Journal of
Water & Climate Change

 Journal of Water and Climate Change Vol 12 No 7, 3264 doi: 10.2166/wcc.2021.115

Influence of multisite calibration on streamflow estimation based on the hydrological model with CMADS inputs

Yongyu Song[a,b], Jing Zhang[a,b,*] and Yuequn Lai[a,b]

[a] Beijing Laboratory of Water Resources Security, Capital Normal University, Beijing 100048, China
[b] Key Laboratory of 3D Information Acquisition and Application of Ministry of Education, Capital Normal University, Beijing 100048, China
*Corresponding author. E-mail: 5607@cnu.edu.cn

ABSTRACT

Due to the spatial heterogeneity, the hydrological model calibration results only at the total outlet of the basin may not represent the whole basin. To more accurately simulate the historical streamflow process within the Qujiang River Basin, we set up three calibration strategies (single-site, S1; multisite simultaneous, S2; and multisite sequential, S3) for four hydrological stations based on the SWAT (Soil and Water Assessment Tool) model driven by CMADS (China Meteorological Assimilation Driving Datasets for the SWAT model). In addition, the implications of these calibration issues are extended to future streamflow projections using multimodel ensemble data in CMIP6 (Coupled Model Intercomparison Project Phase 6). In the model calibration phase, the SWAT model achieved very satisfactory results in the study area. Compared with S1 and S2, S3 can effectively improve the accuracy of streamflow simulation of stations within the basin and reduce the simulation deviation. Especially at the daily scale, the average NSE values of the four stations with S3 increased by 0.26 and 0.07, and the overall deviation decreased by 0.25 and 6.43%, respectively. Parameter sensitivity analysis also shows that spatial heterogeneity can be more adequately considered when using S3 to calibrate the model. As for the results of future streamflow projection, when using the S3, the average annual streamflow of four stations in the three climate scenarios from 2021 to 2050 is about 44.21, 130.00, 321.55 and 713.24 m^3/s, respectively. Correspondingly, the use of S1 and S2 would bring certain risks to future water resource management.

Key words: CMADS, CMIP6, multisite sequential calibration, streamflow projection, SWAT model

HIGHLIGHTS

- Set up single-site (S1), multisite simultaneous (S2) and multisite sequential (S3) calibration strategies to explore the discrepancy of streamflow simulation of four stations.
- The SWAT model is driven by CMADS data input.
- The impact of three strategies on future streamflow projection is discussed with CMIP6.
- Parameter sensitivity analysis shows that spatial heterogeneity can be more adequately considered using S3 to calibrate the model.

1. INTRODUCTION

As an important tool to analyze the hydrological process of a river basin, the hydrological models have shown very strong capabilities in solving many problems such as water resource planning, flood control and disaster reduction, and nonpoint source pollution in recent decades (Boufala *et al.* 2019; Wang *et al.* 2019; Mahmoodi *et al.* 2020; Rudra *et al.* 2020). According to the different physical perfection of model structure and parameters, hydrological models can be divided into lumped conceptual models and distributed models based on physical mechanism. The conceptual models mainly include the Stanford model (Crawford & Linsley 1966), Tank model (Sugawara *et al.* 1974) and Sacramento model (Burnash & Larry Ferra 1979). The distributed models mainly include MIKE SHE (MIKE System Hydrological European model) (Abbott *et al.* 1986), VIC (Variable Infiltration Capacity) model (Liang *et al.* 1994) and SWAT (Soil and Water Assessment Tool) model (Arnold *et al.* 1998).

SWAT has proven to be an effective and fruitful distributed hydrological model in many areas, including water resources management (Lerat *et al.* 2012; Khatun *et al.* 2018; Pandey *et al.* 2020; Abunada *et al.* 2021). Although the SWAT model can divide the basin into different units by the underlying surface attributes for simulation, considering that each sub-basin in a large basin has different hydrological characteristics, the heterogeneity for the spatial attributes of the basin should be considered in the process of the model, thereby reducing the uncertainty of the model and increasing the credibility of the simulation and prediction (Nkiaka *et al.* 2018). To improve the simulation accuracy of the model, multisite calibration of the basin is an important developmental direction of current hydrological simulation research (Jiang *et al.* 2015; Shrestha *et al.* 2016; Desai *et al.* 2021).

Many studies have discussed differences in simulation effects between multisite and single-site calibrations. In most of the studies, multisite calibration can effectively improve the performance of the SWAT model (Chiang *et al.* 2014; Shrestha *et al.* 2016; Nkiaka *et al.* 2018; Dong *et al.* 2019). Chiang *et al.* (2014) indicated that the hydrological process in SWAT would be better understood with multisite calibration, which could provide sounder model performance at sub-basins of different sizes. Similarly, Dong *et al.* (2019) argued that multisite calibration can adequately consider the spatiotemporal heterogeneity within the basin and provide greater predictive capacity compared with single-site calibration strategy. However, several studies have reported that multisite calibration results have almost the same performance as single-site calibration, with no significant improvement (Lerat *et al.* 2012). Moreover, the multisite calibration can be divided into simultaneous calibration and sequential calibration, and the comparison between them is inconclusive. Hasan & Pradhanang (2017) have shown that multisite simultaneous calibration is more effective for estimating the water quantity, while Wi *et al.* (2015) have the opposite conclusion. Although many applications have been implemented, it remains difficult to draw definite conclusions about the comparison of different calibration strategies considering the geographical properties of different basins and the relative positions of selected stations (Arab Amiri & Conoscenti 2017; Arab Amiri *et al.* 2017; Arab Amiri & Mesgari 2018).

In addition, studies on different calibration strategies focus on historical periods, and there rarely are the implications of these issues under future climate change (Wi *et al.* 2015), which is still an important issue all over the world (Simpkins 2017; Chen *et al.* 2019). Since different calibration strategies affect the capacity of hydrological models, exploring their differences under climate change will help to improve the accuracy of basin hydrological process prediction and better cope with the risks associated with climate change. The current related research tends to combine climate model data in the CMIP (Coupled Model Intercomparison Project Phase) with hydrological models (Jha *et al.* 2015; Wang & Kalin 2018; Sonkoué *et al.* 2019; Gao *et al.* 2021). At present, the CMIP has developed to the sixth phase (CMIP6), including 23 MIPs (Model Intercomparison Projects) and more than 100 climate models (https://pcmdi.llnl.gov/CMIP6/ArchiveStatistics/esgf_data_holdings), which employed a new set of concentrations, emissions and land-use scenarios (Eyring *et al.* 2016; Pascoe *et al.* 2020). Among the 23 MIPs included in CMIP6, ScenarioMIP provides for comprehensive research in different fields to better understand the impact of different scenarios on the physical process of the climate system and society (O'Neill *et al.* 2016; Zhao *et al.* 2021). To reduce the uncertainty of future prediction, multimodel ensemble data in ScenarioMIP was selected in this research instead of single-model data to project streamflow with the hydrological model (Ehsan *et al.* 2019).

Qujiang River is the link connecting Sichuan, Chongqing and other cities in southern China and plays a vital role in the development of the Yangtze River economic belt. However, the Qujiang River Basin has been suffering from frequent floods, and the frequency of extreme disastrous weather tends to increase in recent years, so it is more significant to analyze its hydrological process more accurately. CMADS (China Meteorological Assimilation Driving Datasets for the SWAT model), a high-resolution reanalysis of meteorological data, has proven to be effective in reproducing hydrological processes in the Qujiang River Basin (Song *et al.* 2020). CMADS shows stronger precipitation detection ability than the Tropical Rainfall Measuring Mission (TRMM) and the Integrated Multi-satellite Retrievals for GPM (IMERG) in the basin and can reproduce the gauged precipitation more accurately. Furthermore, the hydrological simulation accuracy of CMADS combined with the SWAT model is also higher than that of TRMM and IMERG, and the overall simulation performance is better than the gauged precipitation station data due to its high spatial resolution.

Due to the large spatial heterogeneity of the underlying surface in the Qujiang River Basin, the hydrological simulation only at the total outlet of the basin may not meet the accuracy requirements of water disaster prevention and control. Discussions on the application of different multisite calibration strategies to improve the capacity of hydrological simulation in this basin have been lacking in the past. Moreover, the implications of different calibration strategies on streamflow projections in the context of climate change have been rarely mentioned in multisite calibration studies of distributed hydrological models, especially using CMIP6 data.

Therefore, based on the CMADS-driven SWAT model, the study attempts to explore two aspects: (1) performance discrepancy of single-site calibration and multisite calibration strategies in simulating the hydrological process of Qujiang River Basin and (2) the influence of different strategies on future streamflow prediction with multimodel ensemble data in CMIP6. The first aspect of the work is to discuss the influence of different calibration strategies on the simulation results of hydrological processes in the Qujiang River Basin for the first time. The second aspect is to extend the research of multisite calibration strategy to the future period and use the latest updated CMIP6 data, which can provide a certain reference for the contents of related fields.

2. DATA AND METHODS

2.1. Study area

The Qujiang River Basin, part of Sichuan Basin in China, lies in between 106°16′ and 109° east longitude and 29°59′ and 32°46′ north latitude, covering an area of 38,900 km^2 (Figure 1). The basin has the characteristics of large topographic span, and the underlying surface attributes have obvious spatial heterogeneity. The terrain varies from the high-altitude mountain in the north (maximum 2,669 m) to the low-altitude plains in the south (minimum 187 m). The distribution of soil and land use can refer to Song *et al.* (2020). The mountainous area located in the upstream is mostly grassland and forest, while the plain area in the downstream is mostly agricultural land. The basin is a typical subtropical humid region, where precipitation mainly concentrates from May to September, and the annual average temperature is around 16.6 °C. To take the spatial heterogeneity of the whole basin into consideration as far as possible, four hydrological stations were selected from upstream to downstream of the basin, namely Bixi (BX) station, Qilituo (QLT) station, Fengtan (FT) station and Luoduxi (LDX) station. The streamflow data of the four stations are derived from the hydrological yearbook from 2008 to 2015, with a daily time resolution.

