



Multi-metric calibration of hydrological model to capture overall flow regimes



Yongyong Zhang^{a,*}, Quanxi Shao^b, Shifeng Zhang^{a,*}, Xiaoyan Zhai^{c,d}, Dunxian She^e

^a Key Laboratory of Water Cycle and Related Land Surface Processes, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China

^b CSIRO Computational Informatics, Leeuwin Centre, 65 Brockway Road, Floreat Park, WA 6014, Australia

^c China Institute of Water Resources and Hydropower Research, Beijing 100038, China

^d Research Center on Flood and Drought Disaster Reduction of the Ministry of Water Resources, Beijing 100038, China

^e State Key Laboratory of Water Resources and Hydropower Engineering Science, Wuhan University, Wuhan 430072, China

ARTICLE INFO

Article history:

Received 9 July 2015

Received in revised form 14 April 2016

Accepted 23 May 2016

Available online 1 June 2016

This manuscript was handled by Konstantine P. Georgakakos, Editor-in-Chief, with the assistance of Venkat Lakshmi, Associate Editor

Keywords:

Flow regime characteristics

Hydrological model

Multi-metric calibration

Improved SCE-UA

Hexi Corridor

SUMMARY

Flow regimes (e.g., magnitude, frequency, variation, duration, timing and rating of change) play a critical role in water supply and flood control, environmental processes, as well as biodiversity and life history patterns in the aquatic ecosystem. The traditional flow magnitude-oriented calibration of hydrological model was usually inadequate to well capture all the characteristics of observed flow regimes. In this study, we simulated multiple flow regime metrics simultaneously by coupling a distributed hydrological model with an equally weighted multi-objective optimization algorithm. Two headwater watersheds in the arid Hexi Corridor were selected for the case study. Sixteen metrics were selected as optimization objectives, which could represent the major characteristics of flow regimes. Model performance was compared with that of the single objective calibration. Results showed that most metrics were better simulated by the multi-objective approach than those of the single objective calibration, especially the low and high flow magnitudes, frequency and variation, duration, maximum flow timing and rating. However, the model performance of middle flow magnitude was not significantly improved because this metric was usually well captured by single objective calibration. The timing of minimum flow was poorly predicted by both the multi-metric and single calibrations due to the uncertainties in model structure and input data. The sensitive parameter values of the hydrological model changed remarkably and the simulated hydrological processes by the multi-metric calibration became more reliable, because more flow characteristics were considered. The study is expected to provide more detailed flow information by hydrological simulation for the integrated water resources management, and to improve the simulation performances of overall flow regimes.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Flow simulation has always been one of the hot topics in applied hydrology for water resources management for over a century. It plays a vital role in the design and operation of water resource projects, water supply schemes, water resources planning, flood mitigations and drought control, etc. (Ghumman et al., 2011). Numerous water resource indicators (WRIs) were proposed to depict flow components, such as average monthly, seasonal and annual flows, magnitude and timing of peak or low flows (Shrestha et al., 2013). Moreover, the prediction of extreme

events (floods and droughts) was taken more and more seriously because of their disastrous damages to society, economy and environment, especially in the arid and semi-arid regions (Smakhtin, 2001; Coulibaly et al., 2001; Held et al., 2005; Kumar et al., 2010). However, the simulation performances of WRIs were still far from satisfactory, particularly for the low flow events (Wenger et al., 2010; Staudinger et al., 2011; Pushpalatha et al., 2012; Shrestha et al., 2013). The critical reasons were that the most widely used goodness-of-fitness measures (e.g. mean squared errors, correlation coefficient, coefficient of efficiency) are sensitive to peak and low flows (Krause et al., 2005; Gupta et al., 2009; Pushpalatha et al., 2012) and that certain uncertainties exist in model inputs and structures (Beven, 2006; Staudinger et al., 2011; Najafi et al., 2011).

* Corresponding authors.

E-mail addresses: zhangyy003@igsnr.ac.cn (Y. Zhang), zhangsf@igsnr.ac.cn (S. Zhang).

Along with the rapid development of water related sciences (e.g. environmental hydrology, eco-hydrology), several flow components have been detected to have close relationships with physical, chemical, environmental and biological processes of aquatic systems (Carlisle et al., 2010). For example, natural flow paradigm has implications for several key processes and life history of aquatic organisms including growth, breeding, spawning, migration, recruitment and mortality (Poff et al., 1997; Poff and Zimmerman, 2010b; Bunn and Arthington, 2002; Arthington et al., 2010). Similarly, the simulation performance of low flow events significantly affects the prediction errors of water quality concentrations (van Griensven et al., 2006). Thus, the term “flow regime” was put forward (Langbein and Iseri, 1960), and the ecologically relevant metrics (ERMs) system was gradually formed to describe overall flow paradigm using magnitude, frequency, variability, timing, duration and rating of flow events (Richter et al., 1996; Olden and Poff, 2003; Poff et al., 1997, 2010a; Knight et al., 2008, 2012; Kennard et al., 2010). It is necessary and urgent to enrich WRIs to predict the entire flow regimes, which will provide hydrological foundations for the integrated water resources management, particularly for flood control, river environment improvement and restoration, and water project regulations.

Several existing studies on flow regime simulation have been reported using statistical models or hydrological models. For the statistical model approach, Rajurkar et al. (2002) applied the ANN methodology to model daily flows during monsoon flood events for a large catchment of the Narmada River in Madhya Pradesh, India. Knight et al. (2012) identified several predictive equations for 19 ecologically relevant streamflow characteristics with independent variables (climate, landscape features, regional indicators and land use) using step-backward regression. However, all of the existing statistical models are specific to the study area, and lack robust physical mechanisms to account for hydrological changes induced by human interferences and climate change (Wenger et al., 2010; Knight et al., 2012). Hydrological model is a robust and widely-accepted alternative which can represent critical hydrological processes using physical mechanism equations at different spatial and temporal scales (Wenger et al., 2010; Shrestha et al., 2013). Kennen et al. (2008) integrated TOPMODEL and multiple linear regression models to predict 78 flow variables at 856 sites in New Jersey, U.S., but the results were limited to the input of average daily discharges and withdrawals. Wenger et al. (2010) and Shrestha et al. (2013) applied VIC model to simulate seven ERMs in the Pacific Northwest United States, six WRIs and 32 ERMs in two headwater sub-basins in the Fraser River, Canada, respectively. Zhang et al. (2012) derived 80 hydrological metrics from monthly regulated and unregulated flow series simulated by SWAT in the upper and middle stream of the Huai River Basin. However, all the metric values in the existing studies were calculated from the well calibrated hydrographs using a single evaluation criterion (e.g. mean squared errors, coefficient of efficiency). Not all of the flow regime metrics were well captured, such as low flow magnitude, duration and rating of flow pulses (Shrestha et al., 2013), and frequency of high and low flows (Wenger et al., 2010).

Multi-Objective Optimization (MOO) is an efficient solution to improve the accuracy of overall flow regime simulation by calibrating different flow metrics simultaneously. Note that the flow regimes relate to many hydrological processes, such as flow yield process which directly affects the whole flow regime characteristics; flow routing and storage processes which probably involve the variability, timing and duration of flow events. In term of model calibration, considering the flow regime characteristics rather than the single flow magnitude could relieve the problems caused by parameter uncertainties and equifinality, because the detailed information of an observed hydrograph is efficiently used.

As a result, more reliable hydrological processes would be captured and the model performance would be improved by MOO. However, for current hydrological model calibration, MMO usually focused on the calibrations of different evaluation criteria (Duan et al., 1992), peak and low flow magnitude (Madsen, 2000) or different observed series from different gauges (Bekele and Nicklow, 2007). The major common algorithms were weighted sum approaches (Madsen, 2000; van Griensven and Bauwens, 2003; Kim and De Weck, 2006) and Pareto multi objective optimization algorithms (Das and Dennis, 1998; Deb et al., 2002; Khu and Madsen, 2005). Moreover, the existing studies about Pareto multi-objective optimization algorithms were confined to two or three conflicting objectives, because the current algorithms were difficult to overcome the curse of dimensionality, and were impossible to obtain the reasonable Pareto fronts in practice along with the increasing of objective number (Madsen, 2000; van Griensven and Bauwens, 2003). Therefore, the weighted sum approach would be an effective alternative to implement the current multi-metric calibration.

The objective of this study was to implement the hydrological model coupling with the equally weighted MOO approach to well capture the overall flow regime characteristics. Two headwater watersheds in the arid Hexi Corridor were selected for the case study. The reasonable agreements between observations and simulations of flow regime metrics were obtained simultaneously by considering the multi-objective functions (including flow magnitude, frequency and variation, duration, timing and rating of change). Moreover, model performances were validated by making a comparison with the single evaluation criterion. The study would provide more flow information by hydrological simulation for the integrated water resources management, and promote the further application of ERMs to capture more reliable hydrological processes of mathematics models.

2. Material and methods

2.1. Study area

Hexi Corridor, as the most representative arid and semi-arid region of the Northwest China, is selected as the study area (Fig. 1). Hexi Corridor is located in the northwest of the Yellow River Basin, and includes a long and narrow passage stretching for over 1000 km from east (the steep Wushaoling hillside) to west (the Yumen Pass), and 100–200 km from south (the Qilian Mountain) to north (the North Mountain). The corridor is the most important route from the North China to the Central Asia for traders and militaries, a critical part of the historic Silk Road, and a famous granary in the Northwest China. The region has a continental arid climate with drought weather, dramatic temperature variability and highly frequent sandstorm. The differences in precipitation and dryness are quite remarkable from east to west.

There are three river basins (i.e., Shiyang River Basin, Hei River Basin and Shule River Basin) in Hexi Corridor bounded by the Hei, Kuantai and Dahuang Mountains, respectively. All of these rivers originate from snowmelt and precipitation of the Qilian Mountains, and provide continuous flows for downstream oases and farmlands. However, most rivers disappear in the Gobi desert after flowing out of mountains due to infiltration and irrigation consumption, except in the mainstreams which reach to rump lakes, i.e., Qingtu Lake for Shiyang River, Juyanhai Lake for Hei River and Luobupo Lake for Shule River which is out of Hexi Corridor.

