Hydrology Research

Quantifying multi-source uncertainties in multi-model predictions using the Bayesian Model Averaging scheme

--Manuscript Draft--

Manuscript Number:	Hydrology-D-16-00272R1
Full Title:	Quantifying multi-source uncertainties in multi-model predictions using the Bayesian Model Averaging scheme
Article Type:	Research Paper
Corresponding Author:	Liliang Ren, Ph. D. Hohai University Nanjing, Jiangsu Province CHINA
Corresponding Author's Institution:	Hohai University
Order of Authors:	Shanhu Jiang
	Liliang Ren
	Chongyu Xu
	Shuya Liu
	Fei Yuan
	Xiaoli Yang
Abstract:	In the study, three widely used hydrological models, e.g., the Xinanjiang (XAJ), hybrid rainfall-runoff (HYB) and HYMOD (HYM) models were first calibrated by two parameter optimization algorithms, namely the Shuffled Complex Evolution (SCE-UA) method and the Shuffled Complex Evolution Metropolis (SCEM-UA) method on the Mishui River basin, south China. Then, the input uncertainty was quantified by utilizing a normally distributed error multiplier. Lastly, the ensemble simulation sets calculated from the three models were combined using the Bayesian Model Averaging (BMA) method. The results indicate that: (1) both SCE-UA and SCEM-UA resulted in good and comparable streamflow simulations that have high Nash-Sutcliffe coefficient (NSE) values and small relative bias (BIAS) values. Specifically, the SCEM-UA implied parameter uncertainty and provided the posterior distribution of the parameters. (2) In terms of the precipitation input uncertainty, the precision of streamflow simulations did not improve remarkably. (3) The BMA combination not only improved the precision of streamflow prediction, but also quantified the uncertainty bounds of the simulation. (4) The prediction interval calculated using SCEM-UA based BMA combination approach appears superior to that calculated using SCE-UA based BMA combination for both the high flows and low flows.

Quantifying multi-source uncertainties in multi-model predictions using the Bayesian Model Averaging scheme

Shanhu Jiang[a], Liliang Ren[a,*], Chong-Yu Xu[a,b], Shuya Liu[a], Fei Yuan[a], and Xiaoli Yang[a]

a. State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering, Hohai University, Nanjing 210098, China

b. Department of Geosciences, University of Oslo, N-0316 Oslo 1047 Blindern, Norway

Submitted to *Hydrology Research* for publication

*Corresponding author.

Professor Liliang Ren

State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering, Hohai University

1 Xikang Road, Nanjing 210098, P. R. China

Email: njRLL9999@126.com

ABSTRACT

Sources of prediction uncertainties in hydrologic modeling are commonly itemized and evaluated individually, while a comprehensive assessment of the effects of different sources of uncertainty on the deterministic simulation and probabilistic assessment is limited. This study focuses on a quantitative multi-source uncertainty analysis of multi-model predictions. Sources of uncertainties considered include the rainfall input uncertainty, parameter uncertainty, and model structural uncertainty. In the study, three widely used hydrological models, e.g., the Xinanjiang (XAJ), hybrid rainfall-runoff (HYB) and HYMOD (HYM) models were first calibrated by two parameter optimization algorithms, namely the Shuffled Complex Evolution (SCE-UA) method and the Shuffled Complex Evolution Metropolis (SCEM-UA) method on the Mishui River basin, south China. Then, the input uncertainty was quantified by utilizing a normally distributed error multiplier. Lastly, the ensemble simulation sets calculated from the three models were combined using the Bayesian Model Averaging (BMA) method. The results indicate that: (1) both SCE-UA and SCEM-UA resulted in good and comparable streamflow simulations that have high Nash-Sutcliffe coefficient (NSE) values and small relative bias (BIAS) values. Specifically, the SCEM-UA implied parameter uncertainty and provided the posterior distribution of the parameters. (2) In terms of the precipitation input uncertainty, the precision of streamflow simulations did not improve remarkably. (3) The BMA combination not only improved the precision of streamflow prediction, but also quantified the uncertainty bounds of the simulation. (4) The prediction interval calculated using SCEM-UA based BMA combination approach appears superior to that calculated using SCE-UA based BMA combination for both the high flows and low flows. The overall results suggest that the comprehensive uncertainty analysis concerning model parameter uncertainties and multi-model ensembles by using the SCEM-UA algorithm and BMA method is superior for streamflow predictions and flood forecasting, because this approach can collectively provide more robust streamflow series and more reliable uncertainty bounds both at calibration and validation periods.

Keywords: hydrological prediction, uncertainty analysis, ensemble, parameter optimization, Bayesian Model Averaging

1 Introduction

Hydrological models have been widely used in watershed hydrological processes simulation, flood forecasting and impact study of climate change and land-use change (Hailegeorgis & Alfredsen, 2015; Emam *et al.*, 2016; Jie *et al.*, 2016); and they play important roles in understanding of the complex hydrologic cycle and solving practical hydrologic problems (Singh *et al.*, 2002). Since 1850s, hydrological models have experienced abundant development from empirical models through lumped conceptual models to physically-based distributed models (Todini, 2011). Nowadays, the precision of hydrological prediction has increased with the development of the model structure and improvement of the input data precision. However, in the hydrological processes simulation and flood forecasting, there still inevitably exist different modeling uncertainties, i.e. parameter uncertainty, input uncertainty and model structural uncertainty (Beven *et al.*, 2000). Quantification and reduction of these uncertainties in hydrological modeling remain as challenges for hydrologists.

Numerous studies have recently focused on the itemized analysis of uncertainties of hydrological modeling (Krzysztofowicz, 1999; Kavetski *et al.*, 2006; Duan *et al.*, 2007; McMillan *et al.*, 2011; Liang *et al.*, 2013; Dong *et al.*, 2013; Yen *et al.*, 2014a; Yen *et al.*, 2015a and 2015b; Zhou *et al.*, 2016). They highlighted that input error quantification, parameter optimization, and multi-model ensemble strategies are the three most popular methods used to demonstrate the impacts of hydrological prediction uncertainties. Rainfall is the most important input data for a hydrological model; thus, adequate characterization of rainfall is fundamental for the success of rainfall-runoff modeling. The true value of the amount of watershed rainfall in practice is often unknown because of its high spatial variability and insufficient rain gauge observations. Hence, an accurate statistical representation of watershed rainfall errors is critical for the estimation of uncertainty of rainfall inputs, which affect streamflow simulations. Kavetski *et al*. (2006) introduced a normally-distributed error multiplier to reduce the precipitation input uncertainty. McMillan *et al*. (2011) evaluated the multiplicative error model of rainfall uncertainty and implied the dependence of rainfall error structure on the time-step data. Yen *et al.* (2015a) assessed the effects of the latent variables on the model simulations and implied the improvement of the model results is still limited. In hydrological modeling, model parameters often need to be calibrated based on observed hydrographs. Two main parameter calibration methods are currently used. In the first method, only one optimal parameter set can be obtained for a basin and model, and the typical algorithms are Genetic Algorithm (GA, Wang *et al*., 1991); Shuffled Complex Evolution (SCE-UA, Duan *et al*., 1992) and Dynamically Dimensioned Search (DDS, Tolson and Shoemaker, 2007). In the other method, the model parameter involves one set of random variables that follow a certain joint probability distribution, and the typical algorithms are Generalised Likelihood Uncertainty Estimation (GLUE, Beven and Binley, 1992); Shuffled Complex Evolution Metropolis (SCEM-UA, Vrugt *et al*., 2003) and Differential Evolution Adaptive Metropolis (DREAM, Vrugt *et al*., 2009). Different optimization algorithms demonstrated different convergence speed and behavioral statistics in model parameter calibration and uncertainty analysis (Xu *et al*., 2013; Yen *et al*., 2014a). Among the mentioned optimization algorithms, the SCE-UA and SCEM-UA approaches have been widely used in parameter calibration and uncertainty analysis in the literature, but the effects of the two algorithms on the deterministic simulation and probability prediction still need to be evaluated and compared further. This consideration has motivated our current study.

Different hydrological models have diverse foci in describing hydrological physical processes. No one model can sufficiently describe the principles of watershed rainfall-runoff in all conditions (Chen *et al*., 2013). Hence, an ensemble strategy based on multiple models has been considered as an effective method to reduce the uncertainty of model structures and improve the precision of hydrological predictions. Different model combination methods, such as neural network (Shamseldin *et al*., 1997), fuzzy system (Xiong *et al*., 2001), and Bayesian model averaging (BMA; Raftery *et al*., 2005), have emerged. In which, BMA is the representative method that can consider the weighted average of the individual predictions from various models. It has been widely used in hydrological ensemble prediction studies. For example, Raftery *et al*. (2005) applied BMA to dynamic numerical weather predictions and attained valuable results. Duan *et al*. (2007), Liang *et al*. (2013), Dong *et al*. (2013), Yen *et al.* (2015b), Arsenault *et al.* (2015) and Zhou *et al.* (2016) successfully used BMA to combine multi-model/multi-method simulations to obtain more robust streamflow series and more reliable probability predictions. Jiang *et al*. (2012, 2014) also applied BMA to merge the multi-satellite precipitation-based streamflow simulations to improve the hydrological utility of satellite precipitation products.

There are also some researches on assessment of the effects of different uncertainty sources on the hydrological modeling (Kavetski *et al*., 2006; Ajami *et al*., 2007; Yen *et al.*, 2014b). While the comprehensive assessment of the effects of different uncertainty sources on the deterministic simulation and probability prediction is still limited. Thus, the current study focuses on uncertainty analysis of multi-source and multi-model hydrological prediction. The innovations of the study include: (1) it considers rainfall input uncertainty, parameter estimation uncertainty, and model structural uncertainty by using three models, i.e., Xinanjiang (XAJ), hybrid rainfall-runoff (HYB), and HYMOD (HYM) models; (2) it compares the effects of SCE-UA and SCEM-UA algorithms on the hydrological prediction results; and (3) it investigates the superiority of the BMA multi-model ensemble strategy over the individual modelling approach. The study is conducted in a humid catchment in southern China. The remainder of this paper is organized as follows. Section 2 introduces the study area and data sets used. Section 3 describes the methodology and models. Section 4 discusses the simulation results of different simulation scenarios. Finally, Section 5 draws the conclusions.

2 Methodology

The flowchart for the multi-source uncertainty analysis of multi-model predictions is shown in Fig.1. We adopted three different simulation cases to systematically consider the three sources (i.e., parameter uncertainty, input uncertainty and model structural uncertainty) of hydrological modeling uncertainties. In case I, the model parameter uncertainty (hereafter "Para") using SCE-UA and SCEM-UA algorithms for three hydrological models, i.e., XAJ, HYB, and HYM, was determined. In case II, a normally distributed error multiplier and combined parameter optimization algorithms were introduced to consider the model input and model parameter uncertainties (hereafter "Para+input"). In case III, the simulations calculated from case II were combined using BMA to comprehensively determine the model input, model parameter, and model structure uncertainties (hereafter "Para+input+struc"). The detailed methodologies are as follows.

Figure 1

2.1 Hydrological models

Xinanjiang model, hereinafter referred to as XAJ, is a well-known conceptual hydrological model developed by Zhao in the 1970s in China (Zhao, 1992). In the present study, a sub-basin-structured semi-distributed XAJ model for streamflow simulation was constructed. The simulation was performed by computing the runoff from each sub-basin, and the slope and river network convergence processes were then integrated to obtain the streamflow series of the hydrologic station. A hybrid rainfall-runoff model, hereinafter referred to as HYB, is a modified version of the XAJ model (Hu *et al.*, 2005). Numerous field studies have shown that runoff within a basin is mainly generated by infiltration excess (Horton) runoff and saturation excess (Dunne) runoff (Ren *et al.*, 2008). HYB model combines the two runoff generation mechanisms by introducing spatial distribution curves of soil tension water storage capacity and infiltration capacity. Detailed description of the mechanisms and applications of the HYB model was discussed by Hu *et al*. (2005). HYMOD, hereinafter referred to as HYM, is a simple conceptual lumped hydrological model developed by Moore in the 1980s (Moore, 1985). HYM consists of a simple rainfall excess model, which is connected to two series of linear reservoirs to route surface and subsurface flow. In the present study, an evaporation reduction factor K and a river network routing Muskingum-Cunge model were added to the original HYM. These three hydrological models have different complex model structure and different runoff generation mechanisms. They have been successfully and widely used in different river basins for streamflow simulation and flood forecasting (Ajami *et al.*, 2007; Ren *et al.*, 2008; Najafi *et al.*, 2011; Jie *et al.*, 2016; Xu *et al.*, 2016). Tables 1-3 show the parameters and their prior ranges of the three models.

Table 1

Table 2

Table 3

The models were operated on daily time step within the 15 sub-basins in Mishui basin. Calibration period was from January 2000 to December 2005, and the period from January 2006 to December 2008 was used as validation period. This period of data was considered to be more representative of the current climate and landuse situation of the study region.

2.2 Input error modeling

The main inputs of the hydrological models are the hydro-meteorological data sets, in which precipitation is the most important one (Ajami *et al.*, 2007). In this study, we adopted an error multiplier to determine the precipitation input uncertainty.

$$P_t = \varphi_t \cdot \tilde{P}_t \tag{1}$$

$$\phi_t = N(m, \sigma_m^2) \tag{2}$$

where $\tilde{P}_t$ and P_t are the measured and modified precipitation at time step t, respectively; φ_t is a normal error multiplier with a mean value of m and a variance of σ_m^2 at time step t. Based on the research of Ajami *et al* (2007), we assume that $m \in [0.9, 1.1]$ and

$\sigma_m^2 \in [10^{-5}, 10^{-3}]$.

