Comparison of three global optimization algorithms for calibration of the Xinanjiang model parameters

Dong-mei Xu, Wen-chuan Wang, Kwok-wing Chau, Chun-tian Cheng and Shou-yu Chen

ABSTRACT

The Xinanjiang model, a conceptual rainfall-runoff (CRR) model with distributed parameters, has been successfully and widely applied to flood forecasting of large basins in humid and semi-humid regions of China. With an increasing demand for timely and accurate forecasts in hydrology, how to obtain more appropriate parameters for CRR models has long been an important topic. These models have a large number of parameters which cannot be directly obtained from measurable quantities of catchments characteristics. In this study, three different optimization methods are used to calibrate the Xinanjiang streamflow model: genetic algorithm (GA), shuffled complex evolution of the University of Arizona (SCE-UA) and the recently developed shuffled complex evolution Metropolis algorithm of the University of Arizona (SCEM-UA), using streamflow data of the Shuangpai Reservoir in China. Two different time steps of 1 and 3 hr are used in the analysis. The results indicate that the SCEM-UA algorithm can infer the most probable parameter set and furnish useful information about the nature of the response surface in the vicinity of the optimum. Moreover, there is larger uncertainty for 1 hr forecasting than for 3 hr forecasting. This is significant in assessing risks in likely applications of Xinanjiang models.

Key words | genetic algorithm, global optimization methods, Metropolis algorithm, shuffled complex evolution, time step, Xinanjiang model calibration

Dong-mei Xu
Shou-yu Chen
Faculty of Infrastructure Engineering,
Dalian University of Technology,
Dalian, 116024,
China

Wen-chuan Wang (corresponding author)
Faculty of Water Conservancy Engineering,
North China Institute of Water Conservancy and Hydroelectric Power,
Zhengzhou 450011,
China
E-mail: *wangwen1621@163.com*

Kwok-wing Chau
Department of Civil and Structural Engineering,
Hong Kong Polytechnic University,
Hung Hom, Kowloon,
Hong Kong,
China

Chun-tian Cheng
Institute of Hydropower System & Hydroinformatics,
Dalian University of Technology,
Dalian, 116024,
China

INTRODUCTION

Conceptual rainfall-runoff (CRR) models have become a basic tool for flood forecasting and become increasingly important for catchment basin management. They can provide an approximate, lumped description of the dominant sub-catchment scale processes that contribute to the overall catchment scale hydrologic response of the watershed system. These models, to various degrees of approximation, attempt to simulate catchment water balance dynamics, which are represented by heuristic equations, or as physically based with scientifically accepted principles (Kuczera 1997). Most CRR models usually include 10 or more parameters that link transfer functions of several interconnected water stores. Model parameters are conceptual representations of abstract catchment characteristics and define the behavior of the various conceptual elements and the way they relate to each other, therefore, they cannot be directly obtained from measurable quantities of catchment characteristics.

Setting model parameters is commonly performed by calibration, which is a process of changing parameter values until a satisfactory agreement between simulated and observed catchment behavior is obtained (Sorooshian & Gupta 1995). Traditionally, a process of trial and error parameter adjustment or 'manual approach' is commonly made, and the simulated and observed watershed behavior is compared using visual inspection and different measures of performance. This is a very tedious and time-consuming task, depending on the number of free model parameters

doi: 10.2166/hydro.2012.053

and the degree of parameter interaction. Moreover, it is difficult to explicitly assess the confidence of model simulations because of the subjectivity involved. These shortcomings have provided incentive for automation of the calibration process. This has transformed the calibration problem into an optimization problem, and the main objective is to determine the values of the model parameters within a large multidimensional parameter space which provides the best fit between the simulated and observed flows hydrograph. In the last two decades, major developments in the calibration of CRR models have provided the incentive for automatic calibration, for instance, the use of global optimization methods (GOM) for model parameter estimation (Cooper *et al.* 2007).

As a means to finding optimal solutions, unlike conventional methods, genetic algorithms (GAs) utilize a population of individuals to search the parameters' space, and do not require auxiliary gradient information to solve optimization problems. Since Wang (1991) showed that the GA is a robust and efficient method for the calibration of hydrological conceptual models, GA has become one of the most widely used techniques for model calibration (Franchini 1996; Cooper *et al.* 1997; Franchini & Galeati 1997; Wang 1997; Ndiritu & Daniell 2001). Recently, Cheng *et al.* (2002, 2006) combined a fuzzy optimal model with GAs to solve multi-objective rainfall-runoff model calibration. Furthermore, Cheng *et al.* (2005) proposed a hybrid method that combines a parallel GA with a fuzzy optimal model in a cluster of computers for rainfall-runoff model calibration. Liong *et al.* (2005) used a micro-GA (mGA) search engine to implement the inverse approach for forecasting hydrological time series in an efficient way. Cai and Wang (2006) presented an approach based on GA to calibrate a holistic water resources-economic model. In order to consider multi-objective functions in reservoir system operation, Kim *et al.* (2008) adopted a multi-objective genetic algorithm (NSGA-II) to obtain the optimization results. Bakhtyar & Barry (2009) used GA to determine the optimal design of cascade stilling basins in terms of the height of falls and length of stilling basins. Fazal *et al.* (2005) and Goswami & O'Connor (2007) applied GA to calibrate soil moisture accounting and routing (SMAR) model parameters. Hejazi *et al.* (2008) applied GA to calibrate a watershed simulation model involving human interference. Sharifi *et al.* (2009) used the NSGA-II algorithm to study the lateral variation and value of these parameters for simple trapezoidal channels over a wide range of aspect ratios through the model calibration process. A good account of its applications in rainfall-runoff forecasts can also be found in Anctil *et al.* (2006). GAs do not really use efficient rules. The AMALGAM method of Vrugt & Robinson (2007) combined multiple different search methods at the same time within a self-adaptive learning kernel. Awad & Von (2010) presented a real GA combined with a local search method for calibrating CRR models. Wang *et al.* (2012) presented a novel hybrid algorithm for calibration of Xinanjiang model parameters using hybrid GA based fuzzy optimal model.

The shuffled complex evolution of the University of Arizona (SCE-UA) algorithm, as a global search algorithm, was compared against the multi-start simplex (MSX) method and the adaptive random search (ARS) method on watershed model calibration problems (Duan *et al.* 1992). The results showed that SCE-UA was a much superior method than MSX and ARS method. The SCE-UA has been widely used in various watershed model calibrations (Sorooshian *et al.* 1993; Duan *et al.* 1994; Luce & Cundy 1994; Gan & Biftu 1996; Yapo *et al.* 1996; Cooper *et al.* 1997; Kuczera 1997; Franchini *et al.* 1998; Abdulla *et al.* 1999; Thyer *et al.* 1999; Eckhardt & Arnold 2001). Recently, the SCE-UA has also been applied with success to Soil and Water Assessment Tool (SWAT) for hydrologic parameters (Eckhardt & Arnold 2001) and hydrologic and water quality parameters (van Griensven & Bauwens 2003). Cooper *et al.* (2007) applied the global optimization SCE-UA method with the established hydrologic process-based constraints to calibrate the Tank Model. It is found that performances of the SCE and GA are better than simulated annealing. More recently, the SCE, simple genetic algorithm (SGA) and micro-genetic algorithm (μGA), are applied in the parameter calibration of a grid-based distributed rainfall-runoff model (GBDM) and their performances are compared (Wang *et al.* 2010). Goswami & O'Connor (2007) applied SCE-UA to calibrate SMAR model parameters. These studies demonstrate that the SCE-UA method is a robust, effective and efficient search algorithm.

The shuffled complex evolution Metropolis algorithm of the University of Arizona (SCEM-UA) is a general-purpose, stochastic global optimization algorithm that provides efficient estimation of the most likely parameter set and its underlying posterior probability distribution within a single optimization run. The SCEM-UA algorithm has been successfully applied to calibrate the five-parameter CRR model HYMOD for a 1,944 km^2 watershed (Vrugt *et al.* 2003b), to identify rainfall interception model parameters from measurements of throughfall and forest canopy storage (Vrugt *et al.* 2003a), to estimate vadose zone properties for a small watershed using the distributed fully coupled surface vadose zone groundwater model MODHMS (Vrugt *et al.* 2004) and to calibrate the parameters in the Sacramento soil moisture accounting (SAC-SMA) model using historical data from the Leaf River in Mississippi (Vrugt *et al.* 2006). Feyen *et al.* (2007) employed the SCEM-UA global optimization algorithm to calibrate the LISFLOOD model against daily discharge observations.

The major aim of the study presented in this paper is to investigate three different GOMs for calibration and analysis of Xinanjiang model parameters in the Shuangpai Reservoir. According to the national standard for flood forecasting in China, three statistical ratios of acceptable criteria relative to the peak discharge, peak time and total runoff volume among the calibrated and validated historical flood events, respectively, are employed to evaluate the parameter obtained by different calibration methods performance for Xinanjiang model. The posterior mean, standard deviation (SD), coefficient of variation (CV), and correlation structure induced between the parameters of the Xinanjiang model in the high probability distribution (HPD) region of the posterior probability distribution with SCEM-UA algorithm are evaluated for different time steps. Furthermore, results comparison was made.

SELECTED OPTIMIZATION METHODS

The GA is a powerful global search algorithm that is based on the concepts of natural selection and genetics (Holland 1975, 1992). An extensive description of GA can be found in Goldberg (1989). GA relies on the collective learning process within a population of individuals, each of which represents a search point in the space of potential solutions. GA differs from other traditional optimization methods in that it searches among a population of points and works with a coding of the parameter set rather than the parameter values themselves. GA uses probabilistic transition rules, not deterministic rules. In a GA, first an initial population is generated randomly. A fitness value is associated with each individual based on a measure of optimality of the objective function evaluated at the individual it represents. Then, it evaluates the population and operates on the population using reproduction, crossover and mutation operators to produce new and hopefully better solutions. Reproduction is designed to use fitness to guide the evolution of chromosomes. Crossover is the process by which chromosomes selected from a source population are combined to form offspring which are potential members of a successor population. It is hoped that good parents may produce good solutions. Mutation is an operator which is used to maintain the diversity in the population and allows the algorithm to avoid local minima by preventing the individuals in a population from becoming too similar. In the initial implementation of GA, the variables were encoded as strings of binary digits, i.e. zero and one. But if the binary coded GAs are applied to problems having a large search space and seeking high precision, they will spend a considerable time performing encoding and decoding processes due to encoding parameters as finite length strings. Therefore, a real coded GA is adopted in this study. The rules of reproduction, crossover and mutation employed in this work are well described in Cheng *et al.* (2002).