The current studies of hydrological simulation in the arid Hexi Corridor always focused on small catchments in the upper and middle regions because of data sparsity and complicated hydrological mechanism in the downstream region (Wang et al.,

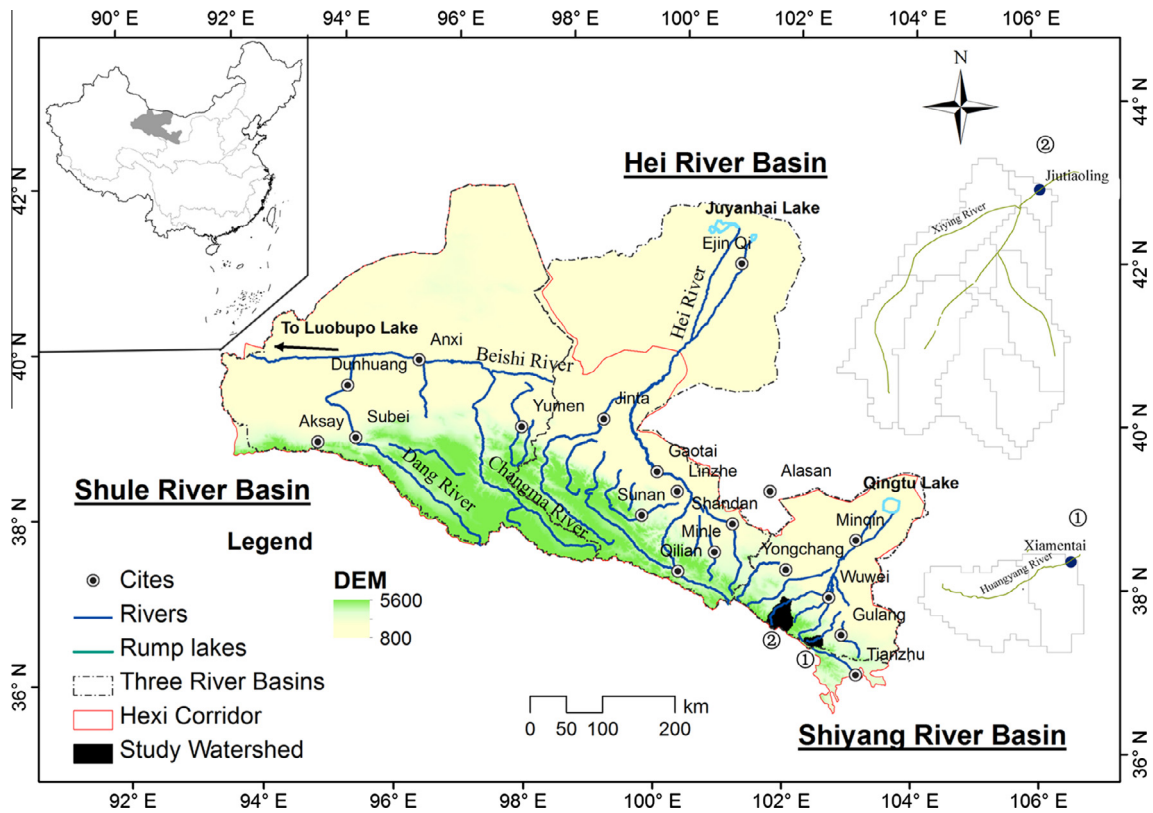


Fig. 1. The location of study area and two representative watersheds.

Table 1
The flow regime metrics selected to describe overall flow paradigm.

No	Groups	Flow regimes	Hydrologic metrics	Abbreviation	Units	Time scale
1	Magnitude	Middle flow	Middle flow (25th–75th percentile)	MMDF	m ³ /s	Daily
2		Low flow	Mean annual minimum flow	MinF	log-m ³ /s	Annual
3			Low flow discharge (75th percentile)	Low75	log-m ³ /s	Daily
4	Frequency and variation	High flow	Mean annual maximum flow	MaxF	m ³ /s	Annual
5			High flow discharge(25th percentile)	Hig25	m ³ /s	Daily
6		Average flow	CV of daily flow	CVDF	–	Annual
7	Duration	Low flow	Low flow spell count (<75th percentile)	LowC75	–	Annual
8		High flow	High flow spell count (>25th percentile)	HigC25	–	Annual
9		Low flow	Low flow spell duration (<75th percentile)	LowS75	Days	Annual
10	Timing	Low flow	Number of zero-flow days	ZeroN	Days	Annual
11		High flow	High flow spell duration (>25th percentile)	HigS25	Days	Annual
12		Low flow	Julian day of annual minimum	JMinF	–	Annual
13	Rating	High flow	Julian day of annual maximum	JMaxF	–	Annual
14		Low flow	Number of positive changes in flow from one day to the next	RLF	–	Annual
15		High flow	Number of negative changes in flow from one day to the next	RHF	–	Annual
16		Reversals	Number of negative and positive changes in flow from one day to the next	NFLH	–	Annual

2008, 2012; Yang et al., 2015). Therefore, in this study, two head-water watersheds in the Shiyang River Basin were also selected as the case study areas. One was Xiamentai watershed in the Huangyang River. The total drainage area was 273 km² with the dominant land covers of forest (47.3%) and unused land (33.2%). The other was Jiutiaoling watershed in the Xiyang River. Its drainage area was 1120 km² with the dominant land covers of grassland (40.5%) and unused land (39.4%).

2.2. Distributed water system model (HEQM)

HEQM is a semi-distributed and well-integrated model that couples multi-scale water related processes including hydrology, biogeochemistry, environment and ecology, as well as human

interferences at basin scale. The model provides a scientific tool to solve severe water crises faced globally (Zhang et al., 2016). Specifically, **H** indicates Hydrological submodel (i.e., hydrological cycle module), **E** is used to indicate Ecological submodel (i.e., soil biochemical module and crop growth module) and **Q** indicates water Quality submodel (i.e., soil erosion module, overland and instream water quality modules). Additionally, dam regulation module and parameter analysis tool were developed for human interferences evaluation and model calibration, respectively.

The Time Variant Gain Model (TVGM) was applied in HEQM to calculate surface water yield (Xia, 1991). TVGM performance was satisfactory especially in the arid and semi-arid regions (Xia et al., 2005; Wang et al., 2009; Li et al., 2010). Potential evapotranspiration was calculated using Hargreaves method (Hargreaves

Table 2
The sensitive parameters of HEQM detected by LH-OAT.

Type	Name	Min	Max	Definition	Relative importance (%)	
Distributed parameters	WM	0.45	0.75	Saturation moisture capacity of soil	24.12	
	WM_c	0.20	0.45	Field capacity of soil layer	22.02	
	KET_p	0.01	3.00	Adjustment factor of evapotranspiration	18.49	
	g_2	0.01	3.00	Influence coefficient of soil moisture	17.17	
	g_1	0.01	3.00	Basic surface flow coefficient	10.88	
	k_r	0.00	1.00	Interflow yield coefficient	1.58	
	T_g	0	365	Delay time for aquifer recharge	0.84	
	f_c	0	1000	Steady state infiltration rate	0.75	
	kf_{mx}	1.00	2.00	Ratio of maximum saturated soil infiltration rate to steady infiltration rate	0.73	
	k_{rg}	0.00	1.00	Base flow yield coefficient	0.45	
	Lumped parameters	SMF_{MX}	0.00	100.00	Melt factor for snow on June 21	0.20
		n	0.02	0.15	Coefficient of roughness in the riverbed	0.69
		SMT_B	0.46	2.00	Snow melt base temperature	0.47
SMF_{LF}		0.00	1.00	Snow pack temperature lag factor	0.34	
Total relative importance					98.73	

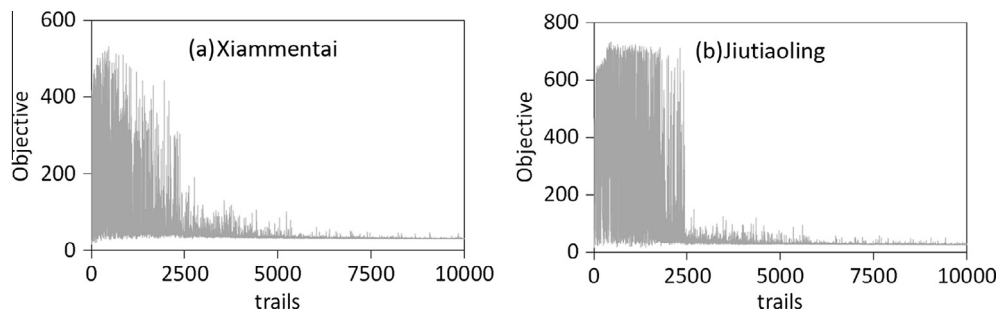


Fig. 2. Convergence of objective function value by multi-metric calibration at the Xiammentai (a) and Jiutiaoling (b) stations.

and Samani, 1982) and river flow routing used the Muskingum method.

Two and nine sub-basins were divided in Xiammentai and Jiutiaoling watersheds based on DEM and digital river systems, respectively. The area of sub-basins ranged from 37 km² to 198 km². The daily series of precipitation, minimum and maximum temperatures from 1990 to 2005 at nearby stations were interpolated to each sub-basin using the inverse distance weighting method. The parameter sensitivity analysis was performed in advance by Latin Hypercube One factor At a Time technique (LH-OAT) (van Griensven et al., 2006) in order to improve the efficiency of auto-calibration. The observed daily streamflow series at Xiammentai and Jiutiaoling stations were used for flow regime calibration. The calibration and validation periods were from 1990 to 1999, and from 2000 to 2005, respectively. The daily streamflows were simulated by the well-calibrated HEQM.

2.3. Flow regime metrics and evaluation criteria

The flow metrics were usually extracted from the daily or monthly streamflows (Poff, 1996; Kennard et al., 2010; Zhang et al., 2012, 2015a,b). In this study, the overall flow regime was divided into high, middle and low flow events based on the daily flow duration curve and was characterized by 16 flow metrics including magnitude (5), frequency and variation (3), duration (3), timing (2) and rating (3) of flow events (Richter et al., 1996; Olden and Poff, 2003; Poff et al., 1997, 2010a; Knight et al., 2008, 2012; Kennard et al., 2010) (Table 1). All the 16 observed and simulated metric values were extracted from the observed and simulated daily streamflow series in the entire period (1990–2005), respectively.

The root mean square error (*RMSE*) was selected as the primary criterion to evaluate model performance and the equation was given as

$$RMSE = \sqrt{\sum_{i=1}^N (O_i - S_i)^2 / N} \quad (1)$$

where O_i and S_i are the i th observed and simulated values, respectively; N is the length of series. The *RMSE* value ranges from zero to positive infinity. A smaller *RMSE* indicates a better simulation performance, and the optimal *RMSE* value is zero. Moreover, in order to overcome the insensitivity of low flow events, the *RMSE* values of low flow metrics (MinF and Low75) were calculated after a logarithmic transformation of O_i and S_i (Staudinger et al., 2011; Pushpalatha et al., 2012). Finally, *RMSE* values of 16 metrics were calculated for multi-objective calibration, while *RMSE* values of the flow magnitude was calculated for single objective calibration.

The bias (*bias*) was also used to measure the average deviation between the observed and simulated series after calibration.

$$bias = \frac{\sum (O_i - S_i)}{\sum O_i} \quad (2)$$

A positive value of *bias* indicates an underestimation, while a negative value represents an overestimation. The optimal value of *bias* is also zero.