2.3 Parameter optimization

SCE-UA is an effective and efficient global optimization algorithm proposed by Duan *et al*. (1992). It has been widely used in hydrological model parameter optimization. SCE-UA combines the direction searching of deterministic, non-numerical methods and the robustness of stochastic, non-numerical methods. It adopts the competition evolution theory, concepts of controlled random search, complex shuffling method, and downhill simplex procedures to obtain a global optimal estimation. Detailed calculation steps of SCE-UA are found in the study of Duan *et al*. (1992).

SCEM-UA was built upon the principles of SCE-UA. Vrugt *et al*. (2003) combined the strengths of the Monte Carlo Markov Chain sampler with the concept of complex shuffling from SCE-UA to form the SCEM-UA algorithm, which not only provides the most probable parameter set, but also estimates the uncertainty associated with estimated parameters. SCEM-UA can simultaneously identify the most likely parameter set and its associated posterior probability distribution in every model run (Ajami *et al*., 2007). SCEM-UA has been successfully used in hydrologic and climate applications, such as rainfall-runoff model parameter calibration and uncertainty analysis (Ajami *et al*., 2007; Jiang *et al*., 2014). Detailed calculation steps of SCEM-UA are found in the work of Vrugt *et al*. (2003). In the present study, initial samples were obtained and then computations using SCEM-UA were performed using datasets with 5,000 and 10,000 samples.

2.4 BMA

BMA is a scheme for model combination that derives consensus predictions from competing predictions using likelihood measures as model weights. BMA has been primarily used to generalize linear regression applications. Raftery *et al*. (2005) successfully applied BMA to dynamic numerical weather predictions. Duan *et al*. (2007) and Ajami *et al*. (2007) used the BMA scheme to combine multiple models for hydrologic ensemble prediction that can provide more skillful and reliable predictions. The advantage of BMA is that the weights are directly bound with individual model simulation, that is, a well performing model can receive a higher weight than a poorly performing one. A more stable result can be obtained when BMA method is used to combine different simulations. In the present study, we use BMA to merge the streamflow simulations from the three different hydrological models. Detailed calculation steps of the BMA method are found in the studies of Duan *et al*. (2007) and Ajami *et al*. (2007). For the sake of completeness, a brief description of the essence of the BMA scheme is presented as follows.

Consider y is BMA prediction, $D=\left[X,Y\right]$ are observed data sets (in which X denotes input forcing data and Y is observed streamflow data) and $f=[\mathrm{f}_1,\mathrm{f}_2,\cdots,\mathrm{f}_k]$ is the ensemble of the *K*-member predictions. The posterior distribution of the BMA prediction y is given as

$$p(y\,|\,D)=\sum_{k=1}^{K}p(\mathrm{f}_k\,|\,D)\cdot p_k(y\,|\,\mathrm{f}_k,D) \tag{3}$$

Where $p(f_k|D)$ is the posterior probability of the prediction f_k given the input data D, and it reflects how well model f_k fits Y. Actually $p(f_k|D)$ is the BMA weight w_k, and better performing predictions receive higher weights than the worse performing ones, and all weights are positive and should add up to 1. $p_k(y|f_k,D)$ is the conditional probability density function (PDF) of the prediction y conditional on f_k and D. Thus, the posterior mean and variance of the BMA prediction could be expressed as:

$$E[y\,|\,D]=\sum_{k=1}^{K}w_k f_k \tag{4}$$

$$Var[y\,|\,D]=\sum_{k=1}^{K}w_k\left(f_k-\sum_{i=1}^{K}w_i f_i\right)^2+\sum_{k=1}^{K}w_k\sigma_k^{\,2} \tag{5}$$

Where $\sigma_k^{\,2}$ is the variance associated with model prediction f_k with respect to observation D. Compared with the deterministic multi-model combination method, BMA can better describe the uncertainty of analog variable. In this study, we use the expectation-maximization (EM) algorithm to estimate the BMA weight w_k and model prediction variance σ_k^2 (Ajami *et al*., 2007).

2.5 Prediction uncertainty interval

For SCE-UA-based simulation, the BMA weights and the variances of each model in the combination process were calculated, and then Monte Carlo Markov Chain sampling method was used to calculate the prediction uncertainty interval (Duan *et al*., 2007). Based on the repeated sampling experiments, we set the sampling times as 1000. For SCEM-UA-based simulation, 15000 streamflow series in the BMA combination process were simulated, and then normal population interval estimation method was used to calculate the prediction uncertainty interval (Ajami *et al*., 2007).

2.6 Evaluation statistics

The validation statistical indices Nash-Sutcliffe coefficient (NSE), relative bias (BIAS), and root mean square error (RMSE) were employed to evaluate hydrologic model performance based on the observed and simulated streamflow series. These three indices jointly measured the consistency of the simulated and observed streamflow series both in terms of temporal distribution and amount. The formulas for NSE, BIAS and RMSE are given as

$$\mathrm{NSE} = 1 - \frac{\sum_{i=1}^{n}(Q_{oi} - Q_{si})^2}{\sum_{i=1}^{n}(Q_{oi} - \overline{Q_o})^2} \tag{6}$$

$$BIAS = \frac{\sum_{i=1}^{n} Q_{si} - \sum_{i=1}^{n} Q_{oi}}{\sum_{i=1}^{n} Q_{oi}} \times 100\% \tag{7}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Q_{si} - Q_{oi})^2} \tag{8}$$

Where Q_{oi} and Q_{si} are the observed and simulated runoff at time step i, respectively, $\overline{Q_o}$ and $\overline{Q_s}$ are the mean values of the observed and simulated streamflow values, respectively, and n is the number of simulation days.

Other validation statistical indices comprising containing ratio (CR), average bandwidth (B), and average deviation amplitude (D) were adopted to evaluate the prediction bounds of the hydrological models (Xiong *et al.*, 2009). CR, expressed as percentage, denotes the ratio of the number of observed streamflows enveloped by prediction bounds to the total number of observed hydrographs. B represents the average bandwidth of the whole prediction bounds. With a certain confidence level, a lower B value indicates a better prediction bound. D denotes the actual discrepancy between the trajectories consisting of the middle points of the prediction bounds and the observed hydrograph. It also shows the symmetry with respect to the observed discharges and the middle point of the prediction bounds. The formulas for CR, B, and D are given as

$$CR = \frac{n_c}{n} \times 100\% \tag{9}$$

$$B = \frac{1}{n}\sum_{i=1}^{n}(q_{ui} - q_{li}) \tag{10}$$

$$D = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{2}\left|(q_{ui} + q_{li}) - Q_{oi}\right| \tag{11}$$

Where n_c is the number of observed streamflows enveloped by prediction bounds, n is the total number of observed hydrographs, and q_{ui} and q_{li} are the upper and low boundaries of the prediction bounds at time step i, respectively.

3 Study area and Data

3.1 Study area

Mishui basin, a tributary of the Xiangjiang River, with a drainage area of 9, 972 km^2 above the Ganxi hydrologic station, was selected as the study area (Figure 2). The basin is located southeast of Hunan Province in Southern China and extends from longitudes 112.85°E to 114.20°E and latitudes 26.00°N to 27.20°N. The basin has a complex topography, with elevations ranging from 49 m to 2093 m above sea level. The climate is of humid subtropical monsoon type, with annual average temperature of approximately 18.0 °C and mean annual precipitation of approximately 1561.0 mm. Temporal and spatial distributions of precipitation in the study region are uneven because of atmospheric circulation and most of the annual precipitation occurs between April and September. During these months, particularly in June, basin-wide heavy rains continuously occur, thereby resulting in flash floods. This multi-model ensemble prediction method can reduce the streamflow prediction and flood forecasting uncertainties, thus it is important to decision support system for such river basins to prevent flood disasters and reduce flood damages.

Figure 2

3.2 Data

The daily precipitation data from 2000 to 2008 were obtained from 35 rain gauge stations in the Mishui basin. For the same period, daily streamflow and potential evapotranspiration data were collected from the Ganxi hydrologic station and Wulipai evaporation station, respectively. This period of data was considered to be more representative of the current climate and landuse situation of the study region. The inverse distance weighting of the three nearest rain gauges was used to obtain the spatially distributed precipitation database of 15 sub-basins for the Mishui basin. The 30 arc-second global digital elevation model data were obtained from the U.S. Geological Survey. The vegetation-type data obtained from the International Geosphere-Biosphere Program were calculated and showed the land use distribution in the basin as forest and shrubs (54.4%), grasslands (33.5%), cropland (11.8%), and urban and water (0.3%).

4 Results and Discussions

4.1 Parameter uncertainty analysis

The model parameters' prior ranges are defined in Tables 1-3 according to the physical meanings of the parameters and the actual hydro-climatic conditions of the

Mishui basin. The SCE-UA algorithm gives a set of optimal solution of the model parameters, while the SCEM-UA algorithm estimates the posteriori probability density functions (PDFs) of the model parameters, which can reflect the effect of the model parameters uncertainty on simulation result. Extraction 10000 group model parameters after convergence of the SCEM-UA algorithm to plot the parameter frequency histograms, in which the peak value of the posterior PDFs of the parameters is the optimal parameter value for all samples. The marginal posterior probability distribution of the XAJ parameters estimated by SCEM-UA in case I was shown in Figure 3 and the statistical indices of the posterior probability distribution of the parameters estimated by SCEM-UA and the optimal parameters estimated by SCE-UA in case I were shown in Table 4. The histograms of XAJ parameters suggested that 12 parameters such as Kc, WDM, and so on (including all the sensitive parameters) approximately follow the normal distribution or the log-normal distribution. While the rest of the two parameters such as WLM and EX have two or more modal values, and this will increase the uncertainty of parameters optimization. Table 4 shows that the parameters WDM, EX and CS0 have large CV values, implying that the mean value of the three parameters has poor representative power and big uncertainty. Some optimal parameters estimated by SCE-UA and SCEM-UA have some differences, and the possible reason may be due to the correlation between parameters and the “equifinality concept” that different parameter sets may produce similar hydrologic behaviors (Beven and Binley, 1992). Similar to the XAJ model results, most parameters of the HYB model and all parameters of the HYM model approximately follow the normal distribution or the log-normal distribution, which explaining the effectiveness of the SCEM-UA optimization algorithm. Generally, the HYM model has less number of parameters, which are easy to obey normal distribution. The XAJ and HYB models have more parameters, for the influence of the correlation between parameters, their parameters’ uncertainty is larger than HYM model.

Figure 3

Table 4

In order to consider the parameter and input uncertainty together, two rain input error modeling parameters *m* and σ_m^2 are added to model parameter sets and further estimate the posterior PDFs simultaneously in case II. Figure 4 shows the marginal posterior probability distribution of the XAJ parameters estimated by SCEM-UA in case II. Table 5 demonstrates the statistical indices of the posterior probability distribution of the parameters estimated by SCEM-UA and the optimal parameters estimated by SCE-UA in case II. Comparing the parameter posterior PDFs of case II with that in case I, it can be concluded that the boundary of the models' parameters posterior distribution moves to a much more reasonable direction, and their posterior distributions are much more closer to normal distribution. The rain input parameter σ_m^2 is hard to concentrate to a single value, and it is difficult to optimize its value. This proved that there were rain input errors in the modeling, and the rain input error multiplier can describe the input errors at a certain extent. While the two rain input parameters may introduce some new parameter estimating uncertainty and increase the difficult of parameter optimization.

Figure 4

Table 5

4.2 Streamflow comparison between BMA ensemble and Single model

For comprehensive consideration of the model input, model parameter, and model structure uncertainties, we used the BMA to combine the three models' simulations at case II. Figure 5 displays the weight estimates of different models calculated using the BMA method. For the SCE-UA-based simulations, the weights of the XAJ, HYB and HYM models are 0.36, 0.31 and 0.33, respectively. For the SCEM-UA-based simulations, the mean values of the weights of the XAJ, HYB and HYM models are 0.35, 0.32 and 0.33, respectively. The weight of the BMA method is directly bound to individual model simulation, that is, a well performing model can receive a higher weight than a poorly performing one in theory. In this study, the XAJ model got the highest weight value, and then followed by the HYM model and the HYB model. The HYM model got a higher weight value than that of the HYB model, which may be due to the similar model structure of the XAJ model and the HYB model (Ren *et al.*, 2008). By using the BMA combination, we can obtain deterministic streamflow series and probability predictions, which comprehensively considered the multi-source uncertainties.

