RESEARCH ARTICLE

Multi-objective synchronous calibration and Pareto optimality of runoff and sediment parameters in an arid and semi-arid watershed

Lei Wu[1,2,3,4] · Xia Liu[5] · Junlai Chen[1,4] · Xiaoyi Ma[1,4]

Abstract

Calibration methodologies must extract as much information as possible from available data, but it is not well understood in investigating the multi-objective synchronous calibration strategy by using multiple sources of information and by exploiting the data in better ways. The non-dominated sorting genetic algorithm II (NSGA-II) is introduced to study the calibration performance of runoff and sediment parameters under nine targeted scenarios, which considers the best choice to obtain high-cost performance results for decision makers through multi-objective optimization and the calculation of Pareto-optimal front with a high precision. (i) SWAT model has good adaptability in the runoff simulation of the Yanhe River watershed. Both the calibration results of NSGA-II and sequential uncertainty fitting approach-version 2 (SUFI-2) can meet the accuracy requirements of runoff simulation. Particularly, the NSGA-II based on multiple objective functions not only has strong applicability but also can better constrain the parameter process, making the calibrated model more in line with the physical conditions of the watershed. (ii) The two-site synchronous calibration of runoff or sediment can make full use of data information of different sites, reduce the impact of spatial heterogeneity on model parameters, and improve the calibration efficiency of the model. The single-site synchronous calibration of runoff and sediment based on NSGA-II not only has high calibration efficiency but also can avoid the tedious steps of calibrating runoff and sediment separately. (iii) The two-site synchronous calibration of runoff and sediment based on NSGA-II combines the advantages of the above synchronous calibration strategies, which can get Pareto-optimal front and represent the best trade-offs among different objectives, and its applicability is stronger than the traditional single-site or single-element calibration strategy. This study provides new and competing ways to evaluate hydrological models and their performance, and the multiple criteria approach for watershed modeling is one of the focuses in future research extensions.

Keywords SWAT Model · NSGA-II · Synchronization calibration · Multi-objective · Pareto-optimal front · Yanhe River Watershed

Introduction

Parameter calibration in a hydrological model is a process of constantly trying and adjusting the parameter values by various means to make the calculated flow, as closely

Responsible Editor: Marcus Schulz

✉ Lei Wu
conquer2006@126.com

1 Key Laboratory of Agricultural Soil and Water Engineering in Arid and Semiarid Areas, Ministry of Education, Northwest A&F University, Yangling 712100, Shaanxi, China

2 Blackland Research and Extension Center, Texas A&M AgriLife Research, Texas A&M University, Temple, TX 76502, USA

3 State Key Laboratory of Soil Erosion and Dryland Farming On the Loess Plateau, Northwest A&F University, Yangling 712100, Shaanxi, China

4 College of Water Resources and Architectural Engineering, Northwest A&F University, Yangling 712100, Shaanxi, China

5 Department of Construction, Northwest A&F University, Yangling 712100, Shaanxi, China

and consistently as possible, appropriate to the measured flow of the watershed (Duan et al. 1992; Rui 2017). From a mathematical point of view, parameter calibration is an optimization problem (Yapo et al. 1998; Fan et al. 2015). The calibration method consists of objective function and constraints (Yen et al. 2019). At present, most of the parameter calibration methods of hydrological model studied in the world are single-objective calibration procedures, such as the simplex method and the shuffled complex evolution algorithm (SCE-UA) (Daggupati et al. 2015). The sequential uncertainty fitting approach-version 2 (SUFI-2), generalized likelihood uncertainty estimation (GLUE), particle swarm optimization (PSO), parametric solutions (parasol), and Markov chain Monte Carlo (MCMC) embedded in SWAT calibration and uncertainty programs (SWAT-CUP) are also single-objective algorithms (Jiang et al. 2017; Liu et al. 2019). The fundamental challenge stems from the effective design of multiple objective calibration strategies to find Pareto-optimal solutions in a single run and, therefore, eliminate the need for running a sequence of single-objective optimization problems.

The calibration methodology based on a single optimization objective only considers one aspect of the complex hydrological process, cannot fully mine other characteristic information contained in hydrological data (Chen et al. 2008; Li et al. 2010), and may not be enough to capture all the aspects of the system response that the model is supposed to reproduce. Therefore, several criteria need to be considered simultaneously, so that the multi-objective automatic calibration algorithms have been widely used in the hydrological models (Yang and Wang 2010). For example, Cao (2010) proposed the multi-objective differential evolution adaptive metropolis (MODREAM) and applied it to the runoff simulation of catchment moisture deficit loss equation-three parallel linear reservoirs (CMD-3PAR) model and found the automatic calibration algorithm could better reflect the actual hydrological characteristics of the watershed and showed higher simulation accuracy than the traditional single objective algorithm. Guo (2013) proposed the multi-objective culture shuffled complex differential evolution (MOCSCDE), analyzed its performance in the calibration of lumped Xin'anjiang model, distributed Xin'anjiang model and support vector regression model, and considered the MOCSCDE calibration method could reflect the hydrological characteristics in different periods, effectively avoid the "homogenization effect" produced by the single-objective algorithm, and significantly improve the simulation performance.

The non-dominated sorting genetic algorithm II (NSGA-II), developed by Ercan and Goodall (2016), is a fast and efficient multi-objective genetic algorithm (MOGAS) (Anderton et al. 2002). NSGA-II algorithm has been applied in the calibration of hydrological model and is an effective tool for

the multi-objective calibration of watershed model parameters (Chen et al. 2018). For example, Bekele and Nicklow (2007) used NSGA-II to optimize the runoff parameters of different hydrological stations at the same time, which made the simulation performance higher than other single-objective scenarios. Guo et al. (2013) found that the NSGA-II method could obtain better simulation results in Xin'anjiang model than the single-objective method by reasonably selecting the type and number of objective functions. Overall, multiple objective optimization is used to solve coupled inverse problems that can extract information from multiple data sources or multi-sites, or both to reduce uncertainty of the inverse solution. Examples of such multiple objectives include the objective function value such as NSE, BIAS, and R^2 calculated separately on low and high flows, timing errors, and error in reproducing water balance (Chilkoti et al. 2018). More importantly, the models constituting the Pareto in criteria space can be seen as the best models, since there are no models better than these on all criteria (Vallerio et al. 2015). This is critical to test hypotheses regarding the need to design additional criteria to determine a Pareto-optimal solution for the multiple objective optimization.

The objectives of this study are to (i) compare the runoff calibration performance between the non-dominated sorting genetic algorithm-II (NSGA-II) and the single-objective SUFI-2 method, (ii) evaluate the single and synchronous calibration efficiency of runoff and sediment by NSGA-II at different sites, and (iii) demonstrate the possible advantages of multiple objective synchronous calibration over single-site or single-element calibration for the SWAT model. The findings of this study can provide suggested possible remedies for the certain deficiencies of single-objective function and single-element calibration.