The SCE-UA algorithm combines the direct search method of the simplex downhill descent procedure (Nelder & Mead 1965) with the concept of a controlled random search by a systematic evolution of points in the direction of global improvement, competitive evolution (Holland 1975) and the concept of complex shuffling. In brief, the SCE-UA method involves the initial selection of a 'population' of points distributed randomly throughout the feasible parameter space. The population is portioned into several 'complexes' that consist of $2n+1$ points. Each complex evolves independently using the simplex downhill descent algorithm. The complexes are periodically shuffled to form new complexes in order to share information between the complexes. The SCE-UA algorithm includes various algorithmic parameters. The number of complexes p is the most important parameter. Sensitivity tests indicate that the dimensionality of the calibration problem (number of calibration parameters) is the primary factor determining the proper choice of p (Duan *et al.* 1994). In general, the larger the chosen value of p the higher the probability of converging into the global optimum but at the expense of a larger number of model simulations (the number of model simulations is virtually proportional to p), and vice versa. In order to reduce the chance of premature termination, Kuczera (1997) suggested setting p equal to the number of calibration parameters. In the application example presented below p was set equal to the number of calibration parameters, and the recommended values given by Duan *et al.* (1994) were applied for the other algorithmic parameters.

The SCEM-UA algorithm is an adaptive evolutionary Monte Carlo Markov Chain (MCMC) method inspired by the SCE-UA global optimization algorithm, and uses the Metropolis Hastings (MH) strategy (Metropolis *et al.* 1953; Hastings 1970) instead of the downhill simplex method for population evolution. It not only provides the most probable parameter set, but also estimates the uncertainty associated with estimated parameters. Thus SCEM-UA in every model run is able to simultaneously identify both the most likely parameter set and its associated posterior probability distribution. In brief, the SCEM-UA algorithm starts by randomly selecting initial population of points throughout the feasible parameter space. In the absence of any information about the posterior distribution, a uniform sampling distribution is used. For each parameter set θ, the posterior density, or likelihood for describing the observed data y can be computed by SCEM-UA using the equation specified by Box & Tiao (1973):

$$L(\theta|y) = \exp\left[-\frac{1}{2}\sum_{j=1}^{n}\left|\frac{e(\theta)_j}{\sigma}\right|^2\right] \tag{1}$$

in which n represents the number of measurements; the SD parameter σ denotes the measurement error deviation of the observations and Box & Tiao (1973) showed the influence of σ can be integrated out by assuming a noninformative prior density of the form $p(\theta)\infty\sigma^{-1}$, leading to the following form of the posterior density of $\theta^{(t)}$:

$$p(\theta^{(t)}|y) \,\infty\, \left[M(\theta^{(t)})\right]^{-\frac{1}{2}N} \tag{2}$$

where

$$M(\theta^{(t)}) = \sum_{i=1}^{N}(\theta^{(t)})_i^2 \tag{3}$$

and e represents the error residuals between model and observation, and it is expressed as follows:

$$e_i = y_i - \hat{y}_i \ (i = 1, 2, \cdots N) \tag{4}$$

where y_i and $\hat{y}_i$ are the observed and simulated value, respectively. Then, the population of parameter sets is partitioned into a number of complexes, and in each complex a parallel sequence is launched from the point that exhibits the highest posterior density. Subsequently, a new candidate point in each sequence is generated using a multivariate normal distribution either centered on the current draw of the sequence or the mean of the points in the complex augmented with the covariance structure induced between the points in the complex. In order to test whether the candidate point is added to the current sequence, and the Metropolis-annealing (Metropolis *et al.* 1953) criterion is employed. Finally, the new candidate point is shuffled into the original population of complexes. The algorithmic steps are performed until the Gelman–Rubin convergence (Gelman & Rubin 1992) diagnostic for each of the parameters demonstrates convergence to a stationary posterior target distribution. A detailed description and explanation of the SCEM-UA can be found in Vrugt *et al.* (2003b).

XINANJIANG MODEL

The Xinanjiang model was developed to forecast flows to the Xinanjiang reservoir by Zhao *et al.* (1980). As a CRR model with distributed parameters, the model has been successfully and widely applied to large basins in the humid and semi-humid regions of China for flood forecasting (Zhao 1992). The main hypothesis used in the model development is the concept of runoff formation on repletion of storage. The original Xinanjiang model includes a runoff generating component and a runoff routing component. For a large catchment, it can be divided into a set of sub-areas and runoff is first transformed into discharge by a linear system

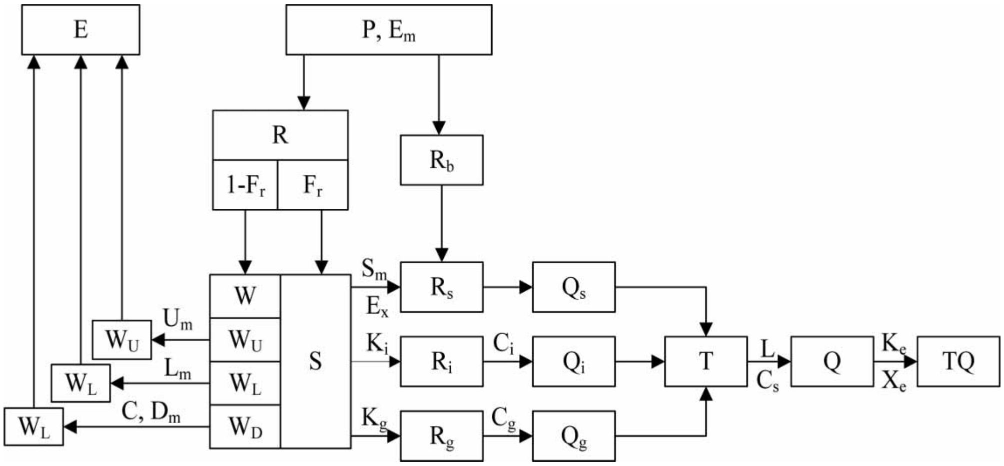

Figure 1 | Flow chart of the Xinanjiang model.

calculated from the runoff generating component. The flow chart of the Xinanjiang model is shown in Figure 1. It has 17 parameters that include seven runoff generating component parameters (U_m, L_m, D_m, B, I_m, K, C) and 10 runoff routing parameters (S_m, E_x, K_g, K_i, C_g, C_i, C_s, K_e, X_e, L). The runoff routing component parameter L, denoting the lag time of routing for each sub-area and being an empirical value which is mainly dependent on the length and slope of a stream, can be estimated by relating to observable characteristics of the watershed. The other 16 parameters are abstract conceptual representations of non-measurable watershed characteristics that have to be calibrated by an optimization method. This is mainly because the manual calibration can be a rather tedious and time-consuming task. The physical descriptions of these parameters are listed in Table 1. The value of each parameter is usually within a certain range according to physical and mathematical constraints, information about watershed characteristics, and from modeling experiences. The Muskingum method is used for the routing of discharge from the outlet of each sub-area to the outlet of the main catchment. The parameters of Muskingum method must satisfy the constraints as follows for each channel of the sub-basin.

$$2K_eX_e \leq \Delta t \leq 2(K_e - K_eX_e)$$

where K_e and X_e are the Muskingum coefficients, K_e is a storage constant having the dimension of time, X_e is a dimensionless constant for the reach of the channel, and Δt is the time step. For a detailed description and explanation of the Xinanjiang model, please refer to Zhao *et al.* (1980) and Zhao (1992).

CASE STUDY

Study area and data

As shown in Figure 2, the study area is located in Hunan province of southern China and at the downstream of the Xiaoshui Stream, which is one of the tributary rivers of the Xiangjiang River. The investigated GOMs are used to estimate and analyze parameter values of the Xinanjiang model for flood forecasting in the Shuangpai Reservoir. The reservoir, with a drainage area of 10,594 km^2 and a water holding capacity of up to 373.8 million cubic meters, is used for power generation and flood control, as well as for irrigation purposes. The length of the main stream is 154.9 km with an averaged slope of 0.61%. The area is in a subtropical monsoon zone with rich rainfalls and good vegetation cover. The annual rainfall is 1,500 mm; the averaged depth of runoff is 893 mm and the averaged discharge is 300 $m^3 s^{-1}$. However, the temporal distribution of the rainfall during a given year is significantly heterogeneous in this area. The flood events in this area are mainly due to the thunderstorms. Of the total rainfall, 45.9% falls between April and June, and 34% between September and October, which are referred to as the high flow periods.

Table 1 | Parameters of the Xinanjiang model

	Parameter	Physical description	Unit	Initial range
		Runoff generating parameter		
1	K	Ratio of potential evapotranspiration to pan evaporation	[–]	0.1–1.2
2	U_{m}	Averaged soil moisture storage capacity of the upper layer	[mm]	10–40
3	L_{m}	Averaged soil moisture storage capacity of the lower layer	[mm]	50–90
4	D_{m}	Averaged soil moisture storage capacity of the deep layer	[mm]	10–80
5	C	Coefficient of the deep layer that depends on the proportion of the basin area covered by vegetation with deep roots	[–]	0.1–0.3
6	B	Exponential parameter with a single parabolic curve, which represents the non-uniformity of the spatial distribution of the soil moisture storage capacity over the catchment	[–]	0.1–0.9
7	I_{m}	Percentage of impervious and saturated areas in the catchment	[%]	0.0–0.1
		Runoff routing parameter		
8	S_{m}	Areal mean free water capacity of the surface soil layer, which represents the maximum possible deficit of free water storage	[mm]	10–50
9	E_x	Exponent of the free water capacity curve influencing the development of the saturated area	[–]	1.1–1.4
10	K_{g}	Outflow coefficients of the free water storage to groundwater relationships	[–]	0.1–0.4
11	K_{i}	Outflow coefficients of the free water storage to interflow relationships	[–]	0.2–0.5
12	C_{i}	Recession constants of the lower interflow storage	[–]	0.1–0.99
13	C_{g}	Recession constants of the groundwater storage	[–]	0.7–0.99
14	C_{s}	Recession constant in the lag and route method for routing through the channel system with each sub-basin	[–]	0.01–0.4
15	K_e	Parameter of the Muskingum method	[–]	1.5–2
16	X_e	Parameter of the Muskingum method	[–]	0.3–0.5
17	L	Lag in time	[–]	

The region is divided into 12 sub-areas, each of which has the same set of model parameters. Each sub-area is represented by a rain gauge station. Table 2 summarizes the rain gauge stations covered in this study, the representing area and the corresponding weighting for each rain gauge. Historical flooding data for two different time steps are employed to determine appropriate Xinanjiang model parameters. First, a total of 34 historical floods with 3 hr time step for 12 years between 1984 and 1995 are employed to calibrate the model parameters, whilst 11 historical floods for 2 years between 1999 and 2000 are utilized to verify these parameters. Then, a total of 36 historical floods with a 1 hr time step for 5 years between 2000 and 2004 are employed to calibrate the model parameters, whilst 12 historical floods for 2 years between 2005 and 2006 are used to verify these parameters. Table 1 lists the initial ranges of parameter values for the Shuangpai Reservoir.