2.2. CMADS data

CMADS is a public dataset developed by Xianyong Meng from China Agricultural University using various reanalysis fields in the world and atmospheric assimilation system technology. This dataset covers East Asia, including precipitation, air temperature, air pressure, humidity and wind velocity data, which can be directly used for establishing SWAT models without any

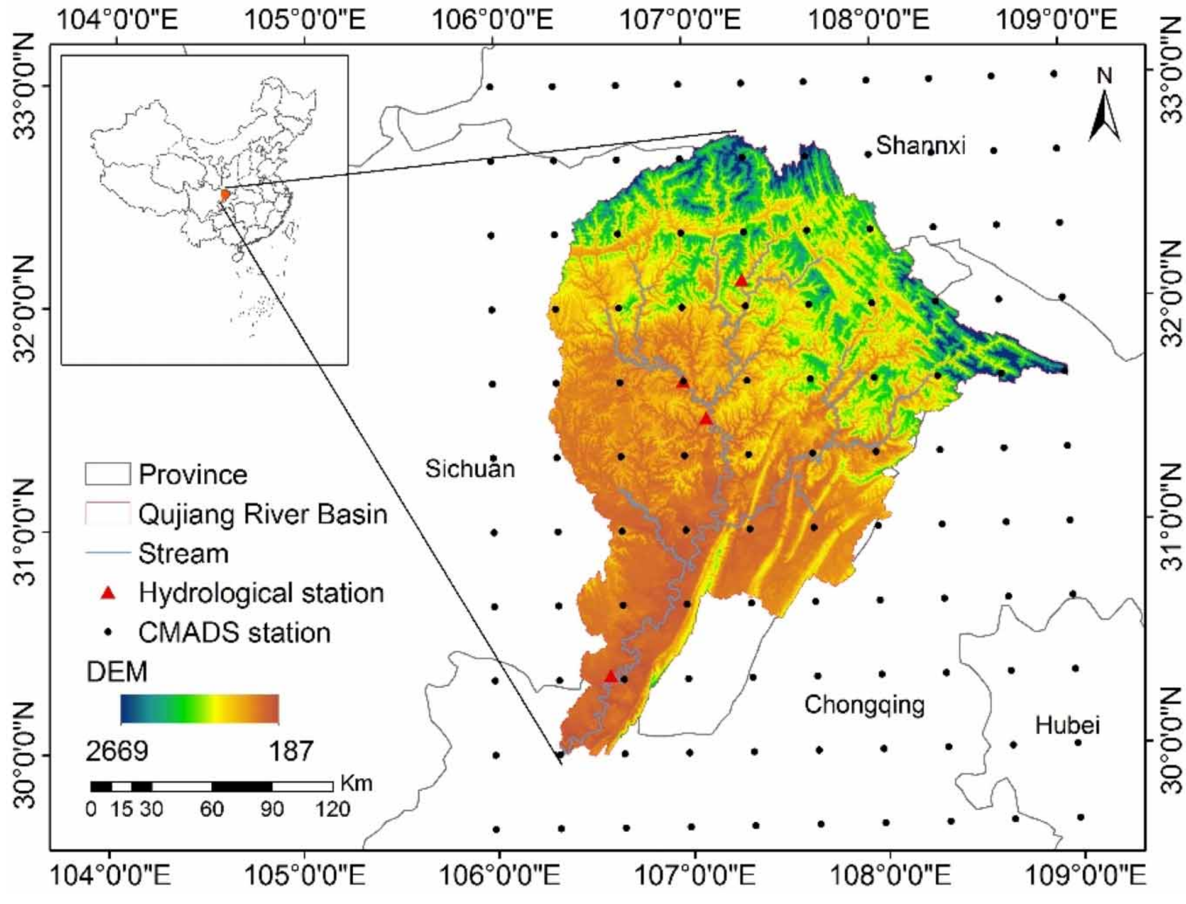

Figure 1 | Geographical location and distribution of hydrological stations and CMADS meteorological stations in Qujiang River Basin.

format conversion (Meng & Wang 2017). The v1.0 version of the dataset was used in this study, with a spatial resolution of 1/3° and a time span of 2008–2016. Considering the limited length of streamflow data obtained, the data of 110 sites from 2008 to 2015 were finally selected. The average annual precipitations of the four sub-basins were 1,254.78, 1,170.26, 1,177.27 and 1,134.46 mm, respectively, of which the upstream precipitation was slightly higher than that of the downstream.

2.3. Climate model data

For scenario MIP, we select three sets of available climate model data with high spatial resolution, namely BCC-CSM2-MR (Wu *et al.* 2019), EC-Earth3 (Massonnet *et al.* 2019) and MPI-ESM1-2-HR (Gutjahr *et al.* 2019), the basic information of which is shown in Table 1. Since the second aspect of this study is mainly to compare the differences of the three calibration strategies in streamflow prediction, the daily scale precipitation, minimum and maximum temperatures of the three climate models for 2021–2050 are analyzed. And three new future pathways of societal development called Shared Socioeconomic Pathways (SSPs) are considered: SSP1-2.6 (low-forcing scenario), SSP2-4.5 (medium-forcing scenario) and SSP5-8.5 (high-forcing scenario). Note that the three forcing scenarios are similar to the RCP2.6, RCP4.5 and RCP8.5 scenarios, respectively, in CMIP5 models (Riahi *et al.* 2017).

2.4. SWAT and SWAT-CUP

The SWAT model is a semi-distributed model, which can divide the study area into Hydrologic Research Units (HRUs) according to land use, soil and slope. Hydrological elements in the model will be calculated separately on each HRU, and then the results will be given separately at each sub-basin outlet. Finally, the total calculation amount of the whole basin can be obtained at the total outlet (Arnold 2012). Based on the river network generated based on the Digital Elevation Model (DEM), four hydrological stations (BX, QLT, FT and LDX) were manually added as the outlet, and finally, the Qujiang River Basin was divided into 28 sub-basins, and 733 HRUs were divided in combination with the classification of land use, soil and slope. The locations of the four sub-basins with four stations as outlets are shown in Figure 2. After loading meteorological data, SWAT models with daily and monthly scales were established, with 2008 as the warm-up period, 2009–2012 as the calibration period and 2013–2015 as the validation period.

SWAT-CUP is a semi-automatic calibration software, through which users can set key parameters and their ranges in the SWAT model according to their requirements and can be calibrated separately or simultaneously for different sub-basins (Abbaspour 2012). Based on this, the model calibration method used in this study is the SUFI-2 algorithm in SWAT-CUP (Abbaspour *et al.* 2004). The SWAT-CUP software can directly give the results of some evaluation indexes, among which R^2 (coefficient of determination), NSE (Nash–Sutcliffe efficiency) and PBIAS (percentage bias) are the most widely used to evaluate the results of the model (Song *et al.* 2020).

R^2 can well reflect the consistency between the observed value and the simulated value, which ranges from 0 to 1. The higher the value, the closer the result is to the best-fit line (Ghoraba 2015). NSE is a normalized processing method to reflect the relative quantity between observed values and simulated values, and it is a good objective function to reflect the overall fitting degree of hydrological process lines, with an optimal value of 1 (Van Liew *et al.* 2007). PBIAS measures the average trend of the simulated value relative to the observed value to judge the overestimation or underestimation of the model average, with low absolute values indicating accurate model simulation (Gupta *et al.* 1999). The good R^2 and NSE indicate that the model might capture the pattern of flow but do not always prove the model to be performing well, so PBIAS should be combined to understand the magnitude of deviation of the simulated streamflow from the observed. Moriasi *et al.* (2007) pointed out that for the SWAT model, when NSE >0.5 and PBIAS $\pm 25\%$, the simulation results could be considered satisfactory.

Table 1 | Basic information of climate models used in this study

Model name	Institute	Country	Resolution (lon. × lat.)
BCC-CSM2-MR	Beijing Climate Center	China	1.1° × 1.1°
EC-Earth3	EC Earth Consortium	Europe	0.7° × 0.7°
MPI-ESM1-2-HR	Max Planck Institute for Meteorology	Germany	0.9° × 0.9°

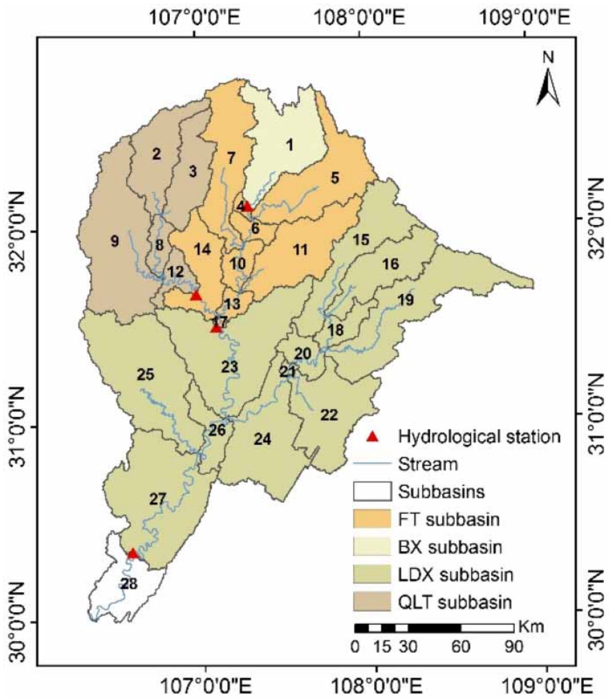

Figure 2 | Sub-basin number of Qujiang River Basin.

2.5. Model calibration strategies

Among the four stations selected in this study, BX, FT and LDX stations are all located on the mainstream of Qujiang River. The BX station is located in the mountainous area of the upper reaches and only controls the No. 1 sub-basin; the FT station is located in the transitional zone between the mountainous area and the plain in the middle reaches; the LDX station is located in the downstream plain area close to the total water outlet of the basin. The QLT station is also located in the middle reaches of the basin, close to the FT station, but it is located in the tributary and relatively independent of the BX station. Both QLT and BX sub-basins are nested in the FT sub-basin and further nested in the LDX sub-basin. The selection of station location in this way can comprehensively consider the spatial heterogeneity of the underlying surface of the basin, and the QLT station can be used to discuss the influence of a multi-station calibration strategy under different hydraulic connections. For the calibration of the four sub-basins in the SWAT model constructed in the Qujiang River Basin, three calibration strategies are set up by selecting different calibration stations with the help of SWAT-CUP software. The initial ranges of parameters of the three strategies are the same.