Materials and method

Study region

Yanhe River Watershed (108°39′E ~ 110°29′E and 36°22′N ~ 37°20′N) is located in the north of Shaanxi Province, belonging to the loess hilly and gully region (Fig. 1). The total area of Yanhe River Watershed is about 7725 km². The altitude is about 471 ~ 1213 m. The terrain is high in the northwest and low in the southeast, with an average slope angle of 17°. From the headwater to the estuary, the Yanhe River is divided into upper, middle, and lower reaches with Huaziping and Ganguyi as the dividing point. The upper, middle, and lower reaches are the loess hilly and gully area with dominated ridges and hills, the loess hilly and gully area with dominated hills, and the crushing table plateau, respectively. The interannual difference of precipitation in Yanhe River watershed is obvious, and the distribution

Fig. 1 Study region: **a** the relative location of the Yanhe River watershed in China, and **b** digital elevation model (DEM), longitude and latitude coordinates, and the upper reaches of Ganguyi and Yan'an hydrological stations in the Yanhe River watershed
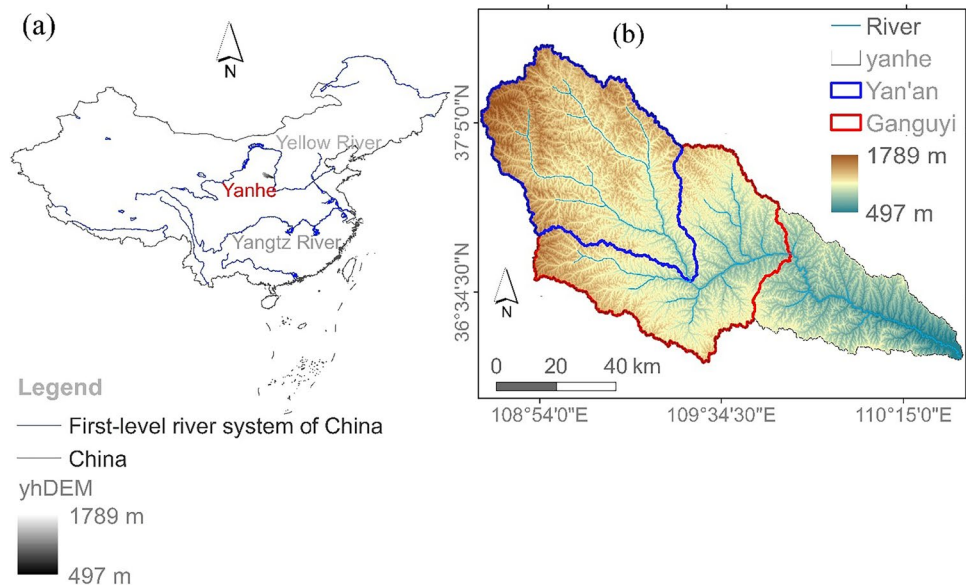
within a year is very uneven. The annual average temperature is about 9.3 ℃, and the annual average precipitation for many years is 482.3 mm, which is mostly concentrated in June to August, accounting for 68.55% of the annual precipitation (Liu et al. 2010). Its runoff accounts for 99.95~100% of the annual runoff, and the scouring amount of sediment accounts for 94~100% of the whole year (Wu et al. 2020). The month with monthly precipitation frequency greater than 50% is from May to October, which can be considered as the wet season, and the runoff in wet season accounts for more than 90% of the whole year. Loessial soil is the most widely distributed soil type in the Yanhe River watershed, accounting for more than 85% of the total area. The loessial soil is loose, and the particle size is mainly silt, belonging to cultivated soil. It is very easy to be eroded by rainfall

runoff and is widely distributed in ridge and hill tops, slopes, tablelands, and ditches.

The basic data used for the establishment of SWAT model of runoff and sediment mainly include digital elevation model (DEM), land use types, soil types, meteorological data, runoff, and sediment (Table 1).

Construction of SWAT model

In the modeling process using ArcSWAT 2012, the river network is generated through DEM, and the catchment area threshold is set as the default value (15095 hm^2). In addition, two hydrological stations, Ganguyi and Yan'an, were added to the river channel for hydrological calibration. By setting the total outlet of the watershed to obtain

Table 1 Data types, descriptions, and sources of the SWAT model in the Yanhe River watershed

Data types	Descriptions	Data sources	Notes
DEM	Digital elevation model	Geospatial Data Cloud http://www.gscloud.cn/	30 m resolution
Land use	Dryland, grassland, forestland, etc	Geographical Information Monitoring Cloud Platform http://www.dsac.cn/	100 m resolution (1995)
Soil	Soil mechanical composition, organic matter, density, water conductivity, etc	Cold and Arid Regions Sciences Data Center at Lanzhou (http://westdc.westgis.ac.cn) China Soil Map Based Harmonized World Soil Database(v1.1)	1:1,000,000
Meteorological data	Precipitation, maximum and minimum temperature, solar radiation, wind speed and relative humidity in Yan'an, Ansai, Luochuan, Suide, Wuqi, Xixian, and Zichang weather stations	National Meteorological Information Center (http://www.nmic.cn/) and Shaanxi Provincial Meteorological Burea (http://sn.cma.gov.cn/)	1965–2012
Runoff	Annual and monthly runoff of Ganguyi hydrological station	National Science & Technology Infrastructure of China, Data Sharing Infrastructure of Earth System Science (http://loess.geodata.cn)	m^3/s (1979–1997)
Sediment	Annual and monthly sediment of Ganguyi hydrological station		kg/m^3 (1979–1997)

a total of 27 sub-watersheds. In this study, the preheating period of the model is 1989–1990, the calibration period is 1991–1995, and the validation period is designed as 1996–2000 or 1996–1997. First, the NSGA-II and SUFI-2 algorithms were adopted with SWAT to compare the calibration performance of runoff parameters (Wu et al. 2022a).

In this study, twelve parameters related to runoff and eight parameters related to sediment are selected with the help of relevant literatures to study the calibration performance of runoff and sediment under different scenarios. The parameters, parameter range, optimal parameters, and the physical meaning of parameters are shown in Table 2.

Among them, the parameter range is based on SWAT input and output manual (Abbaspour et al. 2007).

Calibration scenario design

This study explores and evaluates the calibration strategy and the efficiency of NSGA-II through nine designed scenarios.

Firstly, the scenarios of S1, S2, and S3 (Table 3) were, respectively, designed to calibrate and validate the runoff parameters of SWAT model by NSGA-II and to compare the simulation results of single and synchronous calibration in the Yanhe River watershed. The initial population was set to

Table 2 Variable name and definition, minimum, maximum, and optimal parameters by the NSGA-II calibration method

