. J. Hydrol. 320(1), 62-83.

Pechlivanidis, I.G., McIntyre, N., Wheater, H.S., 2017. The significance of spatial variability of rainfall on simulated runoff: an evaluation based on the Upper Lee catchment, UK. Hydrol. Res. 48(4), 1118-1130.

Refsgaard, J.C., Havnø, K., Ammentorp, H.C., Verwey, A., 1988. Application of hydrological models for flood forecasting and flood control in India and Bangladesh. Adv. Water Resour. 11, 101-105.

Ruelland, D., Ardoin-Bardin, S., Billen, G., Servat, E., 2008. Sensitivity of a lumped and semi-distributed hydrological model to several methods of rainfall interpolation on a large basin in West Africa. J. Hydrol. 361(1), 96-117.

Sattari, M-T, Rezazadeh-Joudi, A., Kusiak, A., 2017. Assessment of different methods for estimation of missing data in precipitation studies. Hydrol. Res. 48(4), 1032-1044.

St-Hilaire, A., Ouarda, T. B., Lachance, M., Bobée, B., Gaudet, J., Gignac, C., 2003. Assessment of the impact of meteorological network density on the estimation of basin precipitation and runoff: a case study. Hydrol. Process. 17(18), 3561-3580.

Su, B. N., Huang, H., Zhu, W., 2018. An urban pluvial flood simulation model based on diffusive wave approximation of shallow water equations. Hydrol. Res. DOI: 10.2166/nh.2017.233

Thiboult, A., Anctil, F., 2015. On the difficulty to optimally implement the ensemble kalman filter: an experiment based on many hydrological models and catchments. J. Hydrol. 529, 1147-1160.

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

Villarini, G., Mandapaka, P. V., Krajewski, W. F., Moore, R. J., 2008. Rainfall and sampling uncertainties: A rain gauge perspective. J. Geophys. Res-Atmos. 113, D11102, doi:10.1029/2007JD009214.

Vrugt, J. A., Gupta, H. V., Bouten, W., Sorooshian, S., 2003. A Shuffled Complex Evolution Metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters. Water Resour. Res. 39(8), 1201, doi:10.1029/2002WR001642.

Xu, C. Y., Gong, L., Jiang, T., Chen, D., Singh, V. P., 2006. Analysis of spatial distribution and temporal trend of reference evapotranspiration and pan evaporation in Changjiang (Yangtze River) catchment. J. Hydrol. 327(1), 81-93.

Xu, C. Y., Tunemar, L., Chen, Y. D., Singh, V. P., 2006. Evaluation of seasonal and spatial variations of lumped water balance model sensitivity to precipitation data errors. J. Hydrol. 324(1), 80-93.

Xu, C.Y., Seibert, J., Halldin, S., 1996. Regional water balance modelling in the NOPEX area: development and application of monthly water balance models. J. Hydrol. 180, 211-236.

Xu, H., Xu, C. Y., Chen, H., Zhang, Z., Li, L., 2013. Assessing the influence of rain gauge density and distribution on hydrological model performance in a humid region of China. J. Hydrol. 505, 1-12.

Xu, H., Xu, C. Y., Sælthun, N. R., Xu, Y., Zhou, B., Chen, H., 2015a. Entropy theory based multi-criteria resampling of rain gauge networks for hydrological modelling–A case study of humid area in southern China. J. Hydrol. 525, 138-151.

Xu, H., Xu, C. Y., Sælthun, N. R., Zhou, B., Xu, Y., 2015b. Evaluation of reanalysis and satellite-based precipitation datasets in driving hydrological models in a humid region of Southern China. Stoch. Env. Res. Risk. A. 29(8), 2003-2020.

Yan, R.H., Huang, J.C., Wang, Y., Gao, J.F., Qi, L.Y., 2016. Modeling the combined impact of future climate and land use changes on streamflow of Xinjiang Basin, China. Hydrol. Res. 47(2), 356-372, doi:10.2166/nh.2015.206

Yang, J., Reichert, P., Abbaspour, K. C., Yang, H., 2007. Hydrological modelling of the Chaohe Basin in China: Statistical model formulation and Bayesian inference. J. Hydrol. 340(3), 167-182.

Younger, P. M., Freer, J. E., Beven, K. J., 2009. Detecting the effects of spatial variability of rainfall on hydrological modelling within an uncertainty analysis framework. Hydrol. Process. 23(14), 1988-2003.

Zeng, Q., Chen, H., Xu, C. Y., Jie, M. X., Hou, Y. K., 2016. Feasibility and uncertainty of using conceptual rainfall-runoff models in design flood estimation. Hydrol. Res. 47(4), 701-717.

Zhang, W., Wang, Y., Peng, H., Li, Y., Tang, J., Wu, K. B., 2009. A coupled water quantity–quality model for water allocation analysis. Water Resour. Manag. 24(3), 485-511.

Zhao, R.J., 1992. The Xinanjiang model applied in China. J. Hydrol. 135: 371–381.

Zhao, R.J., Liu, X.R., 1995. The Xinanjiang model, in Singh, V.P. (Ed.), Computer Models of Watershed Hydrology. Water Resources Publication, Colorado, pp. 215–232, Chapter 7.

Zhao, R.J., Zhuang, Y.L., Fang, L.R., Liu, X.R., Zhang, Q.S., 1980. The Xinanjiang Model. In: Hydrological Forecasting, IAHS Publication No. 129. IAHS Press, Wallingford: 351–356.

Zhuan, M.J., Chen, J., Shen, M.X., Xu, C-Y, Chen, H., Xiong, L.H., 2018. Timing of human-induced climate change emergence from internal climate variability for hydrological impact studies. Hydrol. Res. DOI: 10.2166/nh.2018.059.

Tables

Table 1 Summary of papers related to the impacts of imperfect precipitation inputs on hydrological modelling. The relation of these papers to the objectives of this study is assessed in one or all of three aspects: (a) assessed the precipitation errors related to rain gauge density and gauge distribution, (b) assessed impacts of precipitation errors on hydrological modelling in a deterministic sense, (c) assessed impacts of precipitation errors on hydrological modelling considering uncertainty. NG indicates the information is not given in the paper.