(1) *Single-site calibration strategy (S1)*. In this strategy, only the LDX station is selected for model calibration and validation, and a set of optimal model parameter values representing the total outlet of the basin are obtained. Then, the optimal parameter values are substituted into the other three stations in the basin to simulate their streamflow from 2009 to 2012 so as to compare with the results of the other two calibration strategies.
(2) *Multisite simultaneous calibration strategy (S2)*. All four hydrological stations are calibrated and validated with the same set of parameter values. Finally, a set of optimal parameter values that can make the simulation results of the four stations achieve the good performance simultaneously are obtained.
(3) *Multisite sequential calibration strategy (S3)*. To fully consider the spatial heterogeneity within the basin, this strategy separately calibrates and validates the streamflow of hydrological stations from upstream to downstream and obtains four sets of parameter values corresponding to four hydrological stations. Because the basic conditions of the model remain unchanged, the streamflow simulation results of the three calibration strategies are only related to the parameter values, so the differences of the three calibration strategies are mainly reflected in the optimal parameter values and the corresponding streamflow results of the four stations.

3. RESULTS AND DISCUSSION

3.1. Daily and monthly streamflow results

The evaluation index can judge the fitting degree between the model result and the observed value overall, while the streamflow hydrograph can reflect the difference in the time series more intuitively. Therefore, the evaluation of SWAT model results was mainly carried out in these two aspects. Table 2 lists all the evaluation index values of three strategies in the calibration period.

It is obvious from the daily scale simulation results that S3 performs better than the other two strategies. The simulated and observed values of the four stations of S3 all have a high degree of fittings, with R^2 and NSE between 0.65 and 0.72. In contrast, the R^2 of S2 are in the range of 0.56–0.7, and NSE are in the range of 0.52–0.69. Since the LDX sub-basin was calibrated independently in both S1 and S3, the results of this sub-basin are consistent under the two strategies. However, the results of the three sub-basins in the middle and upper reaches of the basin are not satisfactory under S1, with a low degree of fitting with the measured values, and R^2 and NSE are all below 0.5. Meanwhile, comparing S1 with S2, it can be found that when the four sub-basins are calibrated simultaneously, the results of the LDX sub-basin become worse, but the results of the other three sub-basins are improved. This indicated that in the calibration process of S2, the model would adjust the parameters according to the results of the four sub-basins at the same time and consider the hydraulic connection between the four stations so as to reduce the weight of LDX stations and correspondingly strengthen the weight of the other three stations. In terms of simulation deviation, the three strategies have different degrees of overestimation on the whole. Among them, the deviations of four stations of S3 are the least obvious, with PBIAS between 2.5 and 11.35%, which are nearly half of those of S2.

The SWAT model has obtained very satisfactory results at the monthly scale in the study area. As shown in Table 2, R^2 and NSE of three strategies are above 0.93 and 0.92, respectively, and PBIAS values are also within a relatively small range. The performance differences of the three strategies in the four sub-basins are consistent with the daily scale. S3 can achieve the highest R^2 and NSE and the lowest PBIAS. Meanwhile, S2 has a slight improvement in all three stations except LDX comparing with S1. However, due to very good simulation results at the monthly scale, the discrepancies between the three strategies are not as obvious as those at the daily scale.

In terms of streamflow hydrograph (Figures 3 and 5), streamflow and precipitation keep a high consistency, and three strategies can basically capture the hydrological process. The differences are mainly reflected in the gap between simulated and observed values. Since the measured data from 2009 to 2012 were used to verify the model at FT, QLT and BX stations in S1 strategy, there are no streamflow simulation results of S1 in the validation period (2013–2015) of hydrological process line, and only the results of S2 and S3 are displayed.

It can be seen from Figure 3 that in daily scale simulation, regardless of the calibration and validation periods, the SWAT model has insufficient simulation ability when using three calibration strategies in the peak part. The result of S3 is closer to

Table 2 | Evaluation of streamflow simulation in the calibration period of daily and monthly scales

	S1			S2			S3		
	R^2	NSE	PBIAS (%)	R^2	NSE	PBIAS (%)	R^2	NSE	PBIAS (%)
Daily									
LDX	0.66	0.65	11.50	0.56	0.52	19.50	0.66	0.65	11.50
FT	0.43	0.32	14.60	0.68	0.66	19.50	0.72	0.72	10.70
QLT	0.43	0.40	9.40	0.65	0.57	13.40	0.67	0.66	7.80
BX	0.50	0.34	2.00	0.70	0.69	5.80	0.71	0.70	2.50
Monthly									
LDX	0.96	0.96	7.40	0.96	0.92	19.60	0.96	0.96	7.40
FT	0.97	0.96	10.60	0.98	0.96	18.30	0.98	0.98	2.50
QLT	0.94	0.94	4.10	0.95	0.92	12.60	0.95	0.95	3.80
BX	0.93	0.93	−0.60	0.94	0.94	4.40	0.95	0.95	0.40

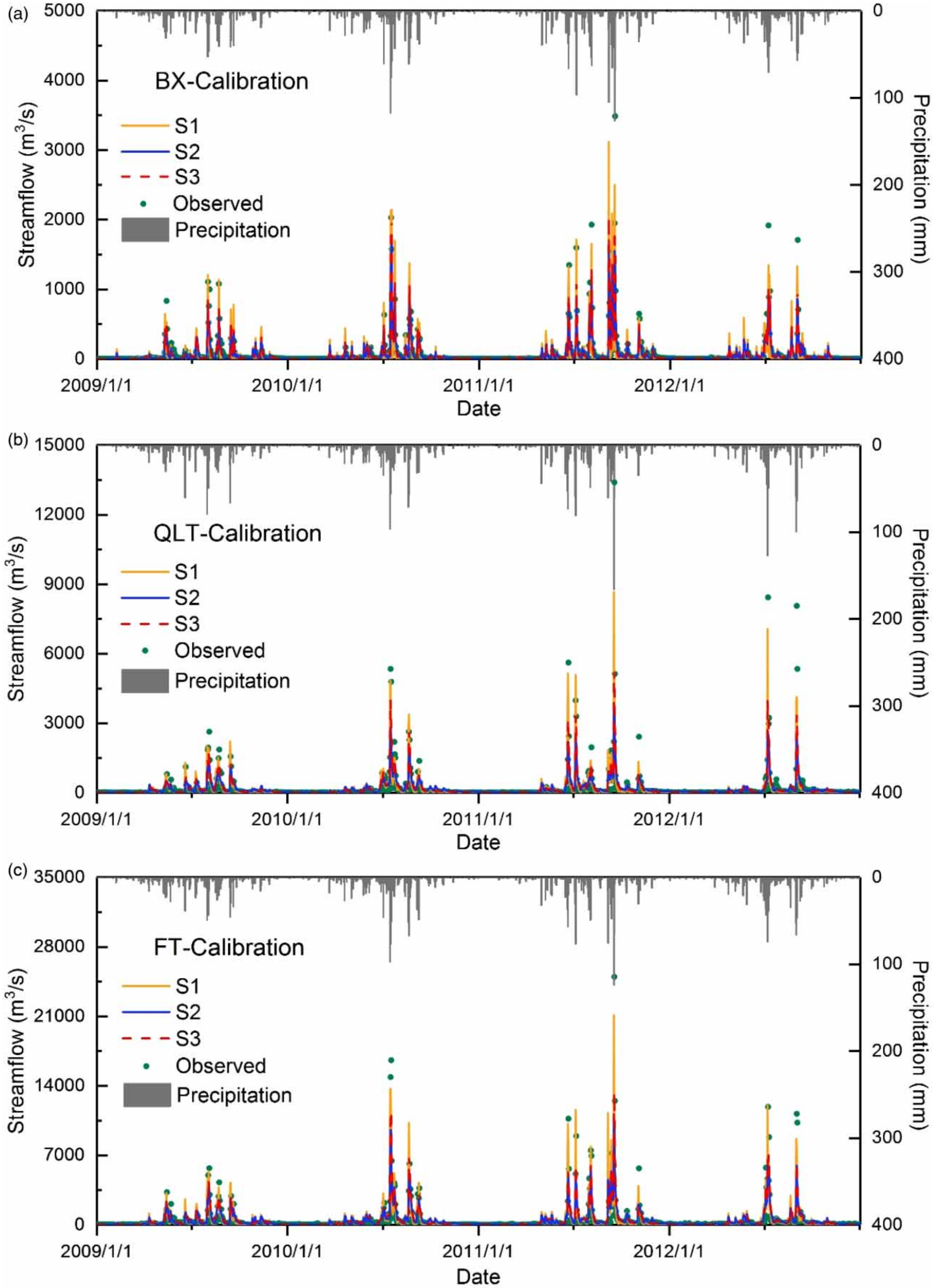

Figure 3 | Hydrograph of the daily streamflow and precipitation in the calibration and validation periods: (a)–(d) calibration period; (e)–(h) validation period. (*continued.*)