Number	Parameter change way	Variable name	Minimum	Maximum	Optimal parameters	Definition
1	r	CN2.mgt	−0.25	0.25	0.17549	Moisture condition II curve number
2	v	ESCO.hru	0	1	0.384314	Soil evaporation compensation coefficient
3	v	EPCO.hru	0	1	0.4	Plant uptake compensation factor
4	v	CANMX.hru	0	10	0.54902	Maximum amount of water that can be trapped in the canopy when the canopy is fully developed (mm H_2O)
5	v	TRNSRCH.bsn	0	1	0.372549	Fraction of transmission losses partitioned to the deep aquifer
6	v	SURLAG.bsn	0.05	24	19.30392	the surface runoff lag coefficient
7	v	CH_N2.rte	0.01	0.3	0.251098	Manning's "n" value for the main channel
8	v	CH_K2.rte	0.01	500	128.2498	Effective hydraulic conductivity in main channel alluvium (mm/hr)
9	v	ALPHA_BF.gw	0	1	0.705882	Bank flow recession constant or constant of proportionality
10	r	SOL_K().sol	−0.25	0.25	−0.04412	Saturated hydraulic conductivity (mm/hr)
11	r	SOL_AWC().sol	−0.25	0.25	0.208824	Available water capacity of the soil layer (mm H_2O/ mm soil)
12	r	SOL_BD().sol	−0.25	0.25	0.234314	Moist bulk density (g/cm^3)
13	r	USLE_P.mgt	−0.25	0.25	0.118627	USLE equation support practice factor
14	r	USLE_C.plant.dat	−0.25	0.25	−0.03824	Minimum value of USLE_C factor for water erosion applicable to the land cover/plant
15	r	USLE_K.sol	−0.25	0.25	0.042157	USLE equation soil erodibility (K) factor
16	v	PRF_BSN.bsn	0	2	1.639216	Peak rate adjustment factor for sediment routing in the main channel
17	v	ADJ_PKR.bsn	0.5	2	1.929412	Peak rate adjustment factor for sediment routing in the sub-basin (tributary channels)
18	v	SPCON.bsn	0.0001	0.01	0.000255	Linear parameter for calculating the maximum amount of sediment that can be reentrained during channel sediment routing
19	v	CH_COV1.rte	−0.05	0.6	0.55302	Channel erodibility factor
20	v	CH_COV2.rte	−0.001	1	0.960745	Channel cover factor

"r" means the initial value of the parameter multiplied by (1 + calibration value); "v" means the initial value of the parameter is replaced by the calibration value

Table 3 Different scenario designs of single and synchronous calibration in runoff and sediment parameters of Ganguyi and Yan'an hydrological stations by NSGA-II

Scenarios	S1	S2	S3	S4	S5	S6	S7	S8	S9
Observed element	Runoff			Runoff and sediment					
Calibration strategy	Single	Single	Synchronization	Single	Single	Synchronization	Synchronization	Synchronization	Synchronization
Calibration tips	Gangguyi	Yan'an	Ganguyi and Yan'an	Runoff in Ganguyi or Yan'an	Sediment in Ganguyi or Yan'an	Runoff and sediment in Ganguyi or Yan'an	Runoff and sediment in Ganguyi	Runoff and sediment in Yan'an	Runoff and sediment in Ganguyi and Yan'an
Simulated elements	Runoff in Ganguyi and Yan'an	Runoff in Ganguyi and Yan'an	Runoff in Ganguyi and Yan'an	Runoff and sediment in Ganguyi or Yan'an	Runoff and sediment in Ganguyi or Yan'an	Runoff and sediment in Ganguyi or Yan'an	Runoff and sediment in Ganguyi and Yan'an	Runoff and sediment in Ganguyi and Yan'an	Runoff and sediment in Ganguyi and Yan'an

1000, the population size of offspring is 50, and the number of generations is 50. The objective functions of the single-site calibration were set as NSE (1) and PBIAS (2), and the objective functions of the two-site calibration are $\overline{NSE}$ (3) and $\overline{PBIAS}$ (4), respectively.

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

$$PBIAS = \frac{\sum_{i=1}^{n}(Q_{oi}(m) - Q_{si}(m))}{\sum_{i=1}^{n}Q_{oi}(m)} \times 100 \tag{2}$$

$$\overline{NSE} = \frac{1}{m}(NSE(1) + NSE(2) + ... + NSE(m)) \tag{3}$$

$$\overline{PBIAS} = \frac{1}{m}(PBIAS(1) + PBIAS(2) + ... + PBIAS(m)) \tag{4}$$

where Q_{si} is the simulated value (m³/s), Q_{oi} is the observed value (m³/s), Q_o is the average observed value (m³/s), and m is the site number, $m = 2$.

Secondly, the scenarios of S4, S5, and S6 (Table 3) were, respectively, designed to calibrate and validate the runoff and sediment parameters of SWAT model by NSGA-II and to compare the simulation results of single and synchronous calibration in the Yanhe River watershed. The initial population was set to 1000, the population size of offspring is 50, and the number of generations is 50. The objective functions of S4 and S5 were set as NSE (5) and PBIAS (6), and the objective functions of S6 are $\overline{NSE}$ (7) and $\overline{PBIAS}$ (8), respectively.

$$NSE = 1 - \frac{\sum_{i=1}^{n}(Q_{si}(k) - Q_{oi}(k))^2}{\sum_{i=1}^{n}(Q_{si}(k) - \overline{Q_o}(k))^2} \tag{5}$$

$$PBIAS = \frac{\sum_{i=1}^{n}(Q_{oi}(k) - Q_{si}(k))}{\sum_{i=1}^{n}Q_{oi}(k)} \times 100 \tag{6}$$

$$\overline{NSE} = \frac{1}{2}(NSE(1) + NSE(2)) \tag{7}$$

$$\overline{PBIAS} = \frac{1}{2}(PBIAS(1) + PBIAS(2)) \tag{8}$$

where Q_{si} is the simulated value (runoff, m³/s; sediment, t), Q_{oi} is the observed value (runoff, m³/s; sediment, t), $\overline{Q_o}$ is the average observed value (runoff, m³/s; sediment, t), "$k = 1$" indicates the runoff data, and "$k = 2$" indicates the sediment data.

Thirdly, the scenarios of S7, S8, and S9 (Table 3) were respectively designed to calibrate and validate the runoff and sediment parameters of SWAT model by NSGA-II and to compare the simulation results of the synchronous calibration in single-site and two-site in the Yanhe River watershed. The initial population was set to 1000, the population size of offspring is 50, and the number of generations is 50. The objective functions of S7 and S8 were set as $\overline{NSE}$ (9) and $\overline{PBIAS}$ (10), and the objective functions of S9 are $\overline{NSE}_2$ (11) and $\overline{PBIAS}_2$ (12), respectively.

$$\overline{NSE}(k, m) = 1 - \frac{\sum_{i=1}^{n}(Q_{si}(k, m) - Q_{oi}(k, m))^2}{\sum_{i=1}^{n}(Q_{si}(k, m) - \overline{Q_o}(k, m))^2} \tag{9}$$

Table 4 Comparison of runoff calibration and validation results between NSGA-II and SUFI-2 in different periods

station	Objective function	R^2		NSE		PBIAS (%)	
	Methodology	NSGA-II	SUFI-2	NSGA-II	SUFI-2	NSGA-II	SUFI-2
Ganguyi	Calibration (91–95)	0.75	**0.76**	**0.72**	**0.72**	**2.55**	15.27
	Validation (96–00)	**0.77**	0.75	0.59	**0.64**	**0.44**	18.58
	Validation (96–97)	0.86	**0.88**	0.77	**0.85**	**0.17**	19.18
	Validation (98–00)	**0.56**	0.54	0.03	0.00	**0.68**	18.03
Yan'an	Calibration (91–95)	**0.80**	0.69	**0.76**	0.68	**12.03**	15.22
	Validation (96–00)	0.65	**0.69**	0.58	**0.68**	**11.74**	17.31
	Validation (96–97)	0.82	**0.83**	**0.81**	0.77	**23.76**	28.23
	Validation (98–00)	0.44	0.52	−0.90	**0.09**	**−3.22**	3.73