2.4. Optimization algorithm

The multi-objectives of selected flow regime metrics were aggregated to a single objective by weighted average approach. The equation was given as

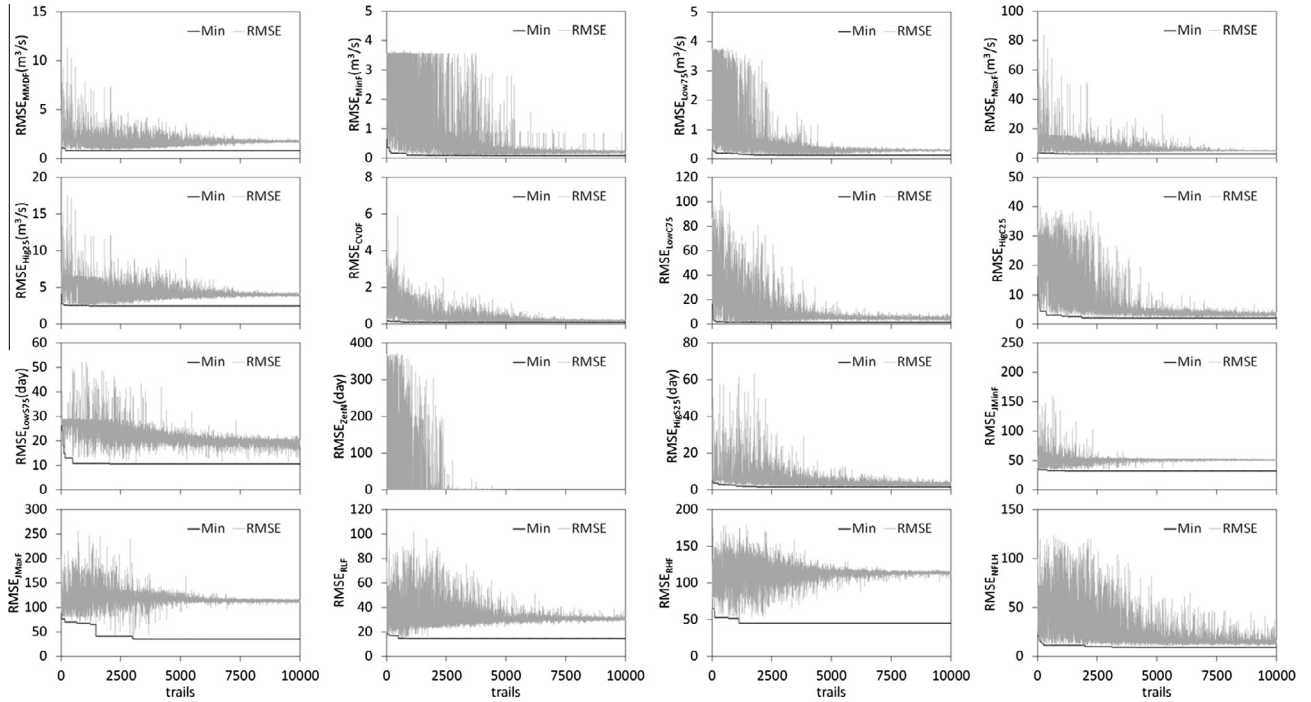


Fig. 3. Convergence of objective function values of individual metrics by multi-metric calibration at Xiamentai Station.

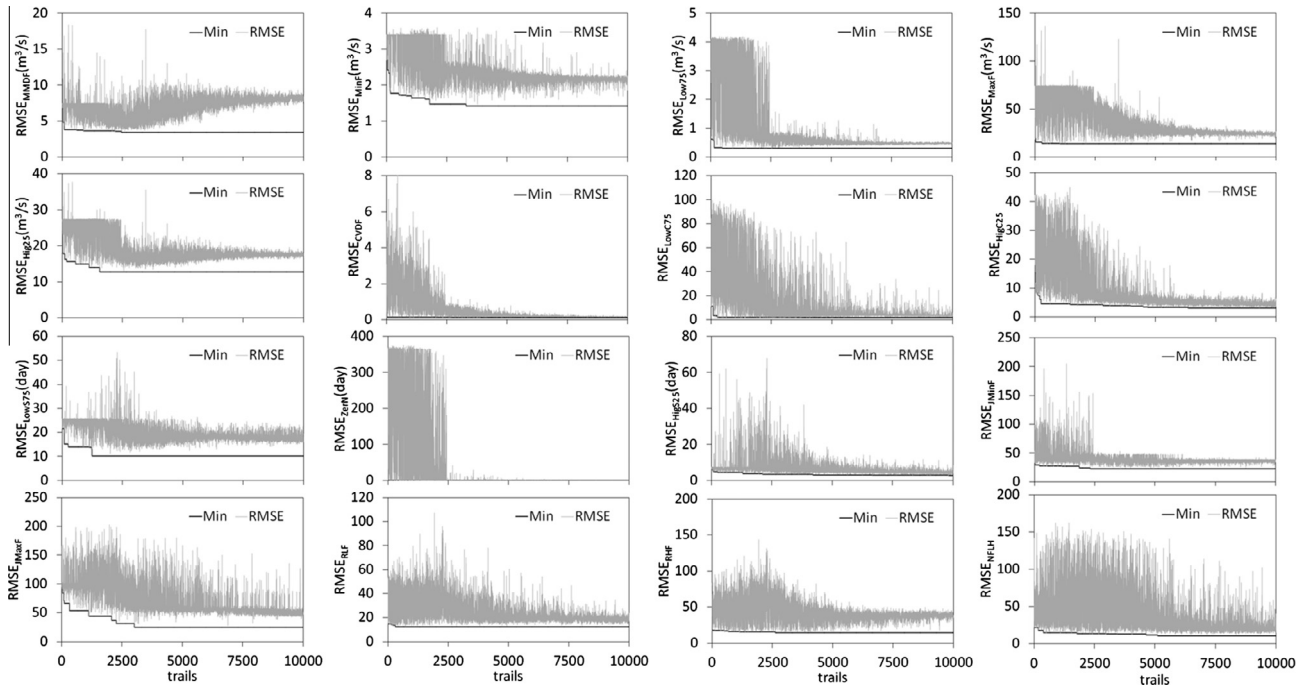


Fig. 4. Convergence of objective function values of individual metrics by multi-metric calibration at Jutiaoling Station.

$$\begin{aligned}
 f_{multi-ob} &= \min(RMSE_1, RMSE_2, \dots, RMSE_M) \\
 &= \min \sum_{i=1}^M (\omega_i \cdot RMSE_i / RMSE_{i,min})
 \end{aligned}
 \tag{3}$$

where $f_{multi-ob}$ is the multi-objective function; M is the number of objectives (i.e., 16 in this study); $RMSE_i$ and $RMSE_{i,min}$ are the root mean square error of metric i and its minimum value, respectively; ω_i is the weight of metric i . In this study, the weights of all the metrics are equal with the values of 1/16.

In practice, a normalization should be performed because different $RMSE$ values have different units in the flow metrics. In this study, we preferred to divide the $RMSE$ by the minimum value of each objective function following [van Griensven and Meixner \(2007\)](#). However, the minimum value of each objective function was initially unknown. Hypothetical minimum values were determined by the previous shuffling step in an optimization algorithm, and kept updating along with the iterations. The objective function tended to the optimal value quickly along the hypothetical values

Table 3
The evaluation results of multi-metric calibration.

Metrics	Stations	Stations							
		Xiamentai				Jiutiaoling			
		Calibration RMSE		$f_{\text{multi-ob}}$	Validation RMSE	Calibration RMSE		$f_{\text{multi-ob}}$	Validation RMSE
		Minimum	Optimum			Minimum	Optimum		
Magnitude (m^3/s)	MMDF	0.82	1.76	2.15	1.53	3.42	7.60	2.22	9.21
	MinF	0.08	0.21	2.63	0.22	1.42	1.84	1.30	2.44
	Low75	0.13	0.29	2.23	0.32	0.31	0.48	1.55	0.41
	MaxF	2.88	5.44	1.89	3.15	14.13	19.45	1.38	19.34
	Hig25	1.08	4.04	3.74	3.52	12.74	16.37	1.28	17.27
Frequency and variation	CVDF	0.11	0.14	1.27	0.24	0.12	0.12	1.00	0.22
	LowC75	1.61	2.60	1.61	5.52	1.87	2.81	1.50	2.61
	HigC25	2.10	3.54	1.69	3.22	3.06	4.27	1.40	5.32
Duration (days)	LowS75	10.58	16.26	1.54	19.46	10.24	19.17	1.87	9.83
	ZeroN	0.00	0.00	1.00	0.00	0.47	0.71	1.51	0.58
	HigS25	1.57	2.63	1.68	3.43	2.78	3.20	1.15	5.65
Timing	JMinF	35.4	121.24	3.42	92.26	24.86	44.31	1.78	59.57
	JMaxF	32.26	47.28	1.47	56.77	22.72	31.59	1.39	51.85
Rating	RLF	14.71	25.06	1.70	31.08	12.26	20.67	1.69	11.72
	RHF	45.18	106.63	2.36	135.12	14.36	44.78	3.12	27.46
	NFLH	9.30	12.18	1.31	15.30	10.75	21.39	1.99	11.88

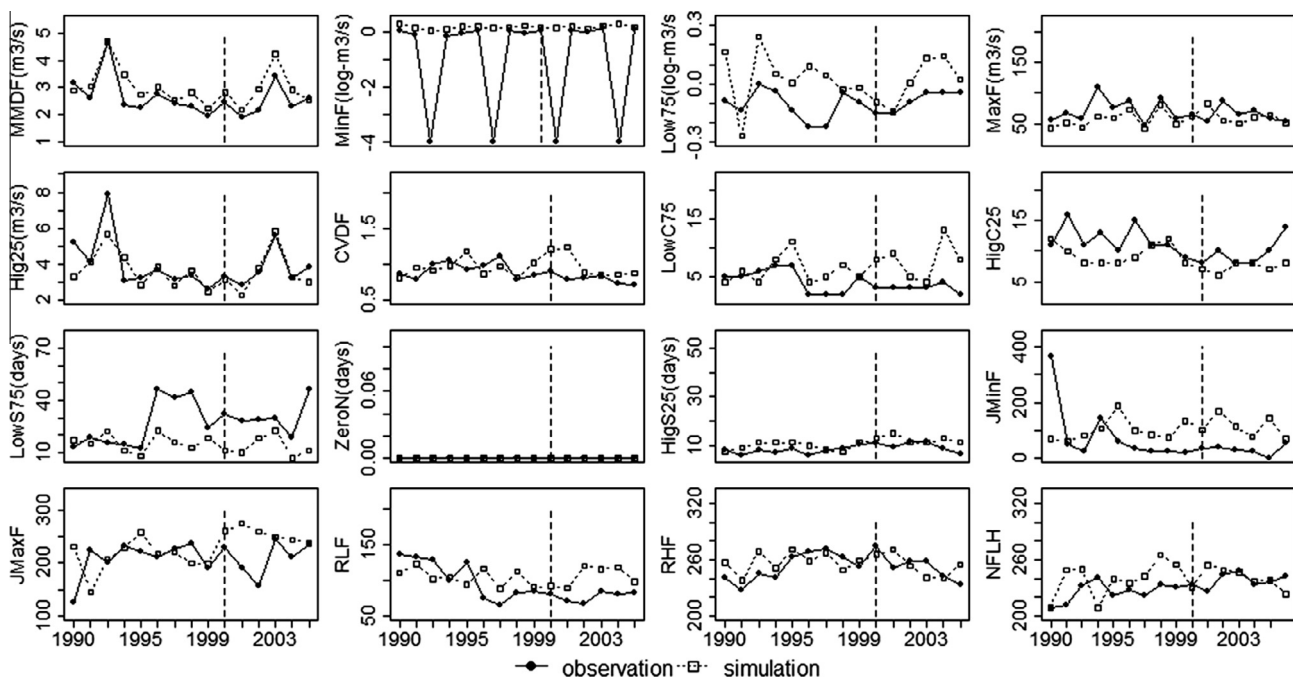


Fig. 5. The observed and simulated annual metrics by multi-metric calibration at Xiamentai Station. The dashed vertical line of each figure separates the calibration period from 1990 to 2000 (before) and validation period (behind) from 2001 to 2005.

converged to the final minima. This normalization approach enhanced the efficiency and effectiveness of the optimization algorithm (van Griensven and Meixner, 2007).