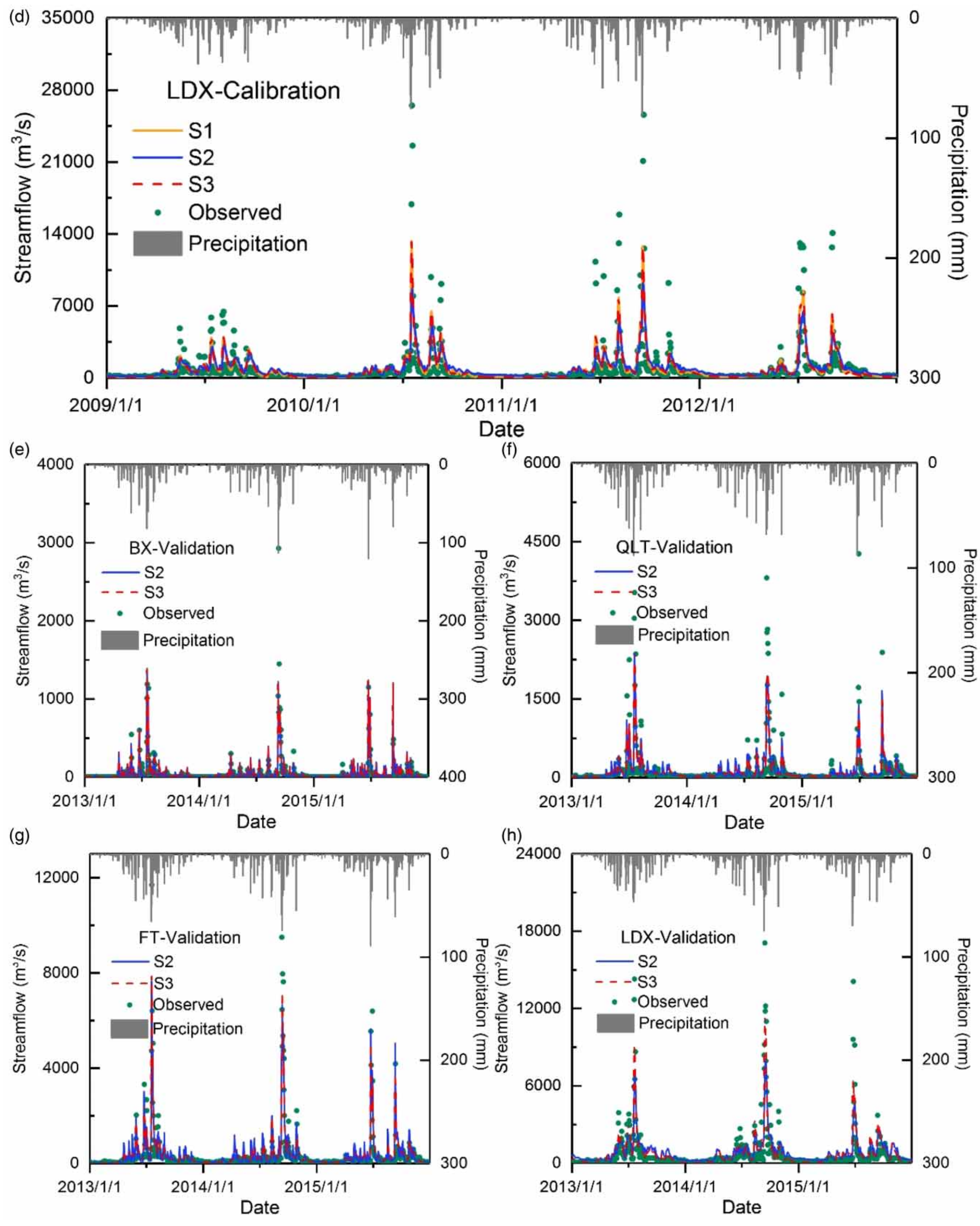

Figure 3 | Continued

the observation value than that of S2, but there are still obvious disparities. Although S1 obtained an unsatisfactory evaluation index value, the recurrence of S1 on the peak part at BX, QLT and FT stations is surprisingly good in 2009–2012. The reason for this phenomenon may be that when the parameter values obtained only in the calibration of LDX stations are applied to the other three stations, the parameter values are more conducive to enhancing the response of streamflow to extreme precipitation due to the differences in hydrological characteristics among sub-basins.

Because of the large disparity between the maximum and minimum streamflow at each station, the daily streamflow hydrographs are not convenient to compare other parts except the peak, so we divide the streamflow results into low-flow, high-flow and middle-flow parts according to the observed values, which are lower than 25th and higher than 75th and 25th–75th percentiles of daily streamflow. The comparison of annual streamflow in different parts using the three strategies is shown in Figure 4. It can be clearly seen that in the middle- and low-flow parts, all the three strategies have overestimated the streamflow, which is consistent with PBIAS values. In the low-flow part, the simulation results with three strategies are close at the BX station, while the streamflow results of S3 are between those of S1 and S2; in the other three stations, the overestimation of S2 is the most obvious, and the simulated value of S3 is closer to the observed value, and the results of S1 are between those of S2 and S3 at QLT and FT stations. In the middle-flow part, the performances of the three strategies are similar to those of the low-flow part.

The results of S3 are still closer to the observed value, but its gap with those of S2 has been narrowed. In the high-flow part, both overestimation and underestimation coexist in the three strategies. The results of S3 are slightly higher than those of the other two strategies, and its simulated values are closest to the observed value at the LDX station. According to the results of the three parts, the performance of S3 is the most stable in the whole process; in other words, it can have the best simulation effect in the middle and low-flow parts and also performs well in the high-flow part. However, S2 performs well only in the high-flow part, while the other parts show obvious overestimation phenomenon; although S1 is stronger than S3 in the simulation peak flow, it is not as good as S3 on the whole. The results of this part are consistent with the results of PBIAS. It can be seen that S3 can not only improve the index value of the model at the daily scale but also reduce the deviation in the whole flow process.

As the evaluation index of the SWAT model at the monthly scale has achieved very satisfactory results, the monthly streamflow hydrographs show a high fitting degree between the simulated value and the measured value, and the simulation

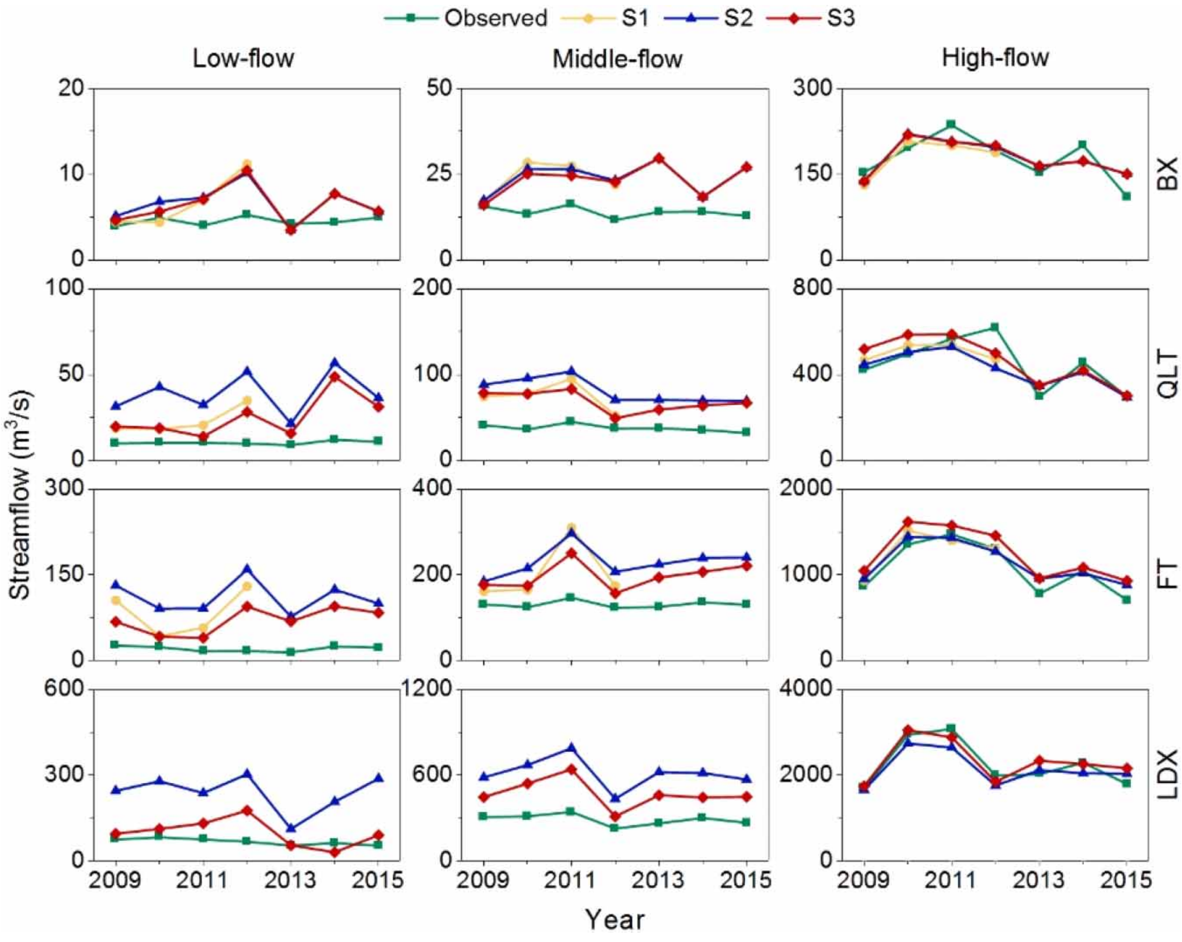

Figure 4 | Comparison of annual streamflow in low-flow, middle-flow and high-flow parts using the three strategies.

performances on the peak part are much better than those at the daily scale (Figure 5). During calibration and validation periods, the gaps between the simulated values and the observed values in the four stations of S3 both are the smallest, which not only have good performance in the peak part but also can more accurately restore the base flow. The results of S2 are most significantly underestimated in the peak part, while they are generally overestimated in the base flow part.