Study	Basin/Basin area (km^2) /Gauge number	Rain gauge network scenario	Precipitation preparation method	Hydrological model/ Time scale	Key results/in relation to the abovementioned three aspects
1-Andréassian et al. (2001)	(1) Yonne, France/ 10700/ 33; (2) Serein, France/ 1120/ 33; (3) Réal Collobrier, France/ 71/ 20;	For each gauge number, randomly select 60 subsets from all the gauges.	Arithmetic mean	(1) Lumped GR3J model; (2) Lumped IHACRES model; (3) Lumped TOPMO model; /daily	(1) Improved accuracy in rainfall input will increase model performance and reduce model efficiency variability; (2) Smaller watersheds need a higher concentration of rain gauges. /(a, b)
2-St-Hilaire et al. (2003)	Five basins in the Mauricie area, Canada/10142, 9320, 3800, 2680, 2770/48	Sparse scenario: 22 gauges; Dense scenario: 27 gauges;	Ordinary Kriging	Lumped HSAMI model; /daily	(1) Denser network relates to better modelling efficiency; (2) Peak flows were better simulated with a denser network. /(a, b)
3-Anctil et al. (2006)	Bas-en-Basset, France/ 3234/ 23	(1) For each of the five rain gauge numbers (20, 15, 10, 5, 2), 50 subsets were randomly selected; (2) 2500 gauge combinations were tested within the genetic algorithm for network optimization	Arithmetic mean	Lumped neural network rainfall-runoff forecasting model. /daily	(1) Ten rain gauges are the minimum requirement; (2) Genetic search can be used to find the optimal gauge combination for hydrological forecasting. /(a,b)
4-Bárdossy and Das (2008)	The upper Neckar with 13 subcatchments, Germany/ 3961/ 51	Seven networks consisting of different rain gauge numbers were selected by simulated annealing algorithm.	Kriging with external drift	Semi-distributed HBV model. /daily	(1) Overall model performance will not be significantly improved by increasing the number of rain gauges over a threshold; (2) Models using different rain gauge networks may need their parameters recalibrated. /(a,b)
5-Xu et al. (2013)	Xiangjiang, China/ 94660/ 181	For each of the six rain gauge numbers (10, 19, 38, 57, 93, 128), 100 subsets are randomly selected.	NG	Lumped Xinanjiang model /daily	(1) The probability of achieving poor model performance is increased when the number of rain gauges falls below a threshold; (2) Better model performance can be achieved with fewer rain gauges if an optimum spatial configuration is provided. /(a,b)
6-Dong et al. (2005)	The upper Qingjiang, China/ 12209/ 26	All possible combinations are enumerated when gauge number ranges from 1 to 7 and from 20 to 26; 5000 subsets are randomly selected for gauge numbers	Arithmetic mean	Lumped HBV model /daily	(1) Five rain gauges are enough for lumped HBV model in this basin; (2) Setting more gauges in the mountainous regions with heavy orographic rainfall will lead to better model performance. /(a,b)

		ranging from 8 to 19;			
7-Xu et al. (2006)	26 watersheds in the Mälaren basin, Sweden/ ranging from 6 to 1293/ 41	Corrupt the precipitation measurement by adding random errors and systematic errors.	Arithmetic mean	Lumped water balance model, NOPEX-6 /monthly	(1) Systematic precipitation errors have significant impacts on model quality; (2) Systematic precipitation errors affect model parameters systematically; (3) Random precipitation errors affect model parameters randomly. /(b)
8-Oudin et al. (2006)	12 watersheds in the United States/ ranging from 1021 to 4421/ NG	Corrupt the precipitation measurement by adding random errors and systematic errors.	NG	(1) Lumped GR4J model; (2) Lumped TOPMO model /daily	(1) Random errors in precipitation significantly affect model performance and model parameters; (2) Systematic errors in precipitation, when large enough, can be detrimental to model performance. /(b)
9-Kavetski et al. (2006b)	(1) French Broad River, the United States/ 2450/ 5; (2) Potomac River, the United States/ 2250/ 8;	Use Gaussian multiplier model as input uncertainty model. The multipliers were inferred by Bayesian theory.	NG	Single-bucket version of the VIC model. /daily	(1) Precipitation errors have considerable effects on the predicted hydrographs (prediction limits) and the calibrated parameters. /(c)
10-Ajami et al. (2007)	Leaf River Basin, the United States/ 1949/ NG	Use random multipliers sampled from Gaussian distribution to represent precipitation errors	NG	(1) Lumped SAC-SMA model; (2) Lumped HYMOD model; (3) Lumped SWB model. /daily	(1) Ignoring input error or model error will cause unrealistic model simulation and incorrect uncertainty bounds. /(c)
11-This study	Four sub-basins of the Xiangjiang Basin, China/ 6053, 9972, 52150, 81638/ 188	For each of six gauge densities, randomly select 1000 subsets from available gauges.	Thiessen polygons	Lumped Xinanjiang model /daily	(1) The impacts of imperfect rainfall estimates on model efficiency can be reduced to some extent through the adjustment of model parameters; (2) modelling uncertainty can be reduced through the increase of rain gauge density or the optimization of rain gauge distribution pattern. /(a,b,c)

Table 2. The drainage area and number of available rain gauges in each sub-basin.

NO.	Gauging Station Name	Drainage Area (km^2)	Available Rain Gauges
1	Xiangxiang	6053	19
2	Ganxi	9972	26
3	Hengyang	52150	120
4	Xiangtan	81638	188

Table 3. The rain gauge number and its spatial density for each density level. The table also shows the sampling number for each density level.

Station Name	Density level	1	2	3	4	5	6	7
	Percentage (%)	5	10	20	30	50	70	100
Xiangxiang	Gauge Number	1	2	4	6	10	13	19
	Sampling Number	19	171	1000	1000	1000	1000	1
	Density (number per 10^3 km^2)	0.17	0.33	0.66	0.99	1.65	2.15	3.14
Ganxi	Gauge Number	1	3	5	8	13	18	26
	Sampling Number	26	1000	1000	1000	1000	1000	1
	Density (number per 10^3 km^2)	0.10	0.30	0.50	0.80	1.30	1.81	2.61
Hengyang	Gauge Number	6	12	24	36	60	84	120
	Sampling Number	1000	1000	1000	1000	1000	1000	1
	Density (number per 10^3 km^2)	0.12	0.23	0.46	0.69	1.15	1.61	2.30
Xiangtan	Gauge Number	9	19	38	56	94	132	188
	Sampling Number	1000	1000	1000	1000	1000	1000	1
	Density (number per 10^3 km^2)	0.11	0.23	0.47	0.69	1.15	1.62	2.30

Table 4. Parameters of the Xinanjiang model.

Number	Parameter	Explanation	Unit	Uncertainty analysis
1	KE	Ratio of potential evapotranspiration to pan evaporation	--	Yes
2	WM	Areal mean tension water storage capacity	mm	No
3	UM	Upper layer tension water storage capacity	mm	No
4	LM	Lower layer tension water storage capacity	mm	No
5	B	Tension water distribution index	--	No
6	IMP	Impermeable coefficient	--	No
7	SM	Areal mean free water storage capacity	mm	No
8	EX	Free water distribution index	--	No
9	KI	Outflow coefficient of free water storage to interflow	day^{-1}	Yes
10	KG	Outflow coefficient of free water storage to groundwater flow	day^{-1}	Yes
11	C	Deep layer evapotranspiration coefficient	--	No
12	CI	Interflow recession coefficient	--	Yes
13	CG	Groundwater recession coefficient	--	Yes
14	N	Parameter of Nash unit hydrograph	--	No
15	NK	Parameter of Nash unit hydrograph	--	Yes

Table 5. The prior distributions of parameters to be inferred.

Parameter	KE	KI	KG	CI	CG	NK	α	σ
Prior distribution	Uniform	Uniform	Uniform	Uniform	Uniform	Uniform	Uniform	Jeffreys
Range	[0.8,1.4]	[0,0.5]	[0,0.5]	[0.7,0.95]	[0.95,1.0]	[2,5]	[0,1)	(0,inf)

Table 6. Measures of model uncertainty with precipitation input estimated from the best gauge distribution of each rain gauge density level for the Xiangtan Basin.