The bold part in the table is the better value

$$\overline{\text{PBIAS}}(k,m) = \frac{\sum_{i=1}^{n}\left(Q_{\text{oi}}(k,m) - Q_{\text{si}}(k,m)\right)}{\sum_{i=1}^{n}Q_{\text{oi}}(k,m)} \times 100 \qquad (10)$$

$$\overline{\text{NSE}}_2 = \frac{1}{4}(\text{NSE}(1,1) + \text{NSE}(1,2) + \text{NSE}(2,1) + \text{NSE}(2,2)) \qquad (11)$$

$$\overline{\text{PBIAS}}_2 = \frac{1}{4}(\text{PBIAS}(1,1) + \text{PBIAS}(1,2) + \text{PBIAS}(2,1) + \text{PBIAS}(2,2)) \qquad (12)$$

where Q_{si} is the simulated value (runoff, m³/s; sediment, t), Q_{oi} is the observed value (runoff, m³/s; sediment, t), $\overline{Q_{\text{o}}}$ is the average observed value (runoff, m³/s; sediment, t), "$k = 1$" indicates the runoff data, and "$k = 2$" indicates the sediment data. "$m = 1$" indicates data from Yan'an station, and "$m = 2$" indicates data from Ganguyi station.

Results

Calibration and validation by NSGA-II and SUFI-2

NSGA-II and SUFI-2 algorithms are, respectively, used to calibrate runoff parameters of Ganguyi catchment and Yan'an catchment (Table 4). During the calibration period from 1991 to 1995, both R^2 and NSE of Ganguyi catchment

reached > 0.7, PBIAS were < 20%, and R^2 and NSE of Yan'an catchment reached > 0.65; PBIAS were also < 20%. These data indicate that the two calibration methodologies can meet the accuracy requirements of hydrological modeling in the Yanhe River watershed. Moreover, R^2 and NSE at both sites are satisfactory in the validation period from 1996 to 1997, although R^2 and NSE decreased significantly or even negative in the validation period from 1998 to 2000. More significantly, NSGA-II algorithm presents an obvious advantage that its PBIAS is better than that of SUFI-2 method in the calibration and validation period of the two sites.

Comparison of single-site and two-site calibration by NSGA-II

After the parameters of S1, S2, and S3 scenarios are calibrated, the simulation performance of each station in the calibration and validation period was statistically shown in Table 5. Generally, NSGA-II method has more constraints than SUFI-2 on the parameter calibration process, and the simulation results are more reasonable. However, the parameters corresponding to S1 and S2 scenarios have obvious spatial limitations. For example, the NSE coefficients of S1 in Ganguyi station are respectively 0.72 and 0.77 in the calibration and validation period (96–97), but the NSE

Table 5 Comparison of runoff calibration and validation results under single-site (S1, S2) and two-site (S3) scenarios

Station	Objective function	R^2			NSE			PBIAS (%)		
	Scenario	S1	S2	S3	S1	S2	S3	S1	S2	S3
Ganguyi	Calibration (91–95)	0.75	0.73	0.73	0.72	0.60	0.63	2.55	31.38	38.23
	Validation (96–00)	0.77	0.62	0.75	0.59	0.19	0.58	0.44	35.98	45.81
	Validation (96–97)	0.86	0.71	0.83	0.77	0.45	0.72	0.17	36.07	42.47
	Validation (98–00)	0.56	0.47	0.56	0.03	-0.65	0.10	0.68	35.90	48.90
Yan'an	Calibration (91–95)	0.67	0.80	0.79	−1.04	0.76	0.75	−126.24	12.03	0.42
	Validation (96–00)	0.68	0.65	0.75	−0.54	0.58	0.74	−131.54	11.74	3.92
	Validation (96–97)	0.84	0.82	0.89	0.42	0.81	0.87	−88.87	23.76	15.30
	Validation (98–00)	0.59	0.44	0.57	−6.62	−0.90	−0.15	−184.61	−3.22	−10.23

coefficients in Yan'an station are reduced to − 1.04 and 0.42 in the calibration and validation period (96–97). The NSE coefficients of S2 in Yan'an station were, respectively, 0.76 and 0.81 in the calibration and validation period (96–97), but in Ganguyi station, the NSE coefficients decreased to 0.60 and 0.45, respectively. This phenomenon indicates that there will be poor NSE and PBIAS, if the S1 scenario parameter is used to Yan'an catchment or the S2 scenario parameter is applied to Ganguyi catchment regardless of the calibration and validation period. On the contrary, the parameters corresponding to S3 scenario obviously have better applicability. The NSE coefficients of S3 in Ganguyi station are 0.63 and 0.72, respectively, in the calibration and validation periods (91–95, 96–97), and those in Yan'an station are, respectively, 0.75 and 0.87. Therefore, the simulation performance of S3 in two hydrological stations is satisfactory in the calibration and validation periods (91–95, 96–97).

Comparison between single and synchronous calibration of runoff and sediment parameters by NSGA-II

After the parameters in S4, S5, and S6 scenarios are calibrated, the simulation performance of each station in the calibration and validation period was statistically shown in Table 6. The calibration and verification effect of runoff and sediment in S6 scenario is relatively good, but there are obvious differences between S4 and S5 scenarios. Firstly, the S4 scenario only uses runoff data for calibration, but the calibration and validation performance of runoff and sediment in these two stations is acceptable. The NSE coefficients of the runoff parameters in Ganguyi station are 0.72 and 0.77, respectively, in the calibration and validation period (91–95, 96–97), and the NSE coefficients of sediment parameters in Ganguyi station are 0.53 and 0.46 in the calibration and validation period (91–95, 96–97). The NSE coefficients of the runoff parameters in Yan'an station are 0.76 and 0.81, respectively, and the NSE coefficients of sediment parameters in Yan'an station

are 0.51 and 0.56, respectively. Secondly, the S5 scenario only uses sediment data for calibration, and the calibration and validation performance of sediment is better than that of S4 scenario, but the runoff effect is relatively poor, even negative. Specifically, the NSE coefficients of S5 in the sediment parameters of Ganguyi station are 0.71 and 0.60, respectively, in the calibration and validation period (91–95, 96–97), but the NSE coefficients of runoff parameters in Ganguyi station are − 0.93 and − 3.26, respectively; the NSE coefficients of S5 in the sediment parameters of Yan'an station are, respectively, 0.64 and 0.65 in the calibration and validation period (91–95, 96–97), but the NSE coefficients of runoff parameters in Yan'an station are − 14.83 and − 6.59, respectively.

Furthermore, the NSE coefficients of S6 scenario in the runoff parameters of Ganguyi station are 0.74 and 0.82, respectively, in the calibration and validation period (91–95, 96–97) and 0.56 and 0.58, respectively, in the sediment parameters of Ganguyi station. The NSE coefficients of S6 scenario in the runoff parameters of Yan'an station are 0.78 and 0.83, respectively, in the calibration and validation period (91–95, 96–97) and 0.56 and 0.60, respectively, in the sediment parameters of Yan'an station.