According to the above results of streamflow simulation in the Qujiang River Basin, the performance of the multisite sequential calibration strategy is significantly superior to that of the single-site calibration and the multisite simultaneous

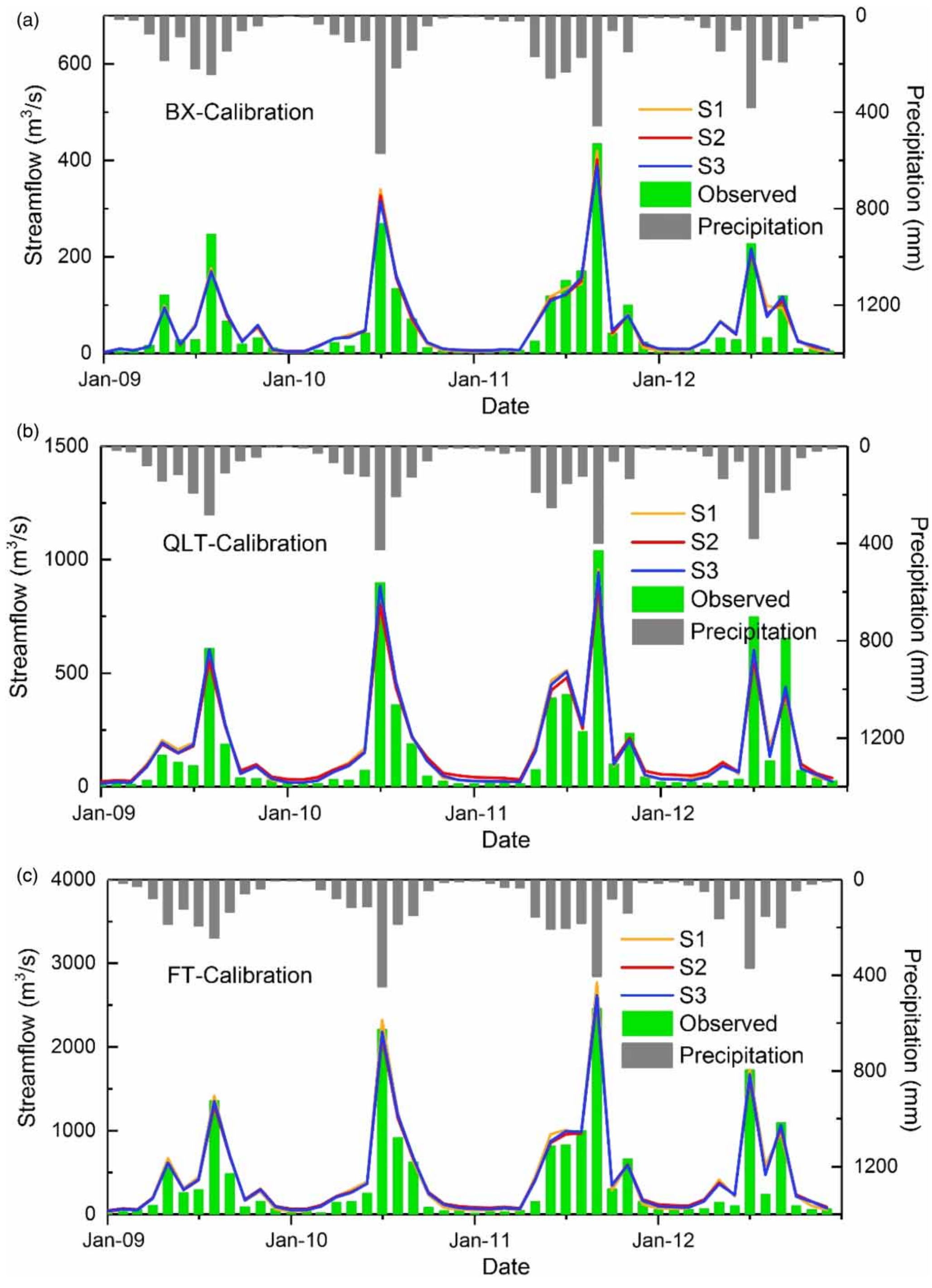

Figure 5 | Hydrograph of the monthly streamflow and precipitation in the calibration and validation periods: (a)–(d) calibration period; (e)–(h) validation period. (*continued*.)

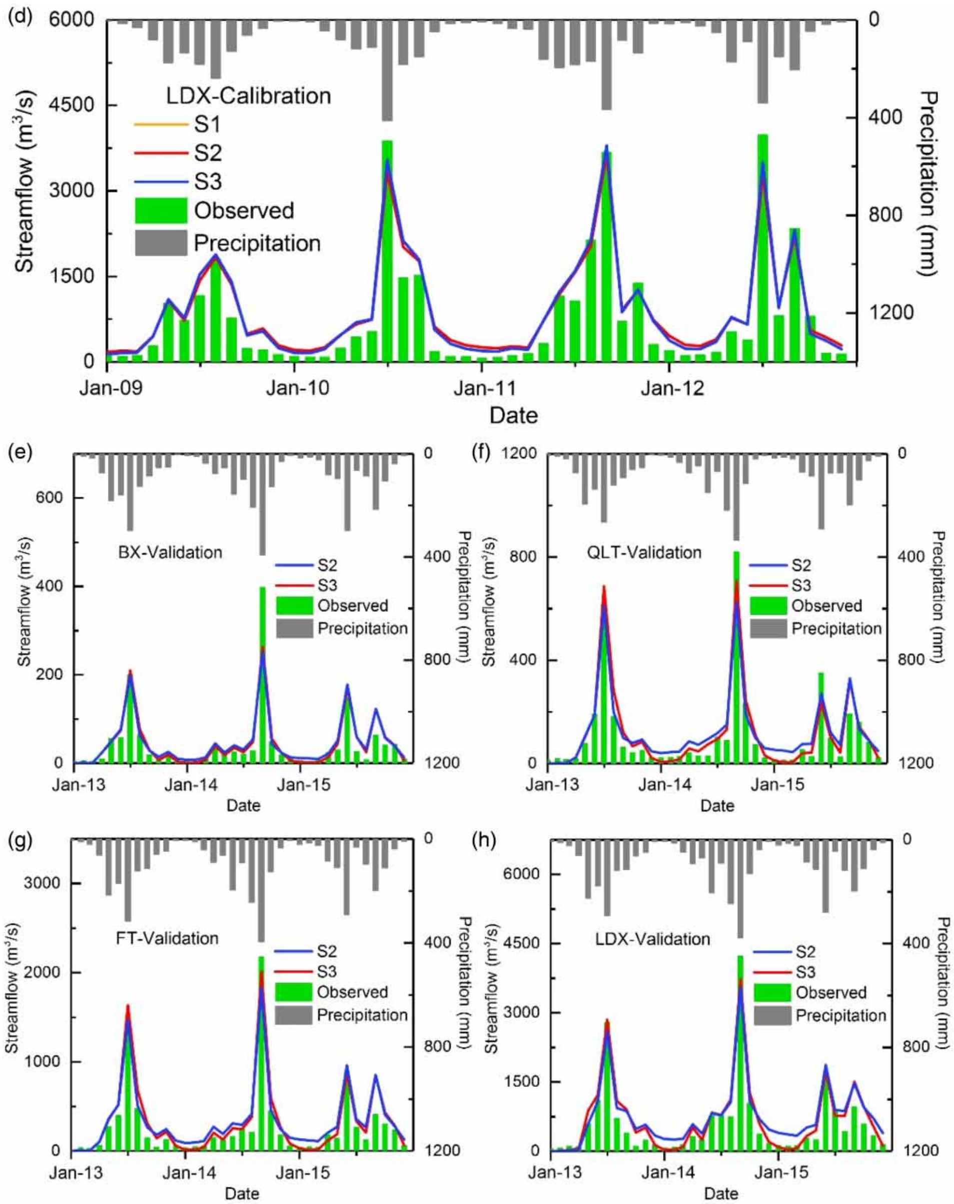

Figure 5 | Continued

calibration strategies, especially at the daily scale, which is consistent with the views of most of the multisite calibration studies (Dong *et al.* 2019; Franco *et al.* 2020). S3 strategy can not only improve the degree of fit between the simulation results and the measured values but also effectively reduce the deviation of the simulation results. Although multisite sequential calibration will cause the variation of the parameters on the basin level in the process of progressive calibration from upstream to downstream (Migliaccio & Chaubey 2007), S3 strategy calibrated each station and obtained the most optimal parameter values so as to more accurately restore the streamflow process of the corresponding sub-basin.

For the single-site calibration strategy which only calibrates the model at the total outlet of the basin, its performance differs greatly from that of the multisite calibration strategies at the daily and monthly scales. At the daily scale, the simulation performances of the three stations within the Qujiang River Basin of S1 strategy are far lower than those of S2 and S3 strategies, which do not reach a satisfactory level. At the monthly scale, simulation performance at some stations when using S1 strategy is better than when using S2 strategy. The reason for this phenomenon may be the differences of hydrological elements between the four sub-basins at the daily scale, which are more obvious than those at the monthly scale, and the requirements for the accuracy of model parameters are higher during daily scale simulation, which makes it more difficult for the calibration parameters at the LDX station to meet the requirements of daily scale streamflow simulation in the internal sub-basins.

In monthly streamflow simulation, the hydrological elements are processed on a monthly average basis, which reduces the differences between the four sub-basins. Coupled with the high accuracy of meteorological data, the single-site calibration strategy can also achieve a good simulation effect within the basin, so the improvement brought by the multisite calibration strategy is relatively insignificant. This suggests that although the single-site calibration strategy can meet the requirements of hydrological process simulation in the basin in some cases, further validation is required for stations within the basin.

The simulation performances of the multisite simultaneous calibration strategy at all four stations are slightly worse than those of the multisite sequential calibration strategy, but the difference between the single-site calibration strategies varies by time scale as mentioned above. In the monthly streamflow simulation, the results of LDX and QLT stations both show a decrease in fitting degree when using S2 strategy, which is shown that the NSE values of the two stations are lower than those of the single-site calibration strategy, and this phenomenon is also reflected in the study of Nkiaka *et al.* (2018). Owing to the multisite simultaneous calibration strategy which will take the average NSE of the four stations as the objective function and increase the weight of the internal sub-basin, the result of the LDX station is worse than that of the single-site calibration strategy. Moreover, the relative deviation of S2 simulation results is the largest among the three strategies, especially for the low-flow part at the daily scale (Figure 4). While some studies pointed out that the multisite simultaneous calibration strategy can take into consideration the basin spatial heterogeneity and effectively enhance the ability of simulating the SWAT model (Wi *et al.* 2015; Hasan & Pradhanang 2017), the results of this study and Nkiaka *et al.* (2018) show that because of the difference of the study area, the improvement of the strategy in monthly scale streamflow simulation still has some limitations.

Although the multisite sequential calibration strategy can achieve the best simulation performance in the Qujiang River Basin and reduce the risk of water disasters caused by model deviation to the basin, the need to calibrate each station successively increases the time cost of simulation. In view of the fact that the multisite simultaneous calibration strategy can calibrate multiple stations at the same time, the hydraulic connection and information exchange between the stations are also considered. Therefore, for the practical application, appropriate model calibration strategy should be selected according to the actual situation and management measures of the basin.

3.2. Parameter sensitivity

The performance of the model is determined by the value of the parameter set. The difference between three calibration strategies is mainly reflected in the selection of parameters, which in this study are the different optimal values of the same parameter. In the model calibration process, SWAT-CUP can not only give the optimal value of parameters but also evaluate the sensitivity of the results to different parameters. In this paper, the differences among the parameters of three strategies are discussed by taking the parameter results of the monthly scale as an example (Table 3).