Optimization objective and performance evaluation

CRR model calibration is a highly complex non-linear optimization problem, and its process is an iterative procedure in which parameter values are adjusted and refined based on comparison of simulated and observed values. Its main objective is to determine the values of the model parameters by a robust calibration procedure, which provides the best agreement between simulated and observed values. To measure the closeness of the model output and observed data, a single criterion selected has been devoted to identify the 'best' criterion and the 'best' optimization (Boyle *et al.* 2000). In general, the root-mean-squared error (RMSE) criterion has been most commonly used to evaluate on either the streamflows or the log of the streamflows in the literature (Boyle *et al.* 2000; Vrugt *et al.* 2003b, 2006; Cooper *et al.* 2007). It is

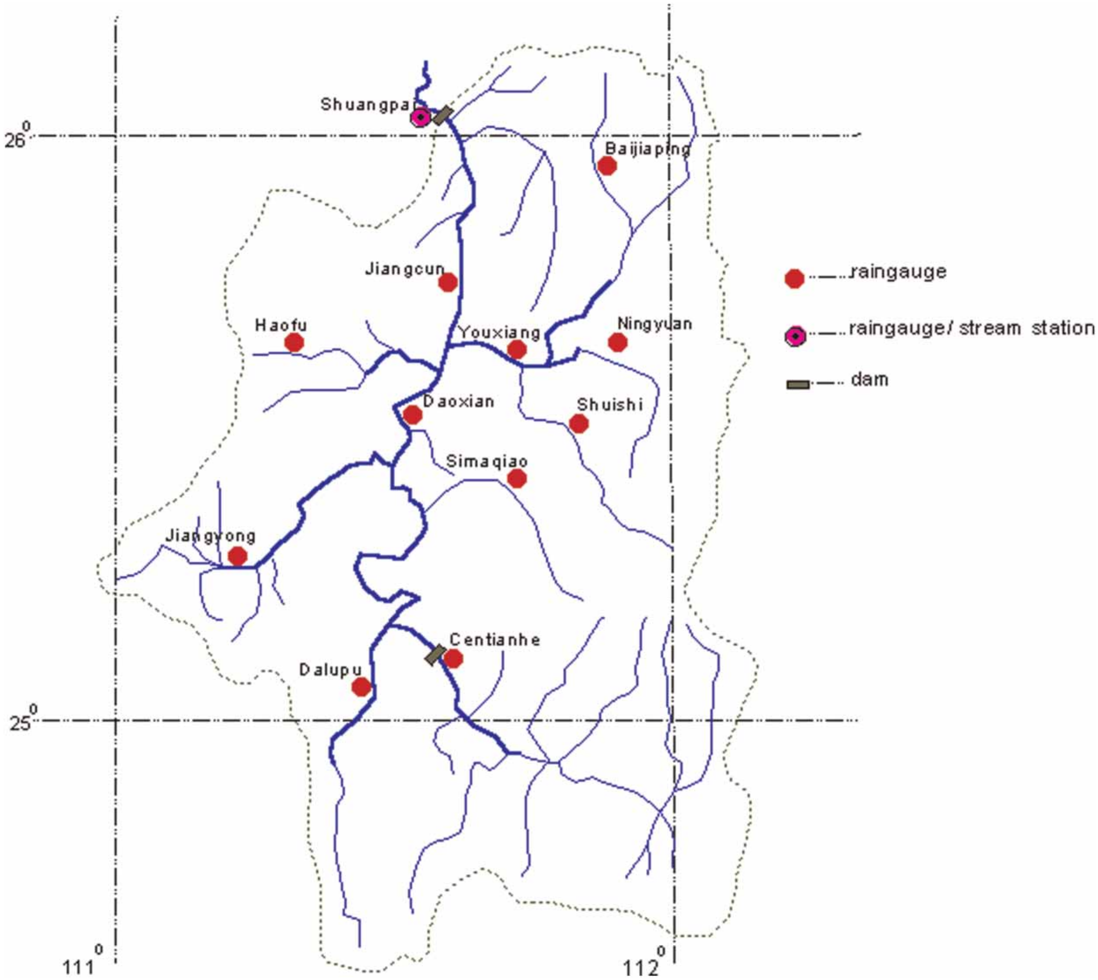

Figure 2 | Map of the Shuangpai area with locations of rain gauge stations (Cheng *et al.* 2002).

Table 2 | Details of rain gauge stations in the study area

Station	Kind of station	Station name	Weighting	Area (km^2)
01	Rainfall	Jiangcun	0.0915	745
02	Rainfall	Daoxian	0.0846	691
03	Rainfall	Haofu	0.0689	562
04	Rainfall	Jiangyong	0.1134	926
05	Rainfall	Dalupu	0.1200	979
06	Rainfall	Centianhe	0.0326	266
07	Rainfall	Simaqiao	0.0651	531
08	Rainfall	Youxiang	0.0762	622
09	Rainfall	Ningyuan	0.0911	744
10	Rainfall	Shuishi	0.0651	532
11	Rainfall	Baijiaping	0.1210	988
12	Rainfall/streamflow	Shuangpai	0.0705	576

expressed as follows:

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(Q_s(i) - Q_o(i))^2} \tag{5}$$

where $Q_o(i)$ and $Q_s(i)$ are the observed and simulated streamflow or log streamflow, respectively, and N is the number of data points considered. In this paper, the RMSE is selected as the optimization objective to obtain the best fit between the simulated and observed flows hydrograph.

The performance of a rainfall-runoff model for accurate flood forecasting heavily depends on choosing more suitable model parameters. In order to evaluate the different model parameters obtained by different optimization methods for rainfall-runoff model performance, three statistical ratios of acceptable criteria recommended by the national criteria in China for flood forecasting (NCHI 2000) are also employed. The result of simulation or forecasting is relative to peak value for this flood if the absolute percentage error of peak discharge and total runoff volume between the simulated and observed floods is less than 20%. The result is relative to peak time if the difference in peak time is within a routing period. Three statistical ratios relative to peak discharge, peak time and total runoff volume among the calibrated and validated historical flood events are expressed as $r_{\text{peak\_discharge}}$, $r_{\text{peak\_time}}$, and r_{runoff}, respectively:

$$\text{Maximize } r_{\text{peak\_discharge}} = \frac{M_{\text{pd}}}{N} \times 100\% \tag{6a}$$

$$\text{Maximize } r_{\text{peak\_time}} = \frac{M_{\text{pt}}}{N} \times 100\% \tag{6b}$$

$$\text{Maximize } r_{\text{runoff}} = \frac{M_{\text{r}}}{N} \times 100\% \tag{6c}$$

where M_{pd}, M_{pt}, and M_{r} represent the total number of floods that satisfy the acceptable criteria relative to the peak discharge, peak time and total runoff volume, respectively, and N is the total number of the calibrated floods or validated ones. If all three ratios are more than 85%, the performances of parameter calibration satisfy the first level standard of flood forecasting calibration or validation. If all three ratios are more than 75% and one of them is more than 85%, the performances of parameter calibration satisfy the second level standard of flood forecasting calibration or validation. Otherwise, the results of the performances of parameter calibration are unsatisfactory for online flood forecasting.

APPLICATION RESULT AND DISCUSSION

Two case studies with different time steps for the same catchment are used to assess and calibrate CRR model parameters by three automatic optimization techniques. The first example with a 3 hr time step has been studied by Cheng et al. (2002, 2005, 2006), and will be employed as a reference benchmark for evaluation and analysis of flood forecasting model parameters. The second example is with a 1 hr time step. Abundant data from a hydrological telemetry system were accumulated in large and medium reservoirs in China since 1998 (Cheng et al. 2004; Cheng & Chau 2004) and 1 hr data about rainfalls and discharges are available. It is extremely significant and valuable work for analysis on the optimization of short-term flood forecasting model parameters based on a continuous time series of hydrological data from a hydrological telemetry system. The Shuangpai Reservoir, as the second batch of reservoir with flood forecasting and operation system in China, has accumulated a large amount of historical floods data since the hydrological telemetry system was established in 2000 (Cheng & Chau 2004). There are a total of 48 historical floods with a 1 hr time step with observed discharge more than 600 $m^3\,s^{-1}$.

In order to obtain good performance of the selected global optimization algorithms, some parameters of algorithms need to be chosen. For GA, it includes a moderate population size (P_{size}), a high crossover probability (P_{c}), a low mutation probability (P_{m}), and the maximum number of generation ($T_{\max}$). Their values have been suggested in some research for the calibration of hydrological models (Cheng et al. 2002, 2006, 2008). P_{size} criticality affects the efficiency and solution quality of the GAs. Generally, P_{size} is set to be a value between 150 and 300. P_{c} controls the frequency of crossover operation. Generally, P_{c} is chosen between 0.5 and 0.8. P_{m} is a critical factor in extending the diversity of the population. Generally, P_{m} is often chosen between 0.001 and 0.1. In this study, $P_{\text{size}} = 300$, $P_{\text{c}} = 0.8$, $P_{\text{m}} = 0.1$, and $T_{\max} = 500$ are employed, and the weighted roulette wheel approach and the ranking schemes are also adopted. For the SCE-UA algorithmic parameters, recommended values given by Duan et al. (1994) are used as follows: $m = 2n + 1$; $q = n + 1$; $\alpha = 1$; $\beta = 2n + 1$; where m is the number of points in a complex; q is the number of points in a sub-complex; α is the number of consecutive offspring generated by each sub-complex; β is the number of evolution steps taken by each complex; p is the number of complexes which depends on the type of the problem. As suggested by Kuczera (1997), p was set equal to the number of calibration parameters to reduce the chance of premature termination of the search

Table 3 | The posterior mean (Mean), standard deviation (SD), coefficient of variation (CV), and correlation coefficients between the Xinanjiang model parameters in highest posterior probability using the SCEM-UA, and the parameters obtained using GA, SCE-UA and the SCEM-UA with highest posterior probability for 3 hr time step