Figure 5

Table 6 shows the statistical performances of the streamflow simulations based on the SCE-UA and SCEM-UA algorithms of the three simulation cases (in which the value set in boldface refers to the optimum performance in the column). Figures 6-8 show the BMA combined streamflow series from the SCE-UA-based simulations and the SCEM-UA-based simulations of the three simulation cases, respectively. From Table 6 and Figures 6-7, we can see that the three models showed a good hydrologic prediction applicability in the Mishui basin, in which the XAJ model performed best, followed by the HYB model, and lastly, the HYM model. Especially for the high flow simulations, the XAJ model and the HYB model performed much better than the HYM model simulation. Generally, both parameter optimization algorithms generated good and comparative streamflow simulations. The SCEM-UA implied parameter uncertainty and provided the posterior distribution of the parameters. Using the 15000 simulation sets, SCEM-UA showed a certain advantage over the SCE-UA algorithm in the calculation of the prediction uncertainty bounds. Given the precipitation input uncertainty in case II, the precisions of the simulated streamflows using the three models were not remarkably enhanced. This phenomenon may have been caused by the relatively small precipitation input uncertainty because of the dense rain gauge observations in the Mishui basin. Moreover, in the model parameters, an evaporation reduction factor parameter K was set, and this parameter could imply some precipitation input uncertainty. Our results are quite consistent with those of Yen *et al.* (2015a), which reported that the use of error multiplier to incorporate input uncertainty might not be the proper alternative choice in terms of generating better results. In case III, for both the SCE-UA and SCEM-UA algorithms, BMA combinations of the simulation sets improved the precision of streamflow predictions, especially during the validation period. This condition was indicated by the high NSE and the small BIAS and RMSE values from BMA combinations compared with those from each single model (see Table 6). The daily NSE, BIAS, and RMSE values of the SCE-UA-based BMA combination in case III for the calibration period were 0.91, 0.04%, and 35.99 m^3/s, respectively; and the corresponding values for the validation period were 0.88, 3.85%, and 56.32 m^3/s. The daily NSE, BIAS, and RMSE values of the SCEM-UA-based BMA combination in case III for the calibration period were 0.92, 0.16%, and 34.66 m^3/s, respectively; and the corresponding values for the validation period were 0.87, 3.49%, and 59.93 m^3/s. Using BMA in combining multiple models to conduct ensemble streamflow simulation can effectively improve the precision of streamflow simulations, especially for the validation period.

Table 6

Figure 6

Figure 7

Figure 8

4.3 Prediction interval comparison between BMA ensemble and Single model

Table 7 shows the reliability performance of the calculated 95% confidence interval of the three simulation cases. Figures 6-8 show the 95% confidence interval from the SCE-UA-based simulations (sampling done 1000 times) and from the SCEM-UA-based simulations of the three simulation cases, respectively. Both parameter optimization algorithms generated a certain precision of prediction uncertainty interval. However, the 95% confidence interval of the SCEM-UA-based simulation was much better than that of the SCE-UA-based simulation. With higher CR and lower D values, SCEM-UA algorithm had an advantage in the estimation of prediction uncertainty bounds compared with the SCE-UA algorithm. Given the precipitation input uncertainty in case II, the performance of the calculated 95% confidence intervals of the three models showed minimal improvement in terms of higher CR values, especially for the validation period. In case III, for both the SCE-UA and SCEM-UA algorithms, the reliability performance of the 95% confidence interval calculated from the BMA combined streamflows was much better than the performance of the interval from each signal model (see Table 6). The daily CR, B, and D values of the SCE-UA-based BMA combination for the calibration period were 90.19%, 315.60 m^3/s, and 56.70 m^3/s, respectively; and the corresponding values for the validation period were 90.97%, 348.56 m^3/s, and 69.74 m^3/s. The daily NSE, BIAS, and RMSE values of the SCEM-UA-based BMA combination for the calibration period were 95.62%, 271.15 m^3/s, and 55.03 m^3/s, respectively; and the corresponding values for the validation period were 95.17%, 303.04 m^3/s, and 66.06 m^3/s. The calculated 95% confidence interval from BMA combination had higher CR and better D values than those of each single model. In addition, it also had a higher B value. The increase in the uncertainty interval CR value was accompanied by the increase in B value, and which has already been discussed by Xiong *et al*. (2009) and Dong *et al*. (2013). Thus, using BMA in combining multiple models to perform the ensemble hydrologic simulations can effectively calculate more reliable uncertainty bounds.

Table 7

Figure 8 compares the BMA-combined streamflow mean values and the calculated 95% confidence interval with the observed hydrograph at the daily time scales from the SCE-UA-based simulation and SECM-UA-based simulation for case III. Both SCE-UA- and SCEM-UA-based BMA combinations generated good streamflow simulations and reliable 95% confidence intervals. The precisions of streamflow simulations of the SCE-UA- and SCEM-UA-based simulations were comparatively good, but the reliability of SCEM-UA-calculated 95% confidence interval was much better than that of SCE-UA in terms of higher CR and lower B and D values (Table 7). Figure 6 also demonstrates that the SCEM-UA-calculated 95% confidence interval can preferably contain the observed high flows and this is very important for the flood control decision-making. For the low flow series, the SCEM-UA-based method can give much better confidence interval than that of the SCE-UA-based method. Thus, the results suggest that determining the model parameter uncertainties using SCEM-UA algorithm can generate more reliable uncertainty bounds than that of the simulation from SCE-UA.

4.4 The different performance in Calibration period and Validation period

For hydrological simulation and forecast, the hydrological model must go through the model parameters calibration and validation stages. The hydrological model can be applied to practical use only on the condition that the calibrated model can also perform well in the validation period (Singh and Woolhiser, 2002). While different hydrological models have different instabilities in the calibration and validation periods for their variant climatic conditions, respectively (Yan *et al.*, 2013; Li *et al.*, 2015). Most models can't have the same performance in the validation period as that in the calibration period. Figure 9 compares the hydrological models simulation performances in the calibration period and validation period for the three different simulation cases. Figure 10 shows the distribution of RMSE for XAJ, HYB and HYM considering different uncertainty sources. From Figure 9 and Figure 10, we can see that both at case I and case II, the three hydrological models have better simulation precision in the calibration period than that in the validation period. While, at case III, by using BMA combination of three hydrological models can effectively improve the precision of streamflow predictions in terms of high NSE value, small BIAS and RMSE values in the validation period. Normally, hydrological modelling has higher uncertainties in the validation period than in the calibration period, while the BMA multi-model ensemble strategy can effectively improve this phenomenon and give a higher skill and reliability forecasting for the future (Vel´azquez *et al.*, 2011; Broderick *et al.*, 2016). Thus, choosing appropriate hydrological models, considering the parameter uncertainties, and using the multi-model ensemble strategy, can improve the accuracy of the hydrological forecasting results.

Figure 9
Figure 10

5 Conclusions and Suggestions

This study performed a multi-source uncertainty analysis of hydrological prediction by using input error quantification, parameter optimization and multi-model ensemble methods in a typical humid watershed in Southern China. The results show that both the SCE-UA and SCEM-UA parameter optimization algorithms can make the XAJ, HYB, and HYM models generate good streamflow simulations with NSE values higher than 0.80 and BIAS values smaller than 7.62%. Specifically, the SCEM-UA can imply parameter uncertainty and provide the posterior distribution of the parameters. Thus, the SCEM-UA algorithm has advantageous in the estimation of model parameter uncertainty and predicting reliable hydrological forecasts. Considering precipitation input uncertainty does not improve the precision of streamflow simulation in the selected Mishui basin, which is probably due to the availability of good quality and dense rain gauge stations. While the BMA combination of the simulation sets calculated from single models not only improves the precision of streamflow predictions in terms of NSE and BIAS values, but also quantifies the uncertainty bounds for the simulation sets in terms of CR values. The improvement of the prediction precision of BMA combination is much more evident in the validation period than in the calibration period. This finding demonstrates that the hydrological modelling has more uncertainties in the validation period, and that the BMA multi-model ensemble can effectively reduce these uncertainties. Comparison of the prediction uncertainty interval from the two different parameter optimization algorithms shows that the calculated 95% prediction interval from SCEM-UA-based BMA simulations is much better than that calculated from SCE-UA-based BMA simulations.

Hence, these results suggest that the comprehensive uncertainty analysis concerning model parameters uncertainties and multi-model ensembles by using the SCEM-UA algorithm and BMA method is advantageous and of practical importance for streamflow predictions and flood forecasting, which can collectively provide more robust streamflow series and more reliable uncertainty bounds.

ACKNOWLEDGMENTS

The current study was jointly supported by the National Key Research and Development Program approved by Ministry of Science and Technology, China (2016YFA0601504); the Programme of Introducing Talents of Discipline to Universities by the Ministry of Education and the State Administration of Foreign Experts Affairs, China (B08048); the National Natural Science Foundation of China (41501017, 51579066) and Natural Science Foundation of Jiangsu Province (BK20150815).

References

Ajami, N. K., Duan, Q. Y. & Sorooshian, S. 2007 An integrated hydrologic Bayesian multimodel combination framework: Confronting input, parameter, and model structural uncertainty in hydrologic prediction. *Water Resources Research* 43, W01403.

Arsenault, R., Gatien, P., Renaud, B., Brissette, F. & Martel, J. L. 2015 A comparative analysis of 9 multi-model averaging approaches in hydrological continuous streamflow simulation. *Journal of Hydrology* 529, 754-767.

Beven, K. & Binley, A. 1992 Thefuture of distributed models: Model calibration and uncertainty prediction. *Hydrological Processes* 6(3), 279-298.

Beven, K. 2000 Rainfall-Runoff modelling. Chichester: Wiley.

Broderick, C., Matthews, T., Wilby, R. L., Bastola, S. & Murphy, C. 2016 Transferability of hydrological models and ensemble averaging methods between contrasting climatic periods. *Water Resources Research* 52, 8343–8373.

Chen, X., Yang, T., Wang, X. Y., Xu, C. Y. & Yu, Z. B. 2013 Uncertainty Intercomparsion of Different Hydrological Models in Simulating Extreme Flows. *Water Resources Management* 27, 1393-1409.

Dong, L. H., Xiong, L. H. & Zheng, Y. F. 2013 Uncertainty analysis of coupling multiple hydrologic models and multiple objective functions in Han River, China. *Water Science and Technology* 68, 506-513.

Duan, Q. Y., Ajami, N. K., Gao, X. G. & Sorooshian, S. 2007 Multi-model ensemble hydrologic prediction using Bayesian model averaging. *Advances in Water Resources* 30, 1371-1386.

Duan, Q. Y., Sorooshian, S. & Gupta, V. 1992 Effective and efficient global optimization for conceptual rainfall-runoff models. *Water Resources Research* 28, 1015–1031.

Emam, A. R., Kappas, M. & Hosseini, S. Z. 2016 Assessing the impact of climate change on water resources, crop production and land degradation in a semi-arid river basin. *Hydrology Research* 46, 854-870.

Hailegeorgis, T.T. & Alfredsen, K. 2015 Comparative evaluation of performances of different conceptualisations of distributed HBV runoff response routines for prediction of hourly streamflow in boreal mountainous catchments. *Hydrology Research* 46, 607-628

Hu, C. H., Guo, S. L., Xiong, L. H. & Peng, D. Z. 2005 A modified Xinanjiang model and its application in Northern China. *Hydrology Research* 36, 175-192.

Jiang, S. H., Ren, L. L., Hong, Y., Yong, B., Yang, X. L., Yuan, F., & Ma, M. W. 2012 Comprehensive Evaluation of Multi-satellite Precipitation Products with a Dense Rain Gauge Network and Optimally Merging their Simulated Hydrological Flows using the Bayesian Model Averaging Method. *Journal of Hydrology* 452-453, 213-225.

Jiang, S. H., Ren, L. L., Hong, Y., Yang, X. L., Ma, M. W., Zhang, Y. & Yuan, F. 2014 Improvement of multi-satellite real-time precipitation products for ensemble streamflow simulation in a middle latitude basin in South China. *Water Resources Management* 28, 2259-2278.

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. *Hydrology Research* 47, 718-735.

Kavetski, D., Kuczera, G. & Franks S W, 2006. Bayesian analysis of input uncertainty in hydrological modeling: 1. Theory. *Water Resources Research* 42, W03407.

Krzysztofowicz, R. 1999. Bayesian theory of probabilistic forecasting via deterministic hydrologic model. *Water Resources Research* 35, 2739–2750.

Liang, Z. M., Wang, D., Guo, Y. & Zhang, Y. 2013 Application of Bayesian Model Averaging Approach to Multi-model ensemble Hydrologic Forecasting. *Journal of Hydrologic Engineering* 18, 1426-1436.

Li, H., Beldring, S. & Xu, C. Y. 2015 Stability of model performance and parameter values on two catchments facing changes in climatic conditions. *Hydrological Sciences Journal* 60, 1317-1330.

McMillan, H., Jackson, B., Clark, M., Kavetski, D. & Woods, R. 2011 Rainfall uncertainty in hydrological modelling: An evaluation of multiplicative error models. *Journal of Hydrology* 400, 83-94.

Moore, R. J. 1985 The probability-distributed principle and runoff production at point and basin scales. *Hydrological Sciences Journal* 30, 273-297.

Najafi, M. R., Moradkhani, H. & Jung, I. W. 2011 Assessing the uncertainties of hydrologic model selection in climate change impact studies. Hydrol. Process. 25, 2814–2826.

Raftery, A. E., Gneiting, T., Balabdaoui, F., & Polakowski, M. 2005 Using Bayesian Model Averaging to Calibrate Forecast Ensembles. *Monthly Weather Review* 133(5): 1155-1174.

Ren, L. L., Zhang, W., Li, C. H. & Yuan, F. 2008 Comparison of Runoff Parameterization Schemes with Spatial Heterogeneity across Different Temporal Scales in Semihumid and Semiarid Regions. *Journal of Hydrologic Engineering* 13, 400–409.

Shamseldin, A. Y., O´Connor, K. M. & Liang, G. C. 1997 Methods for combining the outputs of different rainfall-runoff models. *Journal of Hydrology* 197, 203-229.

Singh, V. P. & Woolhiser, D. A. 2002 Mathematical Modeling of Watershed Hydrology. *Journal of Hydrologic Engineering* 7, 270-292.

Todini, E. 2011 History and perspectives of hydrological catchment modeling. *Hydrology Research* 42, 73-85.