Table 3 | Parameter values and sensitivity rank when using the three strategies

	S2		S3-BX		S3-QLT		S3-FT		S3(S1)-LDX	
Parameter	Value	Rank	Value	Rank	Value	Rank	Value	Rank	Value	Rank
CN2	0.18	2	−0.73	1	0.12	4	0.02	1	0.10	4
ALPHA_BF	0.94	1	0.26	9	0.94	1	0.68	2	0.76	1
SOL_AWC	0.28	9	0.46	7	0.42	8	0.29	7	0.16	8
SOL_K	1.01	4	1.86	6	2.18	6	1.11	9	2.26	9
ESCO	0.58	10	0.79	5	0.87	7	0.43	8	0.66	10
GW_REVAP	0.17	7	0.05	4	0.12	9	0.19	10	0.19	6
SFTMP	2.52	5	−9.80	2	−4.68	3	6.04	4	−3.40	5
SMTMP	4.97	8	−0.36	3	−11.48	5	1.96	5	−0.28	7
CH_K2	17.38	3	46.75	10	0.44	2	43.65	3	28.55	3
CH_N2	0.26	6	0.30	8	0.29	10	0.20	6	0.11	2

Focus on the results of S3, CN2 (initial SCS runoff curve number for moisture condition II), one of the important parameters in the process of model streamflow generation, is the most sensitive parameter in the BX and FT sub-basins, but only ranks fourth in the QLT and LDX sub-basins. Similarly, ALPHA_BF is highly sensitive and has a high value in QLT, FT and LDX sub-basins, which illustrates the rapid response of groundwater flow to recharge. Therefore, the groundwater has a strong impact on the streamflow generation in the midstream and downstream of the basin. In BX sub-basins with higher altitude, the sensitivity of SFTMP and SMTMP ranked second and third, respectively, indicating that snowfall is an important factor in the streamflow yield of this basin. Although the average elevation of the QLT sub-basin is relatively low, the high sensitivity of SFTMP also reflects the importance of snowfall in this basin. In the S1 and S2 strategies, the parameters with high sensitivity are mainly related to groundwater, which reduces the important level of snowfall on the streamflow of BX and QLT sub-basins. The difference of parameter sensitivity and value between four sub-basins also explains the necessity of separate parameterization for each sub-basin, and the same set of parameter values cannot represent the whole basin.

3.3. Future climate and streamflow projection

Due to the different resolutions of the three climate models, precipitation, minimum and maximum temperature data under three forcing scenarios were extracted according to their respective resolutions, and then the three meteorological elements of the four sub-basins were calculated by using the Thiessen polygon method, respectively. The average value of the three climate models in 2021–2050 was taken as the future climate characteristics of the four sub-basins, as shown in Table 4. The comparisons of projection results among the three forcing scenarios in the four sub-basins are consistent; that is, the precipitation and temperature of SSP5-8.5 are higher than those of the other two, followed by SSP1-2.6, which is about 30 mm higher than that of SSP2-4.5, and the air temperature is about 0.11 °C higher. For the comparisons of the four sub-basins, the BX sub-basin has the highest precipitation and the LDX sub-basin has the highest temperature, while the precipitation and temperature of the QLT and FT sub-basins are both very close to each other and lower than those of the other two sub-basins.

After extracting precipitation, minimum and maximum temperature data from different climate models, we loaded them into the SWAT model and projected future (2021–2050) streamflow with parameter values obtained from three calibration strategies (S1, S2 and S3), respectively. In addition, underlying surface properties of the sub-basins remain unchanged. To make the results as reliable as possible, we chose the monthly scale SWAT model with very satisfactory simulation results in the historical period, and the annual projection results are shown in Figure 6. Although the precipitation and temperature in the three forcing scenarios have the same law, the differences of streamflow projection results in the four sub-basins still exist apparently due to the unique underlying surface properties of each sub-basin. However, in general, the lower limit of streamflow results of SSP1-2.6 during 2021–2050 is basically consistent with that of the other two forcing scenarios, and the upper limit is obviously higher, indicating that the study area will have the most abundant water resources under this forcing scenario.

By comparing the streamflow projection results of the three calibration strategies, it can be found that the discrepancies among the three strategies are alike under three forcing scenarios in the same station, while there are diverse comparison relationships in four stations. The similarity phenomenon shown in the four stations is that the streamflow projection result by using S2 is the smallest among the three strategies, and the relative discrepancy between S2 and S1 is basically unchanged. This is because the discrepancies are only related to parameter values, and the same set of parameter values in the four sub-basins were adopted in S1 and S2. Because of S3 taking different parameter set values in different sub-basins, the discrepancies between S3 and the other two strategies vary across the four stations. At the BX station, the highest

Table 4 | The annual mean precipitation, maximum and minimum temperatures of four sub-basins in 2021–2050 under the scenarios of SSP1-2.6, SSP2-4.5 and SSP5-8.5

	Precipitation (mm)			T_{max} (°C)			T_{min} (°C)		
	SSP1-2.6	SSP2-4.5	SSP5-8.5	SSP1-2.6	SSP2-4.5	SSP5-8.5	SSP1-2.6	SSP2-4.5	SSP5-8.5
BX	1,106.88	1,074.39	1,117.25	18.42	18.33	18.56	9.62	9.56	9.90
QLT	1,030.03	1,000.71	1,057.22	18.23	18.12	18.32	9.32	9.22	9.58
FT	1,038.28	1,007.05	1,059.48	18.14	18.04	18.24	8.81	9.11	9.45
LDX	1,047.17	1,018.62	1,056.11	18.74	18.61	18.85	9.92	9.80	10.18

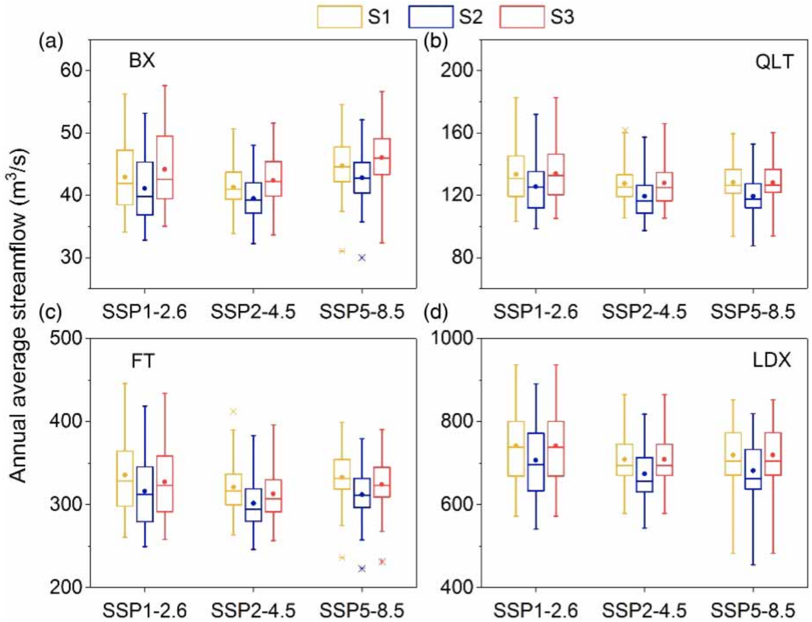

Figure 6 | Annual average streamflow projection results using the three strategies (S1, S2 and S3) in 2021–2050 under SSP1-2.6, SSP2-4.5 and SSP5-8.5. The solid circle represents the mean value, the middle line in the box represents the median value and each box ranges from the lower (25th) to upper quartile (75th).

and average values of S3 are higher than those of S1; at the QLT station, the results of the two strategies are basically the same; at FT stations, the results are opposite to those of the BX station, which means that the maximum and average values of S1 are slightly higher than those of S3; at the LDX station, for S1 and S3 having the same parameter values, the streamflow projection results are completely consistent.

In addition, Figure 7 shows the projected monthly streamflow results in 2021–2050 under three forcing scenarios using the three strategies. For the same forcing scenario, the annual distributions of the four stations are very similar. Under SSP1-2.6 and SSP2-4.5 scenarios, the streamflow mainly concentrated from May to August, and the average streamflow in May is higher than that in other months. Under the SSP1-2.6 scenario, the maximum streamflow value in May increases suddenly compared with that in April and then decreases month by month. Meanwhile, the minimum streamflow value in May is also significantly higher than that in other months. Under the SSP2-4.5 scenario, the minimum streamflow value in May is also significantly higher than that in other months, but the maximum streamflow occurs in July. Under the SSP5-8.5 scenario, the streamflow is mainly concentrated from May to September, and the average streamflow in these months is basically the same, while the maximum and minimum streamflow vary with different stations.

As with annual mean streamflow, the discrepancies among the three calibration strategies are consistent under different forcing scenarios. As can be seen from the above, the annual average streamflow projection results by using S2 are the smallest. In terms of monthly average streamflow, this phenomenon is reflected in that the monthly average streamflow from April to October is slightly lower than that using S1 and S3, and the same is true for the maximum and minimum streamflow values. At the BX station, the differences within the year between S1 and S3 are mainly reflected from June to September, indicating that the streamflow of the former is slightly lower than that of the latter, while the streamflow in other months is basically flat. At the QLT station, when S1 and S3 are used, the monthly average streamflow is basically the same, which is consistent with the annual average streamflow. At FT and LDX stations, the months in which the streamflow using S1 is greater than that using S3 are mainly from April to August, while the streamflow in other months is basically flat.