Numerous optimization algorithms were available (Arsenault et al., 2013), such as genetic algorithm (GA) (Holland, 1975), shuffled complex evolution (SCE-UA) (Duan et al., 1992), adaptive simulated annealing (ASA) (Ingber, 1993), particle swarm optimization (PSO) (Kennedy and Eberhart, 1995), covariance matrix adaptation evolution strategy (CMAES) (Hansen and Ostermeier, 1996) and dynamically dimensioned search (DDS) (Tolson and Shoemaker, 2007). A comparison of different optimization algorithms was given in Arsenault et al. (2013).

SCE-UA is the most often-used algorithm due mostly to open source and the fact that SCE-UA was the first algorithm aimed expressly at calibrating hydrological models (Duan et al., 1992; Sorooshian et al., 1993; Madsen, 2000; Eckhardt and Arnold, 2001; Ajami et al., 2004; Khakbaz et al., 2012). Several multi-objective algorithms have also been developed based on SCE-UA, such as MOCOM-UA (Yapo et al., 1998), MOSCEM-UA (Vrugt et al., 2003) and weighted SCE-UA (van Griensven and Meixner, 2007). In our study, the multi-metric calibration was carried out according to the framework of the improved SCE-UA (van Griensven and Bauwens, 2003; van Griensven and Meixner, 2007).

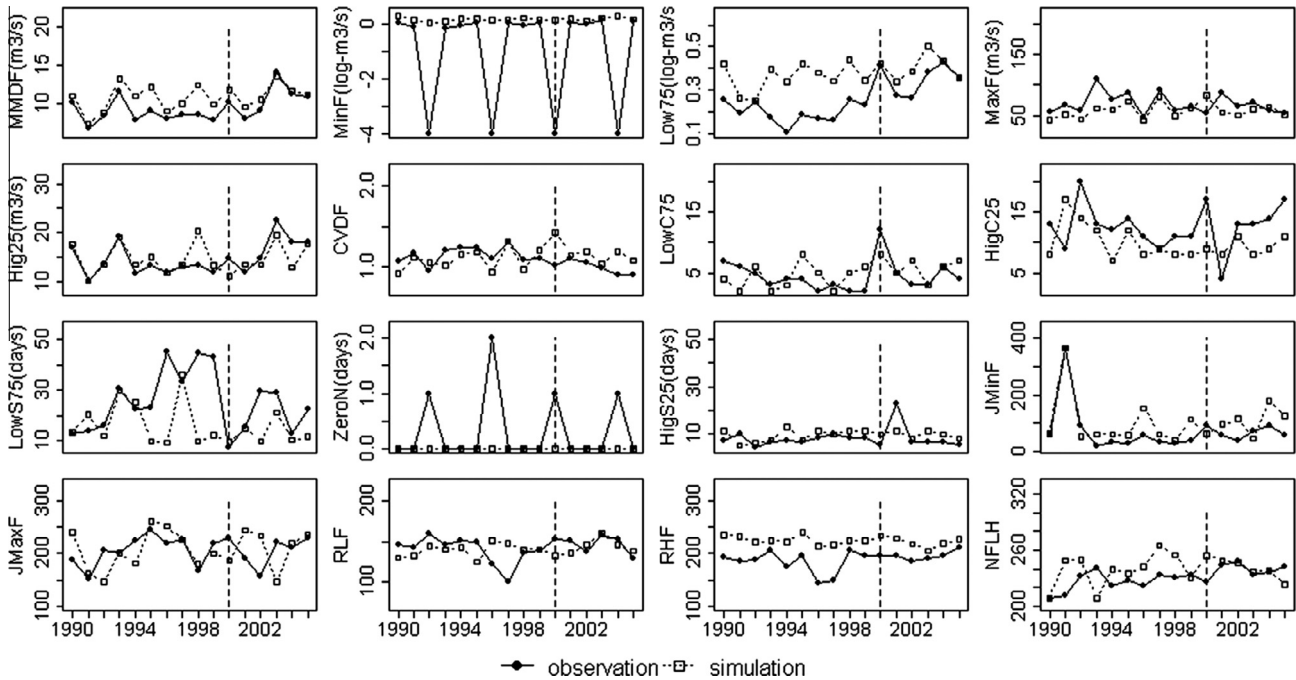


Fig. 6. The observed and simulated annual metrics by multi-metric calibration at Jiutiaoling Station. The dashed vertical line of each figure separates the calibration period from 1990 to 2000 (before) and validation period (behind) from 2001 to 2005.

Table 4
The evaluation results of single objective calibration.

Stations	Calibration period (1990–1999)		Validation period (2000–2005)	
	RMSE (m ³ /s)	bias	RMSE (m ³ /s)	bias
Xiamentai	1.73	0.01	1.15	-0.07
Jiutiaoling	6.60	-0.04	6.52	0.03

3. Results

3.1. Parameter sensitivity analysis

Fourteen sensitive parameters were detected and shown in Table 2. All of these parameters represented 98.73% of flow regime variation, and were categorized into 10 distributed parameters (97.03% of variation) for sub-basins and four lumped parameters (1.70% of variation) for the whole basin.

In the distributed parameters, the soil related parameters (WM , WM_c , k_r , f_c , kf_{mx}) were the most important parameter set explaining 49.20% of flow regime variation, followed by surface flow parameters (g_1 and g_2 : 28.05% of variation). The adjustment factor of evapotranspiration (KET_p : 18.49% of variation) and baseflow parameters (k_{rg} and T_g : 1.29% of variation) were at the third and last ranks, respectively. In the lumped parameters, the first rank was snowmelt related parameters (1.01% of variation) including SMF_{MX} , SMT_B and SMF_{LF} , and the second rank was runoff routing parameter (n : 0.69% of variation).

The explanations were that the soil water processes and evapotranspiration determined the main hydrological processes including surface flow, interflow and baseflow (Zhai et al., 2014). The surface flow usually occupied a large proportion of the total runoff, even in the arid and semi-arid regions (Yang et al., 2015). Moreover, the snowmelt and freezing were the dominant hydrological processes in the winter and early spring in our study area, which played important roles to recharge surface water in the channels and groundwater (Chen et al., 2003).

3.2. Multi-metric calibration

The values of $f_{multi-ob}$ were convergent after nearly 2500 iterations at both Xiamentai and Jiutiaoling stations (Fig. 2). For the individual metrics, both actual and minimum RMSE values of all the 16 metrics were also convergent after iterations at Xiamentai and Jiutiaoling stations (Figs. 3 and 4). However, the actual RMSE values of all the flow metrics were still difficult to converge to the minimum values (Table 3). The convergence patterns were also slightly different between these two stations. The RMSE values of MMDF, Hig25, RLF and RHF increased gradually from the minimum values to certain values after iterations, while the others converged to their minimum values. Therefore, the equally weighted method also captured the performance tradeoffs among different metrics. The slight degradation of model performances of MMDF, Hig25, RLF and RHF improved the model performances of other metrics. The convergence speeds of most metrics did not have great differences, except LowC75 and JMaxF at Jiutiaoling Station, and NFLH at Xiamentai Station. The probable reason was that the applied hydrological model was still disadvantageous to simulate the frequency of low flow events, the timing of maximum flow and the changes of flow events.

At Xiamentai Station, the simulated MaxF (magnitude), CVDF (frequency and variation), ZeroN (duration) and NFLF (rating) showed better agreements with the observations (Fig. 5). The simulated MinF (magnitude), JMinF (timing) and RHF (rating) poorly matched with the observations, as well as the daily series of MMDF, Low75 and Hig25. At Jiutiaoling Station, the simulated

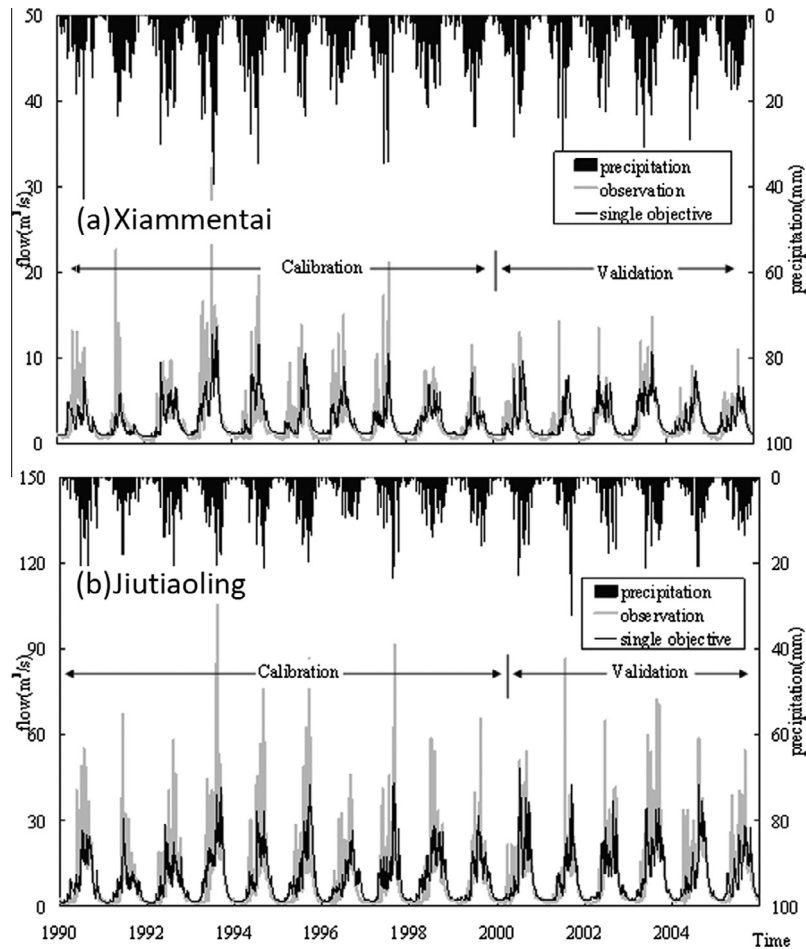


Fig. 7. The observed and simulated hydrographs by single objective calibration at the Xiamantai (a) and Jiutiaoling (b) stations.

Table 5

The comparisons between the single objective and multi-metric calibration.