Tolson, B. A. & C. A. Shoemaker. 2007 Dynamically dimensioned search algorithm for computationally efficient watershed model calibration. Water Resour. Res. 43, W01413.

Vel´azquez, J. A., Anctil, F., Ramos, M. H. & Perrin, C. 2011 Can a multi-model approach improve hydrological ensemble forecasting? A study on 29 French catchments using 16 hydrological model structures. *Advances in Geosciences* 29, 33-42.

Vrugt, J. A., Gupta, H. V., Bouten, W., Bouten, W. & Sorooshian, S. 2003 A Shuffled Complex Evolution Metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters. *Water Resources Research* 39, 1201.

Vrugt, J. A., Braak, C. J. F., Diks, C. G. H., Robinson, B. A., Hyman, J. M. & Higdon, D. 2009 Accelerating Markov chain Monte Carlo simulation by differential evolution with self-adaptive randomized subspace sampling. Int. J. Nonlinear Sci. Numer. Simul. 10, 271–288.

Wang, Q. J. 1991 The Genetic Algorithm and Its Application to Calibrating Conceptual Rainfall-Runoff Models. *Water Resources Research* 27, 2467–2471.

Xiong, L. H., Shanseldin, A. Y. & O´Connor, K. M. 2001 A non-linear combination of the forecasts of rainfall-runoff models by the first-order Takagi-Sugeno fuzzy system. *Journal of Hydrology* 245, 196-217.

Xiong, L. H., Wan, M., Wei, X. J. & O´Connor, K. M. 2009 Indices for assessing the prediction bounds of hydrological models and application by generalized likelihood uncertainty estimation. *Hydrological Science Journal* 54, 852-871.

Xu, D. M., Wang, W. C., Chau, K. W. & Cheng, C. T. 2013 Comparison of three global optimization algorithms for calibration of the Xinanjiang model parameters. *Journal of Hydroinformatics* 15, 174-193.

Xu, H. L., Xu, C. Y., Chen, S. D. & Chen, H. 2016 Similarity and difference of global reanalysis datasets (WFD and APHRODITE) in driving lumped and distributed hydrological models in a humid region of China. *Journal of Hydrology 542*, 343–356.

Yan, D. H, Yuan, Z., Wang, H. & Yang, Z. Y. 2013 Progress of certainty and uncertainty methods of hydrologic and the framework of ensemble analysis. *SHUILI XUEBAO* 44, 73-82. (in Chinese).

Yen, H., Jeong, J., Tseng, W. H., Kim, M. K., Records, R. M. & Arabi, M. 2014a Computational Procedure for Evaluating Sampling Techniques on Watershed Model Calibration. J. Hydrol. Eng. 20, 04014080-1.

Yen, H., Wang, X. Y., Fontane D. G., Harmel, R. D. & Arabi, M. 2014b A framework for propagation of uncertainty contributed by input data, parameterization, model structure, and calibration/validation data in watershed modeling. Environ. Modell. Softw. 54, 211–221.

Yen, H., Jeong, J., Feng, Q. Y. & Deb, D. 2015a Assessment of Input Uncertainty in SWAT Using Latent Variables. Water Resources Management 29, 1137-1153.

Yen, H., White, M. J., Jeong, J., Arabi, M. & Arnold, J. G. 2015b Evaluation of alternative surface runoff accounting procedures using the SWAT model. Int J Agric & Biol Eng 8, 54-68.

Zhao, R. J. 1992 The Xinanjiang model applied in China. *Journal of Hydrology* 135, 371-381.

Zhou, Y. L., Guo, S. L., Xu, C. Y., Chen, H., Guo, J. L. & Lin, K. R. 2016 Probabilistic prediction in ungauged basin (PUB) based on regional parameter estimation and Bayesian model averaging. *Hydrology Research* 47, 1087-1103.

LIST OF FIGURES

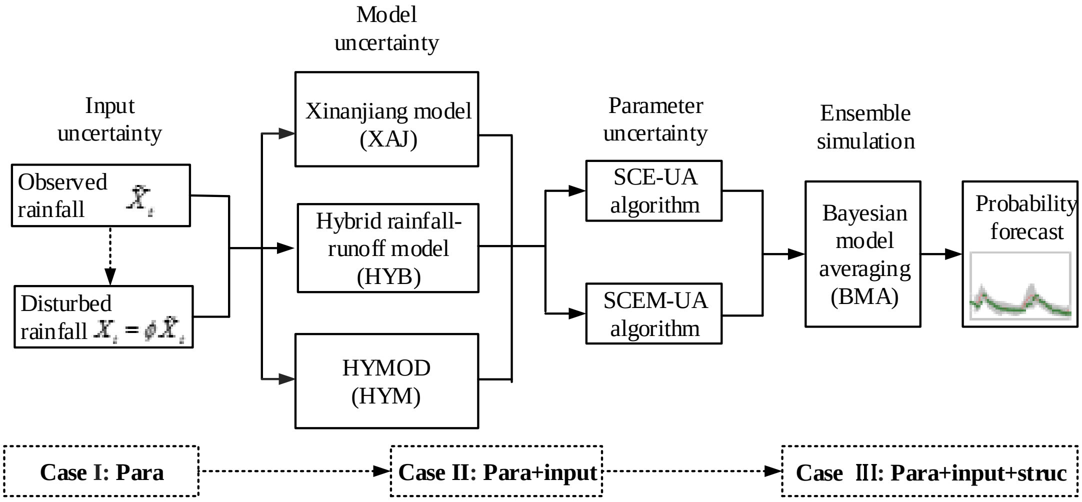

Fig. 1 The flowchart of the multisource hydrological prediction uncertainty analysis. Para indicates considering model parameter uncertainty in case I, Para+input+struc means considering model input, parameter, and structure uncertainties in case III.

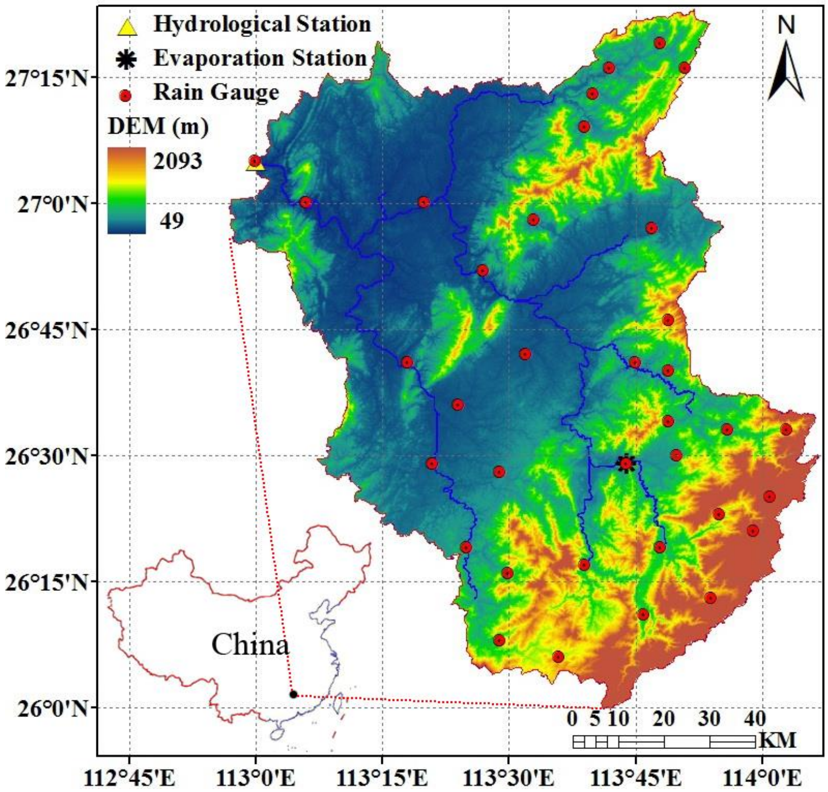

Fig. 2 Location of Mishui basin in South China

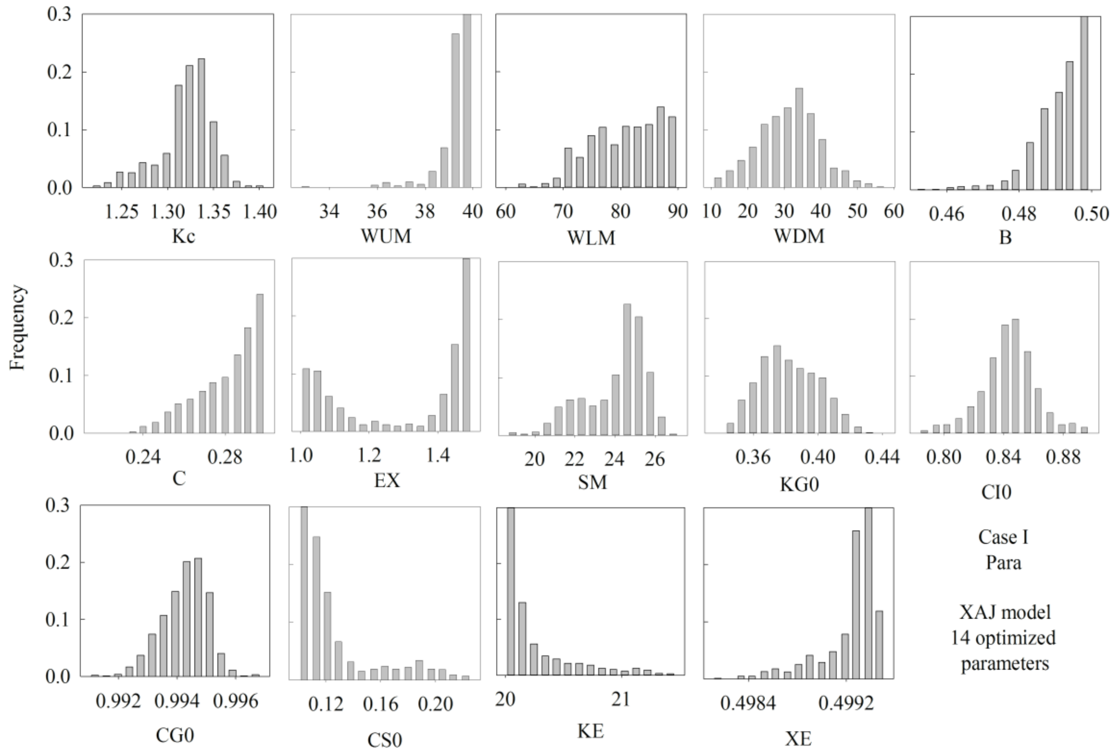

Fig. 3 Marginal posterior probability distribution of the XAJ parameters for case I, using 10 000 samples generated after the SCEM-UA algorithm convergence

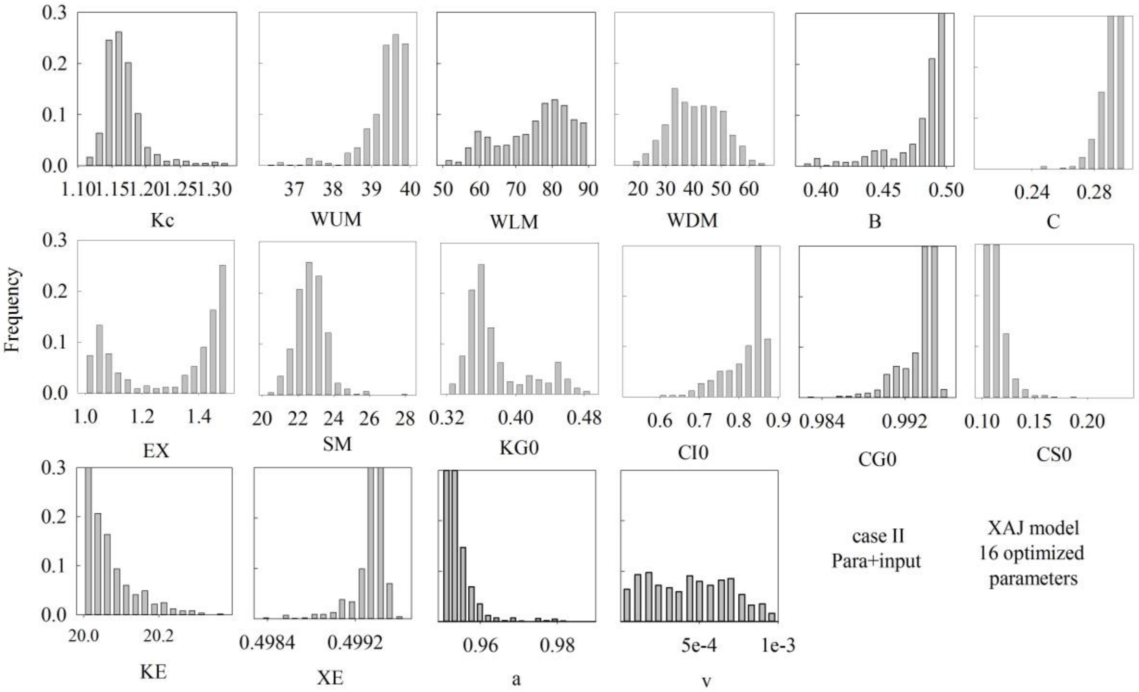

Fig. 4 Marginal posterior probability distribution of the XAJ parameters for case II, using 10 000 samples generated after the SCEM-UA algorithm convergence

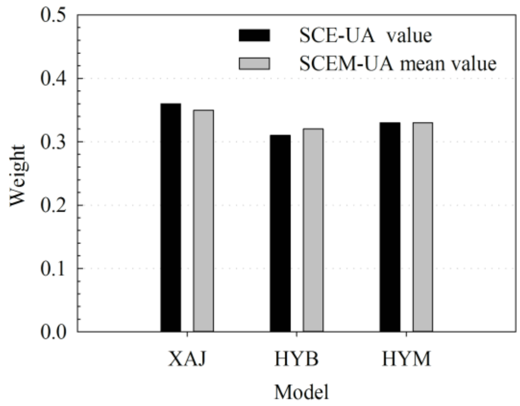

Fig. 5 Histogram of the BMA weights for the different models. SCE-UA value means the BMA weights for the SCE-UA based model simulations. SCEM-UA mean value indicates the BMA weights for the SCEM-UA based model simulations.