Gauge Number	9	19	38	56	94	132	188
$CI95_P$ (%)	15.5	14.6	14.4	15.0	14.9	14.8	15.0
$ARIL_P$	0.09	0.08	0.08	0.08	0.08	0.08	0.08
$CI95_{PM}$ (%)	94.8	94.7	95.1	94.8	94.8	94.8	94.8
$ARIL_{PM}$	1.06	0.99	0.95	0.94	0.95	0.95	0.95
Min NS	0.906	0.910	0.923	0.923	0.923	0.925	0.922
Max NS	0.923	0.929	0.935	0.936	0.936	0.937	0.934

Indices with subscript P measure the 95% confidence interval due to parameter uncertainty; indices with subscript PM measure the 95% confidence interval due to parameter uncertainty and model uncertainty.

Table 7. The standard deviation of model parameters sampled from their posterior distributions under different rain gauge density levels for the Xiangtan Basin.

Gauge Number	KE_F	KE_R	KI_F	KI_R	KG_F	KG_R	CI_F	CI_R	CG_F	CG_R	NK_F	NK_R
9	19.4	100.4	10.4	18.3	6.0	12.3	4.86	8.77	0.44	0.65	37.0	97.4
19	17.8	68.5	10.3	13.4	5.7	8.2	4.62	6.59	0.43	0.56	36.9	58.1
38	16.8	47.5	10.2	12.3	5.5	6.8	4.68	6.03	0.42	0.55	36.8	42.7
56	17.9	37.1	9.8	11.9	5.5	6.3	4.66	5.57	0.39	0.54	36.0	38.9
94	16.7	28.4	10.0	11.7	5.4	6.1	4.72	5.69	0.41	0.55	36.0	37.3
132	16.6	22.4	10.0	11.5	5.3	5.9	4.66	5.26	0.41	0.51	36.0	36.8

Subscript F represents model parameters with precipitation input estimated from the best gauge distribution pattern of each density level; subscript R represents model parameters estimated with 1000 randomly selected precipitation samples as separate model inputs for each density level. For illustration convenience, the standard deviations of all the parameters listed in the table are multiplied by a factor of 1000. Units of the parameters are shown in Table 4.

Table 8. Measures of model uncertainty with the 1000 precipitation samples as separate model inputs for each rain gauge density for the Xiangtan Basin.

Gauge Number	9	19	38	56	94	132
$CI95_P$ (%)	70.2	54.6	41.3	34.6	26.8	21.8
$ARIL_P$	0.49	0.32	0.23	0.19	0.14	0.11
$CI95_{PM}$ (%)	97.4	96.5	95.8	95.5	95.2	95.0
$ARIL_{PM}$	1.26	1.08	1.01	0.98	0.96	0.96
Min NS	0.723	0.819	0.883	0.884	0.885	0.905
Max NS	0.921	0.928	0.934	0.935	0.936	0.936

Figures

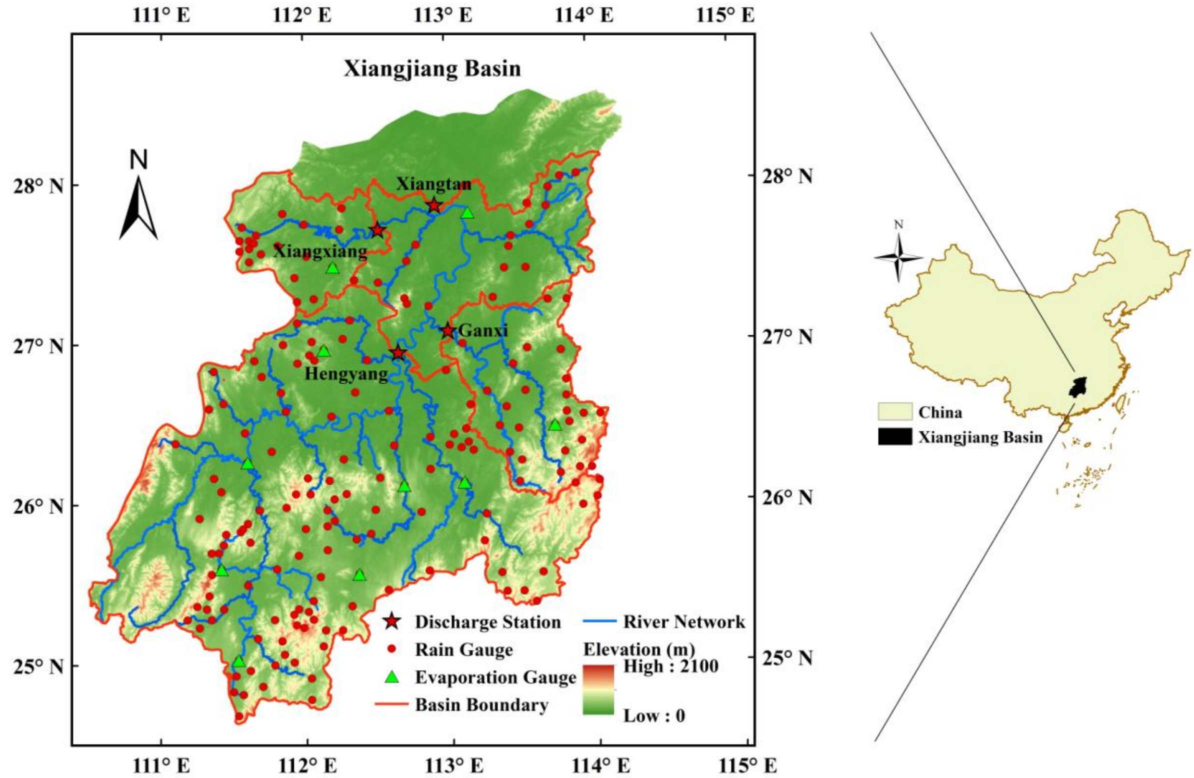


Figure 1. Distribution diagram of discharge stations, evaporation gauges and rain gauges in the Xiangjiang Basin.