Metrics	Stations	Stations											
		Xiamantai						Jiutiaoling					
		Single-objective		Multi-metric		Change (%)		Single-objective		Multi-metric		Change (%)	
		bias (%)	RMSE	bias (%)	RMSE	bias	RMSE	bias (%)	RMSE	bias (%)	RMSE	bias	RMSE
Magnitude (m^3/s)	M MDF	-13.7	1.02	-23.8	1.70	10.10	66.67	-28.0	4.39	-42.3	8.13	14.30	85.19
	MinF	97.8	0.45	43.9	0.17	-53.90	-62.22	120.5	2.23	113.3	2.15	-7.20	-3.59
	Low75	114.3	0.34	98.4	0.26	-15.90	-23.53	-305.0	0.45	-318.6	0.47	13.60	4.44
	MaxF	36.7	6.94	-2.8	5.03	-33.90	-27.52	50.5	40.12	19.9	24.13	-30.60	-39.86
	Hig25	15.8	2.66	1.8	4.04	-14.00	51.88	23.1	12.42	8.0	17.59	-15.10	41.63
Frequency and variation	CVDF	18.2	0.20	-9.5	0.18	-8.70	-10.00	23.1	0.28	-0.6	0.17	-22.50	-39.29
	LowC75	-25.4	2.75	-71.2	4.03	45.80	46.55	17.5	2.96	-11.2	2.72	-6.30	-8.11
	HigC25	47.9	6.43	21.2	3.42	-26.70	-46.81	49.3	7.42	19.1	4.45	-30.20	-40.03
Duration (days)	LowS75	21.8	21.34	45.9	17.61	24.10	-17.48	16.5	15.75	33.9	15.80	17.40	0.32
	ZeroN	0.0	0.00	0.0	0.00	0.00	0.00	100.0	0.67	100.0	0.67	0.00	0.00
	HigS25	-132.7	16.07	-23.0	2.98	-109.70	-81.46	-107.2	12.61	-16.1	4.06	-91.10	-67.80
Timing	JMinF	-60.7	106.61	-68.0	110.56	7.30	3.71	-1.3	83.18	0.2	50.73	-1.10	-39.01
	JMaxF	-6.1	45.19	-9.1	51.29	3.00	13.50	42.2	33.49	-40.0	38.59	-2.20	15.23
Rating	RLF	25.9	34.99	-13.0	27.62	-12.90	-21.06	39.2	57.54	0.9	16.99	-38.30	-70.47
	RHF	-102.2	152.96	-77.5	118.85	-24.70	-22.30	-48.5	92.83	-19.6	39.97	-28.90	-56.94
	NFLH	-12.8	35.89	-3.4	13.52	-9.40	-62.33	-17.0	41.75	-3.8	21.85	-13.20	-47.66

Note: the bias change was the absolute change of bias values between multi-metric and single-objective calibrations; the RMSE change was the relative changes between multi-metric and single-objective calibrations, i.e., the difference of RMSE values and divided by the RMSE of single-objective calibration.

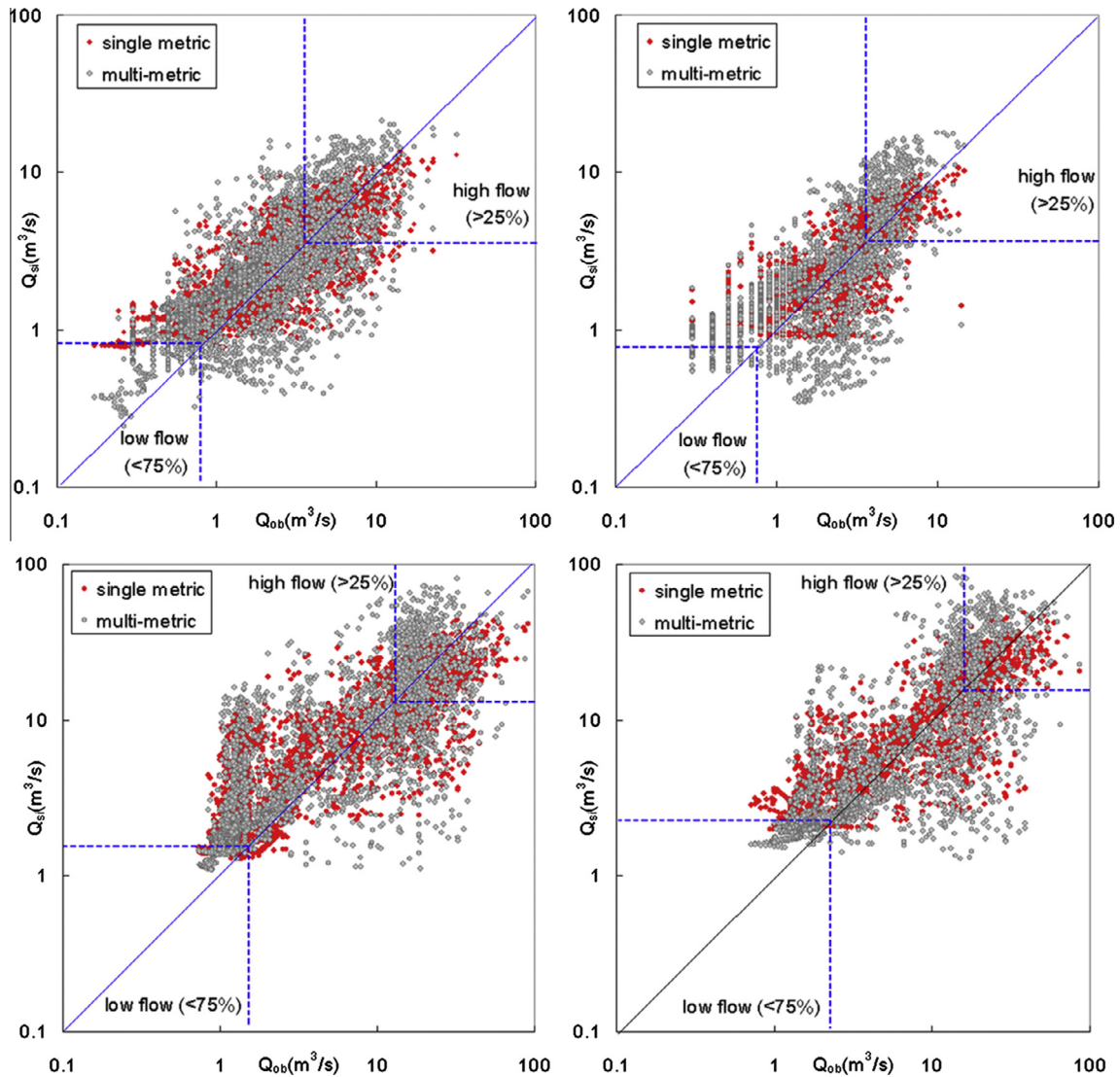


Fig. 8. Simulation versus observation of low, middle and high daily flow regimes in the calibration (left) and validation (right) period at the Xiamentai (upper) and Jiutiaoling (lower) stations. The results of different calibration methods are indicated by symbol colors: red dots for the simulation using single objective calibration, grey dots for the simulation using multi-metric calibration. The 1:1 line is shown as a solid line. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

MaxF (magnitude), CVDF (frequency and variation), HigS25 (duration), JMaxF (timing) at annual scale, and the simulated Hig25 (magnitude) at daily scale well matched with the observations (Fig. 6). The poorly simulated metrics were MMDF and MinF in the magnitude group and RHF in the rating group.

3.3. Single objective calibration

At Xiamentai Station, the optimal value of $RMSE$ was $1.73 \text{ m}^3/\text{s}$, and flow magnitude was slightly underestimated ($bias = 0.01$) in the calibration period, while $RMSE$ was $1.15 \text{ m}^3/\text{s}$ and flow magnitude was overestimated ($bias = -0.07$) in the validation period. At Jiutiaoling Station, the optimal values of $RMSE$ were $6.60 \text{ m}^3/\text{s}$ and $6.52 \text{ m}^3/\text{s}$ in the calibration and validation periods, respectively. The flow magnitude was slightly negative biased ($bias = -0.04$) in the calibration period, and slightly positive biased ($bias = 0.03$) in the validation period (Table 4). However, both the low and high flow events obviously disagreed with the observations (Fig. 7). The low flow events were overestimated while the

high flow events were underestimated, particularly for peak flow (see Table 5).

3.4. Improvement of flow regime simulation by multi-metric calibration

Compared with the traditional single objective calibration, the simulation performances for both the low and high flow events were improved by the multi-metric calibration, particularly for the peak flows, but the simulation performance for the middle flow was slightly worsened. The aggregated scatter plot between simulated and observed series tended to the 1:1 line for both the low and high flow events (Fig. 8). However, the deviation of the middle flow (MMDF) was more obvious by the multi-metric calibration than that of the single objective calibration.

Furthermore, in comparison with the single objective calibration, at Xiamentai Station, the $RMSE$ values in 11 out of 16 (68.8%) metrics were decreased by the multi-metric calibration with the relative reductions from 0.00 (ZeroN) to 81.46% (HigS25);

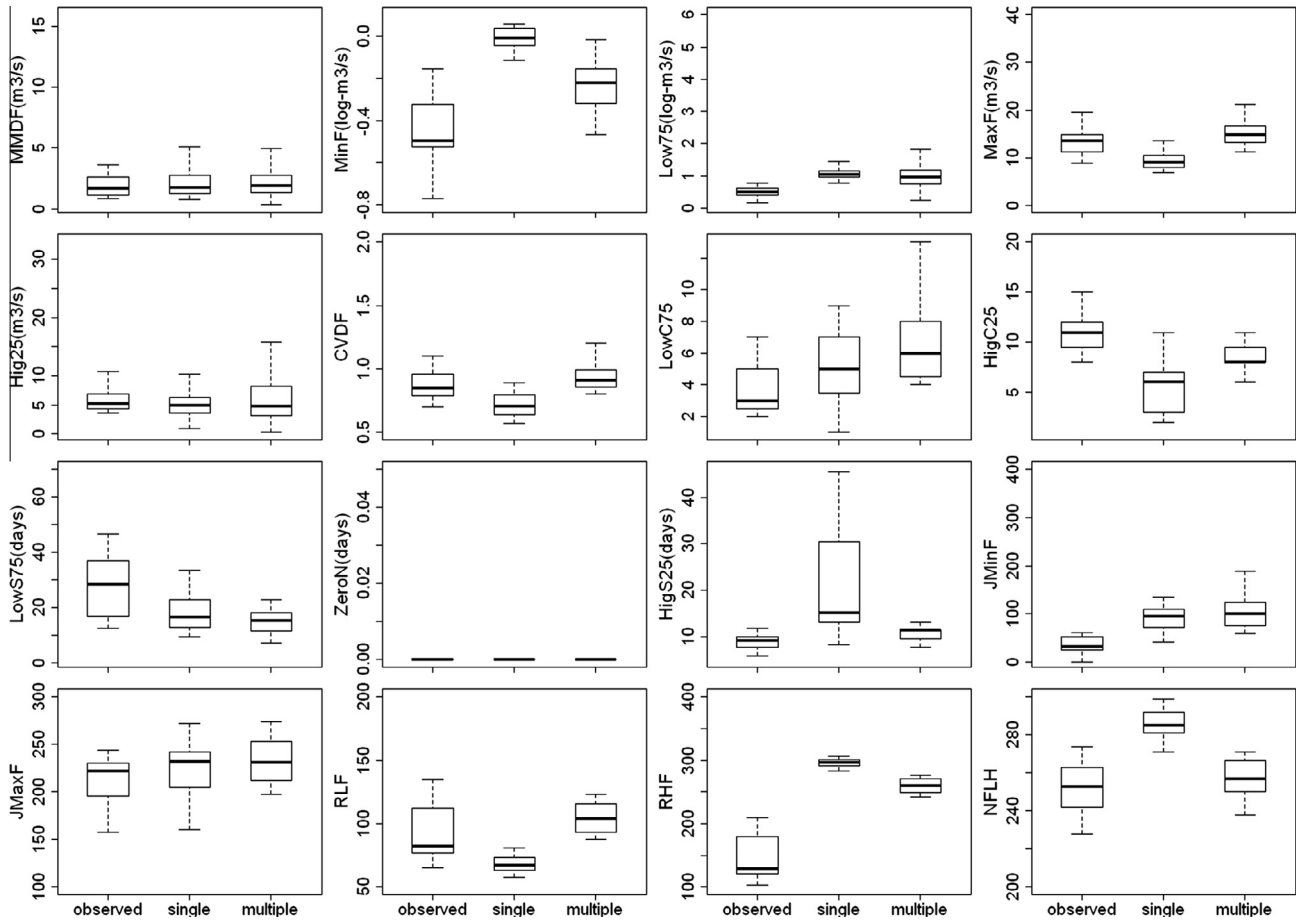


Fig. 9. Observations, simulated metrics by single objective and multi-metric calibration for Xiamentai Station. Each box plot illustrates the 25th, 50th and 75th percentile values, and the vertical bars (whiskers) define the 10th and 90th percentile value.