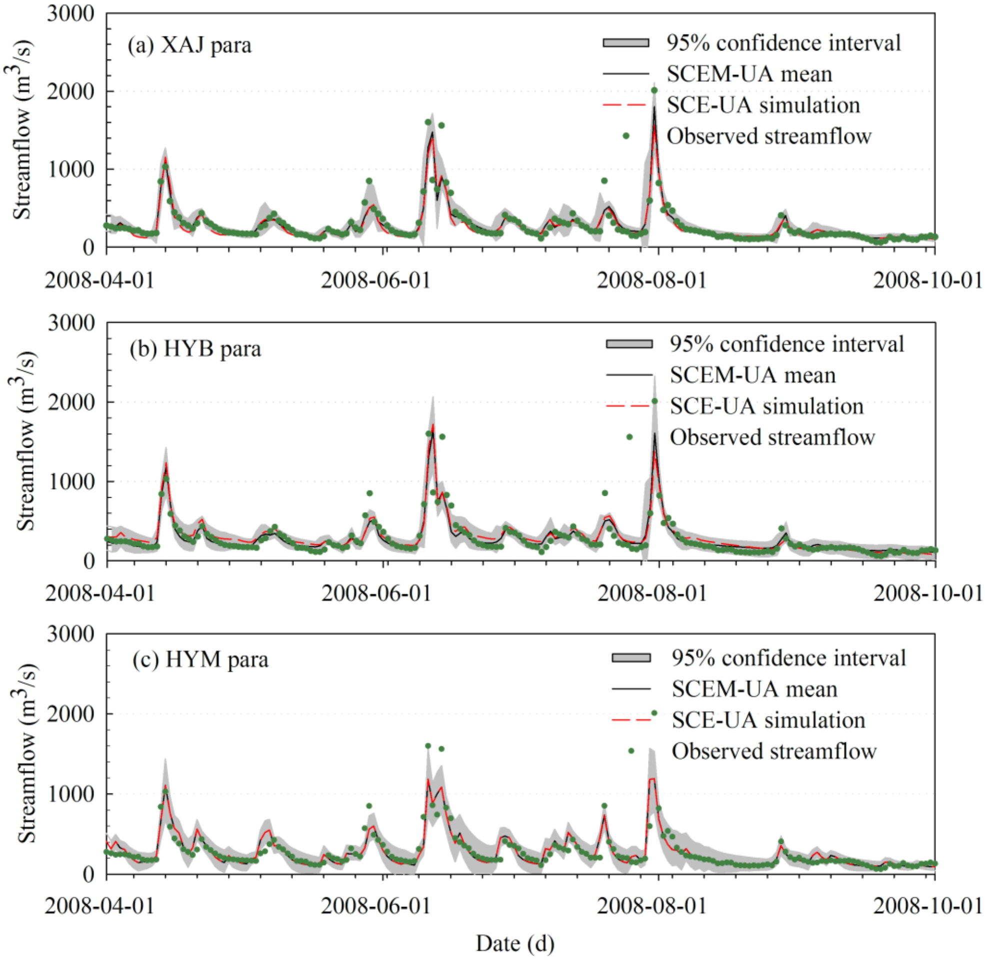

Fig. 6 The streamflow series and the 95% confidence interval of the three hydrological models, (a)-(b) from the SCE-UA-based simulations (sampling done 1000 times) and (c)-(d) from the SCEM-UA-based simulations for case I.

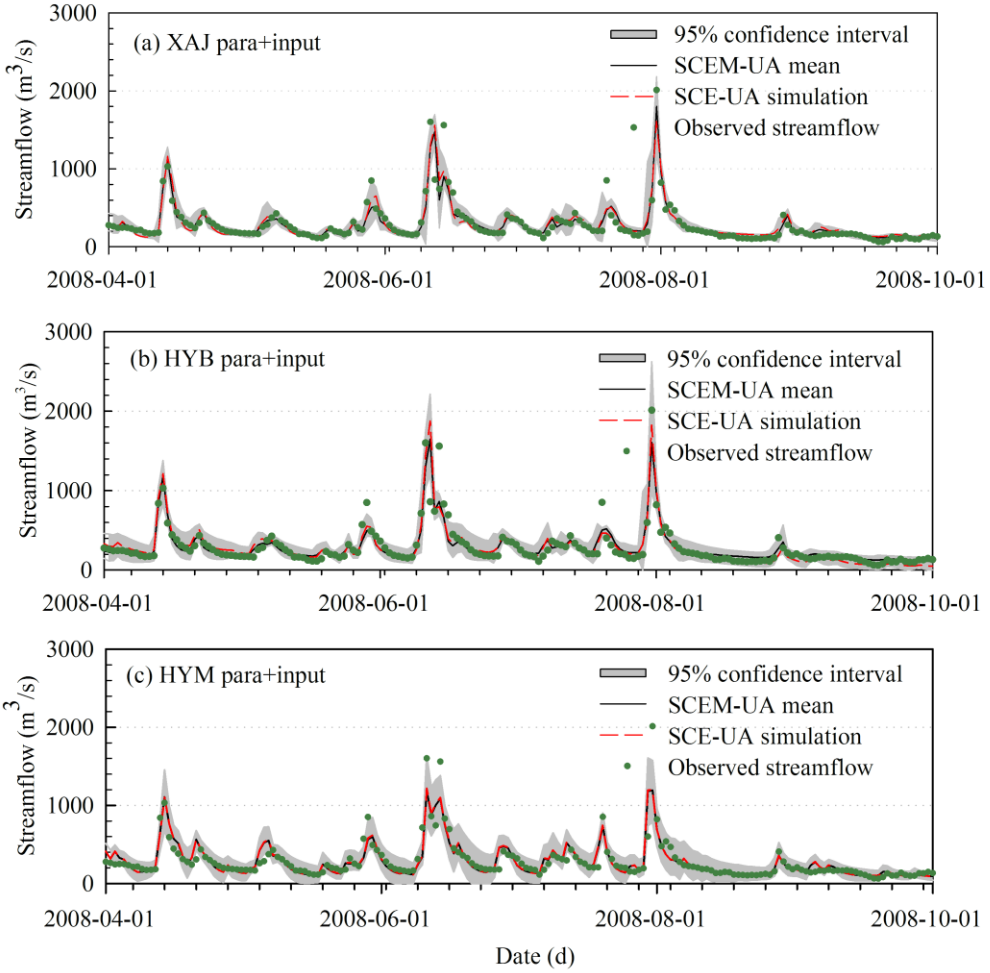

Fig. 7 The streamflow series and the 95% confidence interval of the three hydrological models, (a)-(b) from the SCE-UA-based simulations (sampling done 1000 times) and (c)-(d) from the SCEM-UA-based simulations for case II.

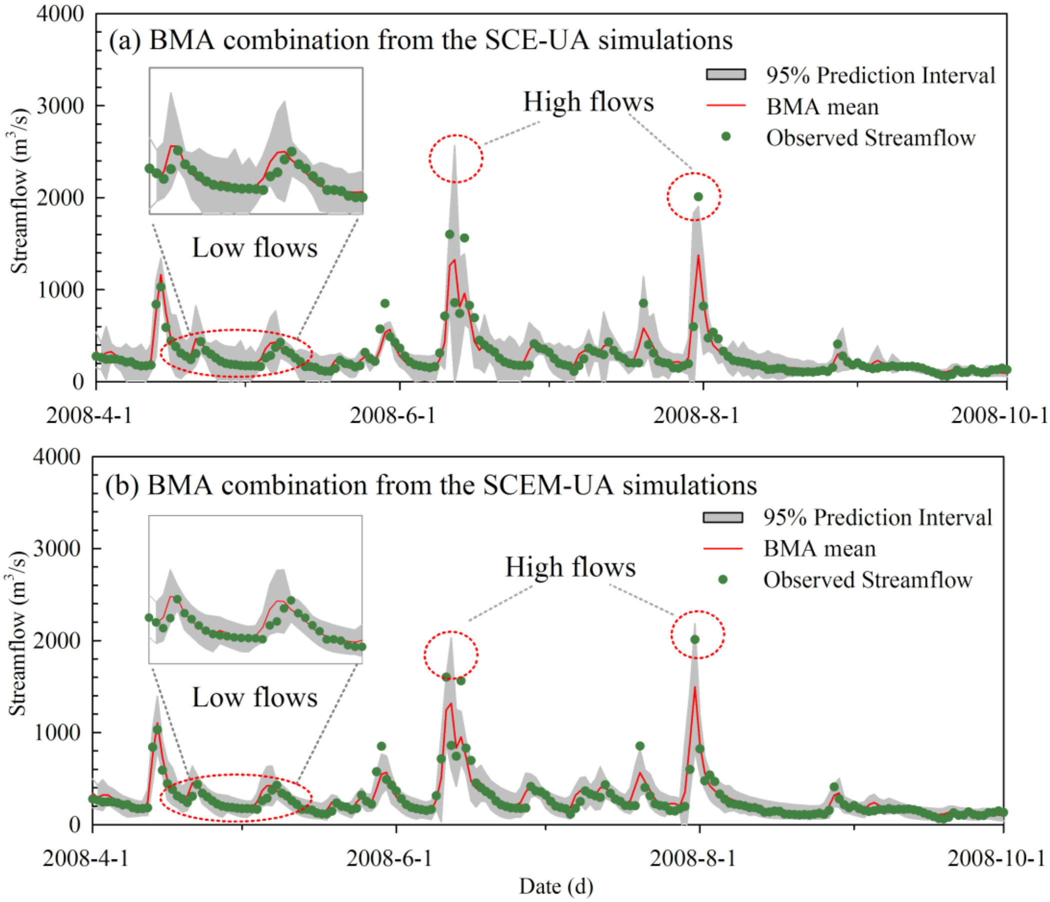

Fig. 8 BMA combined streamflow series and the 95% confidence interval, (a) from the SCE-UA-based simulations (sampling done 1000 times) and (b) from the SCEM-UA-based simulations for case III.

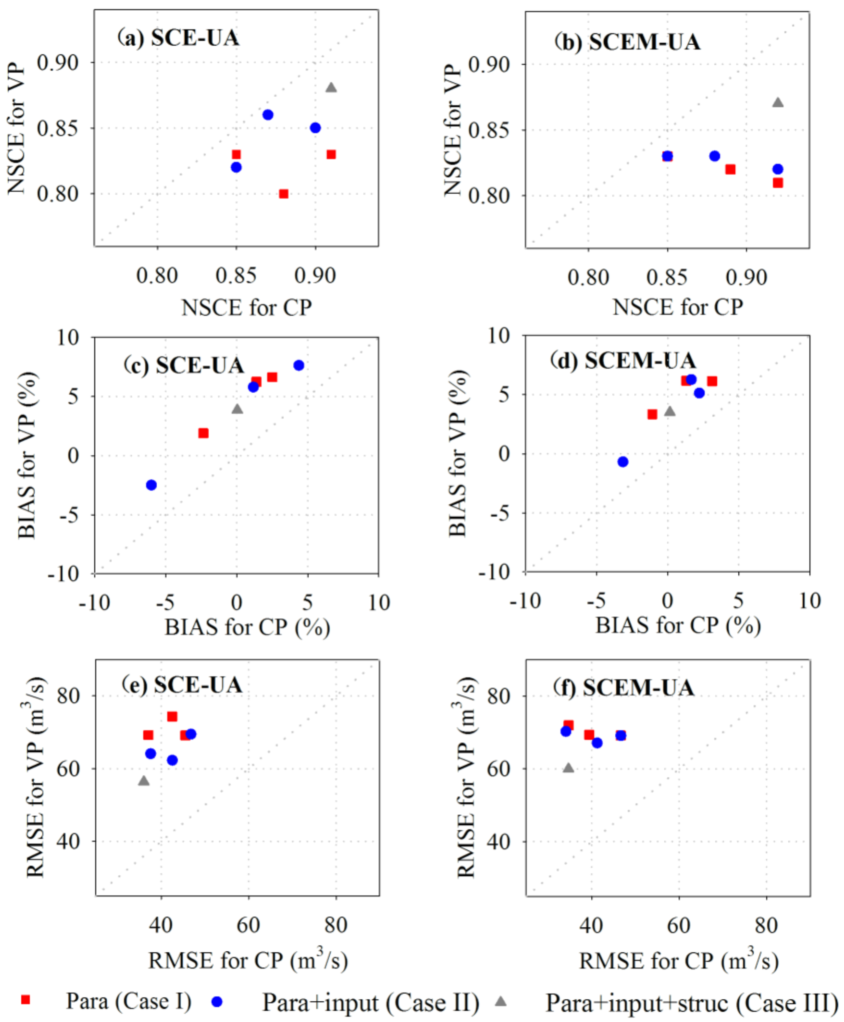

Fig. 9 Comparison of the hydrological model simulation performances in calibration period (CP) and validation period (VP) for the three different cases. Para indicates considering model parameter uncertainty in case I, Para+input means considering model input and parameter uncertainties in case II, Para+input+struc means considering model input, parameter, and structure uncertainties in case III.