Synchronous calibration of runoff and sediment parameters in two sites

Like S1–S3 scenarios, the parameters corresponding to S7 and S8 scenarios still have obvious spatial limitations on runoff (Table 7). When S7 scenario parameter is applied to the runoff in Yan'an catchment or S8 scenario parameter to the runoff in Ganguyi catchment, the simulation effect is obviously poor in the calibration and validation period. For example, the NSE coefficients of S7 in runoff parameters of Ganguyi station are 0.74 and 0.82, respectively, in the calibration and validation period (91–95, 96–97), but the NSE coefficients in Yan'an station are reduced to − 0.87 and 0.37 in the calibration and validation period (91–95,

Table 6 Comparison of single and synchronous calibration and validation results of runoff and sediment in different periods

Station	Hydrological elements	Objective function Scenario	NSE			PBIAS (%)		
			S4	S5	S6	S4	S5	S6
Ganguyi	Runoff	Calibration (91–95)	0.72	− 0.93	**0.74**	2.55	2.55	**2.05**
		Validation (96–97)	0.77	− 3.26	**0.82**	0.17	0.17	**0.17**
	Sediment	Validation (91–95)	0.53	0.71	**0.56**	− 10.42	− 4.59	**3.13**
		Validation (96–97)	0.46	0.60	**0.58**	− 22.64	− 16.13	**− 15.91**
Yan'an	Runoff	Calibration (91–95)	0.76	− 14.83	**0.78**	12.03	− 344.84	**6.31**
		Validation (96–97)	0.81	− 6.59	**0.83**	23.76	− 277.67	**22.18**
	Sediment	Validation (91–95)	0.51	0.64	**0.56**	8.33	0.34	**− 4.46**
		Validation (96–97)	0.56	0.65	**0.60**	5.51	2.17	**5.14**

Values in bold represent the scenarios where the model calibration results are relatively good

Table 7 Comparison of results of single and synchronous calibration of runoff and sediment at two stations

Station	Hydrological elements	Objective function	NSE			PBIAS (%)		
		Scenario	S7	S8	S9	S7	S8	S9
Ganguyi	Runoff	Calibration (91–95)	**0.74**	0.54	0.67	2.05	49.8	**2.05**
		Validation (96–97)	**0.82**	0.68	0.52	0.17	53	**0.17**
	Sediment	Calibration (91–95)	0.56	0.5	**0.67**	3.13	− 50.88	**3.13**
		Validation (96–97)	0.58	0.48	**0.64**	− 15.91	− 51.19	**− 15.91**
Yan'an	Runoff	Calibration (91–95)	− 0.87	**0.78**	0.70	− 108.98	6.31	**− 4.44**
		Validation (96–97)	0.37	0.83	**0.89**	− 75.68	22.18	**8.48**
	Sediment	Calibration (91–95)	**0.6**	0.56	0.56	10.81	**− 4.46**	17.68
		Validation (96–97)	**0.69**	0.6	0.55	**− 1.77**	5.14	27.26

Values in bold represent the scenarios where the model calibration results are relatively good

96–97). The NSE coefficients of S8 in runoff parameters of Yan'an station are 0.70 and 0.89, respectively, in the calibration and validation period (91–95, 96–97), but the NSE coefficients in Ganguyi station are reduced to 0.67 and 0.52. The above results indicate again that using a single station to calibrate runoff will make the application of the model have a large deviation. By contrast, S9 has a good runoff simulation effect in the calibration and verification period (91–95, 96–97) at Ganguyi and Yan'an stations, the two-site method based on NSGA-II can improve the adaptability of parameters in runoff simulation.

Discussion

Comparison of the NSGA-II and SUFI-2 algorithms

The monthly simulated runoff obtained by the NSGA-II and SUFI-2 algorithms is basically consistent with the measured runoff at the two stations (Fig. 2), indicating that the SWAT model has good adaptability in the runoff

simulation of Yanhe River watershed. Furthermore, the simulation effect of the calibrated model in the wet season is significantly better than that in the dry season at the two stations. Many scholars have similar results that there will be significant differences in the simulation effects in wet season and dry season when SWAT is applied to watersheds with extreme uneven in inner-annual flow (Gao 2018). Firstly, SWAT model uses soil conservation service curve number (SCS-CN) method to calculate surface runoff, and dry season factors are not fully considered in the equation (Liew and Garbrecht 2003). Secondly, the objective functions, such as NSE and PBIAS, are more sensitive to the peak value, because they mainly characterize the overall simulation effect and may not reflect the change in dry season (Zhang et al. 2015).

R^2, NSE, and PBIAS during the validation period from 1996 to 2000 were partly worse than that of the calibration period, which was related to complex and severe human activities after 1996 (Chen et al. 2010). Particularly, the decline of validation effect after 1998 may be related to the effective implementation of large-scale "Grain for Green"

Fig. 2 Comparison of two calibration algorithms in different hydrological stations: **a** Ganguyi and **b** Yan'an
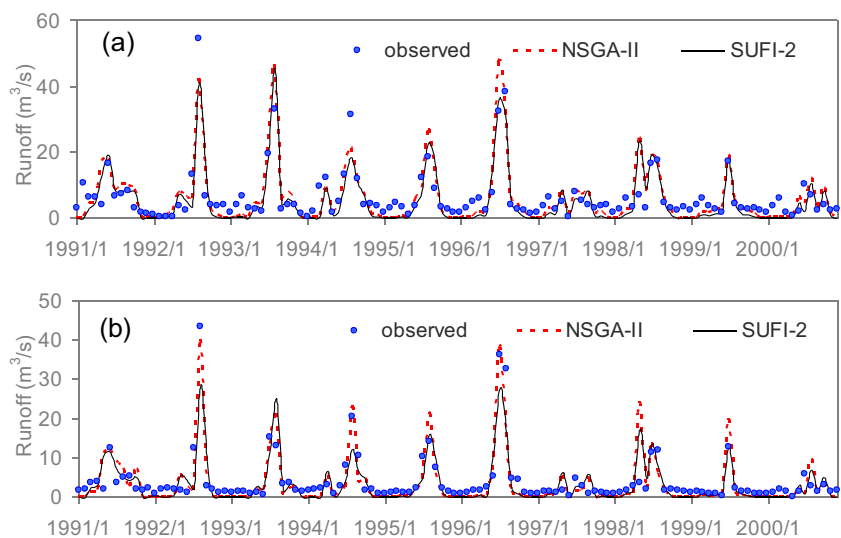
projects in the late 1990s in the Yanhe River Watershed (Wu et al. 2019), such as the project of "beautiful mountains and rivers" implemented in 1997 and the project of "returning farmland to forest/grass" implemented in 1999 (Yu 2008). These soil conservation practices have greatly changed the underlying surface conditions of the watershed, but the model does not take this change into account and only uses the land use data in 1995, which may be the other important reason for the poor simulation performance after 1998.