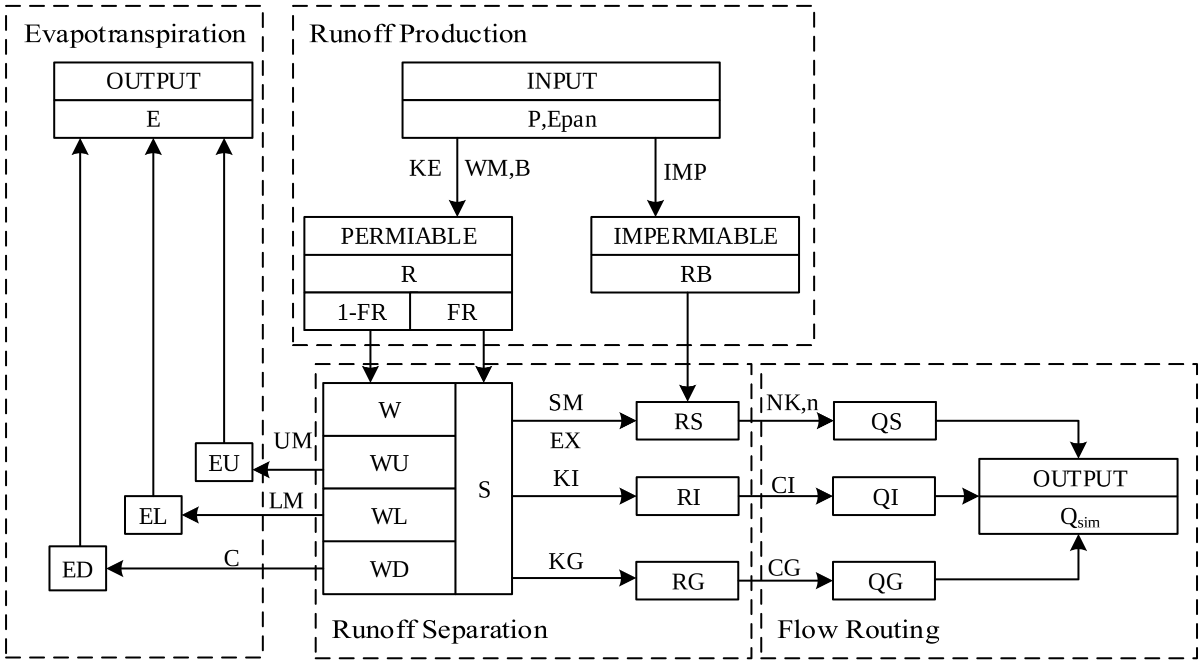


Figure 2. The flowchart of the Xinanjiang model.

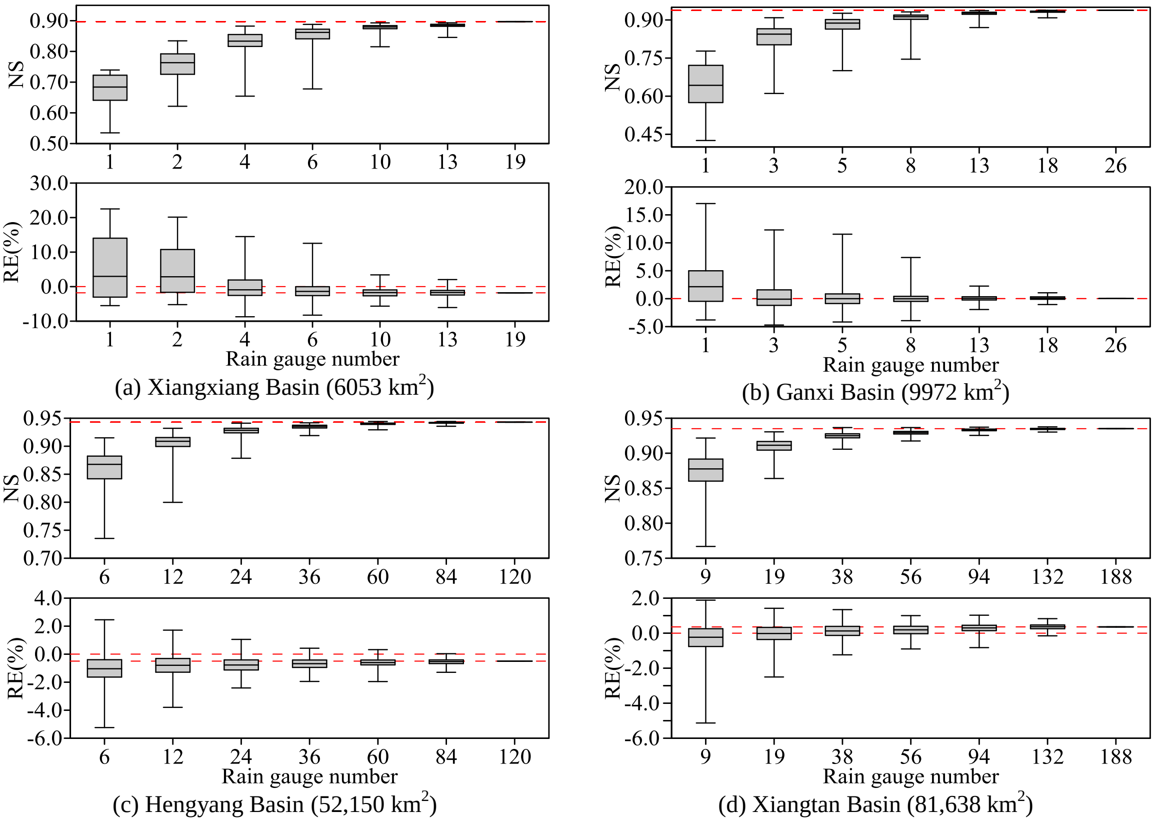
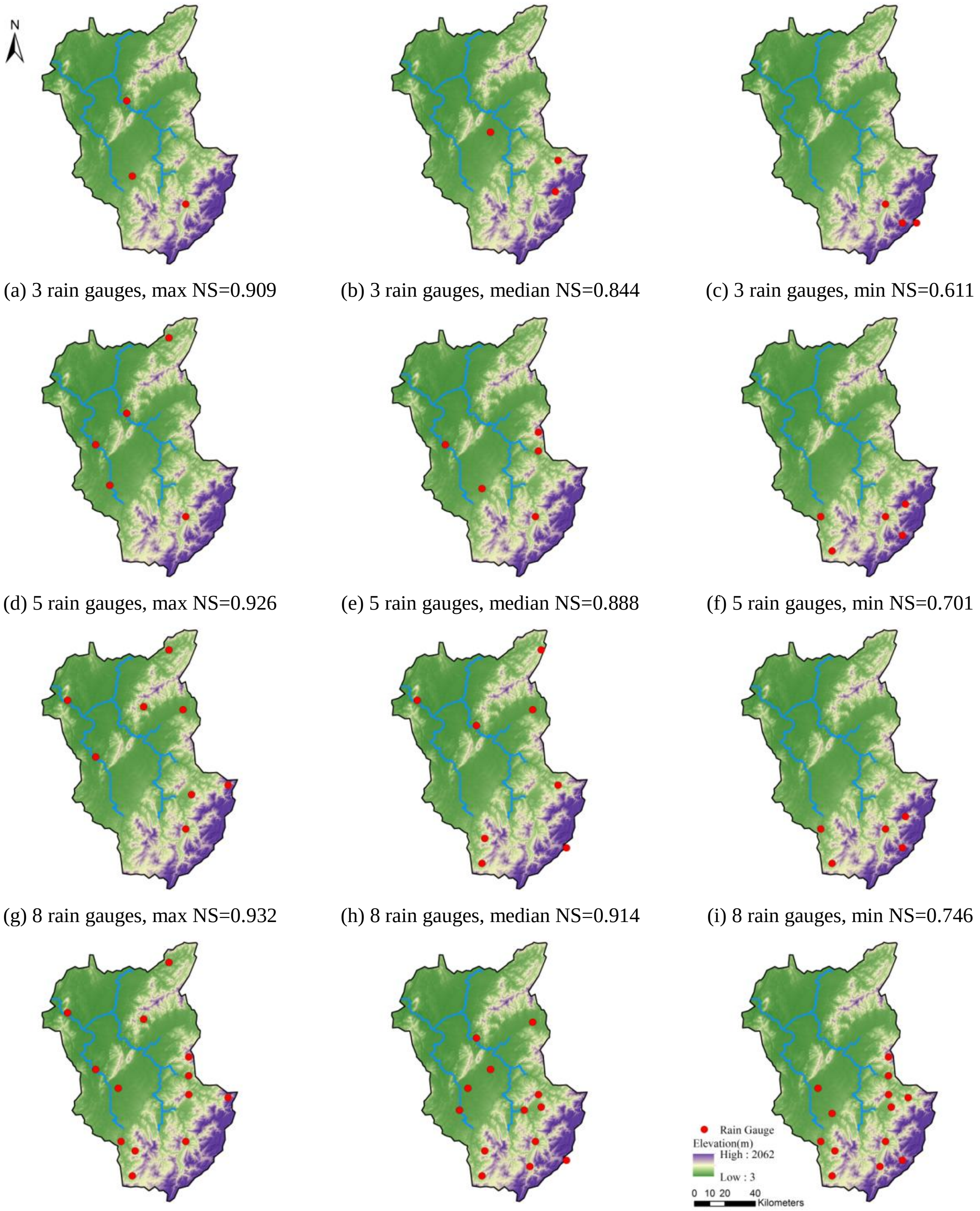