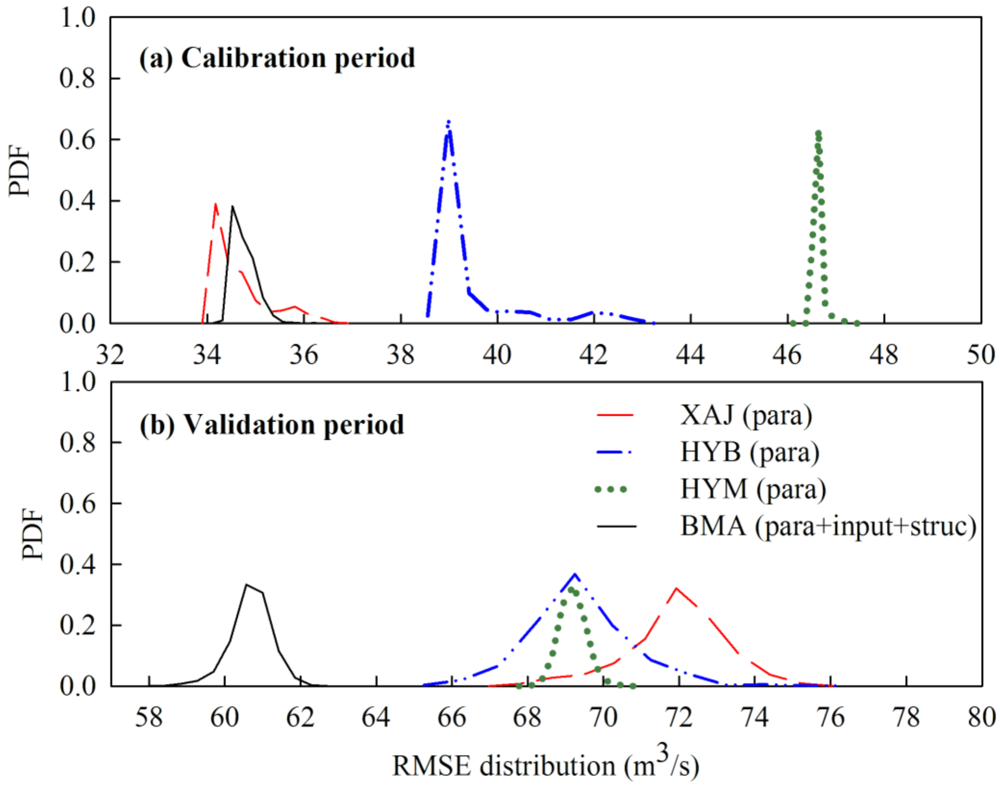

Fig. 10 Distribution of RMSE for XAJ, HYB and HYM considering different uncertainty sources. Para indicates considering model parameter uncertainty in case I, Para+input+struc means considering model input, parameter, and structure uncertainties in case III. The mean values and PDFs of RMSE in case II are similar to that in case I (see Table 6).

LIST OF TABLES

Table 1 Parameters of the XAJ model and their prior ranges

Parameter	Physical meaning	Prior Range
Kc	ratio of potential evapotranspiration to pan evaporation	0.5-1.5
WUM	water capacity in the upper soil layer	10-40
WLM	water capacity in the lower soil layer	50-90
WDM	Water capacity in the deeper soil layer	10-70
B	exponent of the tension water capacity curve	0.1-0.5
C	coefficient of deep evapotranspiration	0.1-0.3
EX	exponent of the free water capacity curve	1-1.5
SM	the free water capacity of the surface soil layer	10-60
KI0	outflow coefficients of the free water storage to interfolw	KI+KG=0.7
KG0	outflow coefficients of the free water storage to groundwater	0.1-0.5
CI0	recession constant of the lower interflow storage	0.1-0.9
CG0	daily recession constant of groundwater storage	0.9-0.999
CS0	recession constant for channel routing	0.1-0.5
KE	Slot storage coefficient	20-24
XE	Flow proportion factor	0.1-0.5

Table 2 Parameters of the HYB model and their prior ranges

Parameter	Physical meaning	Prior Range
Kc	ratio of potential evapotranspiration to pan evaporation	0.5-1.5
WUM	water capacity in the upper soil layer	10-40
WLM	water capacity in the lower soil layer	50-90
WDM	Water capacity in the deeper soil layer	10-70
B	exponent of the tension water capacity curve	0.1-0.5
bx	Infiltration capacity distribution curve index	0.1-2
f0	The average maximum infiltration capacity	5-30
fc	The average stability infiltration capacity	0.1-10
k	Infiltration capacity attenuation coefficient	0.1-0.9
CS	recession constant for channel routing	0.1-0.5
CG	daily recession constant of groundwater storage	0.9-0.999
C	coefficient of deep evapotranspiration	0.1-0.3
KE	Slot storage coefficient	20-24
XE	Flow proportion factor	0.1-0.5

Table 3 Parameters of the HYM model and their prior ranges

Parameter	Physical meaning	Prior Range
Kc	ratio of potential evapotranspiration to pan evaporation	0.5-1.5
Cmax	Max height of soil moisture accounting tank	1-1000
bexp	Distribution function shape parameter	0.1-2
Alpha	Quick-slow split parameter	0.1-0.99
Nq	Number of quick-flow routing tanks	1-8
Rs	Slowflow routing tanks rate parameter	0.001-0.1
Rq	Quick-flow routing tanks rate parameter	0.1-0.99
KE	Slot storage coefficient	20-24
XE	Flow proportion factor	0.1-0.5

Table 4 The posterior probability distribution parameters with SCEM-UA and the optimal parameters estimated by SCE-UA and SCEM-UA for the XAJ model in Case I

Parameter	Kc	WUM	WLM	WDM	B	C	EX
Mean	1.32	39.43	80.76	31.17	0.49	0.28	1.31
SD	0.03	0.67	6.03	8.22	0.01	0.02	0.19
CV	0.02	0.02	0.07	**0.26**	0.01	0.05	**0.15**
SCE-UA	1.49	39.99	50.01	10.06	0.47	0.20	1.42
SCEM-UA	1.34	39.21	86.06	40.20	0.49	0.29	1.46
Parameter	SM	KG0	CI0	CG0	CS0	KE	XE
Mean	24.08	0.38	0.84	0.99	0.13	20.17	0.50
SD	1.53	0.02	0.02	0.01	0.03	0.27	0.01
CV	0.06	0.05	0.02	0.00	**0.23**	0.01	0.00
SCE-UA	36.64	0.47	0.81	0.99	0.16	20.08	0.50
SCEM-UA	25.22	0.35	0.85	0.99	0.10	20.01	0.50

Notes: In the table, SD indicates standard deviation, CV means variable coefficient, SCE-UA and SCEM-UA mean the optimal parameter values of the two algorithms, respectively.

Table 5 The posterior probability distribution parameters with SCEM-UA and the optimal parameters estimated by SCE-UA and SCEM-UA for the XAJ model in Case II

Parameter	Kc	WUM	WLM	WDM	B	C	EX	SM
Mean	1.17	39.38	75.38	40.84	0.48	0.29	1.30	22.70
SD	0.03	0.55	9.61	9.22	0.02	0.01	0.19	0.83
CV	0.03	0.01	0.13	**0.23**	0.05	0.03	0.14	0.04
SCE-UA	1.35	36.59	67.72	54.05	0.50	0.17	1.13	21.15
SCEM-UA	1.16	39.14	85.56	28.32	0.50	0.29	1.43	22.59
Parameter	KG0	CI0	CG0	CS0	KE	XE	a	v
Mean	0.38	0.82	0.99	0.11	20.07	0.50	0.954	0.0004
SD	0.04	0.05	0.00	0.01	0.07	0.00	0.004	0.0003
CV	0.10	0.06	0.00	0.09	0.00	0.00	0.004	**0.5800**
SCE-UA	0.34	0.86	0.99	0.12	20.00	0.50	0.950	0.0002
SCEM-UA	0.35	0.86	0.99	0.11	20.02	0.50	0.951	0.0003

Notes: In the table, SD indicates standard deviation, CV means variable coefficient, SCE-UA and SCEM-UA mean the optimal parameter values of the two algorithms, respectively.

Table 6 Precision performance of the streamflow simulation series at different simulation cases

Cases		SCE-UA			SCEM-UA		
		NSE	BIAS (%)	RMSE (m^3/s)	NSE	BIAS (%)	RMSE (m^3/s)
CP	XAJ (Para)	**0.91**	-2.36	37.05	**0.92**	3.13	34.68
	XAJ (Para+input)	0.90	4.37	37.58	**0.92**	2.23	**34.05**
	HYB (Para)	0.88	2.50	42.49	0.89	-1.08	39.41
	HYB (Para+input)	0.87	-6.00	42.53	0.88	-3.41	41.27
	HYM (Para)	0.85	1.38	46.51	0.85	1.31	46.63
	HYM (Para+input)	0.85	1.17	46.79	0.85	1.67	46.69
	BMA (Para+input+struc)	**0.91**	**0.04**	**35.99**	**0.92**	**0.16**	34.66
VP	XAJ (Para)	0.83	**1.90**	69.23	0.81	6.14	71.95
	XAJ (Para+input)	0.85	7.62	64.03	0.82	5.12	70.23
	HYB (Para)	0.80	6.64	74.37	0.82	3.35	69.35
	HYB (Para+input)	0.86	-2.50	62.26	0.83	**-0.70**	67.04
	HYM (Para)	0.83	6.25	69.10	0.83	6.17	69.19
	HYM (Para+input)	0.82	5.79	69.42	0.83	6.26	69.08
	BMA (Para+input+struc)	**0.88**	3.85	**56.32**	**0.87**	3.49	**59.93**

Notes: In the table, Para indicates considering model parameter uncertainty in case I, Para+input means considering model input and model parameter uncertainties in case II, Para+input+struc revels considering model input, model parameter, and model structure uncertainties in case III. The value set in boldface refers to the optimum performance in the column.

Table 7 Reliability performance of the calculated 95% confidence interval at different simulation cases

Cases		SCE-UA (Sampling 1000 times)			SCEM-UA		
		CR%	B(m^3/s)	D(m3/s)	CR%	B(m^3/s)	D(m3/s)
CP	XAJ (Para)	59.31	**152.87**	58.20	78.65	**169.17**	52.31
	XAJ (Para+input)	74.86	200.15	60.78	79.06	169.78	**51.79**
	HYB (Para)	75.05	258.66	74.08	80.34	222.25	64.41
	HYB (Para+input)	81.07	273.01	65.39	78.97	225.23	67.60
	HYM (Para)	71.40	225.70	63.08	85.26	237.97	62.91
	HYM (Para+input)	68.57	212.49	63.29	87.68	254.06	64.01
	BMA (Para+input+struc)	**90.19**	315.60	**56.70**	**95.62**	271.15	55.03
VP	XAJ (Para)	62.32	**183.64**	71.21	80.47	**188.68**	64.41
	XAJ (Para+input)	73.81	220.50	74.63	81.48	190.31	**63.07**
	HYB (Para)	71.99	289.95	82.84	80.66	244.40	74.13
	HYB (Para+input)	82.66	285.44	71.71	80.38	249.24	77.14
	HYM (Para)	68.61	270.26	77.88	86.77	261.23	76.62
	HYM (Para+input)	69.16	252.23	76.84	88.96	278.24	77.31
	BMA (Para+input+struc)	**90.97**	348.56	**69.74**	**95.17**	303.04	66.06

Notes: In the table, Para indicates considering model parameter uncertainty in case I, Para+input means considering model input and model parameter uncertainties in case II, Para+input+struc revels considering model input, model parameter, and model structure uncertainties in case III. The value set in boldface refers to the optimum performance in the column.

Response to review on manuscript: No. Hydrology-D-16-00272
Title: Quantifying multi-source uncertainties in multi-model predictions using the Bayesian Model Averaging scheme

Journal: *Hydrology Research*
Authors: Shanhu Jiang, Liliang Ren*, Chong-Yu Xu, Shuya Liu, Fei Yuan, and Xiaoli Yang
Corresponding Author: Liliang Ren
Email: njrll9999@126.com

Dear editor,

We appreciate your favorable consideration and the reviewers' constructive comments. According to your suggestions and the reviewers' comments, we have revised our manuscript (No. Hydrology-D-16-00272) carefully and thoroughly. In the revised version, red colored text represents text that has been revised or relocated. Some minor language corrections are not marked with color. The page and line numbers in the reviewers' comments refer to the original version of the manuscript, and in the response (if applicable) they refer to the revised version. We hereby provide our point by point responses to each of the reviewer's comments.

Thanks and appreciate your time

Liliang Ren

State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering, College of Hydrology and Water Resources, Hohai University, Nanjing 210098, China

Response for Reviewers:

Reviewer #1: Quantification and reduction of model input, model parameter and model structural uncertainties in hydrological modeling remain as challenges for hydrologists. The current study focuses on uncertainty analysis of multi-source and multi-model hydrological prediction. The research contents are rich and the conclusions are reliable. Also, there are some innovations, including, comparing the effects of SCE-UA and SCEM-UA algorithms on the hydrological prediction results; and investigating the superiority of the BMA multi-model ensemble strategy over the individual modelling approach. Basically I enjoyed reading your paper and I think with some minor revisions it will be worthful published in Hydrology Research. And, the following comments and corrections should be made:

Reply: thank you very much for your positive evaluation of our paper, and below is our point by point response to your valuable comments.