More importantly, PBAIS is the average deviation between the simulated value and the measured value. The small PBIAS by NSGA-II algorithm implies that this method can not only characterize the consistency of the average annual runoff between simulation and observation but also can better weaken the impact of peak flow (Gebremariam et al. 2014). Therefore, the NSGA-II algorithm based on multiple objective functions can better constrain the parameter calibration process and make the calibrated model more in line with the physical conditions of the watershed.

## Calibration of single- and multi-hydrological elements

Using NSGA-II algorithm to calibrate runoff or sediment alone can effectively improve the simulation performance of runoff or sediment. However, using only runoff data to calibrate model parameters may improve the calibration performance of sediment to a certain extent according to the results of S4 and S5 scenarios, while using only sediment data to calibrate model parameters has no obvious effect on runoff performance. This result proves the rationality of first calibrating runoff and then calibrating sediment in the model calibration. This is because runoff parameters will not only affect runoff but also affect sediment processes (Abbaspour et al. 2007). Although the runoff simulation of S5 scenario is relatively poor, the simulation effect of sediment is significantly better than that of S4 scenario, which can be attributed to the differential impact of some parameters on runoff and sediment simulation (Ghasemizade et al. 2017). The effect of S6 scenario is better than S4 in the runoff simulation of these two stations during the calibration and validation period and is equivalent to S5 scenario in the sediment simulation of these two stations. Therefore, using runoff and sediment data to calibrate hydrological parameters may be more reasonable than using runoff data only. It is feasible to use the NSGA-II methodology and two data sets to calibrate runoff and sediment synchronously, which can not only reduce the number of parameter iteration (Ercan and Goodall 2016) but also improve the simulation effects of runoff and sediment. More importantly, the synchronous calibration of runoff and sediment parameters avoids the cumbersome steps

of calibrating runoff and sediment, respectively, and can make full use of runoff and sediment data information and improve the calibration efficiency of SWAT model in hydrological parameters.

Synchronous calibration of runoff and sediment at multi-sites

Single-site calibration is more reliable in small watersheds, but it will ignore the spatial heterogeneity of parameters and affect the effectiveness of the application of the model in the whole large-scale watershed (Anderton et al. 2002). Many studies have shown that multi-site calibration methodology is an effective way to solve this problem (Bai et al. 2017; Molina-Navarro et al. 2017; Gong et al. 2012), which can make full use of the existing data and reduce the uncertainty of hydrological model of medium- and large-scale watersheds with high spatial heterogeneity (Bekele and Nicklow 2007). Some researchers have proved that better simulation results can be obtained by using multi-site data in the calibration and validation processes (Zhang 2008; Ercan and Goodall 2016). For example, Zhang et al. (2013) proposed a multi-algorithm genetically adaptive multi-objective (AMALGAM) optimization algorithm, which could not only effectively avoid the spatial heterogeneity of parameters but also improved the model calibration efficiency and the reliability of simulation results.

There are some differences between the synchronous calibration of runoff and sediment in single-site and the synchronous calibration of runoff and sediment under multi-sites (Table 7). The applicability of calibrated parameters in sediment simulation may be wider than runoff in the Yanhe River watershed. For example, the NSE coefficients of S7 in the sediment parameters of Ganguyi station are 0.56 and 0.58, respectively, in the calibration and validation period (96–97), and even better in Yan'an station, with the NSE coefficients of 0.60 and 0.69, respectively; the NSE coefficients of S8 in the sediment parameters of Yan'an station are 0.56 and 0.60, respectively, in the calibration and validation period (91–95, 96–97), and 0.50 and 0.48 in Ganguyi station, respectively. Generally, sediment is usually inseparable from runoff when studying sediment yield in a watershed, and it is meaningful to assess the adaptability of parameters to sediment only when they first have good adaptability to runoff. Therefore, both S7 and S8 scenarios still have spatial limitations under certain circumstances. This is because single-site calibration may produce local optimal solutions, and there are obvious adaptability problems in the application of different watershed scales (Huo et al. 2020). S9 scenario not only realizes the applicability of parameters but also its sediment simulation effect is also basically equivalent to that of S7 and S8. The NSE coefficients of monthly runoff in Ganguyi station are 0.67 and 0.52, respectively, in
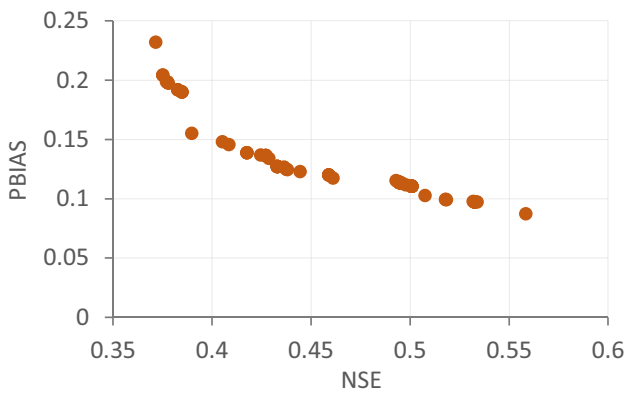
Fig. 3 Pareto-optimal front of NSGA-II calibration methodology in objective function space

the calibration and validation period and 0.70 and 0.89 in Yan'an station; the NSE coefficients of monthly sediment in Ganguyi station are 0.67 and 0.64, respectively, and 0.56 and 0.55 in Yan'an station. The NSE coefficients of runoff and sediment in Ganguyi and Yan'an stations are all greater than 0.5, indicating that the synchronous calibration of runoff and sediment parameters in these two stations based on NSGA-II can determine a Pareto-optimal solution for the multiple objective optimization scheme in a single run and obtain a unique parameter set with wide applicability, which can not only improve the calibration performance of model but also can provide a robust model basis for watershed management.

The set of all noninferior solutions form the Pareto front. The collection of all Pareto-optimal solutions often forms an L-shaped curve in the objective space (Jia and Ierapetritou 2007). Figure 3 shows the Pareto front of a bicriterion minimization problem, where the dots correspond to Pareto-optimal solutions. The parameter set with the optimal average NSE (0.65) is selected as the final solution, meanwhile, with the average PBIAS of 0.19%. Practically, the solution

of a bicriterion optimization problem is a process of making tradeoffs between Pareto-optimal solutions (Abdalla et al. 2023). By using the Pareto-solution set from multi-site calibration, one can generate a Pareto-ensemble of simulated outputs for different sites, so that the uncertainty in the model simulations due to different ways of trading-off the model and data errors can be examined (Salazar and Rocco 2007).

The calibration and validation results of runoff and sediment corresponding to S9 scenario can be respectively presented in Figs. 4 and 5. On the whole, the simulation effect in the wet season is better than that in dry season, with NSE and R^2 of 0.67 and 0.77 in Ganguyi station and 0.77 and 0.78 in Yan'an station, while the simulation effect in the dry season is relatively poor. The reason is that the runoff generation in SWAT model is characterized by the SCS-CN method and a set of parameters associated with that structure, which primarily considers the precipitation and physical factors of the watershed (Liew and Garbrecht 2003). Furthermore, the validation effect of runoff in Yan'an station is better than that of Ganguyi station, but the simulation effect of sediment in Yan'an station is relatively poor. This is because the same set of parameters cannot fully characterize the spatial heterogeneity of the actual underlying surfaces in different catchments (Wu et al. 2022b). The noninferior solution obtained by the NSGA-II method is not the optimal solution, and the influence of parameters on hydrological elements at different stations is not completely consistent, and it also has a certain variability and uncertainty (Leta et al. 2017).