the bias values in 11 out of 16 (68.8%) metrics were also improved with the absolute reductions from 0.00 (ZeroN) to 109.70% (HigS25). At Jiutiaoling Station, the RMSE values in 10 out of 16 (62.5%) metrics decreased with the relative reductions from 0.00 (ZeroN) to 70.47% (RLF); the bias values in 13 out of 16 (81.3%) metrics were improved with the absolute reductions from 0.00 (ZeroN) to 91.10% (HigS25). In summary, the simulation performances for most metrics were improved by the multi-metric calibration and the most improved metrics for both RMSE and bias were the low and high flow magnitude, frequency and variation, duration and rating. However, the simulation performances of the other metrics (only 18.7–37.5% of all the metrics) were worsened at the Xiamentai and Jiutiaoling stations. The main metrics were MMDF and Hig25 at daily scale, both of which were usually well captured by the traditional single objective calibration. The prediction performances of these metrics were slightly degraded by the multi-metric calibration in order to get a better simulation of the overall flow regimes.

Figs. 9 and 10 showed the statistical distributions of all the metric values calculated from the observations, the simulations by the single objective and multi-metric calibrations, respectively. At Xiamentai Station, the values of 12 metrics by the multi-metric calibration distributed more closely to the observations than those by the single objective calibration, except Hig25 (magnitude), LowC75 (frequency and variation), JMinF (timing) and RLF (rating). The well-predicted metrics by the multi-metric calibration were MaxF (magnitude), CVDF (frequency and variation), ZeroN and

HigS25 (duration), and NFLH (rating). At Jiutiaoling Station, 14 metric values distributed more closely to the observations, except LowS75 (duration) and JMinF (timing), and the well predicted metrics were MaxF (magnitude), CVDF (frequency and variation), HigS25 (duration), RLF and NFLH (rating).

3.5. Parameter comparisons

For the selected parameters, the calibrated values by the multi-metric calibration were also compared with those by the single objective calibration (Fig. 11). In Xiamentai watershed, eight parameters values increased (f_c , k_{rg} , SMT_B , n , WM , SMF_{LF} , T_g and WM_C), while the other six parameter values decreased (KET_p , kf_{mx} , SMF_{MX} , g_2 , g_1 and k_r). In Jiutiaoling watershed, the increasing parameters were f_c , k_{rg} , g_1 , WM , SMT_B , T_g , SMF_{MX} , n and SMF_{LF} , while the decreasing parameters were KET_p , g_2 , k_r , kf_{mx} and WM_C . The most significantly increasing parameter was f_c with the relative change of 88.8% and 79.2% in the Xiamentai and Jiutiaoling watersheds, respectively. The most obviously decreasing parameter was k_r with the relative change of -39.7% and -42.5%, respectively.

As showed in Table 6, the soil related parameters (WM , WM_C , f_c , k_r and kf_{mx}) directly controlled the soil moisture processes and interflow. Specifically, WM , k_r and f_c determined the redistribution of high and low flow magnitudes. WM_C mainly controlled the frequency and variation, duration, timing and rating of flow events, and kf_{mx} influenced the overland flow routing

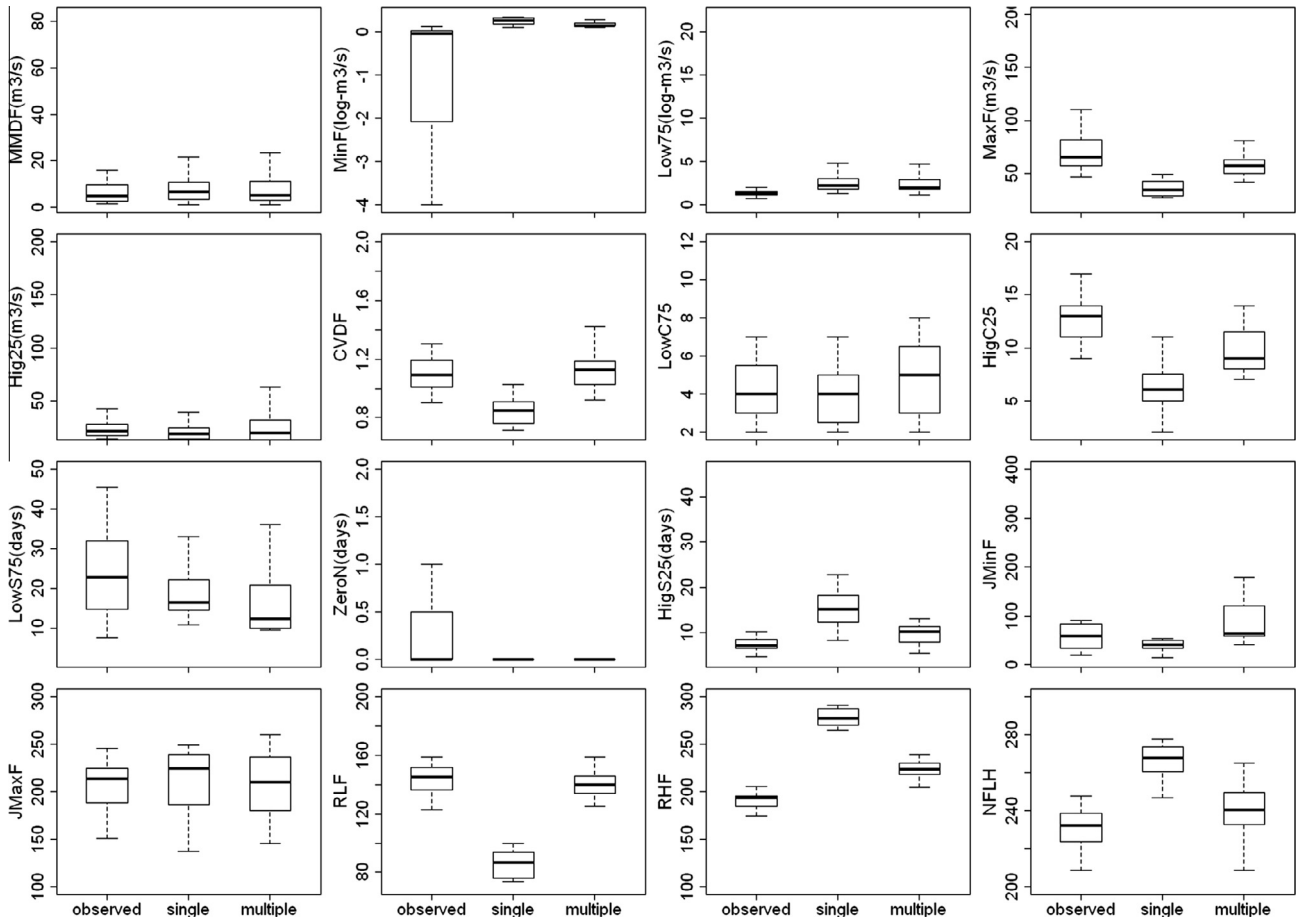


Fig. 10. Observations, simulated metrics by single objective and multi-metric calibration for Jiutiaoling Station. Each box plot illustrates the 25th, 50th and 75th percentile values, and the vertical bars (whiskers) define the 10th and 90th percentile value.

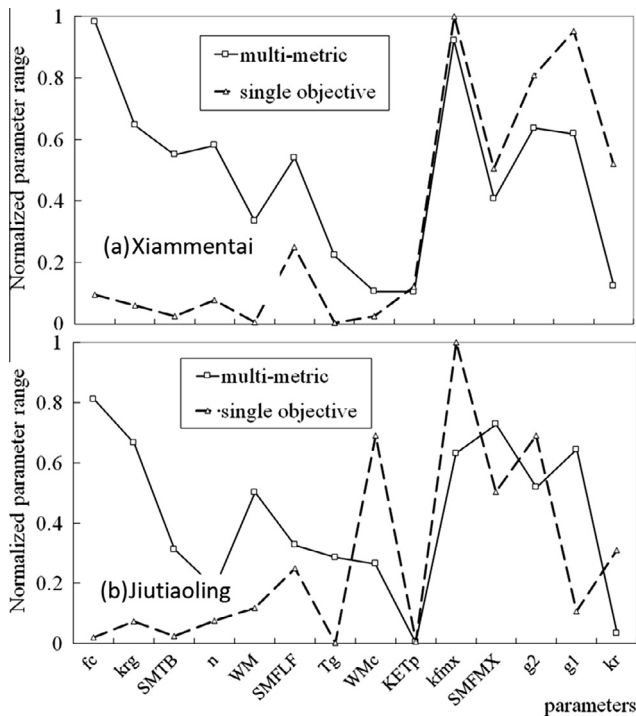


Fig. 11. The parameter variations comparing the multi-metric calibration with the single objective calibration for the Xiammentai (a) and Jiutiaoling (b) watersheds.

process. The surface water parameters (g_1 and g_2) and the evapo-transpiration parameter (KET_p) directly affected the magnitude and duration of flow events. The baseflow parameters (k_{rg} and T_g) were closely related with the magnitude, variation, duration and timing of flow events by controlling baseflow. The snowmelt related parameters (SMF_{MX} , SMT_B and SMF_{LF}) determined the magnitude, frequency and timing of flow events during winter and spring, although they had no significant impacts on the entire flow regime. The runoff routing parameter (n) slightly influenced the flow magnitude.

Therefore, the increase of parameter values (WM , f_c , k_{rg} , SMT_B , n , SMF_{LF} and T_g) raised the magnitude of low flow events (MinF and Low75), the frequency of flow events (LowC75 and HigC25) and the rating of positive changes of flow events (RLF), but reduced the flow variation (CVDF) and delayed the timing of high flow events (JMaxF). However, KET_p , g_2 , k_r and k_{fmX} had the opposite effects on these flow metrics. The variations of both flow metrics and parameters indicated the improvements of model performance by the multi-metric calibration. Moreover, the middle and high flow magnitudes (MIMDF, MaxF and High25) were negatively correlated with six parameters (WM , KET_p , g_2 , k_r , T_g and f_c). The values of parameters (KET_p , g_2 and k_r) decreased for both Xiammentai and Jiutiaoling watersheds. These flow magnitudes still increased because the decreased parameters (37.24% of variation) were more sensitive to the runoff simulation than the increased parameters (25.71% of variation).

Table 6

The relationship between sensitive parameters and flow regime metrics and their variations in the Xiamentai and Jiutiaoling watersheds.