Parameters	K	U_m	L_m	D_m	C	B	I_m	S_m	E_x	K_g	K_i	C_i	C_g	C_s	K_e	X_e
Mean	0.554	26.854	81.948	21.534	0.134	0.159	0.046	13.898	1.157	0.104	0.207	0.227	0.9987	0.332	1.824	0.009
SD	0.015	4.587	5.442	6.212	0.023	0.021	0.010	0.492	0.039	0.003	0.005	0.044	0.0002	0.008	0.034	0.007
CV	2.627	17.081	6.641	**28.849**	17.475	13.167	**21.490**	3.527	3.323	2.835	2.513	19.332	0.0241	2.355	1.882	**76.751**
K	1.000															
U_m	−0.164	1.000														
L_m	0.025	0.013	1.000													
D_m	−0.279	0.243	−0.577	1.000												
C	−0.065	−0.213	0.223	0.207	1.000											
B	− **0.613**	−0.132	−0.006	0.111	0.165	1.000										
I_m	0.164	−0.172	− **0.657**	0.325	−0.491	−0.196	1.000									
S_m	−0.172	0.097	−0.055	0.104	−0.037	−0.010	0.304	1.000								
E_x	0.069	−0.314	0.051	0.109	−0.328	−0.156	**0.564**	0.123	1.000							
K_g	0.003	−0.058	0.049	−0.082	0.062	0.070	−0.142	−0.469	0.007	1.000						
K_i	0.035	−0.006	−0.029	0.004	−0.041	−0.021	0.147	0.253	0.031	−0.066	1.000					
C_i	0.058	−0.252	−0.053	0.044	0.066	−0.074	−0.014	− **0.563**	−0.009	0.221	0.013	1.000				
C_g	−0.103	−0.095	−0.034	0.037	0.056	0.024	−0.080	−0.106	−0.065	−0.012	0.001	0.083	1.000			
C_s	−0.029	0.234	0.070	−0.035	−0.093	0.087	−0.182	−0.015	−0.009	−0.076	−0.139	−0.714	0.039	1.000		
K_e	−0.059	0.046	0.052	0.019	−0.036	0.011	0.220	0.447	0.102	−0.063	0.012	**-0.526**	−0.067	0.033	1.000	
X_e	0.020	0.111	−0.040	−0.004	−0.117	−0.035	−0.004	−0.108	0.105	0.024	−0.039	0.0505	0.130	0.081	−0.127	1.000
GA	0.536	22.967	66.069	25.389	0.121	0.199	0.047	9.307	1.205	0.122	0.327	0.977	0.887	0.352	1.715	0.001
SCE-UA	0.558	33.253	70.703	23.525	0.134	0.163	0.038	8.552	1.31	0.123	0.321	0.357	0.875	0.352	1.803	0.001
SCEM-UA	0.558	27.431	77.324	24.834	0.142	0.152	0.046	13.592	1.307	0.102	0.429	0.262	0.945	0.328	1.801	0.004

algorithm. The SCEM-UA algorithm is run with a population size $s = 250$ and $q = 10$ complexes.

Of those with a 3 hr time step, 34 historical flood data (1984–1995) are applied to the calibration of the Xinanjiang model parameters in the Shuangpai Reservoir by using three different GOMs. Table 3 presents the posterior mean, SD, CV, and correlation structure induced between the parameters of the Xinanjiang model in HPD region of the

Table 4 | Performance of the calibrated parameters with highest posterior probability by SCEM-UA for 3 hr time step

Floods	Observed (m^3/s^{-1})	Simulated (m^3/s^{-1})	Percentage error (%)	Observed peak-time (yyyy-mm-dd hh)	Simulated peak-time (yyyy-mm-dd hh)	Error (number)	Total volume error (%)
19840416	1,130	1,317.0	16.55	1984-04-17 14	1984-04-17 08	−1	−2.48
19840530	4,310	4,192.8	−2.72	1984-06-01 17	1984-06-01 14	−1	−4.65
19850411	1,200	1,433.4	19.45	1985-04-12 23	1985-04-12 23	0	−1.18
19850527	5,770	5,017.8	−13.04	1985-05-27 23	1985-05-27 23	0	0.24
19850610	1,370	980.2	− **28.46**	1985-06-11 23	1985-06-11 20	− **2**	−12.61
19860706	2,560	2,289.5	−10.57	1986-07-06 20	1986-07-06 20	0	6.53
19870404	1,790	1,814.1	1.34	1987-04-05 20	1987-04-05 20	0	5.61
19870515	1,720	1,151.4	− **33.06**	1987-05-16 05	1987-05-16 08	1	−5.31
19870521	1,840	1,494.9	−18.76	1987-05-22 02	1987-05-22 02	1	−4.25
19870606	1,080	1,072.5	−0.70	1987-06-07 11	1987-06-07 14	1	6.88
19870614	1,350	1,793.1	**32.83**	1987-06-14 23	1987-06-14 23	–1	19.34
19870722	1,740	2,037.1	17.07	1987-07-23 11	1987-07-23 14	1	**22.24**
19870729	2,170	1,891.7	−12.83	1987-07-29 20	1987-07-29 20	0	2.07
19880903	1,960	2,329.8	18.87	1988-09-05 08	1988-09-05 02	**–3**	6.71
19890511	2,870	2,785.6	–2.94	1989-05-13 14	1989-05-13 11	–1	11.06
19890522	1,890	2,110.4	11.66	1989-05-22 23	1989-05-22 20	0	15.68
19890529	1,880	1,855.6	−1.30	1989-05-31 05	1989-05-31 02	–1	–5.14
19900530	2,770	2,675.8	–3.40	1990-06-01 02	1990-06-01 02	0	10.76
19900607	2,110	2,480.7	17.57	1990-06-08 11	1990-06-08 11	0	8.42
19910616	1,010	601.0	− **40.49**	1991-06-16 23	1991-06-16 20	–1	-15.67
19920423	2,320	2,544.8	9.69	1992-04-24 11	1992-04-24 14	1	10.50
19920516	3,020	2,395.0	−20.70	1992-05-17 17	1992-05-17 14	−1	−2.96
19920705	3,760	4,504.4	19.80	1992-07-06 20	1992-07-06 17	−1	5.65
19930513	2560	2,313.9	−9.61	1993-05-14 14	1993-05-14 14	0	−0.22
19930607	2,010	1,877.5	−6.59	1993-06-09 11	1993-06-09 14	1	0.51
19930615	1,940	2,082.1	7.32	1993-06-16 02	1993-06-15 23	−1	5.43
19940421	5,070	4,556.0	−10.14	1994-04-23 20	1994-04-23 17	−1	−2.78
19940525	2,700	2,559.2	−5.21	1994-05-26 14	1994-05-26 17	1	−5.29
19940614	3,330	2,746.0	−17.54	1994-06-18 08	1994-06-17 23	− **5**	−5.75
19940723	5,810	5,183.7	−10.78	1994-07-24 05	1994-07-24 02	−1	−4.12
19940805	2,660	2,821.9	6.09	1994-08-06 11	1994-08-06 14	1	0.77
19950425	2,040	1,918.8	−5.94	1995-04-26 11	1995-04-26 08	−1	1.14
19950526	1,400	1,728.3	**23.45**	1995-05-26 17	1995-05-26 20	1	12.18
19950614	3,880	3,896.6	0.43	1995-06-17 02	1995-06-17 05	0	−6.72

Notes: The total number of floods is **34,** which are qualificatory relative to the error of peak discharge, is **28** and the ratio of qualifying simulation is **82.35%,** which are qualificatory relative to the error of peak time, is **32** and the ratio of qualifying simulation is **94.12%,** which are qualificatory relative to the error of total runoff volume, is **33** and the ratio of qualifying simulation is **97.06%.**

Table 5 | Performances of validated parameter with highest posterior probability by SCEM-UA for 3 hr time step

Floods	Observed (m^3/s^{-1})	Simulated (m^3/s^{-1})	Percentage error (%)	Observed peak-time (yyyy-mm-dd hh)	Simulated peak-time (yyyy-mm-dd hh)	Error (number)	Total volume error (%)
19990426	2,550	2,246.7	– 11.88	1999-04-25 20	1999-04-25 23	1	– 7.80
19990526	4,893	4,778.6	– 2.34	1999-05-26 17	1999-05-26 17	0	3.24
19990618	1,690	1,451.5	– 14.13	1999-06-18 20	1999-06-18 23	1	– 5.52
19990627	1,111	1,257.5	13.19	1999-06-25 17	1999-06-25 17	0	2.83
19990831	1,965	2,118.6	7.82	1999-08-31 23	1999-08-31 23	0	7.66
20000403	1,628	2,070.8	**27.24**	2000-04-02 23	2000-04-02 23	0	**23.39**
20000410	1,798	1,799.5	0.08	2000-04-09 20	2000-04-09 17	– 1	– 2.19
20000426	909	837.7	– 7.84	2000-04-26 20	2000-04-26 20	0	16.65
20000510	1,039	885.9	– 14.73	2000-05-10 02	2000-05-10 02	0	– 10.82
20000528	3,163	3,333.4	5.39	2000-05-28 14	2000-05-28 17	1	19.14
20000611	1,096	1,095.7	– 0.05	2000-06-12 05	2000-06-12 02	– 1	– 15.23

Notes: The total number of floods is **11**, which are qualificatory relative to the error of peak discharge, is **10** and the ratio of qualifying simulation is **90.91%**, which are qualificatory relative to the error of peak time, is **11** and the ratio of qualifying simulation is **100%**, which are qualificatory relative to the error of total runoff volume, is **10** and the ratio of qualifying simulation is **90.91%**.

posterior probability distribution with SCEM-UA algorithm. The parameter estimates obtained by GA, SCE-UA and the SCEM-UA with the highest posterior probability are also shown in Table 3. It can be observed that the CV of the parameter D_{m}, I_{m}, X_e are very high. This demonstrates distributions of these parameters are highly dispersed, i.e., there is no well-defined region in the sense of a compact region for a 3 hr time step. Moreover, the absolute values of some correlation coefficients are higher than 0.5 in the parameter correlation matrix, indicating the model parameter values are unstable, when the Xinanjiang model is calibrated. According to the three statistical ratios of acceptable national criteria relative to the peak discharge, peak time and total runoff volume among the calibrated and validated historical flood events for flood forecasting in China, statistics showing the detailed performance of the calibrated parameters with the highest posterior probability using SCEM-UA for a 3 hr time step are listed in Tables 4 and 5.

Table 6 gives the performance statistics obtained by different optimization methods for the 3 hr time step during calibration and validation. From Table 6, using the SCEM-UA algorithm with the highest posterior probability calibrated parameters, the values of three criteria: $r_{\mathrm{peak\_discharge}}$, $r_{\mathrm{peak\_time}}$, and $r_{\mathrm{runoff,}}$ are 82.35, 94.12 and 97.06%, respectively, for the calibrated result, and 90.91, 100 and 90.91%, respectively, for the validated result. The corresponding results calibrated by GA are 79.41, 91.18 and 94.12%, respectively, and the validated results are 90.91, 100 and 90.91%, respectively. The results calibrated by SCE-UA are 82.35, 91.18 and 97.06%, respectively, and the validated results are 90.91, 100 and 90.91%, respectively. For the validated floods, the qualifying ratios of the peak discharge, peak time and runoff total volume are all more than 85%. It can be concluded that the results of SCE-UA and SCEM-UA are basically similar and slightly better than GA.