Conclusions

This study focuses on the NSGA-II algorithm to compare and analyze the impact of nine designed scenarios such as single- and two-site calibration, separate and synchronous calibration of runoff and sediment on the simulation

Fig.4 Calibration (1991–1995) and validation (1996–1997) results of runoff corresponding to scenario S9 in **a** Ganguyi and **b** Yan'an hydrological stations
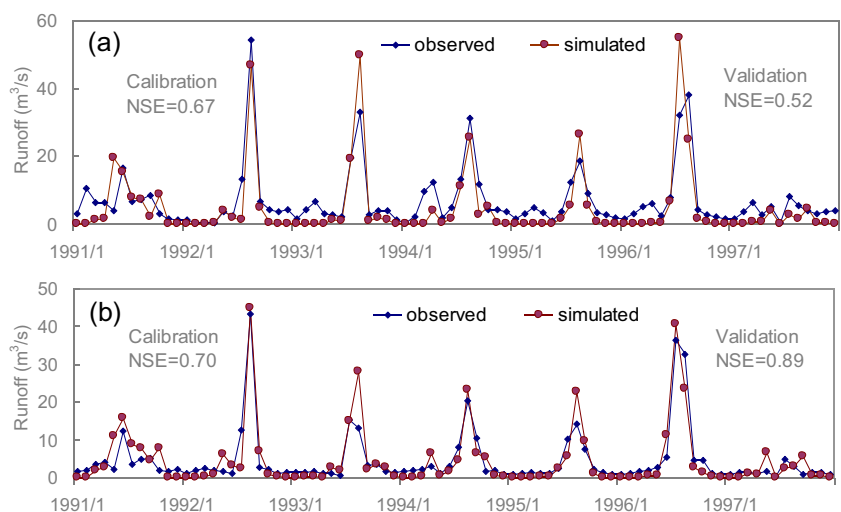
Fig.5 Calibration (1991–1995) and validation (1996–1997) results of sediment corresponding to scenario S9 in **a** Ganguyi and **b** Yan'an hydrological stations
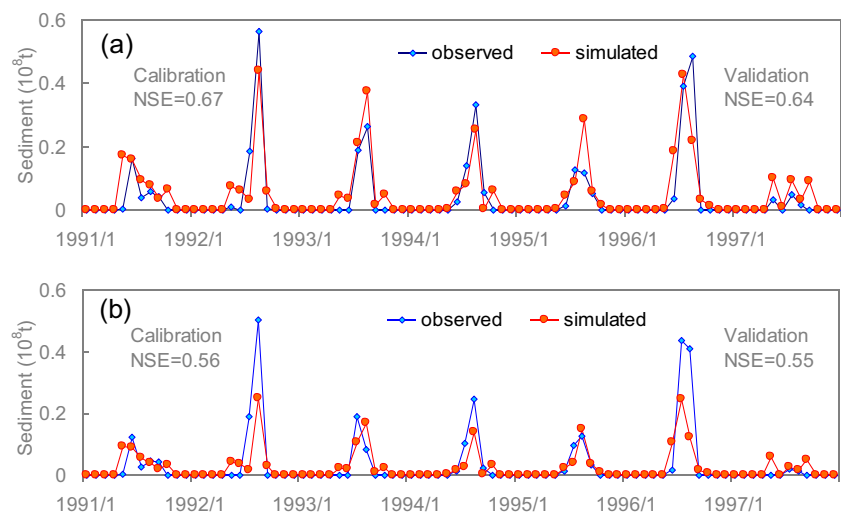
performance of SWAT model in an arid and semi-arid watershed. Results indicate that (i) both NSGA-II and SUFI-2 algorithms can meet the accuracy requirements of parameter calibration in watershed modeling, but NSGA-II based on multiple objective functions can better constrain the parameter calibration process and obtain a hydrological model more in line with the physical conditions of the watershed. (ii) The two-site calibration based on NSGA-II can effectively reduce the impact of spatial heterogeneity on model parameters and obtain an optimal parameter set with wide applicability. The synchronous calibration of runoff and sediment parameters based on NSGA-II can not only improve the calibrating performance but also enhance the calibrating efficiency of SWAT model. (iii) The multi-objective two-site synchronization calibration of runoff and sediment based on NSGA-II combines the advantages of two-site and two-element synchronization calibration, which can identify a Pareto-region of the parameter space and obtain optimal parameters in different locations and can also characterize the trade-offs that can be made between different "optimal" ways of constraining the model to be consistent with the data in the presence of model and data error. The finding aims to provide insights on the practical implication of multi-objective optimization for model calibration and leading to the elusive goal of finding a unique optimal parameter set more efficiently. It is recommended that multi-optimization criteria that measure different aspects of system behavior can be used to investigate model uncertainty and performance, prior to broad use of a watershed water quality model.

Author contribution Dr. Lei Wu wrote the manuscript and is the first author of the paper. Engineer Xia Liu statistically analyzed the data and is the first co-author of the paper. Master Junlai Chen did some modeling work, and Professor Xiaoyi Ma reviewed the paper.

Declarations

Ethical approval and consent to participate This is not applicable.

Consent to publication This is not applicable.

Competing interests The authors declare no competing interests.

References

Abbaspour KC, Yang J, Maximov I, Siber R, Bogner K, Mieleitner J, Zobrist J, Srinivasan R (2007) Modelling hydrology and water quality in the pre-alpine/alpine Thur watershed using SWAT. J Hydrol 333(2):413–430

Abdalla EMH, Alfredsen K, Muthanna TM (2023) On the use of multi-objective optimization for multi-site calibration of extensive green roofs. J Environ Manage 326:116716

Anderton S, Latron J, Gallart F (2002) Sensitivity analysis and multi-response, multi-criteria evaluation of a physically based distributed model. Hydrol Process 16(2):333–353

Bai J, Shen Z, Yan T (2017) A comparison of single- and multi-site calibration and validation: a case study of SWAT in the Miyun Reservoir watershed, China. Front Earth Sci 11(3):592–600

Bekele EG, Nicklow JW (2007) Multi-objective automatic calibration of SWAT using NSGA-II. J Hydrol 341(3):165–176

Cao FF (2010) Study on parameter optimization and uncertainty analysis for conceptual hydrological model based on MCMC method. Doctoral Dissertation, Zhejiang University, 866 Yuhangtang Rd, Hangzhou 310058

Chen JF, Zhang WC, Wu B (2008) Multi-objective calibration with predictive uncertainty analysis for conceptual hydrological models. Bull Soil Water Conserv 28(3):107–112

Chen Q, Gou S, Qin DT, Zhou ZH (2010) A high efficiency auto-calibration method for SWAT model. J Hydraul Eng 41(1):113–119

Chen Y, Chen Y, Chen X, Chen X, Xu C, Xu C, Zhang M, Zhang M, Liu M, Liu M, Gao L, Gao L (2018) Toward improved calibration of SWAT using season-based multi-objective optimization: a case study in the Jinjiang Basin in Southeastern China. Water Resour Manage 32(4):1193–1207