(1) Page 4, Introduction: I suggest the authors add some comments about the recently research progresses of the model input, model parameter and model structural uncertainties analysis.

Response:

Thanks for your good suggestion. In the revised manuscript, we have added some comments about the recently research progresses of the model input, model parameter and model structural uncertainties analysis. To improve the description of the state-of-the-art of the topic, following relevant and excellent references have been cited and discussed.

Xu, D. M., Wang, W. C., Chau, K. W. & Cheng, C. T. 2013 Comparison of three global optimization algorithms for calibration of the Xinanjiang model parameters. Journal of Hydroinformatics 15, 174-193.

Yen, H., Jeong, J., Tseng, W. H., Kim, M. K., Records, R. M. & Arabi, M. 2014a Computational Procedure for Evaluating Sampling Techniques on Watershed Model Calibration. J. Hydrol. Eng. 20, 04014080-1.

Yen, H., Wang, X. Y., Fontane D. G., Harmel, R. D. & Arabi, M. 2014b A framework for propagation of uncertainty contributed by input data, parameterization, model structure, and calibration/validation data in watershed modeling. Environ. Modell. Softw. 54, 211–221.

Yen, H., Jeong, J., Feng, Q. Y. & Deb, D. 2015a Assessment of Input Uncertainty in SWAT Using Latent Variables. Water Resources Management 29, 1137-1153.

Yen, H., White, M. J., Jeong, J., Arabi, M. & Arnold, J. G. 2015b Evaluation of alternative surface runoff accounting procedures using the SWAT model. Int J Agric & Biol Eng 8, 54-68.

(2) Page 6, Hydrological models: the authors should add some explanations why you select those three hydrological models.

Response:

Thanks for your good suggestion. In the revised manuscript, we have added the explanations why we select those three hydrological models.

"These three hydrological models have different complex model structures and different runoff generation mechanisms. They have been successfully and widely used in different river basins for streamflow simulation and flood forecasting (Ajami et al., 2007; Ren et al., 2008; Najafi et al., 2011; Jie et al., 2016; Xu et al., 2016). Tables 1-3 show the parameters and their prior ranges of the three models."

(3) Page 7, Hydrological models: the data you used for calibrating and validating model is not very long, please give an explanation.

Response:

Thanks for your careful reading. We have made a mistake in the submitted manuscript, the data used for the study were from 2000 to 2008 instead of from 2003 to 2008 as was written in the original version. This period of data was considered to be more representative of the current climate and landuse situation of the study region. We have clarified it in the revised manuscript.

(4) Page 15, in Figure 4 and 5: The streamflow series and the 95% confidence interval of the three hydrological models for case I and II. There are some differences for the 95% confidence intervals and the high flow simulations, the authors should indicate them and give some discussions in your manuscript to further support your research conclusions.

Response:

Thanks for your good suggestion. In the revised manuscript, we have demonstrated that:

"From Table 6 and Figures 6-7, we can see that the three models showed a good hydrologic prediction applicability in Mishui basin, in which the XAJ model performed best, followed by the HYB model, and lastly, the HYM model. Especially for the high flow simulations, the XAJ model and HYB model performed much better than the HYM model."
"Given the precipitation input uncertainty in case II, the performance of the calculated 95% confidence intervals of the three models showed a minimal improvement in terms of higher CR values, especially for the validation period."

(5) Page 16, Different performance in Calibration period and Validation period: this

part analysis is interesting, I suggest that the authors add some relevant references to support your conclusions.

Response:

Thanks for your good suggestion. In the revised manuscript, we have added the following relevant references to support our conclusions.

"Li, H., Beldring, S. & Xu, C. Y. 2015 Stability of model performance and parameter values on two catchments facing changes in climatic conditions. Hydrological Sciences Journal 60, 1317-1330.

Singh, V. P. & Woolhiser, D. A. 2002 Mathematical Modeling of Watershed Hydrology. Journal of Hydrologic Engineering 7, 270-292.

Vel´azquez, J. A., Anctil, F., Ramos, M. H. & Perrin, C. 2011 Can a multi-model approach improve hydrological ensemble forecasting? A study on 29 French catchments using 16 hydrological model structures. Advances in Geosciences 29, 33-42.

Broderick, C., Matthews, T., Wilby, R. L., Bastola, S. & Murphy, C. 2016 Transferability of hydrological models and ensemble averaging methods between contrasting climatic periods. Water Resources Research 52, 8343–8373."

(6) Finally, the English needs to be further improved.

Response:

Thanks for your good suggestion. In the revised manuscript, the co-author Prof. Xu (who is a professor in the University of Oslo) has improved the English and grammar of the manuscript again and we hope it is to the satisfaction of the journal.

Reviewer #2: In this manuscript, two sources (input and structural) of uncertainty were incorporated to three different models (XAJ, HYB, and HYM) on hydrological models. Overall, the manuscript was well-written and I'm in favor of your work. However, there are many significant issues before I can recommend for publication at this point.

1. SCE-UA (developed in 1992) has been shown to be substantially less efficient than other optimization approaches (DDS, DREAM, and others) in recent years (Yen et al., 2014a). Why authors were still using SCE-UA. It does not mean authors have to redo everything but proper discussion/justification is needed in the manuscript.

- Yen, H., Jeong, J., Tseng, W., Kim, M., Records, R., and Arabi, M., 2014a. Computational Procedure for Evaluating Sampling Techniques on Watershed Model Calibration. J. Hydrol. Eng., 20(7). DOI: 10.1061/(ASCE)HE.1943-5584.0001095 , 04014080.

Response:

Thanks for your positive evaluation of our paper and for your advice on the important issue that has been overlooked in our paper. In the introduction section of the revised version, we have introduced and discussed the DDS and DREAM approaches and their applications in accordance with your advice. The following text has been added:

"In hydrological modeling, model parameters often need to be calibrated based on observed hydrographs. Two main parameter calibration methods are currently used. In the first method, only one optimal parameter set can be obtained for a basin and model, and the typical algorithms are Genetic Algorithm (GA, Wang et al., 1991); Shuffled Complex Evolution (SCE-UA, Duan et al., 1992) and Dynamically Dimensioned Search (DDS, Tolson and Shoemaker, 2007). In the other method, the model parameter involves one set of random variables that follow a certain joint probability distribution, and the typical algorithms are Generalised Likelihood Uncertainty Estimation (GLUE, Beven and Binley, 1992); Shuffled Complex Evolution Metropolis (SCEM-UA, Vrugt et al., 2003) and Differential Evolution Adaptive Metropolis (DREAM, Vrugt et al., 2009). Different optimization algorithms demonstrated different convergence speed and behavioral statistics in model parameter calibration and uncertainty analysis (Xu et al., 2013; Yen et al., 2014a). Among the mentioned optimization algorithms, the SCE-UA and SCEM-UA approaches have been widely used in parameter calibration and uncertainty analysis in the literature, but the effects of the two algorithms on the deterministic simulation and probability prediction still need to be evaluated and compared further. This consideration has motivated our current study"

2. Similar work (input, structural uncertainty using BMA) has been done before. For example, the framework developed in this study was already developed by Yen et al. (2014b) (Integrated Parameter Estimation and Uncertainty Analysis Tool, IPEAT). More details of the BMA applications with structural uncertainty can also be found in

Yen et al. (2015). However, it was not mentioned/discussed or cited anywhere in the manuscript.

- Yen, H., Wang, X., Fontane, D. G., Harmel, R. D., Arabi, M., 2014b. A framework for propagation of uncertainty contributed by parameterization, input data, model structure, and calibration/validation data in watershed modeling, Environmental Modelling and Software, 54, pp. 211-221, doi: 10.1016/j.envsoft.2014.01.004.

- Yen H, White M J, Jeong J, Arabi M, Arnold J G. Evaluation of alternative surface runoff accounting procedures using the SWAT model. Int J Agric & Biol Eng, 2015; 8(3): 54-68. doi: 10.3965/j.ijabe.20150803.833.

Response:

Thanks for introducing above good works which have been overlooked in our original version of the manuscript. In the introduction section of the revised version, we have added the introduction of these two works to enhance the literature review and knowledge gained in the research field.

"Duan et al. (2007), Liang et al. (2013), Dong et al. (2013), Yen et al. (2015b) and Arsenault et al. (2015) successfully used BMA to combine multi-model/multi-method simulations to obtain more robust streamflow series and more reliable probability predictions."

"There also are some researches about comprehensive assessment of the effects of different uncertainty sources on the hydrological modeling (Ajami et al., 2007; Yen et al., 2014b)."

3. Line 14-17 (page 3), Line 3-4 (page 5): Interestingly, there are plenty of others work cited in this manuscript but not as important. For example, Her et al. 2016 is not really serving the primary purpose in this manuscript. Not just this one, I would suggest removing some of them.

Response:

Thanks for your careful reading. In the revised manuscript, we have deleted some not closely related literatures (i.e., Zeng et al., 2016; Her et al., 2016) and added some more relevant references.

4. Input uncertainty was incorporated in the hydrological model by using the approach proposed by Ajami et al. (2007). However, it was reported that modeling results and input uncertainty may not necessarily be improved accordingly (Yen et al. 2015). How would you compare and explain your results in the discussion? In addition, values of latent variables were not reported in the manuscript. I would suggest adding values of latent variables in an independent table.

- Yen, H., J. Jeong, Q. Feng, D. Deb, 2015. Assessment of Input Uncertainty in

SWAT Using Latent Variables. Water Resources Management, 29(4), pp. 1137-1153. DOI: 10.1007/s11269-014-0865-y

Response:

Thanks for your good suggestions. In the revised manuscript, we have added a Figure and a Table of marginal posterior probability distribution of latent variables for the XAJ model.

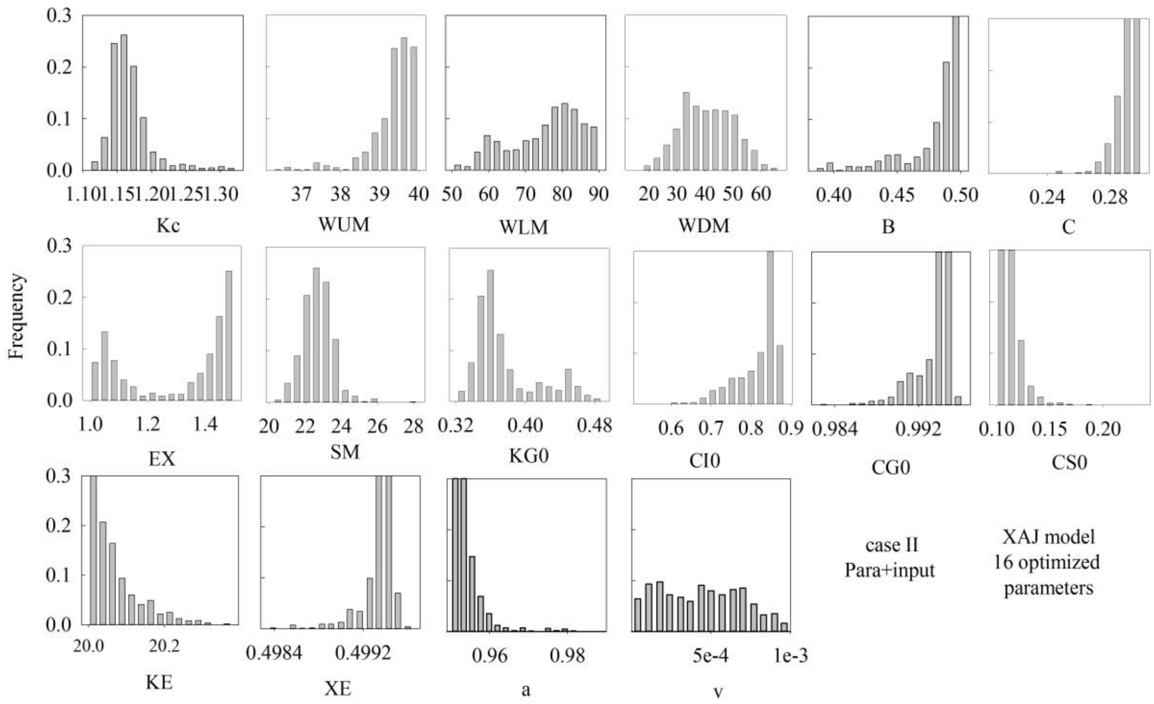

Fig. 4 Marginal posterior probability distribution of the XAJ parameters for case II, using 10 000 samples generated after the SCEM-UA algorithm convergence

Table 5 The posterior probability distribution parameters with SCEM-UA and the optimal parameters estimated by SCE-UA and SCEM-UA for the XAJ model in Case II

Parameter	Kc	WUM	WLM	WDM	B	C	EX	SM
Mean	1.17	39.38	75.38	40.84	0.48	0.29	1.30	22.70
SD	0.03	0.55	9.61	9.22	0.02	0.01	0.19	0.83
CV	0.03	0.01	0.13	**0.23**	0.05	0.03	0.14	0.04
SCE-UA	1.35	36.59	67.72	54.05	0.50	0.17	1.13	21.15
SCEM-UA	1.16	39.14	85.56	28.32	0.50	0.29	1.43	22.59
Parameter	KG0	CI0	CG0	CS0	KE	XE	a	v
Mean	0.38	0.82	0.99	0.11	20.07	0.50	0.954	0.0004
SD	0.04	0.05	0.00	0.01	0.07	0.00	0.004	0.0003
CV	0.10	0.06	0.00	0.09	0.00	0.00	0.004	**0.5800**

SCE-UA	0.34	0.86	0.99	0.12	20.00	0.50	0.950	0.0002
SCEM-UA	0.35	0.86	0.99	0.11	20.02	0.50	0.951	0.0003

Notes: In the table, SD indicates standard deviation, CV means variable coefficient, SCE-UA and SCEM-UA mean the optimal parameter values of the two algorithms, respectively.