Table 6 | Comparison results of the SCEM-UA, SCE-UA and GA algorithm for 3 hr time step

	Ratio of qualifying of calibration (%)			Ratio of qualifying of validation (%)		
Algorithm	Peak discharge	Peak time	Total runoff	Peak discharge	Peak time	Total runoff
GA	79.41	91.18	94.12	90.91	100	90.91
SCE-UA	82.35	91.18	97.06	90.91	100	90.91
SCEM-UA	82.35	94.12	97.06	90.91	100	90.91

The performances of all simulated hydrographs of rainfall-runoff process with 3 hr time step by different calibration methods from 1984 to 1995 during the calibration and from 1999 to 2000 during the validation are shown in Figures 3 and 4, respectively. In Figures 3 and 4, the cyan region shows the model prediction uncertainty region associated with the 95% total error in terms of

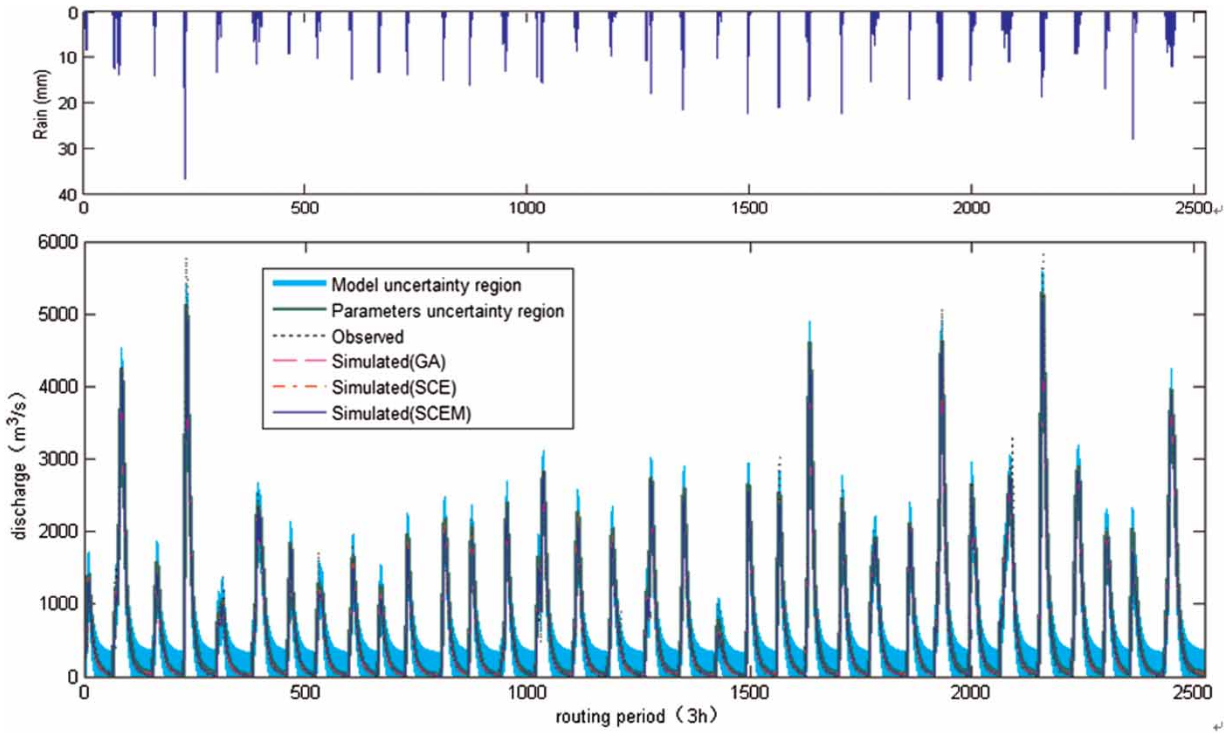

Figure 3 | The rain, model prediction uncertainty that results from model and measurement uncertainty, discharge prediction uncertainty associated with the most probable parameter set derived using the SCEM algorithm, observed hydrographs and simulated hydrographs (GA, SCE, SCEM) for 34 historical floods (1984–1995) with a 3 hr time step during calibration. (The colour version of this figure is available in the online version of the paper at http://www.iwaponline.com/jh/toc.htm.)

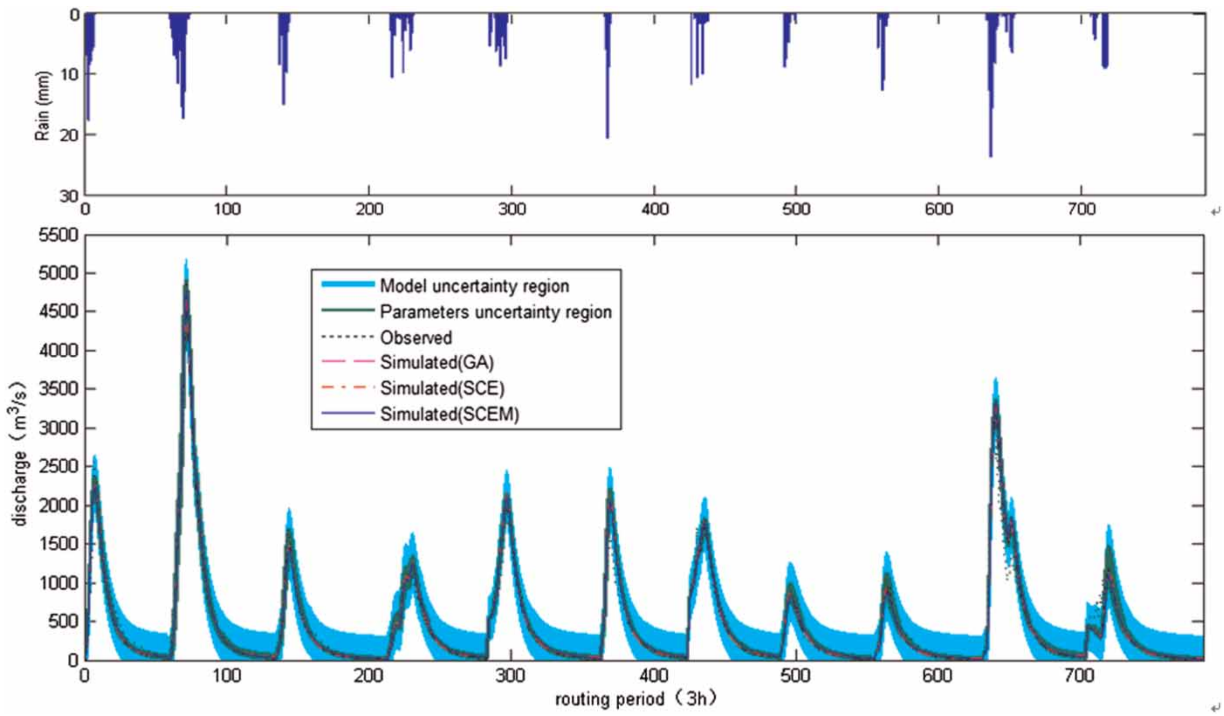

Figure 4 | The rain, model prediction uncertainty that results from model and measurement uncertainty, discharge prediction uncertainty associated with the most probable parameter set derived using the SCEM algorithm, observed hydrographs and simulated hydrographs (GA, SCE, SCEM) for 11 historical floods (1999–2000) with a 3 hr time step during validation. (The colour version of this figure is available in the online version of the paper at http://www.iwaponline.com/jh/toc.htm.)

Table 7 | The posterior mean (Mean), standard deviation (SD), coefficient of variation (CV), and correlation coefficients between the Xinanjiang model parameters in highest posterior probability using the SCEM-UA, and the parameters obtained using GA, SCE-UA and the SCEM-UA with highest posterior probability for 1 hr time step

Parameters	K	U_m	L_m	D_m	C	B	I_m	S_m	E_x	K_g	K_i	C_i	C_g	C_s	K_e	X_e
Mean	0.647	20.899	69.093	60.561	0.142	0.963	0.039	38.859	1.200	0.297	0.376	0.106	0.983	0.208	1.997	0.169
SD	0.014	1.219	3.609	7.119	0.021	0.029	0.014	0.792	0.081	0.008	0.010	0.005	0.004	0.005	0.003	0.024
CV	2.170	5.831	5.223	50.685	14.817	2.975	34.501	2.038	6.721	2.823	2.708	4.617	0.412	2.527	0.145	14.138
K	1.000															
U_m	−0.038	1.000														
L_m	−0.291	−0.425	1.000													
D_m	0.022	−0.433	0.238	1.000												
C	0.071	−0.274	0.105	0.159	1.000											
B	−0.191	0.295	−0.316	−0.631	−0.073	1.000										
I_m	0.199	0.318	−0.316	− **0.666**	−0.284	0.123	1.000									
S_m	−0.024	−0.073	−0.063	0.060	0.127	−0.155	0.086	1.000								
E_x	0.065	−0.200	0.230	**0.666**	0.235	−0.340	− **0.798**	0.164	1.000							
K_g	− **0.613**	0.040	0.081	−0.107	−0.231	0.286	−0.245	−0.218	−0.122	1.000						
K_i	**0.633**	0.117	−0.261	−0.231	0.039	−0.035	0.448	0.301	−0.193	−0.345	1.000					
C_i	0.026	0.010	−0.053	0.021	0.035	0.090	−0.010	−0.141	0.057	0.070	0.232	1.000				
C_g	0.059	−0.158	−0.003	0.133	0.307	−0.160	0.040	−0.009	0.215	− **0.533**	−0.012	0.092	1.000			
C_s	−0.373	−0.080	0.283	0.143	−0.046	0.089	−0.392	− **0.582**	0.169	0.278	− **0.535**	−0.196	0.184	1.000		
K_e	0.074	0.016	−0.019	−0.094	-0.054	0.077	0.009	0.028	-0.015	-0.046	0.031	-0.025	−0.024	−0.145	1.000	
X_e	−0.102	−0.010	0.285	−0.242	−0.437	0.071	0.112	−0.087	−0.092	0.122	−0.109	−0.059	−0.053	0.245	0.105	1.000
GA	0.635	19.801	78.332	65.58	0.191	0.911	0.012	37.139	1.138	0.331	0.321	0.105	0.979	0.220	2	0.117
SCE-UA	0.646	20.926	67.410	48.920	0.159	0.890	0.1	41.266	1.179	0.285	0.385	0.1	0.988	0.198	2	0.174
SCEM-UA	0.641	21.219	66.508	54.811	0.134	0.989	0.05	38.971	1.117	0.297	0.376	0.103	0.986	0.207	1.998	0.179