Figure 3. Boxplots of model performance indices under different rain gauge densities. The boxplots show the 25th, 5th, and 75th percentiles and the minimum and maximum value of the indices.



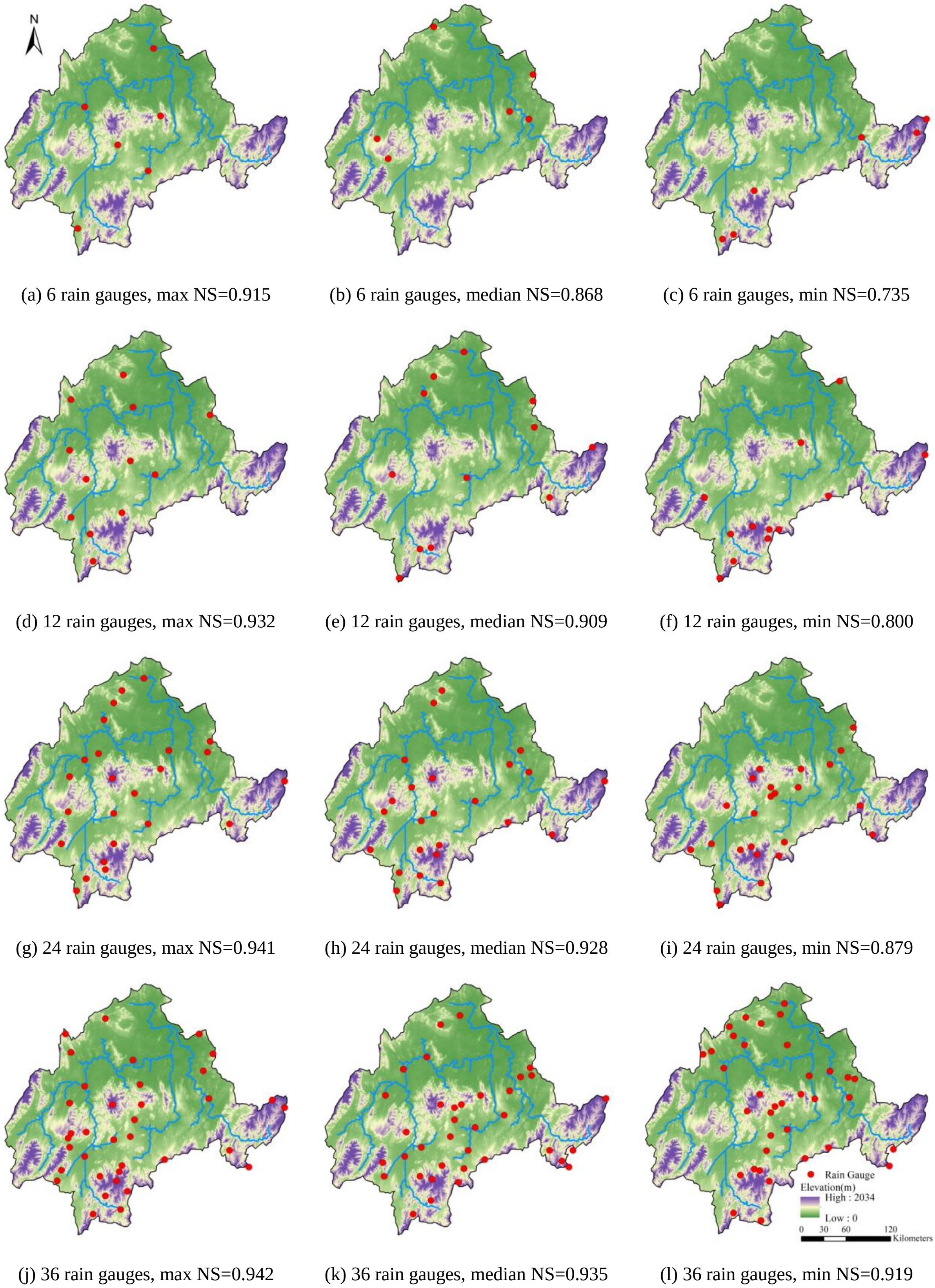
(a) 3 rain gauges, max NS=0.909 (b) 3 rain gauges, median NS=0.844 (c) 3 rain gauges, min NS=0.611

(d) 5 rain gauges, max NS=0.926 (e) 5 rain gauges, median NS=0.888 (f) 5 rain gauges, min NS=0.701

(g) 8 rain gauges, max NS=0.932 (h) 8 rain gauges, median NS=0.914 (i) 8 rain gauges, min NS=0.746

(j) 13 rain gauges, max NS=0.937 (k) 13 rain gauges, median NS=0.924 (l) 13 rain gauges, min NS=0.870

Figure 4. Gauge distributions of 3, 5, 8 and 13 rain gauges with maximum, median and minimum NS values for the Ganxi Basin.