Different calibration strategies will lead to obvious disparities in the future streamflow prediction of the Qunjiang River Basin. On account of S3 having the strongest simulation ability in the historical period, it is deemed that it can most accurately project the future streamflow of four sub-basins. Therefore, under the meteorological conditions of the three climate models, S2 will underestimate the water resources in the Qujiang River Basin, and the distinctions between S1 and S3 also reflect that S1 will affect the correct estimation of future water resources in the basin. Because the average values of

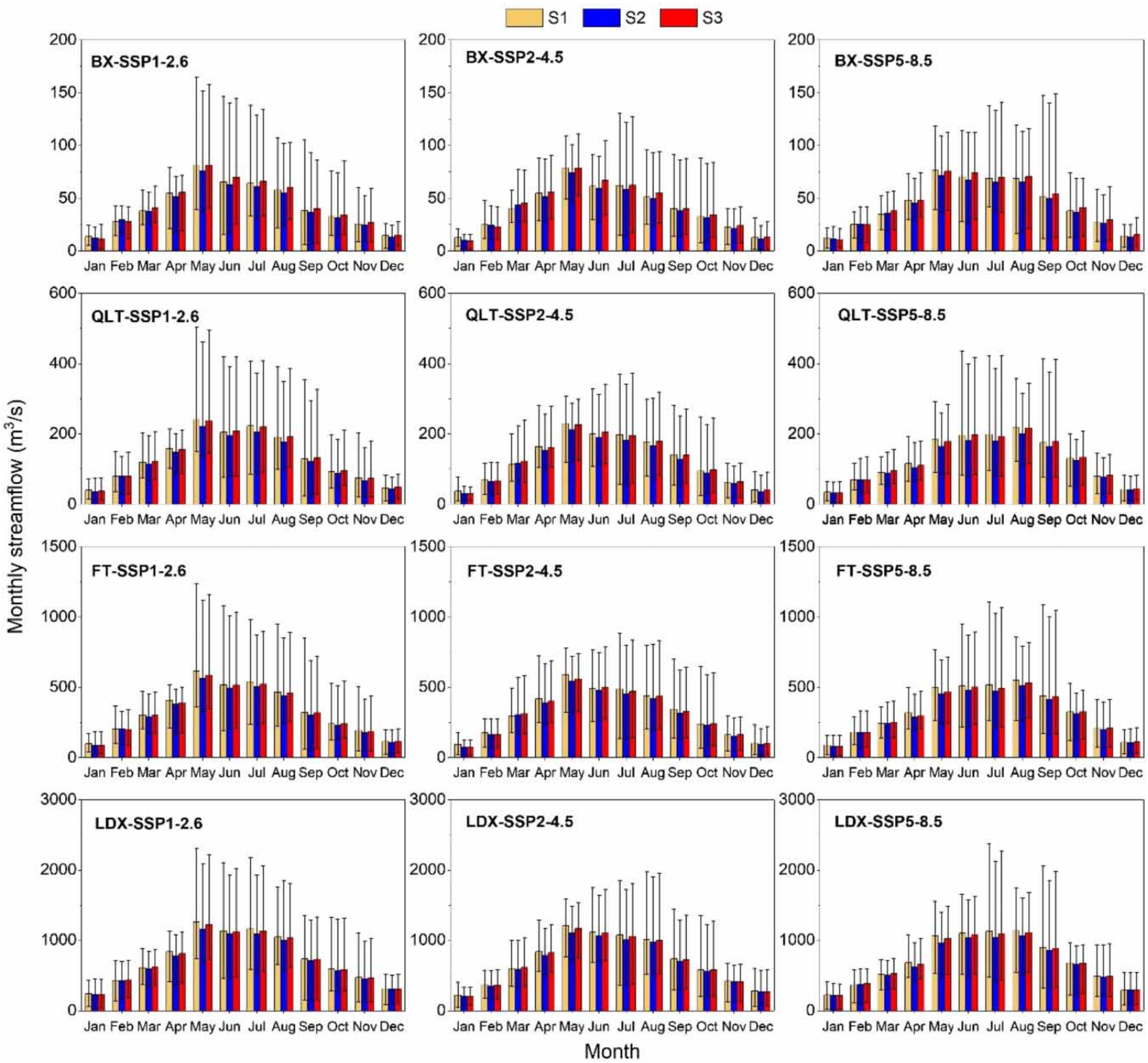

Figure 7 | Monthly average streamflow projection results using the three strategies (S1, S2 and S3) in 2021–2050 under SSP1-2.6, SSP2-4.5 and SSP5-8.5. The error bars represent the streamflow ranges resulting from 2021 to 2050.

three climate model data are selected as future climate data input to the model, the same climate variables are kept when using different calibration strategies to project streamflow, so the differences in results mainly come from the different input values of parameters of the three calibration strategies. In other words, different calibration strategies in the historical period will bring certain risks to the hydrological analysis in the future. Therefore, it is necessary to be cautious when selecting calibration strategy for streamflow projection under future climate conditions.

4. CONCLUSION

Based on the SWAT model driven by CMADS, in this study, we set up single-site calibration, multisite simultaneous calibration and multisite sequential calibration strategies using four hydrological stations in the Qujiang River Basin. The objective is to investigate the discrepancies of the three calibration strategies in streamflow simulation and their effects on streamflow projection in the future period. When using S3, R^2 and NSE of the four stations at the daily scale calibration period can reach 0.65–0.72, while PBIAS is between 2.50 and 11.50%; the R^2 and NSE values at the monthly scale calibration period can reach 0.95–0.98, while PBIAS is between 0.40 and 7.4%. Based on the calibration results of three strategies at the daily scale and the monthly scale, it is suggested that using the S3 can fully consider the spatial heterogeneity of the basin and more accurately figure out the hydrological process of each sub-basin. Although the hydraulic connections between all stations are considered when using S2, the overall performance still lags that of S3 to some extent. S1 takes the downstream station as the highest priority, which cannot fully represent the hydrological process of the internal sub-basins, especially at the daily scale. The parameter sensitivity analysis at the monthly scale also found that the important level of the parameters varied between different sub-basins, while S1 and S2 both ignored this difference. Therefore, it is necessary to conduct separate calibration for different sub-basins to obtain the most accurate hydrological response relationship.

Under the meteorological conditions of the three climate models (BCC-CSM2-MR, EC-Earth3 and MPI-ESM1-2-HR), the distinctions between streamflow projection results of three strategies at four stations are mainly related to the selection of parameter values. This study shows that S2 will underestimate the future water resources in the study area compared with the other two strategies, especially from April to October, and caution is needed when using S1 for streamflow projections. Given the best performance of S3 in the historical period, the use of S3 to project streamflow of sub-basins under climate change conditions may be the least risky.

Although there are still some deficiencies in this study, the comparison of the results is not detailed enough, it can provide a reference for the comparative study of multisite calibration strategies and the improvement of streamflow simulation and projection accuracy of the hydrological model in the basin with complex terrain. Since only three sets of climate model data were used in this study and meteorological data were directly extracted without downscaling, more sets of climate model data should be used for projection analysis in future work and downscaling and correction processing should be carried out to obtain more accurate results.

ACKNOWLEDGEMENTS

This research was funded by the National Key R&D Program of China (No. 2017YFC0406004) and the NSFC (No. 41271004).

DATA AVAILABILITY STATEMENT

All relevant data are included in the paper or its Supplementary Information.

REFERENCES

Abbaspour, K. C., Johnson, C. A. & van Genuchten, M. T. 2004 Estimating uncertain flow and transport parameters using a sequential uncertainty fitting procedure. *Vadose Zone Journal* **3** (4), 1340–1352.

Abbaspour, K. C. 2012 *SWAT-CUP-2012, SWAT Calibration and Uncertainty Program – A User Manual*. Swiss Federal Institute of Aquatic Science and Technology, Duebendorf, Switzerland.

Abbott, M. B., Bathurst, J. C., Cunge, J. A., O'Connell, P. E. & Rasmussen, J. 1986 An introduction to the European Hydrological System – Systeme Hydrologique Europeen, SHE, 1, history and philosophy of a physically-based, distributed modelling system. *Journal of Hydrology* **87** (1), 45–59.

Abunada, Z., Kishawi, Y., Alslaibi, T. M., Kaheil, N. & Mittelstet, A. 2021 The application of SWAT-GIS tool to improve the recharge factor in the DRASTIC framework: case study. *Journal of Hydrology* **592**, 125613.

Arab Amiri, M. & Conoscenti, C. 2017 Landslide susceptibility mapping using precipitation data, Mazandaran Province, north of Iran. *Natural Hazards* **89** (1), 255–273.

Arab Amiri, M., Mesgari, M. S. & Conoscenti, C. 2017 Detection of homogeneous precipitation regions at seasonal and annual time scales, northwest Iran. *Journal of Water and Climate Change* **8** (4), 701–714.

Arab Amiri, M. & Mesgari, M. S. 2018 Analyzing the spatial variability of precipitation extremes along longitude and latitude, northwest Iran. *Kuwait Journal of Science* **45** (1), 121–127.

Arnold, J. G., Srinivasan, R., Muttiah, R. S. & Williams, J. R. 1998 Large area hydrologic modeling and assessment Part I, model development. *Journal of the American Water Resources Association* **34** (1), 73–89.

Arnold, J. G. 2012 SWAT, model use, calibration, and validation. *Transactions of the ASABE* **55** (4), 1491–1508.

Boufala, M. H., El Hmaidi, A., Chadli, K., Essahlaoui, A., El Ouali, A. & Taia, S. 2019 Hydrological modeling of water and soil resources in the basin upstream of the Allal El Fassi dam (Upper Sebou watershed, Morocco). *Modeling Earth Systems and Environment* **5** (4), 1163–1177.

Burnash, R. & Larry Ferra, R. 1979 *A generalized stream-flow simulation system*. Joint Federal State River Forecast Centre, Sacramento, CA.

Chen, C., Kalra, A. & Ahmad, S. 2019 Hydrologic responses to climate change using downscaled GCM data on a watershed scale. *Journal of Water and Climate Change* **10** (1), 63–77.

Chiang, L., Yuan, Y., Mehaffey, M., Jackson, M. & Chaubey, I. 2014 Assessing SWAT's performance in the Kaskaskia River watershed as influenced by the number of calibration stations used. *Hydrological Processes* **28** (3), 676–687.

Crawford, N. H. & Linsley, R. K. 1966 *Digital simulation in hydrology: Stanford Watershed Model IV*. Department of Civil Engineering, Stanford University, CA.

Desai, S., Singh, D. K., Islam, A. & Sarangi, A. 2021 Multi-site calibration of hydrological model and assessment of water balance in a semi-arid river basin of India. *Quaternary International* **571**, 136–149.

Dong, F., Neumann, A., Kim, D., Huang, J. & Arhonditsis, G. B. 2019 A season-specific, multi-site calibration strategy to study the hydrological cycle and the impact of extreme-flow events along an urban-to-agricultural gradient. *Ecological Informatics* **54**, 100993.