Table 8 | Performance of the calibrated parameters with highest posterior probability by SCEM-UA for 1 hr time step

Floods	Observed (m^3/s^{-1})	Simulated (m^3/s^{-1})	Percentage error (%)	Observed peak-time (yyyy-mm-dd hh)	Simulated peak-time (yyyy-mm-dd hh)	Error (number)	Total volume error (%)
20000426	718.5	666.1	−7.29	2000-04-26 22	2000-04-26 23	1	**35.14**
20000430	737.7	746.6	1.20	2000-04-30 12	2000-04-30 09	−3	6.41
20000510	691.3	736.7	6.56	2000-05-10 07	2000-05-10 07	0	19.93
20000527	1,343.4	1,274.0	15.17	2000-05-27 07	2000-05-27 08	1	2.64
20000528	2,445.3	2,514.1	2.81	2000-05-28 16	2000-05-28 18	2	17.45
20001022	1,569.4	1,384.6	− 11.78	2000-10-22 09	2000-10-23 16	**31**	18.79
20010406	1,494.3	1,124.3	− **24.76**	2001-04-06 17	2001-04-06 16	−1	−8.1
20010417	732.7	552.9	− **24.54**	2001-04-17 00	2001-04-16 23	−1	14.66
20010418	899.8	739.9	−17.77	2001-04-19 01	2001-04-18 23	−2	−13.44
20010421	1,298.1	1,121.6	−13.60	2001-04-20 21	2001-04-20 23	2	10.43
20010509	1,249.6	1,165.6	−6.72	2001-05-09 23	2001-05-09 22	−1	4.48
20010613	4,976.6	4,554.7	−8.48	2001-06-13 20	2001-06-13 17	−3	19.98
20010707	1,896.2	1,303.9	− **31.24**	2001-07-07 19	2001-07-07 23	**4**	−2.67
20020313	2,347.1	2,056.6	−12.38	2002-03-13 21	2002-03-13 21	0	12.79
20020411	1,656.6	1,203.0	− **27.38**	2002-04-11 02	2002-04-11 02	0	−9.68
20020426	1,556.6	1,451.6	−6.74	2002-04-26 04	2002-04-26 03	−1	7.28
20020510	1,222.6	1,080.6	−11.61	2002-05-10 06	2002-05-10 04	−2	**43.21**
20020514	2,615.5	2,074.0	− **20.70**	2002-05-14 20	2002-05-14 20	0	9.5
20020618	3,184.9	2,718.2	−14.65	2002-06-18 12	2002-06-18 10	−2	−5.85
20020701	6,245.8	6,172.7	−1.17	2002-07-01 23	2002-07-01 21	−2	−3.14
20020721	1,132.1	1,071.6	−5.34	2002-07-21 06	2002-07-21 07	1	17.12
20020726	3,532.1	2,980.2	−15.63	2002-07-26 13	2002-07-26 14	1	−3.15
20020807	5,230.2	5,443.2	4.07	2002-08-08 00	2002-08-08 00	0	10.28
20020819	3,365.2	2,864.7	−14.87	2002-08-19 21	2002-08-19 21	0	17.72
20021030	3,226.5	3,077.5	−4.62	2002-10-30 08	2002-10-30 06	−3	5.45
20030420	2,196.2	2,229.6	1.52	2003-04-20 16	2003-04-20 17	1	10.16
20030513	1,650.9	1,762.7	6.77	2003-05-13 22	2003-05-13 23	1	**20.21**
20030515	4,490.6	4,464.0	−0.59	2003-05-15 22	2003-05-15 21	−1	16.74
20030607	3,171.1	2,892.9	−8.77	2003-06-07 04	2003-06-07 00	− **4**	17.16
20030629	1,079.3	905.7	−16.08	2003-06-29 03	2003-06-29 01	−2	7.32
20040508	1,267.9	1,266.9	−0.08	2004-05-08 07	2004-05-08 07	0	**26.02**
20040513	1,783.0	1,555.4	−12.77	2004-05-13 04	2004-05-13 03	−1	−11.11
20040517	2,879.7	2,707.9	−5.97	2004-05-16 18	2004-05-16 17	−1	18.52
20040531	1,101.9	1,132.3	2.76	2004-05-31 19	2004-05-31 17	−2	16.87
20040616	2,710.3	3,249.6	19.90	2004-06-16 21	2004-06-16 21	0	**37.09**
20040712	2,364.2	1,973.5	−16.52	2004-07-12 17	2004-07-12 16	−1	9.31

Notes: The total number of floods is **36**, which are qualificatory relative to the error of peak discharge, is **31** and the ratio of qualifying simulation is **86.11%**, which are qualificatory relative to the error of peak time, is **33** and the ratio of qualifying simulation is **91.67%** which are qualificatory relative to the error of total runoff volume, is **31** and the ratio of qualifying simulation is **86.11%**.

modeling residuals in view of model and measurement uncertainty, and the dark-green region illustrates the 95% discharge prediction uncertainty associated only with the posterior distribution of the parameter estimates. (The colour versions of Figures 3 and 4 are available in the online version of the paper at http://www.iwaponline.com/jh/toc.htm.) Note that the total prediction uncertainty ranges (cyan region) bracket the observed discharge during most of 3 hr time step, indicating that the model and/or measurement uncertainty is considerable for the current model structure and the data used. Moreover, the hydrograph prediction uncertainty associated only with the posterior distribution of the parameter estimates (dark-green) is narrow and does not always bracket the observations, suggesting that the model structure or model input data may need further improvement.

Then, the 36 historical floods calibration data (2000–2004) with a 1 hr time step are applied to the calibration of the Xinanjiang model parameters in the Shuangpai Reservoir by using three different GOMs. Table 7 presents the posterior mean, SD, CV, and correlation structure induced between the parameters of the Xinanjiang model in the HPD region of the posterior probability distribution with SCEM-UA algorithm, and the parameter estimates obtained by GA, SCE-UA and the SCEM-UA with highest posterior probability are also shown in Table 7 for a 1 hr time step. From Table 7, it can be seen that the CV of the parameter D_m, I_m, are very high. This demonstrates distributions of these parameters are highly dispersed, i.e., there is no well-defined region in the sense of a compact region for the 1 hr time step calibration. It should be noted that, for the 1 hr time step, the absolute values of more correlation coefficients are higher than 0.5 in the parameter correlation matrix than for the 3 hr time step, indicating that the Xinanjiang model calibration with a 1 hr time step is a more rigorous optimization problem. According to the three statistical ratios of acceptable national criteria relative to the peak discharge, peak time and total runoff volume among the calibrated and validated historical flood events for flood forecasting in China, the statistics of detail performance of the calibrated parameters with the highest posterior probability using SCEM-UA for a 1 hr time step are listed in Tables 8 and 9.

Table 10 gives the performance statistics obtained by different optimization methods for 1 hr time steps during calibration and validation. From Table 10, using the SCEM-UA algorithm with the highest posterior probability calibrated parameters, the values of three criteria: $r_{peak\_discharge}$, r_{peak_time}, and $r_{runoff,}$ are 86.11, 91.67 and 86.11%, respectively, for the calibrated result, and 100,

Table 9 | Performances of validated parameter with highest posterior probability by SCEM-UA for 1 hr time step

Floods	Observed (m^3/s^{-1})	Simulated (m^3/s^{-1})	Percentage error (%)	Observed peak-time (yyyy-mm-dd hh)	Simulated peak-time (yyyy-mm-dd hh)	Error (number)	Total volume error (%)
20050215	1,775.5	1,900.3	7.03	2005-02-15 18	2005-02-15 18	0	13.88
20050420	627.4	581.7	−7.28	2005-04-20 06	2005-04-20 05	−1	**33.83**
20050506	1,041.5	897.4	−13.83	2005-05-06 08	2005-05-06 06	−2	13.03
20050527	1,552.4	1,436.7	−7.45	2005-05-27 22	2005-05-27 20	−2	12.48
20050606	1,579.1	1,532.7	−2.94	2005-06-06 04	2005-06-06 03	−1	18.35
20050622	2,776.0	2,796.6	0.74	2005-06-22 02	2005-06-22 01	−2	16.27
20060527	2,366.0	2,360.9	−0.22	2006-05-27 05	2006-05-27 05	−1	8.81
20060601	971.9	942.2	−3.05	2006-06-01 09	2006-06-01 09	−1	8.12
20060608	2,641.5	2,637.2	−0.16	2006-06-08 01	2006-06-08 01	0	15.12
20060615	2,329.1	2,133.1	−8.41	2006-06-15 02	2006-06-15 02	−1	18.09
20060716	5,020.2	4,295.5	−14.44	2006-07-16 05	2006-07-15 13	**16**	−4.03
20060805	1,877.4	1,567.3	−16.52	2006-08-05 01	2006-08-05 01	0	11.71

Notes: The total number of floods is **12**, which are qualificatory relative to the error of peak discharge, is **12** and the ratio of qualifying simulation is **100%**, which are qualificatory relative to the error of peak time, is **11** and the ratio of qualifying simulation is **91.67%**, which are qualificatory relative to the error of total runoff volume, is **11** and the ratio of qualifying simulation is **91.67%**.

Table 10 | Comparison results of the SCEM-UA, SCE-UA and GA algorithm for 1 hr time step

Algorithm	Ratio of qualifying of calibration (%)			Ratio of qualifying of validation (%)		
	Peak discharge	Peak time	Total runoff	Peak discharge	Peak time	Total runoff
GA	83.33	91.67	80.56	100	91.67	75
SCE-UA	86.11	91.67	83.33	100	91.67	83.33
SCEM-UA	86.11	91.67	86.11	100	91.67	91.67

91.67 and 91.67%, respectively, for the validated result. The results calibrated by GA are 83.33, 91.67 and 80.56%, respectively, and the validated results are 100, 91.67 and 75%, respectively. The results calibrated by SCE-UA are 86.11, 91.67 and 83.33%, respectively, and the validated results are 100, 91.67 and 83.33%, respectively. The qualifying ratios of peak time of the three methods are the same. In the calibration phase, the qualifying ratios of the peak discharge and runoff total volume by SCEM-UA calibration are 2.78 and 5.55% higher than those by GA calibration, respectively. During the validation phase, the qualifying ratio of the runoff total volume by SCEM-UA calibration is 16.67% more than that by GA calibration. Moreover, the qualifying ratio of the runoff total volume by SCEM-UA calibration is 2.78 and 8.43% more than those by SCE-UA during the calibration and validation periods, respectively. This demonstrates that using the SCEM-UA algorithm with the highest posterior probability calibrated parameters is able to obtain better simulation results than GA and SCE-UA for a 1 hr time step.