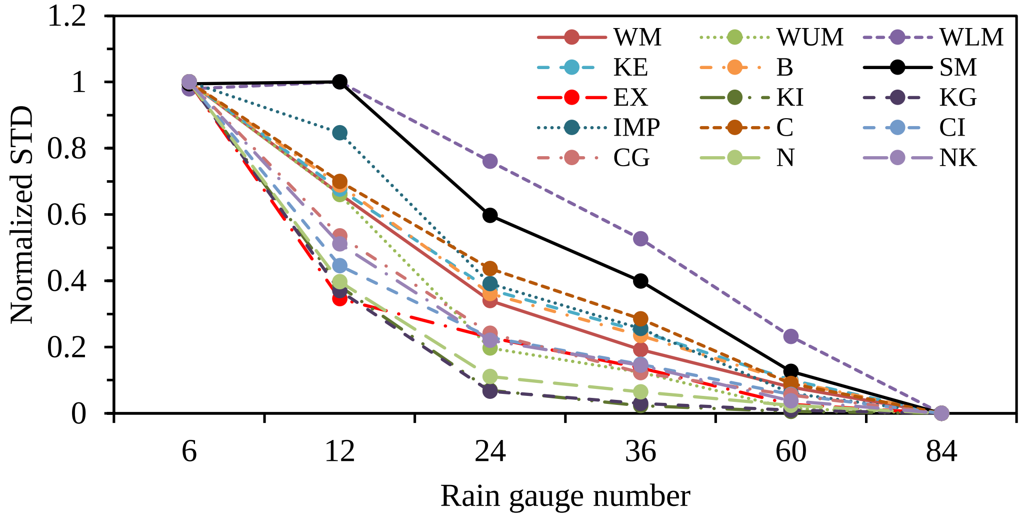
(a) 6 rain gauges, max NS=0.915 (b) 6 rain gauges, median NS=0.868 (c) 6 rain gauges, min NS=0.735

(d) 12 rain gauges, max NS=0.932 (e) 12 rain gauges, median NS=0.909 (f) 12 rain gauges, min NS=0.800

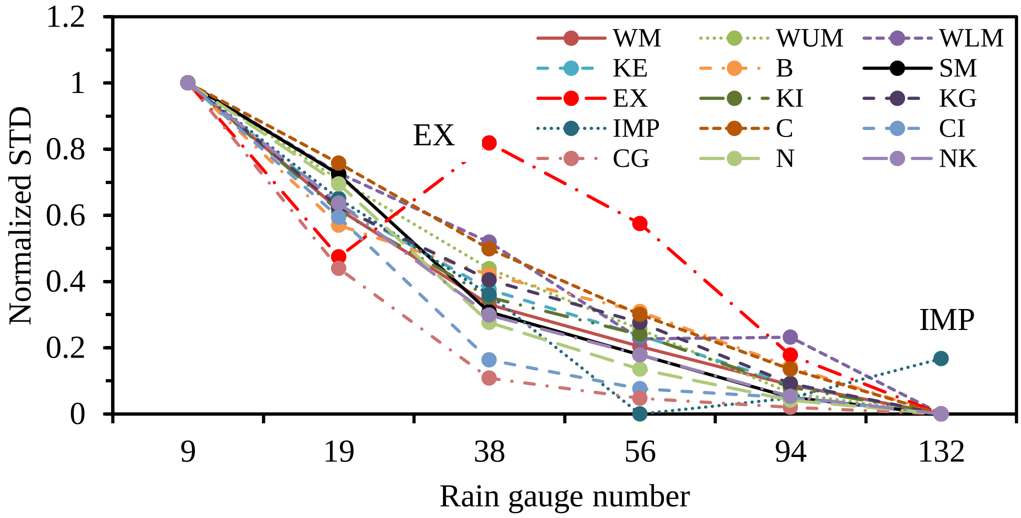
(g) 24 rain gauges, max NS=0.941 (h) 24 rain gauges, median NS=0.928 (i) 24 rain gauges, min NS=0.879

(j) 36 rain gauges, max NS=0.942 (k) 36 rain gauges, median NS=0.935 (l) 36 rain gauges, min NS=0.919

Figure 5. Gauge distributions of 6, 12, 24 and 36 rain gauges with maximum, median and minimum NS values for the Hengyang Basin.



(a) Hengyang Basin



(b) Xiangtan Basin

Figure 6. The normalized STD values of model parameters under different rain gauge numbers. (a) Hengyang Basin; (b) Xiangtan Basin.

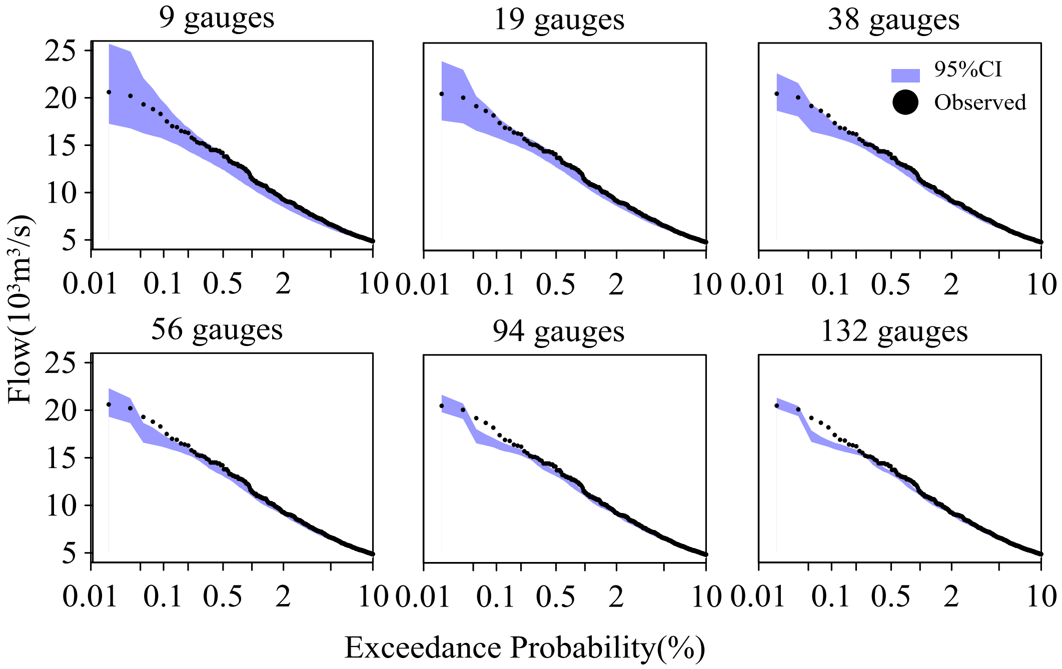


Figure 7. Flow duration curve of the observed runoff and the 95% confidence interval of the 1000 simulated runoffs under different rain gauge densities for the Xiangtan Basin.