Chilkoti V, Bolisetti T, Balachandar R (2018) Multi-objective auto-calibration of SWAT model for improved low flow performance for a small snowfed catchment. Hydrol Sci J 63(9–12):1482–1501

Daggupati P, Yen H, White MJ, Srinivasan R, Arnold JG, Keitzer CS, Sowa SP (2015) Impact of model development, calibration and validation decisions on hydrological simulations in West Lake Erie Basin. Hydrol Process 29(26):5307–5320

Duan Q, Gupta VK, Sorooshian S (1992) Effective and efficient global optimization for conceptual rainfall-runoff models. Water Resour Res 28(4):1015–1031

Ercan MB, Goodall JL (2016) Design and implementation of a general software library for using NSGA-II with SWAT for multi-objective model calibration. Environ Model Softw 84(10):112–120

Fan K, Ma XY, Li ZJ, Feng ZZ, Zhong XM (2015) A comparative study of the SWAT model parameter calibration methods. China Rural Water Hydropower 4:77–81

Gao X (2018) Study on temporal and spatial differences of runoff process parameters in SWAT model. Fujian Normal University, Fuzhou

Gebremariam SY, Martin JF, DeMarchi C, Bosch NS, Confesor R, Ludsin SA (2014) A comprehensive approach to evaluating watershed models for predicting river flow regimes critical to downstream ecosystem services. Environ Model Softw 61:121–134

Ghasemizade M, Baroni G, Abbaspour K, Schirmer M (2017) Combined analysis of time-varying sensitivity and identifiability indices to diagnose the response of a complex environmental model. Environ Model Softw 88:22–34

Gong Y, Shen Z, Liu R, Hong Q, Wu X (2012) A comparison of single- and multi-gauge based calibrations for hydrological modeling of the upper Daning River watershed in China's Three Gorges Reservoir Region. Nord Hydrol 43(6):822–832

Guo J, Zhou JZ, Zou Q, Song LX, Zhang YC (2013) Study on multi-objective parameter optimization of Xinanjiang model. Hydrology 33(1):1–7 (**26**)

Guo J (2013) Studies on watershed hydrological modeling and forecasting. Doctoral Dissertation, Huazhong University of Science & Technology, Luoyu Road 1037, Wuhan

Huo AD, Huang ZK, Cheng YX, Van Liew MW (2020) Comparison of two different approaches for sensitivity analysis in Heihe River basin (China). Water Supply 20(1):319–327

Jia Z, Ierapetritou MG (2007) Generate Pareto optimal solutions of scheduling problems using normal boundary intersection technique. Comput Chem Eng 31(4):268–280

Jiang YN, Wang L, Wei XM, Ding XC (2017) Impacts of climate change on runoff of Jinghe River based on SWAT model. Trans Chin Soc Agric Mach 48(2):262–270

Leta OT, van Griensven A, Bauwens W (2017) Effect of single and multisite calibration techniques on the parameter estimation, performance, and output of a SWAT model of a spatially heterogeneous catchment. J Hydrol Eng 22:05016036

Li X, Weller DE, Jordan TE (2010) Watershed model calibration using multi-objective optimization and multi-site averaging. J Hydrol 380(3–4):277–288

Liew MW, Garbrecht J (2003) Hydrologic simulation of the little Washita River experimental watershed using SWAT. J Am Water Resour Assoc 39(2):413–426

Liu CL, Yang QK, Xie HX (2010) Spatial and temporal distributions of rainfall erosivity in the Yanhe River Basin. Environ Sci 31(4):850–857

Liu N, Zhang X, Zhu XP, Zhao ZH, Wu PL (2019) Runoff simulation using SWAT model and SUFI-2 algorithm in Biliu River Basin. Water Power 45(3):18–22

Molina-Navarro E, Andersen HE, Nielsen A, Thodsen H, Trolle D (2017) The impact of the objective function in multi-site and multi-variable calibration of the SWAT model. Environ Model Softw 93:255–267

Rui XF (2017) Discussion of watershed hydrological model. Advances in Science and Technology of Water Resources 37(4):1–7, 58

Salazar ADE, Rocco SCM (2007) Solving advanced multi-objective robust designs by means of multiple objective evolutionary algorithms (MOEA): a reliability application. Reliab Eng Syst Saf 92(6):697–706

Vallerio M, Vercammen D, Impe JV, Logist F (2015) Interactive NBI and (E)NNC methods for the progressive exploration of the criteria space in multi-objective optimization and optimal control. Comput Chem Eng 82:186–201

Wu L, Li GX, Jiang J, Ma XY (2019) Using vegetation correction coefficient to modify a dynamic particulate nutrient loss model for monthly nitrogen and phosphorus load predictions: a case study in a small loess hilly watershed. Environ Sci Pollut Res 26:32610–32623

Wu L, He Y, Ma XY (2020) Using five long time series hydrometeorological data to calibrate a dynamic sediment delivery ratio algorithm for multi-scale sediment yield predictions. Environ Sci Pollut Res 27:16377–16392

Wu L, Liu X, Chen JL, Yu Y, Ma XY (2022a) Overcoming equifinality: time-varying analysis of sensitivity and identifiability of SWAT runoff and sediment parameters in an arid and semiarid watershed. Environ Sci Pollut Res 29:31631–31645

Wu L, Liu X, Yang Z, Yu Y, Ma XY (2022b) Effects of single- and multi-site calibration strategies on hydrological model performance and parameter sensitivity of large-scale semi-arid and semi-humid watersheds. Hydrol Process 36(6):e14616

Yang BB, Wang WC (2010) Comparison between multi-objective evolutionary algorithms for calibration of Xinanjiang Model. Hydrology 30(3):38–42

Yapo PO, Gupta HV, Sorooshian S (1998) Multi-objective global optimization for hydrologic models. J Hydrol 204(1–4):83–97

Yen H, Park S, Arnold JG, Srinivasan R, Chawanda CJ, Wang R, Feng Q, Wu J, Miao C, Bieger K, Daggupati P, Griensven AV, Kalin L, Lee S, Sheshukov AY, White MJ, Yuan Y, Yeo I-Y, Zhang M, Zhang X (2019) IPEAT+: a built-in optimization and automatic calibration tool of SWAT+. Water 11:1681. https://doi.org/10.3390/w11081681

Yu H (2008) Analysis and trend prediction of runoff and sediment in Yanhe River Basin based on time series. Norwest A&F University, Yangling

Zhang X (2008) Multi-site calibration of the SWAT model for hydrologic modeling. Trans ASABE 51(6):2039–2049

Zhang X, Beeson P, Link R, Manowitz D, Izaurralde RC, Sadeghi A, Thomson AM, Sahajpal R, Srinivasan R, Arnold JG (2013) Efficient multi-objective calibration of a computationally intensive hydrologic model with parallel computing software in Python. Environ Model Softw 46:208–218

Zhang D, Chen X, Yao H, Lin B (2015) Improved calibration scheme of SWAT by separating wet and dry seasons. Ecol Model 301:54–61