No.	Metrics	Parameters	Parameters													
			WM	WM _c	KET _p	g ₂	g ₁	k _r	T _g	f _c	kf _{mx}	k _{rg}	SMF _{MX}	n	SMT _B	SMF _{LF}
1	Magnitude	M MDF	↓	↓	↓	↓	↑	↓	↓	↑	↑	↑	↓	↓	↑	↓
2		MinF	↑	↓	↓	↓	↓	↑	↓	↑	↑	↓	↓	↓	↓	↓
3		Low75	↑	↓	↓	↓	↓	↑	↑	↑	↑	↓	↓	↓	↓	↓
4		MaxF	↓	↑	↓	↓	↓	↓	↓	↓	↑	↑	↑	↑	↑	↑
5		Hig25	↓	↓	↓	↑	↓	↓	↓	↓	↑	↑	↑	↑	↑	↑
6	Frequency and Variation	CVDF	↓	↑	↑	↑	↑	↓	↓	↓	↓	↑	↑	↑	↑	↑
7		LowC75		↑												
8		HigC25		↑												
9	Duration	LowS75														
10		ZeroN		↑	↑	↑	↑	↓	↓	↓	↓	↑		↑	↑	
11		HigS25		↓	↓											
12	Timing	JMinF		↓											↓	
13		JMaxF								↑						
14	Rating	RLF		↑									↓			
15		RHF		↑												
16		NFLH		↑												
Stations	Overall flow regime	Xiamentai	+	+	–	–	–	–	+	+	–	+	–	+	+	+
		Jiutiaoling	+	–	–	–	+	–	+	+	–	+	+	+	+	+

Note: “↑” means that the flow metric value increases with the parameter, while the “↓” means that the flow metric value decreases with the parameter. “+” means the parameter value by the multi-metric calibration increases compared with that by the single objective calibration, while “–” means the corresponding parameter value decreases.

4. Discussion and conclusions

4.1. Multi-objective calibration approaches

The multi-objective calibration of hydrological model was more effective to capture the entire flow regime characteristics than the existing studies that derived flow metrics from well simulated hydrographs based on a single criterion. In the theory, the Pareto algorithm would be the best approach to solve the multi-metric calibration. However, it is quite difficult to obtain the reasonable Pareto fronts by the current generations of Pareto optimization algorithms if the number of objectives was great (van Griensven and Bauwens, 2003), particularly 16 objectives in this study (Yapo et al., 1998; Vrugt et al., 2003; Vrugt and Robinson, 2007; Bekele and Nicklow, 2007).

The equally weighted approach was advantageous and widely accepted to handle numerous objectives because it was not confined to the number of objectives and the correlations among different objectives, although the weight assignment was usually questioned (Efstratiadis and Koutsoyiannis, 2010). The results showed that most of the flow regime characteristics were well captured compared with the single objective calibration, including the low and high flow magnitudes, frequency and variation, duration, and rating. For the optimization algorithm, except for the widely-used SCE-UA, more efficient and more versatile algorithms (e.g., ASA, DDS and CMAES), should be adopted to calibrate hydrological models, particularly for the nonlinear or high parameter dimension models (Arsenault et al., 2013).

4.2. Simulation performance of flow regime metrics

The flow magnitude metrics were overestimated by the multi-metric calibration compared with the single objective calibration, and their corresponding RMSE values also had distinct discrepancies with the minimum optimal values, particularly the daily flow hydrographs. On the one hand, the model still had disadvantages to reproduce the low and middle flow magnitudes at daily scale like most existing models such as VIC (Wenger et al., 2010; Shrestha et al., 2013), landscape stratification models (Carlisle et al., 2010) and regional statistical models (Knight et al., 2012). On the other hand, it was a remarkable tradeoff among the high, middle and

low flow magnitude simulations. Specifically, the traditional single objective calibration usually focused on the average flow magnitude (Wenger et al., 2010; Pushpalatha et al., 2012; Shrestha et al., 2013). Thus, the slight loss of daily middle flow magnitude simulation performance improved the accuracy of annual high and low flow magnitudes.

The estimates of frequency, variation, and duration metrics showed a “goodness of fitness” with the observations by the multi-metric calibration, except LowC75 at Xiamentai Station and LowS75 at Jiutiaoling Station. For the timing metrics, JMaxF was well simulated, but JMinF was not. The explanation might be the poor simulation performance of low flow events, particularly at the headwater streams with frequent interchanges between surface water and groundwater, and violent transformations among glaciers, snow and water (Wang et al., 2008; Yang et al., 2015). For the rating metrics, RLF and NFLH were accurately predicted. Although the model predictions of RHF tended to be systematically higher than the observations, the inter-annual variability was captured. It might be reasonable to use HEQM to predict trends in negative changes in flow events (RHF). By the multi-metric calibration, the simulation accuracies of all these rating metrics were significantly improved, and the variations in flow events were well captured.

4.3. Flow regimes and water resources management

It is widely accepted that flow regimes play a critical role in flood control, domestic and production water supply (Bauwens and Vandewiele, 1989; Lehner et al., 2006), migration and transformation processes of nutrients (Poff et al., 1997; Ahmet et al., 2006; van Griensven and Meixner, 2007), variation of species habitat, as well as biodiversity and life history patterns in aquatic ecosystems (Poff et al., 1997; Bunn and Arthington, 2002; Knight et al., 2008; Arthington et al., 2010). Except for the traditional resource attribute of flow regimes, their environmental and ecological attributes were gradually paid more attention. The magnitude-oriented management of water resources was being transferred to the flow regime management (Poff et al., 1997; Arthington et al., 2010) and risk management of extreme events (Lehner et al., 2006).

It should also be more urgent to implement flow regime management in our study area due to the severe water related issues

(e.g., water shortage, ecological vulnerability, downstream oasis degradation). The prediction of flow magnitude would be determinant to investigate the quantity of water resources, while the predictions of flow variation, frequency, duration, timing and rating would be helpful to reasonably allocate the intra-annual distribution of limited water resources and increase water use efficiency. Moreover, the duration and timing of flow events would be beneficial to improve the ecological conditions along the rivers, inhibit the invasion of desert, as well as guarantee the agricultural productions for residents in the downstream oasis.

In summary, hydrological model coupling with multi-metric calibration algorithm can capture overall flow regime characteristics for the integrated river basin management. The overall flow regime simulation is expected to provide scientific foundations and critical linkages for the studies of environmental hydrology, eco-hydrology and so on. Nonetheless, further investigation of hydrological mechanism, model structure improvements and uncertainty analysis should be prioritized to enhance the simulation performance of the magnitude, timing and rating of flow events, especially for the low flow events (Smakhtin, 2001; Staudinger et al., 2011; Pushpalatha et al., 2012).

Acknowledgements

This study was supported by the National Basic Research Program of China (973 Program) (No. 2012CB955304), the Natural Science Foundation of China (No. 41271005), the China Youth Innovation Promotion Association CAS (No. 2014041), the Program for “Bingwei” Excellent Talents in IGSNRR CAS (No. 2015RC201), the China Visiting Scholar Project from China Scholarship Council, Australian Endeavour Research Fellowship and CSIRO Computational and Simulation Sciences Platform. Thanks to Dr. Jasper A. Vrugt, Dr. Kalyanmoy Deb and Dr. Gupta Hoshin Vijai for their suggestions of multi-objective optimization. The authors also thank to the editor and the anonymous reviewers for their useful comments that improved this manuscript.