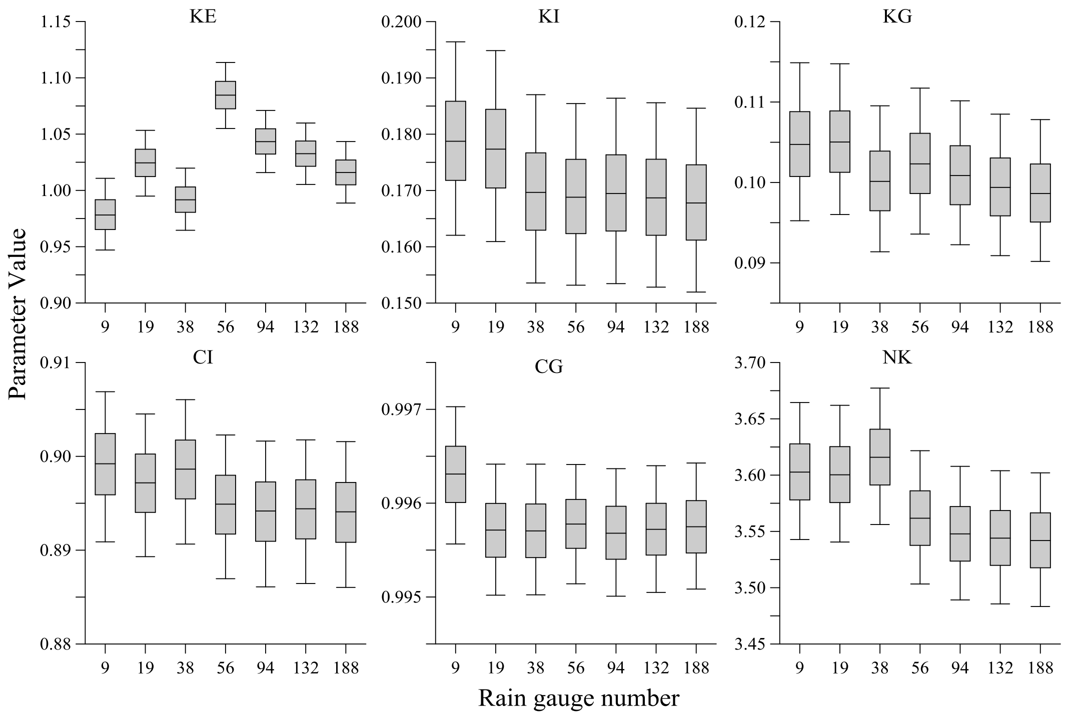


Figure 8. Boxplots of 800,000 model parameter sets sampled from their posterior distributions for the Xiangtan Basin. The boxplots show the 5th, 25th, 50th, 75th, and 95th percentiles of the model parameters.

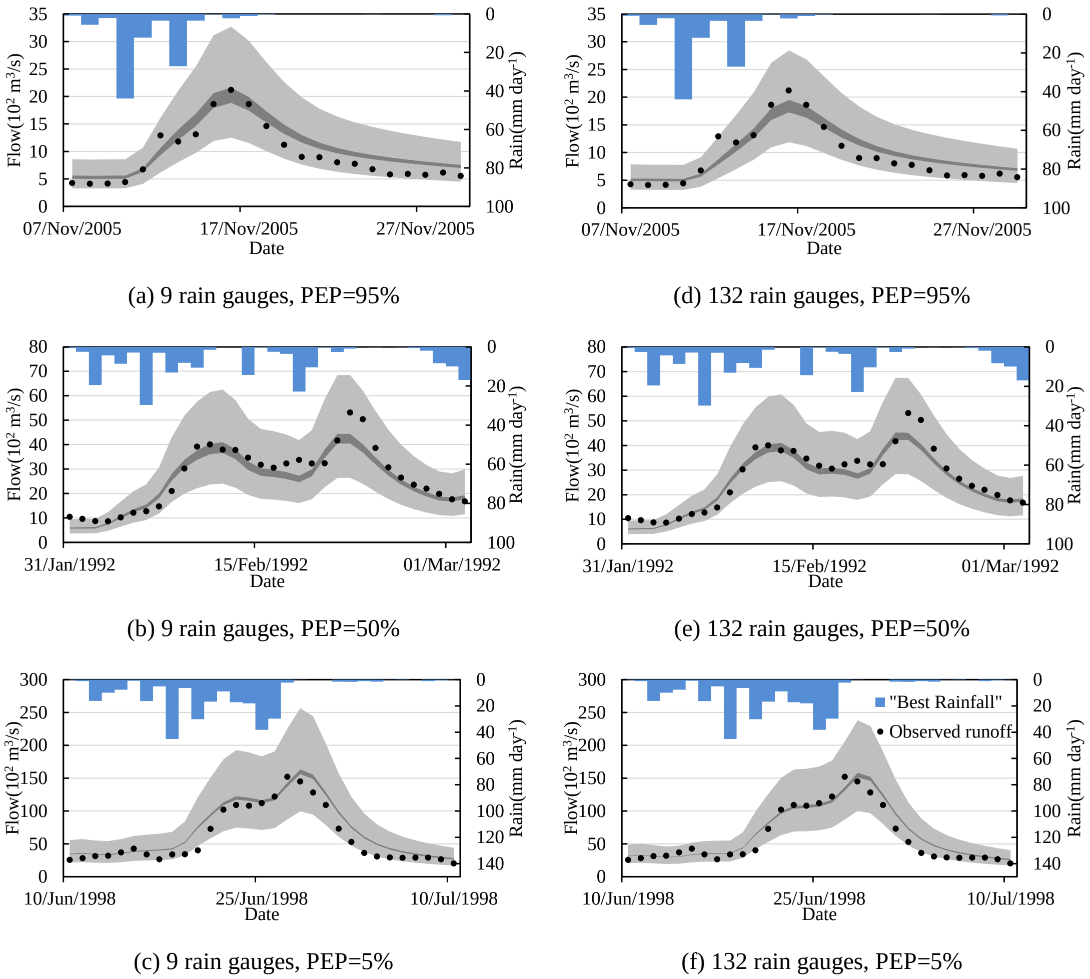


Figure 9. The 95% confidence intervals of simulated runoffs for floods with a peak exceedance probability (PEP) of 5%, 50% and 95% using "best precipitation samples" as model input for density levels with 9 and 132 rain gauges in the Xiangtan Basin. The dark shaded region represents the 95% confidence interval only considering parameter uncertainty; the light shaded region represents the 95% confidence interval considering both parameter uncertainty and model uncertainty.

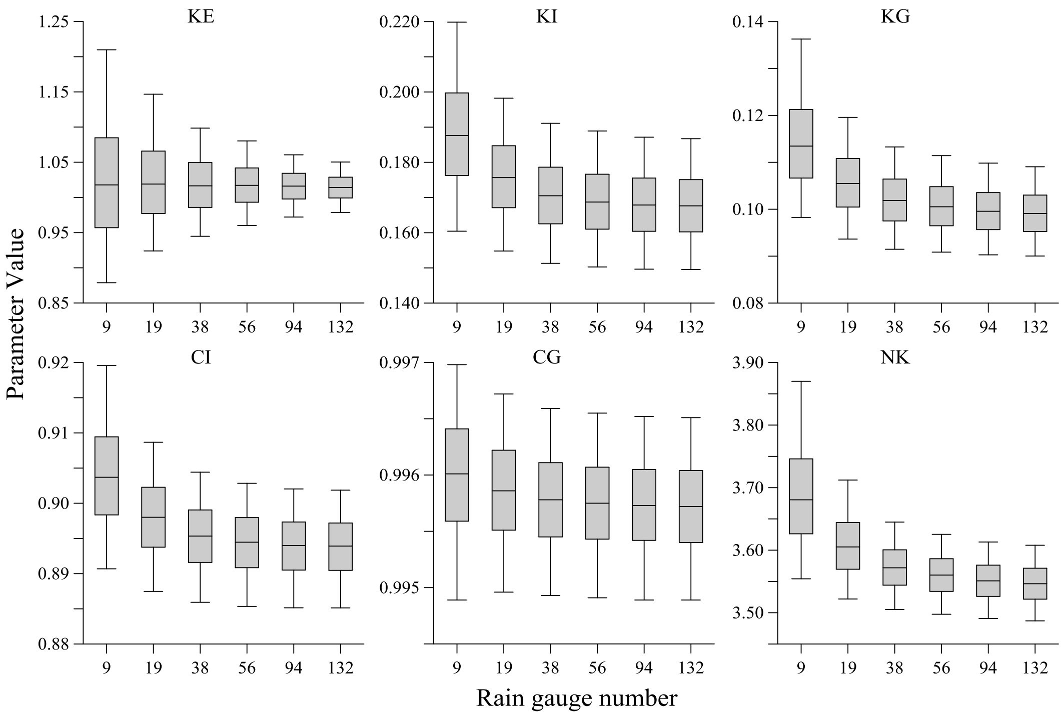


Figure 10. Boxplots of 800,000 model parameter sets sampled from their posterior distributions with the 1000 precipitation samples as model inputs for each rain gauge density level in the Xiangtan Basin. The boxplots show the 5th, 25th, 50th, 75th, and 95th percentiles of the model parameters.