The performances of all simulated hydrographs of rainfall-runoff process with a 1 hr time step period by different calibration methods from 2000 to 2004 during the calibration, and from 2005 to 2006 during the validation are shown in Figures 5 and 6, respectively. In Figures 5 and 6, the cyan region shows model prediction uncertainty region

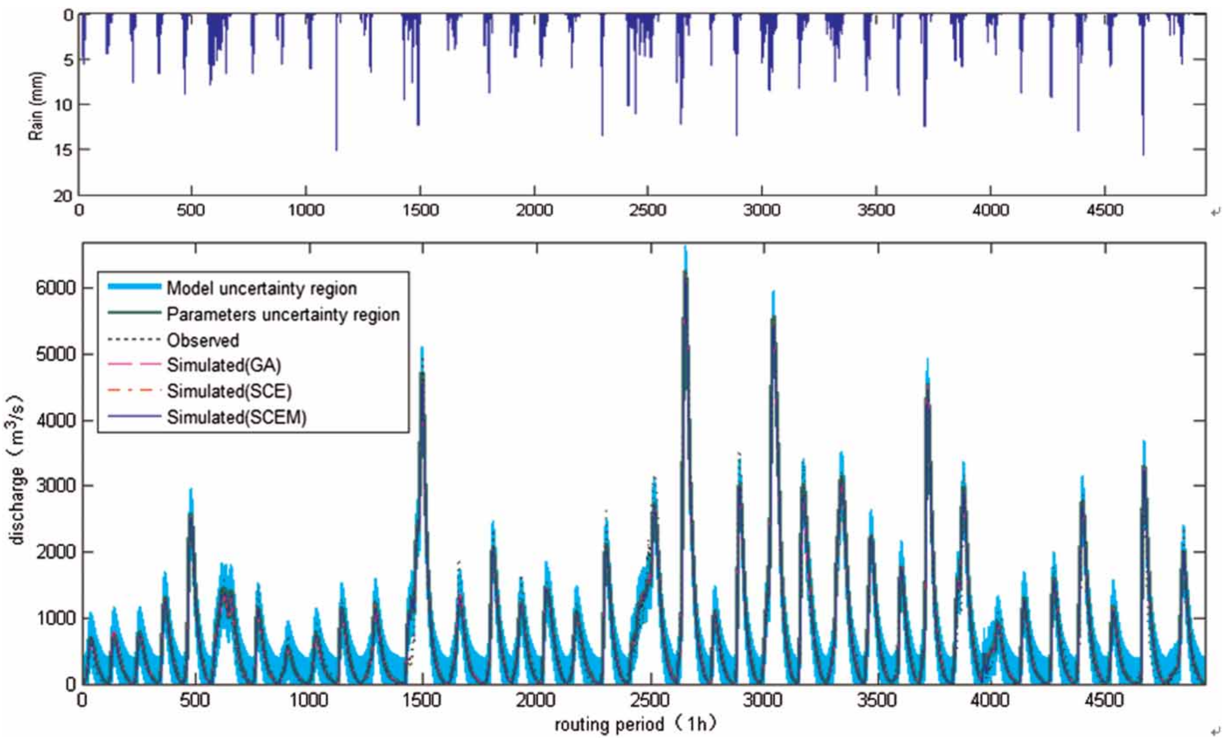

Figure 5 | The rain, model prediction uncertainty that results from model and measurement uncertainty, discharge prediction uncertainty associated with the most probable parameter set derived using the SCEM algorithm, observed hydrographs and simulated hydrographs (GA, SCE, SCEM) for 36 historical floods (2000–2004) with a 1 hr step during calibration. (The colour version of this figure is available in the online version of the paper at http://www.iwaponline.com/jh/toc.htm.)

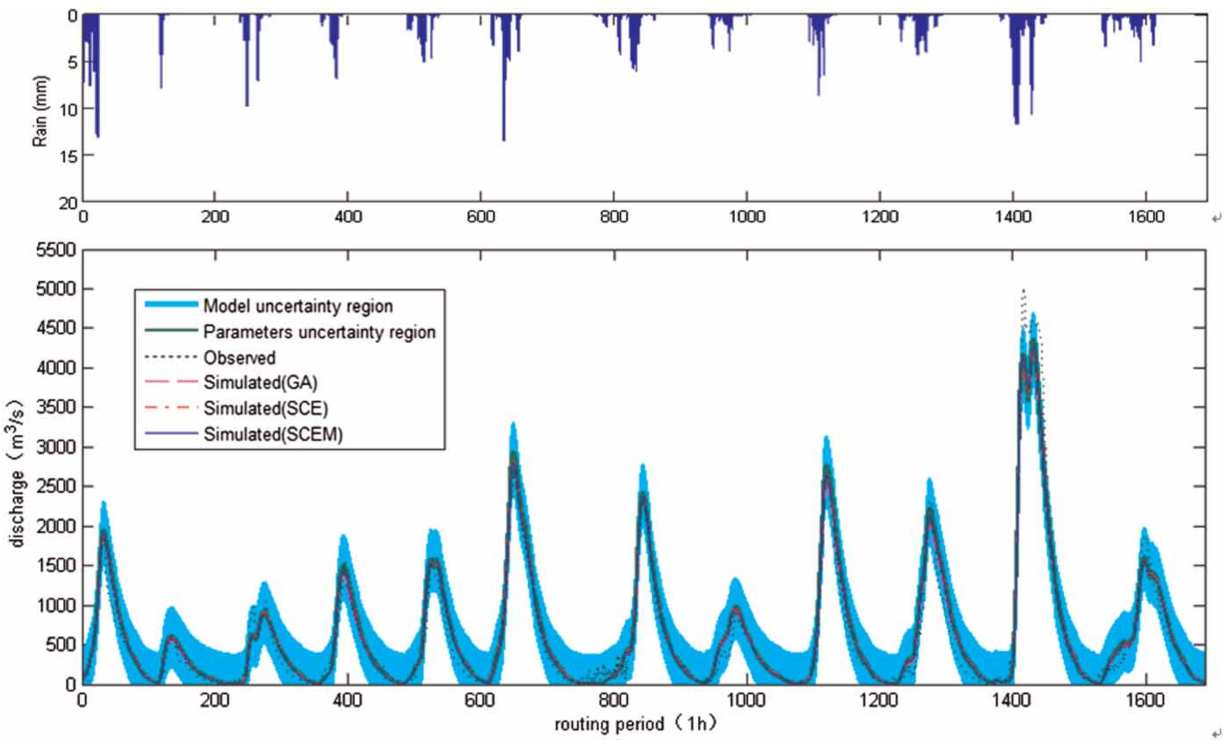

Figure 6 | The rain, model prediction uncertainty that results from model and measurement uncertainty, discharge prediction uncertainty associated with the most probable parameter set derived using the SCEM algorithm, observed hydrographs and simulated hydrographs (GA, SCE, SCEM) for 12 historical floods (2005–2006) with a 1 hr time step during validation. (The colour version of this figure is available in the online version of the paper at http://www.iwaponline.com/jh/toc.htm.)

associated with the 95% total error in terms of modeling residuals in view of model and measurement uncertainty, and the dark-green region illustrates the 95% discharge prediction uncertainty associated only with the posterior distribution of the parameter estimates. (The colour versions of Figures 5 and 6 are available in the online version of the paper at http://www.iwaponline.com/jh/toc.htm.) It should be noted that the total prediction uncertainty ranges (cyan region) bracket the observed discharge during almost the entire 1 hr time step, but are quite large for some periods, indicating that the model and/or measurement uncertainty is considerable for the current model structure and the data used. Moreover, the discharge prediction uncertainty associated only with the posterior distribution of the parameter estimates (dark-green) is narrow and does not always bracket the observations, indicating that the model structure or model input data may need further improvement. Comparing Figures 3 and 5, and Figures 4 and 6, it can be observed that the total model prediction uncertainty ranges with a 1 hr time step is larger than those with a 3 hr time step. This indicates that there is larger uncertainty for 1 hr forecasting than for 3 hr forecasting under the current model structure and field data.

CONCLUSIONS

The Xinanjiang model is a widely used CRR model for flood forecasting in China, as the demand for timely and accurate forecasts has increased. It has a large number of parameters which cannot be directly obtained from measurable quantities of catchments characteristics. Its successful application depends critically on how well the model is calibrated. In this study, an attempt is made to investigate three effective GOMs for calibration and analysis of the Xinanjiang model parameters with a 3 hr time step and a 1 hr time step. The optimization methods investigated include the GA technique, the SCE-UA method and the SCEM-UA algorithm. For the 3 hr time step, 34 historical floods in 12 years (1984–1995) are used for calibration while 11 historical floods in 2 years (1999–2000) are utilized for verification. For the 1 hr time step, 36 historical floods in 5 years (2000–2004) are used for calibration while 12 historical floods in the recent 2 years (2005–2006) are employed for verification. Three statistical ratios of acceptable criteria recommended by the national criteria in China for flood forecasting are employed to evaluate the performances of various optimization methods.

The results obtained in this study show that these methods are all powerful tools to model calibration, and SCEM-UA can give better performance than GA and SCE-UA. Moreover, the SCEM-UA provides both an estimate of the model with the highest posterior probability (the 'best' parameter set) and a sample set of parameter values describing the probabilistic representation of remaining parameter uncertainty. This posterior description of parameter uncertainty is then used to produce probabilistic model forecasts. Thus, the SCEM-UA algorithm not only correctly infers the most likely parameter set, but also generates useful information about the nature of the response surface in the vicinity of the optimum. The result comparison obtained from two different time step case applications indicate that there is greater uncertainty for 1 hr forecasting than for 3 hr forecasting under the current model structure and field data. This is significant for assessment of the risk in likely applications of hydrological models. The narrow discharge prediction uncertainty region with the most probable set obtained using the SCEM-UA algorithm did not always bracket the observed discharge, indicating that improvements in the model or input data may result in more accurate forecasting. It is hoped that future research efforts will focus in these directions.

ACKNOWLEDGEMENTS

This research was supported by the China National Funds for Distinguished Young Scientists (No. 51025934) and the Key Project of Science Technique Research of Henan Educational Committee (No. 12A570005). The authors are grateful for Dr Qingyun Duan and Dr Jasper A. Vrugt for developing code of SCE-UA and SCEM-UA, respectively. The authors would like to express their sincere gratitude to the three anonymous referees for their constructive comments and useful suggestions that helped us improve our paper.