And added the following text:

“Given the precipitation input uncertainty in case II, the precisions of the simulated streamflows using the three models were not remarkably enhanced. This phenomenon may have been caused by the relatively small precipitation input uncertainty because of the dense rain gauge observations in the Mishui basin. Moreover, in the model parameters, an evaporation reduction factor parameter K was set, and this parameter could imply some precipitation input uncertainty. Our results are quite consistent with those of Yen et al. (2015a), which reported that the use of error multiplier to incorporate input uncertainty might not be the proper alternative choice in terms of generating better results.”

5. BMW weights were not reported/discussed in the manuscript. Please add it.

Response:

Thanks for your good suggestion. In the revised manuscript, we have added the report of the BMA weights.

“For comprehensive consideration of the model input, model parameter, and model structure uncertainties, we used the BMA to combine the three models’ simulations at case II. Figure 5 displays the weight estimates of different models calculated using the BMA method. For the SCE-UA-based simulations, the weights of the XAJ, HYB and HYM models are 0.36, 0.31 and 0.33, respectively. For the SCEM-UA-based simulations, the mean values of the weights of the XAJ, HYB and HYM models are 0.35, 0.32 and 0.33, respectively. The weight of the BMA method is directly bound to individual model simulation, that is, a well performing model can receive a higher weight than a poorly performing one in theory. In this study, the XAJ model got the highest weight value, and then followed by the HYM model and the HYB model. The HYM model got a higher weight value than that of the HYB model, which may be due to the similar model structure of the XAJ model and HYB model (Ren et al., 2008). By using the BMA combination, we can obtain deterministic streamflow series and probability predictions, which comprehensively considered the multi-source uncertainties.”

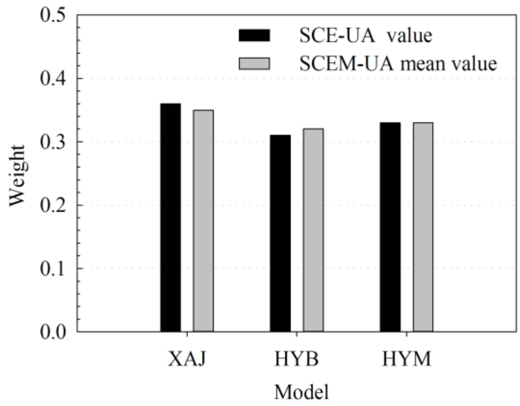

Fig. 5 Histogram of the BMA weights for the different models. SCE-UA value means the BMA weights for the SCE-UA based model simulations. SCEM-UA mean value indicates the BMA weights for the SCEM-UA based model simulations.

6. From above, I know it seems that I'm strongly encouraging you to cite and compare results according to my work. However, it happens that it is exactly the case I had very similar work done in the past few years but not being discussed at all. In addition to that, the novelty of the proposed work can be enhanced by highlighting some local concerns (for example, why do we need this work in Southern China?). I'm looking forward to reviewing the next round revision if you can address all the mentioned issues properly.

Response:

Thanks again for your professional comments and introducing your excellent work to us which not only has greatly enhanced the literature review of this paper, but also will broaden our horizons and beneficial to our future study. Also, based on your suggestions, we have highlighted some local concerns of Southern China.

"Temporal and spatial distributions of precipitation in the study region are uneven because of atmospheric circulation and most of the annual precipitation occurs between April and September. During these months, particularly in June, basin-wide heavy rains continuously occur, thereby resulting in flash floods. This multi-model ensemble prediction method can reduce the streamflow prediction and flood forecasting uncertainties, thus it is important to decision support system for such river basins to prevent flood disasters and reduce flood damages."

Reviewer #3: I think the paper provides good information about the different sources of uncertainty and their interactions in hydrological modelling. While there have been numerous studies addressing this in the past, the way the authors present this case is interesting nonetheless. It was enjoyable to read and very clear english, althoush some small errors remain. I suggest minor revision and do not recommend the paper be sent out for supplemental review once the modifications are brought to the paper. I think it could remain at the editorial board level.

I only have a few comments that the authors should consider addressing.

Also, a suggestion for further submissions: please use continuous line numbering instead of having line numbers start at 1 on each page. It is easier for the reviewers to pinpoint lines in the paper.

Response:

Thanks for your positive evaluation and encouragement. We have improved the paper following your suggestion. Also, in the revised manuscript, we have used continuous line numbering instead of having line numbers start at 1 on each page.

1. General comment: The Case I, Case II and Case III are not easily followed until the figure captions. Perhaps add a section with clear indications of what to expect from each of the three cases.

Response:

Thanks for your careful reading. In the revised manuscript, we have introduced their indications in the methodology section.

"The flowchart for the multi-source uncertainty analysis of multi-model predictions is shown in Fig.1. We adopted three different simulation cases to systematically consider the three sources (i.e., parameter uncertainty, input uncertainty and model structural uncertainty) of hydrological modeling uncertainties. In case I, the model parameter uncertainty (hereafter "Para") using SCE-UA and SCEM-UA algorithms for three hydrological models, i.e., XAJ, HYB, and HYM, was determined. In case II, a normally distributed error multiplier and combined parameter optimization algorithms were introduced to consider the model input and model parameter uncertainties (hereafter "Para+input"). In case III, the simulations calculated from case II were combined using BMA to comprehensively determine the model input, model parameter, and model structure uncertainties (hereafter "Para+input+struc"). The detailed methodologies are as follows."

2. Page 2, line 15: NSCE is not defined. Usual Nash-Sutcliffe value is NSE, so please define NSCE to make clear.

Response:

Thanks for your good suggestion. In the revised manuscript, we have defined and used NSE instead of NSCE.

3. Page 5, line 6-10: Please itemize more clearly, with bullet-points if need be.

Response:

Thanks for your suggestion. In the revised manuscript, we have rewritten it following your suggestion.

“The innovations of the study include: (1) it considers rainfall input uncertainty, parameter estimation uncertainty, and model structural uncertainty by using three models, i.e., Xinanjiang (XAJ), hybrid rainfall-runoff (HYB), and HYMOD (HYM) models; (2) it compares the effects of SCE-UA and SCEM-UA algorithms on the hydrological prediction results; and (3) it investigates the superiority of the BMA multi-model ensemble strategy over the individual modelling approach.”

4. Page 5, line 18: Add a step between 1 and 2 for the hydrological modelling.

Response:

Thanks for your good suggestion. In the revised manuscript, combined with your first suggestion of Case I, Case II and Case III, we have rewritten it as:

“The flowchart for the multi-source uncertainty analysis of multi-model predictions is shown in Fig.1. We adopted three different simulation cases to systematically consider the three sources (i.e., parameter uncertainty, input uncertainty and model structural uncertainty) of hydrological modeling uncertainties. In case I, the model parameter uncertainty (hereafter “Para”) using SCE-UA and SCEM-UA algorithms for three hydrological models, i.e., XAJ, HYB, and HYM, was determined. In case II, a normally distributed error multiplier and combined parameter optimization algorithms were introduced to consider the model input and model parameter uncertainties (hereafter “Para+input”). In case III, the simulations calculated from case II were combined using BMA to comprehensively determine the model input, model parameter, and model structure uncertainties (hereafter “Para+input+struc”). The detailed methodologies are as follows.”

5. Page 11, lines 25-26: How was the normal error multiplier applied? On the station data directly or on the final, inverse-distance weighted average?

Response:

We have made it clearer that the normal error multiplier was applied on the 15 sub-basins of the Mishui basin.

“The inverse distance weighting of the three nearest rain gauges was used to obtain the spatially distributed precipitation database of 15 sub-basins of the Mishui basin.”

6. Page 13, lines 20-30: Please refer to appropriate figures and/or table to support these claims.

Response:

Thanks for your suggestion. In the revised manuscript, we have added appropriate figure and/or table to support these claims.

“In order to consider the parameter and input uncertainty together, two rain input error modeling parameters m and σ_m^2 were added to model parameter sets and further estimate the posterior PDFs simultaneously in case II. Figure 4 shows the marginal posterior probability distribution of the XAJ parameters estimated by SCEM-UA in case II. Table 5 demonstrates the statistical indices of the posterior probability distribution of the parameters estimated by SCEM-UA and the optimal parameters estimated by SCE-UA in case II.”

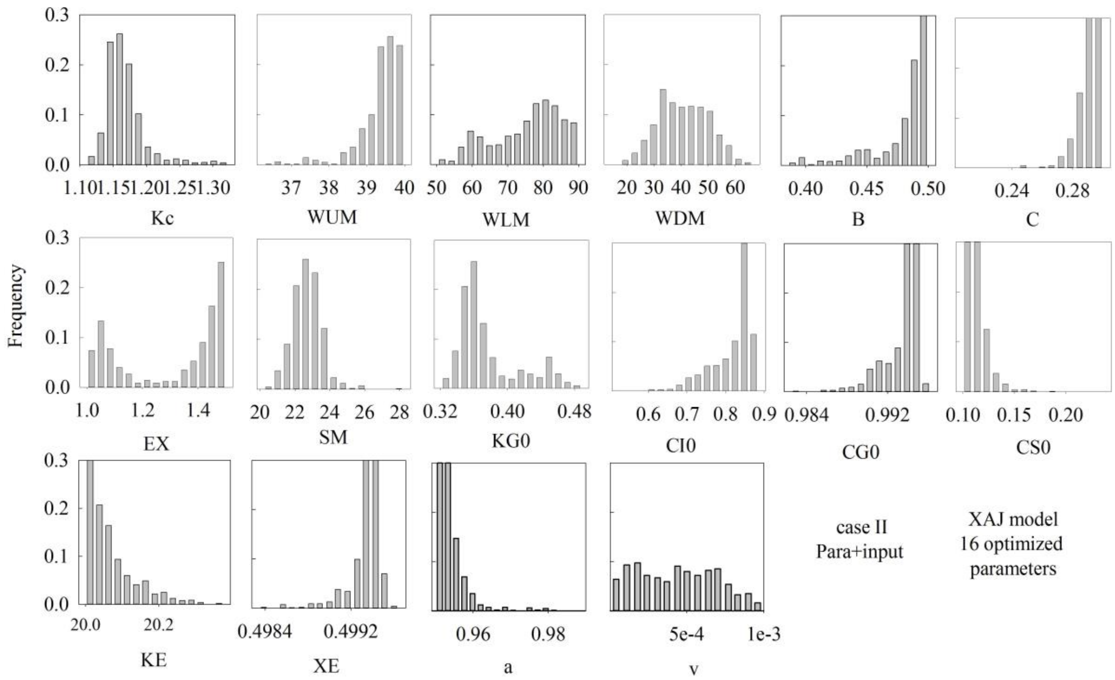

Fig. 4 Marginal posterior probability distribution of the XAJ parameters for case II, using 10 000 samples generated after the SCEM-UA algorithm convergence

Table 5 The posterior probability distribution parameters with SCEM-UA and the optimal parameters estimated by SCE-UA and SCEM-UA for the XAJ model in Case II

Parameter	Kc	WUM	WLM	WDM	B	C	EX	SM
Mean	1.17	39.38	75.38	40.84	0.48	0.29	1.30	22.70

SD	0.03	0.55	9.61	9.22	0.02	0.01	0.19	0.83
CV	0.03	0.01	0.13	0.23	0.05	0.03	0.14	0.04
SCE-UA	1.35	36.59	67.72	54.05	0.50	0.17	1.13	21.15
SCEM-UA	1.16	39.14	85.56	28.32	0.50	0.29	1.43	22.59
Parameter	KG0	CI0	CG0	CS0	KE	XE	a	v
Mean	0.38	0.82	0.99	0.11	20.07	0.50	0.954	0.0004
SD	0.04	0.05	0.00	0.01	0.07	0.00	0.004	0.0003
CV	0.10	0.06	0.00	0.09	0.00	0.00	0.004	0.5800
SCE-UA	0.34	0.86	0.99	0.12	20.00	0.50	0.950	0.0002
SCEM-UA	0.35	0.86	0.99	0.11	20.02	0.50	0.951	0.0003

Notes: In the table, SD indicates standard deviation, CV means variable coefficient, SCE-UA and SCEM-UA mean the optimal parameter values of the two algorithms, respectively.

7. Page 14, line 13: How does the 1000-simulation SCE-UA compare to the 15000-simulation SCEM-UA? there is 14000 simulation difference, was the effect of the difference in simulation numbers explored?

Response:

Thanks for your question. In our manuscript, for SCE-UA-based simulation, the Monte Carlo Markov Chain sampling method was used to calculate the prediction uncertainty interval (Duan et al., 2007). We have conducted different sampling times test and draw the conclusion that 1000 times is the optimal. In the revised manuscript, we have illustrated this point.

"Based on the repeated sampling experiments, we set the sampling times as 1000."

8. Figures 4-6: The legend for "observations" overlaps the figure data, so the legend looks like it is part of the figure! (the green dot).

Response:

Thanks for your careful reading. In the revised manuscript, we have modified the legend of the Figures and made them clearer.

For the rest, I think the methodology is clear and concise, the figures are nice and required, the tables are of interest and everything looks like it is in its place.

Thank you again for your appreciation and encouragement.