Ehsan, M. A., Kucharski, F., Almazroui, M., Ismail, M. & Tippett, M. K. 2019 Potential predictability of Arabian Peninsula summer surface air temperature in the North American multimodel ensemble. *Climate Dynamics* **53** (7–8), 4249–4266.

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J. & Taylor, K. E. 2016 Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development* **9** (5), 1937–1958.

Franco, A. C. L., Oliveira, D. Y. D. & Bonumá, N. B. 2020 Comparison of single-site, multi-site and multi-variable SWAT calibration strategies. *Hydrological Sciences Journal* **65**, 2376–2389.

Gao, H., Feng, Z., Zhang, T., Wang, Y., He, X., Li, H., Pan, X., Ren, Z., Chen, X., Zhang, W. & Duan, Z. 2021 Assessing glacier retreat and its impact on water resources in a headwater of Yangtze River based on CMIP6 projections. *Science of the Total Environment* **765**, 142774.

Ghoraba, S. M. 2015 Hydrological modeling of the Simly Dam watershed (Pakistan) using GIS and SWAT model. *Alexandria Engineering Journal* **54**, 583–594.

Gupta, H. V., Sorooshian, S. & Yapo, P. O. 1999 Status of automatic calibration for hydrologic models: comparison with multilevel expert calibration. *Journal of Hydrologic Engineering* **4**, 135–143.

Gutjahr, O., Putrasahan, D., Lohmann, K., Jungclaus, J. H., von Storch, J., Brüggemann, N., Haak, H. & Stössel, A. 2019 Max Planck Institute Earth System Model (MPI-ESM1.2) for the High-Resolution Model Intercomparison Project (HighResMIP). *Geoscientific Model Development* **12** (7), 3241–3281.

Hasan, M. A. & Pradhanang, S. M. 2017 Estimation of flow regime for a spatially varied Himalayan watershed using improved multi-site calibration of the Soil and Water Assessment Tool (SWAT) model. *Environmental Earth Sciences* **76** (23), 787

Jha, M. K., Gassman, P. W. & Panagopoulos, Y. 2015 Regional changes in nitrate loadings in the Upper Mississippi River Basin under predicted mid-century climate. *Regional Environmental Change* **15** (3), 449–460.

Jiang, S., Jomaa, S., Büttner, O., Meon, G. & Rode, M. 2015 Multi-site identification of a distributed hydrological nitrogen model using Bayesian uncertainty analysis. *Journal of Hydrology* **529**, 940–950.

Khatun, S., Sahana, M., Jain, S. K. & Jain, N. 2018 Simulation of surface runoff using semi distributed hydrological model for a part of Satluj Basin, parameterization and global sensitivity analysis using SWAT CUP. *Modeling Earth Systems and Environment* **4** (3), 1111–1124.

Lerat, J., Andréassian, V., Perrin, C., Vaze, J., Perraud, J. M., Ribstein, P. & Loumagne, C. 2012 Do internal flow measurements improve the calibration of rainfall-runoff models? *Water Resources Research* **48** (2), W02511.

Liang, X., Lettenmaier, D. P., Wood, E. F. & Burges, S. J. 1994 A simple hydrologically based model of land surface water and energy fluxes for general circulation models. *Journal of Geophysical Research* **99** (D7), 14415.

Mahmoodi, N., Kiesel, J., Wagner, P. D. & Fohrer, N. 2020 Integrating water use systems and soil and water conservation measures into a hydrological model of an Iranian Wadi system. *Journal of Arid Land* **12** (4), 545–560.

Massonnet, F., Ménégoz, M., Acosta, M., Yepes-Arbós, X., Exarchou, E. & Doblas-Reyes, F. J. 2019 Replicability of the EC-Earth3 Earth System Model under a change in computing environment. *Geoscientific Model Development* 1–19.

Meng, X. & Wang, H. 2017 Significance of the China meteorological assimilation driving datasets for the SWAT model (CMADS) of East Asia. *Water* **9** (10), 765.

Migliaccio, K. W. & Chaubey, I. 2007 Comment on Cao W, Bowden BW, Davie T, Fenemor A. 2006. 'Multi-variable and multi-site calibration and validation of SWAT in a large mountainous catchment with high spatial variability'. Hydrological Processes 20(5): 1057–1073. *Hydrological Processes* **21**, 3226–3228.

Moriasi, D. N., Arnold, J. G., Liew, M. W. V., Bingner, R. L., Harmel, R. D. & Veith, T. L. 2007 Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Transactions of the Asabe* **50** (3), 885–900.

Nkiaka, E., Nawaz, N. R. & Lovett, J. C. 2018 Effect of single and multi-site calibration techniques on hydrological model performance, parameter estimation and predictive uncertainty, a case study in the Logone catchment, Lake Chad basin. *Stochastic Environmental Research and Risk Assessment* **32** (6), 1665–1682.

O'Neill, B. C., Tebaldi, C., van Vuuren, D. P., Eyring, V., Friedlingstein, P., Hurtt, G., Knutti, R., Kriegler, E., Lamarque, J. F., Lowe, J., Meehl, G. A., Moss, R., Riahi, K. & Sanderson, B. M. 2016 The Scenario Model Intercomparison Project (ScenarioMIP) for CMIP6. *Geoscientific Model Development* **9** (9), 3461–3482.

Pandey, V. P., Dhaubanjar, S., Bharati, L. & Thapa, B. R. 2020 Spatio-temporal distribution of water availability in Karnali-Mohana Basin, Western Nepal: hydrological model development using multi-site calibration approach (Part-A). *Journal of Hydrology, Regional Studies* **29**, 100690.

Pascoe, C., Lawrence, B. N., Guilyardi, E., Juckes, M. & Taylor, K. E. 2020 Documenting numerical experiments in support of the Coupled Model Intercomparison Project Phase 6 (CMIP6). *Geoscientific Model Development* **13** (5), 2149–2167.

Riahi, K., van Vuuren, D. P., Kriegler, E., Edmonds, J., Neill, B. C. O., Fujimori, S., Bauer, N., Calvin, K., Dellink, R., Fricko, O., Lutz, W., Popp, A., Cuaresma, J. C., Samir, K. C., Leimbach, M., Jiang, L., Kram, T., Rao, S., Emmerling, J., Ebi, K., Hasegawa, T., Havlik, P., Humpenöder, F., Da Silva, L. A., Smith, S., Stehfest, E., Bosetti, V., Eom, J., Gernaat, D., Masui, T., Rogelj, J., Strefler, J., Drouet, L., Krey, V., Luderer, G., Harmsen, M., Takahashi, K., Baumstark, L., Doelman, J. C., Kainuma, M., Klimont, Z., Marangoni, G., Lotze-Campen, H., Obersteiner, M., Tabeau, A. & Tavoni, M. 2017 The Shared Socioeconomic Pathways and their energy, land use, and greenhouse gas emissions implications: an overview. *Global Environmental Change* **42**, 153–168.

Rudra, R. P., Mekonnen, B. A., Shukla, R., Shrestha, N. K., Goel, P. K., Daggupati, P. & Biswas, A. 2020 Currents status, challenges, and future directions in identifying critical source areas for non-point source pollution in Canadian conditions. *Agriculture* **10** (10), 468.

Shrestha, M. K., Recknagel, F., Frizenschaf, J. & Meyer, W. 2016 Assessing SWAT models based on single and multi-site calibration for the simulation of flow and nutrient loads in the semi-arid Onkaparinga catchment in South Australia. *Agricultural Water Management* **175**, 61–71.

Simpkins, G. 2017 Increasing river flood risk. *Nature Climate Change* **7** (3), 172.

Song, Y., Zhang, J., Meng, X., Zhou, Y., Lai, Y. & Cao, Y. 2020 Comparison study of multiple precipitation forcing data on hydrological modeling and projection in the Qujiang River Basin. *Water* **12** (9), 2626.

Sonkoué, D., Monkam, D., Fotso-Nguemo, T. C., Yepdo, Z. D. & Vondou, D. A. 2019 Evaluation and projected changes in daily rainfall characteristics over Central Africa based on a multi-model ensemble mean of CMIP5 simulations. *Theoretical and Applied Climatology* **137** (3–4), 2167–2186.

Sugawara, M., Ozaki, E., Watanabe, I. & Katsuyama, Y. 1974 *Tank model and its application to Bird Creek, Wollombi Brook, Bikin River, Kitsu River, Sanaga River and Nam Mune*. National Research Center for Disaster Prevention, Tokyo..

Van Liew, M. W., Veith, T. L., Bosch, D. D. & Arnold, J. G. 2007 Suitability of SWAT for the conservation effects assessment project: comparison on USDA agricultural research service watersheds. *Journal of Hydrologic Engineering* **12**, 173–189.

Wang, R. & Kalin, L. 2018 Combined and synergistic effects of climate change and urbanization on water quality in the Wolf Bay watershed, southern Alabama. *Journal of Environmental Sciences* **64**, 107–121.

Wang, J., Zhang, X., Xu, C., Wang, H., Lei, X., Wang, X. & Li, S. 2019 Development of load duration curve system in data-scarce watersheds based on a distributed hydrological model. *Hydrology Research* **50** (3), 886–900.

Wi, S., Yang, Y. C. E., Steinschneider, S., Khalil, A. & Brown, C. M. 2015 Calibration approaches for distributed hydrologic models in poorly gaged basins, implication for streamflow projections under climate change. *Hydrology and Earth System Sciences* **19** (2), 857–876.

Wu, T., Lu, Y., Fang, Y., Xin, X., Li, L., Li, W., Jie, W., Zhang, J., Liu, Y., Zhang, L., Zhang, F., Zhang, Y., Wu, F., Li, J., Chu, M., Wang, Z., Shi, X., Liu, X., Wei, M., Huang, A., Zhang, Y. & Liu, X. 2019 The Beijing Climate Center Climate System Model (BCC-CSM), the main progress from CMIP5 to CMIP6. *Geoscientific Model Development* **12** (4), 1573–1600.

Zhao, S., Yu, Y., Lin, P., Liu, H., He, B., Bao, Q. & Guo, Y. 2021 Datasets for the CMIP6 Scenario Model Intercomparison Project simulations with the Coupled Model CAS FGOALS-f3-L. *Advances in Atmospheric Sciences* **38**, 329.

First received 17 March 2021; accepted in revised form 26 June 2021. Available online 14 July 2021