REFERENCES

Abdulla, F. A., Lettenmaier, D. P. & Liang, X. 1999 Estimation of the ARNO model baseflow parameters using daily streamflow data. *Journal of Hydrology* **222**, 37–54.

Anctil, F., Lauzon, N., Andreassian, V., Oudin, L. & Perrin, C. 2006 Improvement of rainfall-runoff forecasts through mean areal rainfall optimization. *Journal of Hydrology* **328**, 717–725.

Awad, A. R. & Von Poser, I. 2010 Calibrating conceptual rainfall-runoff models using a real genetic algorithm combined with a local search method. Latest Trends on Computers, Corfu, Greece, pp. 174–181.

Bakhtyar, R. & Barry, D. A. 2009 Optimization of cascade stilling basins using GA and PSO approaches. *Journal of Hydroinformatics* **11**, 119–132.

Box, G. E. P. & Tiao, G. C. 1973 *Bayesian Inference in Statistical Analyses*. Addison-Wesley-Longman, Reading, MA.

Boyle, D. P., Gupta, H. V. & Sorooshian, S. 2000 Toward improved calibration of hydrologic models: combining the strengths of manual and automatic methods. *Water Resources Research* **36**, 3663–3674.

Cai, X. M. & Wang, D. B. 2006 Calibrating holistic water resources-economic models. *Journal of Water Resources Planning and Management-ASCE* **132**, 414–423.

Cheng, C. T. & Chau, K. W. 2004 Flood control management system for reservoirs. *Environmental Modelling and Software* **19**, 1141–1150.

Cheng, C. T., Chau, K. W., Li, X. Y. & Li, G. 2004 Developing flood forecasting system for reservoirs with J2EE. *Hydrological Sciences Journal* **49**, 973–986.

Cheng, C. T., Ou, C. P. & Chau, K. W. 2002 Combining a fuzzy optimal model with a genetic algorithm to solve multi-objective rainfall-runoff model calibration. *Journal of Hydrology* **268**, 72–86.

Cheng, C. T., Wang, W. C., Xu, D. M. & Chau, K. W. 2008 Optimizing hydropower reservoir operation using hybrid genetic algorithm and chaos. *Water Resources Management* **22**, 895–909.

Cheng, C. T., Wu, X. Y. & Chau, K. W. 2005 Multiple criteria rainfall-runoff model calibration using a parallel genetic algorithm in a cluster of computers. *Hydrological Sciences Journal* **50**, 1069–1087.

Cheng, C. T., Zhao, M. Y., Chau, K. W. & Wu, X. Y. 2006 Using genetic algorithm and TOPSIS for Xinanjiang model calibration with a single procedure. *Journal of Hydrology* **316**, 129–140.

Cooper, V. A., Nguyen, V. T. V. & Nicell, J. A. 1997 Evaluation of global optimization methods for conceptual rainfall-runoff model calibration. *Water Science and Technology* **36** (5), 53–60.

Cooper, V. A., Nguyen, V. T. V. & Nicell, J. A. 2007 Calibration of conceptual rainfall-runoff models using global optimization methods with hydrologic process-based parameter constraints. *Journal of Hydrology* **334**, 455–466.

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

Duan, Q. Y., Sorooshian, S. & Gupta, V. K. 1994 Optimal use of the SCE-UA global optimization method for calibrating watershed models. *Journal of Hydrology* **158**, 265–284.

Eckhardt, K. & Arnold, J. G. 2001 Automatic calibration of a distributed catchment model. *Journal of Hydrology* **251**, 103–109.

Fazal, M. A., Imaizumi, M., Ishida, S., Kawachi, T. & Tsuchihara, T. 2005 Estimating groundwater recharge using the SMAR conceptual model calibrated by genetic algorithm. *Journal of Hydrology* **303**, 56–78.

Feyen, L., Vrugt, J. A., Nuallain, B. O., van der Knijff, J. & De Roo, A. 2007 Parameter optimisation and uncertainty assessment for large-scale streamflow simulation with the LISFLOOD model. *Journal of Hydrology* **332**, 276–289.

Franchini, M. 1996 Use of a genetic algorithm combined with a local search method for the automatic calibration of conceptual rainfall-runoff models. *Hydrological Sciences Journal* **41**, 21–39.

Franchini, M. & Galeati, G. 1997 Comparing several genetic algorithm schemes for the calibration of conceptual rainfall-runoff models. *Hydrological Sciences Journal* **42**, 357–379.

Franchini, M., Galeati, G. & Berra, S. 1998 Global optimization techniques for the calibration of conceptual rainfall-runoff models. *Hydrological Sciences Journal* **43**, 443–458.

Gan, T. Y. & Biftu, G. F. 1996 Automatic calibration of conceptual rainfall-runoff models: optimization algorithms, catchment conditions, and model structure. *Water Resources Research* **32**, 3513–3524.

Gelman, A. & Rubin, D. B. 1992 Inference from iterative simulation using multiple sequences. *Statistical Science* **7**, 457–472.

Goldberg, D. E. 1989 *Genetic Algorithm in Search, Optimization and Machine Learning*. Addison-Wesley Publishing Co., Inc., Reading, MA.

Goswami, M. & O'Connor, K. M. 2007 Comparative assessment of six automatic optimization techniques for calibration of a conceptual rainfall-runoff model. *Hydrological Sciences Journal* **52**, 432–449.

Hastings, W. K. 1970 Monte Carlo sampling methods using Markov Chains and their applications. *Biometrika* **57**, 97–109.

Hejazi, M. I., Cai, X. M. & Borah, D. K. 2008 Calibrating a watershed simulation model involving human interference: an application of multi-objective genetic algorithms. *Journal of Hydroinformatics* **10**, 97–111.

Holland, J. H. 1975 *Adaptation in Natural and Artificial Systems*. University of Michigan Press, Ann Arbor, MI.

Holland, J. H. 1992 *Adaptation in Natural and Artificial Systems*. MIT Press Cambridge, MA.

Kim, T., Heo, J. H., Bae, D. H. & Kim, J. H. 2008 Single-reservoir operating rules for a year using multiobjective genetic algorithm. *Journal of Hydroinformatics* **10**, 163–179.

Kuczera, G. 1997 Efficient subspace probabilistic parameter optimization for catchment models. *Water Resources Research* **33**, 177–185.

Liong, S. Y., Phoon, K. K., Pasha, M. F. K. & Doan, C. D. 2005 Efficient implementation of inverse approach for forecasting hydrological time series using micro GA. *Journal of Hydroinformatics* **7**, 151–163.

Luce, C. H. & Cundy, T. W. 1994 Parameter identification for a runoff model for forest roads. *Water Resources Research* **30**, 1057–1069.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N. & Teller, A. H. 1953 Equations of state calculations by fast computing machines. *Journal of Chemical Physics* **21**, 1087–1091.

Ndiritu, J. G. & Daniell, T. M. 2001 An improved genetic algorithm for rainfall-runoff model calibration and function optimization. *Mathematical and Computer Modelling* **33**, 695–706.

Nelder, J. A. & Mead, R. 1965 A simplex method for function minimization. *Computer Journal* **7**, 308–313.

NCHI 2000 *The National Criteria for Hydrological Forecasting*. National Center of Hydrological Information, Hydroelectric Press, Beijing.

Sharifi, S., Sterling, M. & Knight, D. W. 2009 A novel application of a multi-objective evolutionary algorithm in open channel flow modelling. *Journal of Hydroinformatics* **11**, 31–50.

Sorooshian, S. & Gupta, V. K. 1995 Model calibration. In: *Computer Models of Watershed Hydrology*. Water Resources Publications, Highlands Ranch, Colorado, 23–68.

Sorooshian, S., Duan, Q. Y. & Gupta, V. K. 1993 Calibration of rainfall-runoff models: application of global optimization to the Sacramento soil moisture accounting model. *Water Resources Research* **29**, 1185–1194.

Thyer, M., Kuczera, G. & Bates, B. C. 1999 Probabilistic optimization for conceptual rainfall-runoff models: a comparison of the shuffled complex evolution and simulated annealing algorithms. *Water Resources Research* **35**, 767–773.

van Griensven, A. & Bauwens, W. 2003 Multiobjective autocalibration for semidistributed water quality models. *Water Resources Research* **39**, 1348.

Vrugt, J. A., Dekker, S. C. & Bouten, W. 2003a Identification of rainfall interception model parameters from measurements of throughfall and forest canopy storage. *Water Resources Research* **39**, 1251.

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

Vrugt, J. A., Gupta, H. V., Dekker, S. C., Sorooshian, S., Wagener, T. & Bouten, W. 2006 Application of stochastic parameter optimization to the Sacramento Soil Moisture Accounting Model. *Journal of Hydrology* **325**, 288–307.

Vrugt, J. A. & Robinson, B. A. 2007 Improved evolutionary optimization from genetically adaptive multimethod search. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 708–711.

Vrugt, J. A., Schoups, G., Hopmans, J. W., Young, C., Wallender, W. W., Harter, T. & Bouten, W. 2004 Inverse modeling of large-scale spatially distributed vadose zone properties using global optimization. *Water Resources Research* **40**, W06503.

Wang, Q. J. 1991 The genetic algorithm and its application to calibrating conceptual rainfall-runoff models. *Water Resources Research* **27**, 2467–2471.

Wang, Q. J. 1997 Using genetic algorithms to optimize model parameters. *Environmental Modelling & Software* **12**, 27–34.

Wang, W. C., Cheng, C. T., Chau, K. W. & Xu, D. M. 2012 Calibration of Xinanjiang model parameters using hybrid genetic algorithm based fuzzy optimal model. *Journal of Hydroinformatics* **14**, 784–799.

Wang, Y. C., Yu, P. S. & Yang, T. C. 2010 Comparison of genetic algorithms and shuffled complex evolution approach for calibrating distributed rainfall-runoff model. *Hydrological Processes* **24**, 1015–1026.

Yapo, P. O., Gupta, H. V. & Sorooshian, S. 1996 Automatic calibration of conceptual rainfall-runoff models: Sensitivity to calibration data. *Journal of Hydrology* **181**, 23–48.

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

Zhao, R. J., Zhang, Y. L. & Fang, L. R. 1980 The Xinanjiang model. Paper presented at Hydrological Forecasting Proceeding Oxford Symposium, IASH-AISH Publ. no. 129, Washington, DC, pp. 351–356.

First received 17 March 2012; accepted in revised form 24 June 2012. Available online 21 September 2012