References

- Ahmet, K., Kadri, Y., Cengiz, O., 2006. Effects of Kilickaya Dam on concentration and load values of water quality constituents in Kelkit Stream in Turkey. *J. Hydrol.* 317 (1–2), 17–30.
- Ajami, N.K., Gupta, H., Wagener, T., Sorooshian, S., 2004. Calibration of a semi-distributed hydrologic model for streamflow estimation along a river system. *J. Hydrol.* 298 (1), 112–135.
- Arsenault, R., Poulin, A., Côté, P., Brissette, F., 2013. A comparison of stochastic optimization algorithms in hydrological model calibration. *J. Hydrol. Eng.* [http://dx.doi.org/10.1061/\(ASCE\)HE.1943-5584.0000938](http://dx.doi.org/10.1061/(ASCE)HE.1943-5584.0000938).
- Arthington, A.H., Naiman, R.J., McClain, M.E., Nilson, C., 2010. Preserving the biodiversity and ecological services of rivers: new challenges and research opportunities. *Freshw. Biol.* 55, 1–16.
- Bauwens, W., Vandewiele, G.L., 1989. The real time runoff forecast models for the River Dijle. *Water Resour. Manage.* 3, 1–9.
- Bekele, E.G., Nicklow, J.W., 2007. Multi-objective automatic calibration of SWAT using NSGA-II. *J. Hydrol.* 341 (3), 165–176.
- Beven, K., 2006. A manifesto for the equifinality thesis. *J. Hydrol.* 320 (1), 18–36.
- Bunn, S.E., Arthington, A.H., 2002. Basic principles and ecological consequences of altered flow regimes for aquatic biodiversity. *Environ. Manage.* 30 (4), 492–507.
- Carlisle, D.M., Falcone, J., Wolock, D.M., Meador, M.R., Norris, R.H., 2010. Predicting the natural flow regime: models for assessing hydrological alteration in streams. *River Res. Appl.* 26 (2), 118–136.
- Chen, R.S., Kang, E.S., Yang, J.P., Zhang, J.S., 2003. A distributed runoff model for inland mountainous river basin of Northwest China. *J. Geogr. Sci.* 13 (3), 363–372.
- Coulialy, P., Bobée, B., Ancil, F., 2001. Improving extreme hydrologic events forecasting using a new criterion for artificial neural network selection. *Hydrol. Process.* 15 (8), 1533–1536.
- Das, I., Dennis, J.E., 1998. Normal-boundary intersection: a new method for generating the Pareto Surface in nonlinear multicriteria optimization problems. *SIAM J. Optim.* 8 (3), 631–657.
- Deb, K., Pratap, A., Agarwal, S., Meyarivan, T., 2002. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Trans. Evol. Comput.* 6 (2), 182–197.
- Duan, Q., Sorooshian, S., Gupta, V., 1992. Effective and efficient global optimization for conceptual rainfall-runoff models. *Water Resour. Res.* 28 (4), 1015–1031.
- Eckhardt, K., Arnold, J.G., 2001. Automatic calibration of a distributed catchment model. *J. Hydrol.* 251 (1–2), 103–109.
- Efstratiadis, A., Koutsoyiannis, D., 2010. One decade of multi-objective calibration approaches in hydrological modelling: a review. *Hydrol. Sci. J.* 55 (1), 58–78.
- Ghumman, A.R., Ghazaw Yousry, M., Sohail, A.R., Watanabe, K., 2011. Runoff forecasting by artificial neural network and conventional model. *Alexandria Eng. J.* 50, 345–350.
- Gupta, H.V., Kling, H., Yilmaz, K.K., Martinez, G.F., 2009. Decomposition of the mean squared error and NSE performance criteria: implications for improving hydrological modelling. *J. Hydrol.* 377 (1), 80–91.
- Hansen, N., Ostermeier, A., 1996. Adapting arbitrary normal mutation distributions in evolution strategies: the covariance matrix adaptation. In: *Proceeding, 1996 IEEE International Conference on Evolutionary Computation*. IEEE Press, NY, pp. 312–317.
- Hargreaves, G.H., Samani, Z.A., 1982. Estimating potential evapotranspiration. *J. Irrigat. Drain. Div.* 108 (3), 225–230.
- Held, I.M., Delworth, T.L., Lu, J., Findell, K.U., Knutson, T.R., 2005. Simulation of Sahel drought in the 20th and 21st centuries. *PNAS* 102 (50), 17891–17896.
- Holland, J., 1975. *Adaptation in Natural and Artificial Systems*. Univ. of Michigan Press, Oxford, England, pp. 1–183.
- Ingber, L., 1993. *Adaptive Simulated Annealing (ASA)*. Lester Ingber Research, McLean, VA.
- Kim, I.Y., De Weck, O.L., 2006. Adaptive weighted sum method for multiobjective optimization: a new method for Pareto front generation. *Struct. Multidiscip. Optimiz.* 31 (2), 105–116.
- Kennard, M.J., Mackay, S.J., Pusey, B.J., Olden, J.D., Marsh, N., 2010. Quantifying uncertainty in estimation of hydrologic metrics for ecohydrological studies. *River Res. Appl.* 26 (2), 137–156.
- Kennedy, J., Eberhart, R., 1995. Particle swarm optimization. *Proceeding, IEEE International Conference on Neural Networks*, vol. 4. IEEE Press, Piscataway, NJ, pp. 1942–1948.
- Kennen, J.G., Kauffman, L.J., Ayers, M.A., Wolock, D.M., Colarullo, S.J., 2008. Use of an integrated flow model to estimate ecologically relevant hydrologic characteristics at stream biomonitoring sites. *Ecol. Model.* 211, 57–76.
- Khakbaz, B., Imam, B., Hsu, K., Sorooshian, S., 2012. From lumped to distributed via semi-distributed: calibration strategies for semi-distributed hydrologic models. *J. Hydrol.* 418, 61–77.
- Knight, R.R., Gregory, M.B., Wales, A.K., 2008. Relating streamflow characteristics to specialized insectivores in the Tennessee River Valley: a regional approach. *Ecohydrology* 1, 394–407.
- Knight, R.R., Gain, W.S., Wolfe, W.J., 2012. Modelling ecological flow regime: an example from the Tennessee and Cumberland River basins. *Ecohydrology* 5 (5), 613–627.
- Krause, P., Boyle, D.P., Bäse, F., 2005. Comparison of different efficiency criteria for hydrological model assessment. *Adv. Geosci.* 5, 89–97.
- Kumar, R., Samaniego, L., Attinger, S., 2010. The effects of spatial discretization and model parameterization on the prediction of extreme runoff characteristics. *J. Hydrol.* 392 (1), 54–69.
- Khu, S.T., Madsen, H., 2005. Multiobjective calibration with Pareto preference ordering: an application to rainfall-runoff model calibration. *Water Resour. Res.* 41 (3).
- Langbein, W.B., Iseri, K.T., 1960. General introduction and hydrologic definitions. *Manual of Hydrology, Part 1. General Surface-Water Techniques: US Geological Survey Water-Supply Paper 1541-A*, 29.
- Lehner, B., Döll, P., Alcamo, J., Henrichs, T., Kaspar, F., 2006. Estimating the impact of global change on flood and drought risks in Europe: a continental, integrated analysis. *Climatic Change* 75 (3), 273–299.
- Li, L., Xia, J., Xu, C.Y., Singh, V.P., 2010. Evaluation of the subjective factors of the GLUE method and comparison with the formal Bayesian method in uncertainty assessment of hydrological models. *J. Hydrol.* 390 (3–4), 210–221.
- Madsen, H., 2000. Automatic calibration of a conceptual rainfall-runoff model using multiple objectives. *J. Hydrol.* 235 (3), 276–288.
- Najafi, M.R., Moradkhani, H., Jung, I.W., 2011. Assessing the uncertainties of hydrologic model selection in climate change impact studies. *Hydrol. Process.* 25 (18), 2814–2826.
- Olden, J.D., Poff, N.L., 2003. Redundancy and the choice of hydrologic indices for characterizing streamflow regimes. *River Res. Appl.* 19, 101–121.
- Poff, N.L., 1996. A hydrogeography of unregulated streams in the United States and an examination of scale-dependence in some hydrological descriptors. *Freshw. Biol.* 36, 71–91.
- Poff, N.L., Allan, J.D., Bain, M.B., Karr, J.R., Prestegard, K.L., Richter, B.D., Sparks, R.E., Stromberg, J.C., 1997. The natural flow regime: a paradigm for river conservation and restoration. *Biosciences* 47, 769–784.
- Poff, N.L., Richter, B.D., Arthington, A.H., Bunn, S.E., Naiman, R.J., Kendy, E., Acreman, M., Apse, C., Bledsoe, B.P., Freeman, M.C., Henriksen, J., Jacobson, R.B., Kennen, J.G., Merritt, D.M., O'keeffe, J.H., Olden, J.D., Rogers, K., Tharme, R.E., Warner, A., 2010a. The ecological limits of hydrologic alteration (ELOHA): a new framework for developing regional environmental flow standards. *Freshw. Biol.* 55, 147–170.
- Poff, N.L., Zimmerman, J.K.H., 2010b. Ecological responses to altered flow regimes: a literature review to inform environmental flows science and management. *Freshw. Biol.* 55, 194–205.
- Pushpalatha, R., Perrin, C., Le Moine, N., Andréassian, V., 2012. A review of efficiency criteria suitable for evaluating low-flow simulations. *J. Hydrol.* 420–421, 171–182.

- Rajurkar, M.P., Kothiyari, U.C., Chaube, U.C., 2002. Artificial neural networks for daily rainfall–runoff modelling. *Hydrol. Sci. J.* 47 (6), 865–877.
- Richter, B.D., Baumgartner, J.V., Powell, J., Braun, D.P., 1996. A method for assessing hydrologic alteration within ecosystems. *Conserv. Biol.* 10, 1–12.
- Shrestha, R.R., Daniel, L.P., Markus, A.S., 2013. Evaluating the ability of a hydrologic model to replicate hydro-ecologically relevant indicators. *Hydrol. Process.* <http://dx.doi.org/10.1002/hyp.9997>.
- Smakhtin, V.U., 2001. Low flow hydrology: a review. *J. Hydrol.* 240 (3), 147–186.
- Sorooshian, S., Duan, Q., Gupta, V.K., 1993. Calibration of conceptual rainfall-runoff models using global optimization: application to the Sacramento soil moisture accounting model. *Water Resour. Res.* 29 (4), 1185–1194.
- Staudinger, M., Stahl, K., Seibert, J., Clark, M.P., Tallaksen, L.M., 2011. Comparison of hydrological model structures based on recession and low flow simulations. *Hydrol. Earth Syst. Sci.* 15 (11), 3447–3459.
- Tolson, B.A., Shoemaker, C.A., 2007. Dynamically dimensioned search algorithm for computationally efficient watershed model calibration. *Water Resour. Res.* 43 (1), W01413.
- van Griensven, A., Bauwens, W., 2003. Multiobjective autocalibration for semidistributed water quality models. *Water Resour. Res.* 39 (12).
- van Griensven, A., Meixner, T., Grunwald, S., Bishop, T., Diluzio, M., Srinivasan, R., 2006. A global sensitivity analysis tool for the parameters of multi-variable catchment models. *J. Hydrol.* 324 (1), 10–23.
- van Griensven, A., Meixner, T., 2007. A global and efficient multi-objective auto-calibration and uncertainty estimation method for water quality catchment models. *J. Hydroinform.* 9 (4), 277–291.
- Vrugt, J.A., Gupta, H.V., Bastidas, L.A., Bouten, W., Sorooshian, S., 2003. Effective and efficient algorithm for multiobjective optimization of hydrologic models. *Water Resour. Res.* 39 (8).
- Vrugt, J.A., Robinson, B.A., 2007. Improved evolutionary optimization from genetically adaptive multimethod search. *PNAS* 104, 708–711.
- Wang, S., Kang, S., Zhang, L., Li, F., 2008. Modelling hydrological response to different land-use and climate change scenarios in the Zamu River basin of Northwest China. *Hydrol. Process.* 22 (14), 2502–2510.
- Wang, G., Xia, J., Chen, J., 2009. Quantification of effects of climate variations and human activities on runoff by a monthly water balance model: a case study of the Chaobai River basin in northern China. *Water Resour. Res.* 45, W00A11. <http://dx.doi.org/10.1029/2007WR006768>.
- Wang, Z.G., Ficklin, D.L., Zhang, Y.Y., Zhang, M.H., 2012. Impact of climate change on streamflow in the arid Shiyang River Basin of northwest China. *Hydrol. Process.* 26, 2733–2744.
- Wenger, S.J., Luce, C.H., Hamlet, A.F., Isaak, D.J., Neville, H.M., 2010. Macroscale hydrologic modeling of ecologically relevant flow metrics. *Water Resour. Res.* 46 (9).
- Xia, J., 1991. Identification of a constrained nonlinear hydrological system described by Volterra Functional Series. *Water Resour. Res.* 27 (9), 2415–2420.
- Xia, J., Wang, G.S., Tan, G., Ye, A.Z., Huang, G.H., 2005. Development of distributed time-variant gain model for nonlinear hydrological systems. *Sci. China: Earth Sci.* 48 (6), 713–723.
- Yang, D.W., Gao, B., Jiao, Y., Lei, H.M., Zhang, Y.L., Han, B.Y., Cong, Z.T., 2015. A distributed scheme developed for eco-hydrological modeling in the upper Heihe River. *Sci. China: Earth Sci.* 58, 36–45.
- Yapo, P.O., Gupta, H.V., Sorooshian, S., 1998. Multi-objective global optimization for hydrologic models. *J. Hydrol.* 204 (1), 83–97.
- Zhang, Y.Y., Arthington, A.H., Bunn, S.E., Mackay, S., Xia, J., Kennard, M., 2012. Classification of flow regimes for environmental flow assessment in regulated rivers: the Huai River Basin, China. *River Res. Appl.* 28, 989–1005.
- Zhang, Y.Y., Fu, G.B., Sun, B.Y., Zhang, S.F., Men, B.H., 2015a. Simulation and classification of the impacts of projected climate change on flow regimes in the arid Hexi Corridor of Northwest China. *J. Geophys. Res.* 120, 7429–7453.
- Zhang, Y.Y., Shao, Q.X., Ye, A.Z., Xing, H.T., Xia, J., 2016. Integrated water system simulation by considering hydrological and biogeochemical processes: model development, with parameter sensitivity and autocalibration. *Hydrol. Earth Syst. Sci.* 20, 529–553.
- Zhang, Y.Y., Zhai, X.Y., Shao, Q.X., Yan, Z.Q., 2015b. Assessing temporal and spatial flow regime alterations in the regulated Huai River Basin, China. *J. Hydrol.* 529, 384–397.
- Zhai, X.Y., Zhang, Y.Y., Wang, X.L., Xia, J., Liang, T., 2014. Non-point source pollution modeling using soil and water assessment tool and its parameter sensitivity analysis in Xin'anjiang catchment, China. *Hydrol. Process.* 28 (28), 1627–1640.