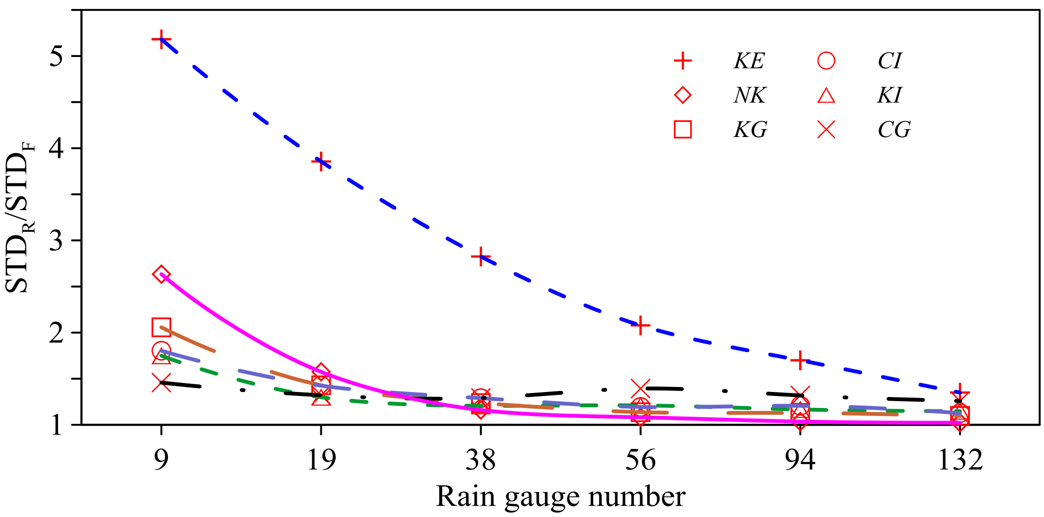
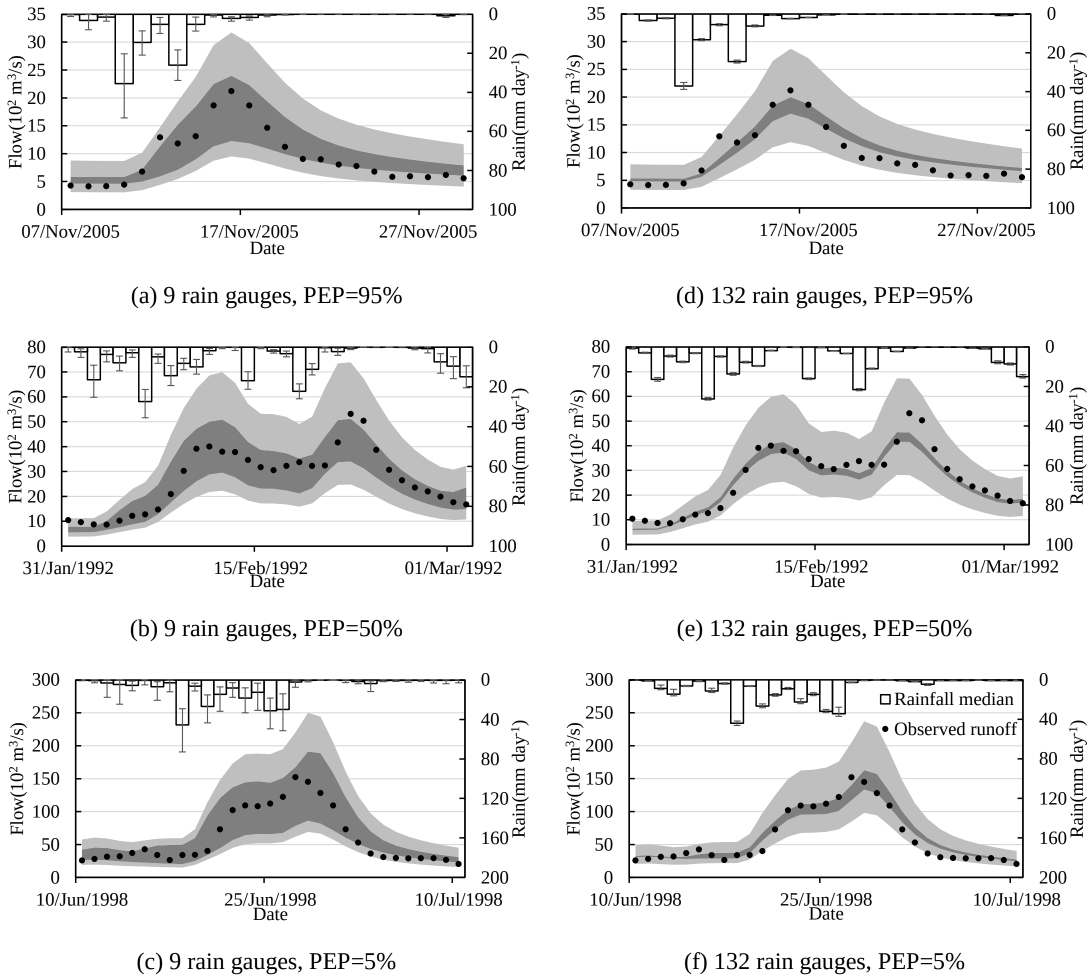


Figure 11. The ratio between parameter standard deviations of random rainfall input and fixed rainfall input in the Xiangtan Basin.



(a) 9 rain gauges, PEP=95%

(d) 132 rain gauges, PEP=95%

(b) 9 rain gauges, PEP=50%

(e) 132 rain gauges, PEP=50%

(c) 9 rain gauges, PEP=5%

(f) 132 rain gauges, PEP=5%

Figure 12. The 95% confidence intervals of simulated runoffs for floods with a peak exceedance probability (PEP) of 5%, 50% and 95% using 1000 randomly selected precipitation samples as separate model inputs for density levels with 9 and 132 rain gauges in the Xiangtan Basin. The dark shaded region represents the 95% confidence interval only considering parameter uncertainty; the light shaded region represents the 95% confidence interval considering both parameter uncertainty and model uncertainty; the white bar graphs with whiskers represent the 2.5%, 50%, and 97.5% percentiles of the 1000 precipitation samples at different times.