Water Cycle 7 (2026) 219–233

Contents lists available at ScienceDirect

Water Cycle

journal homepage: www.keaipublishing.com/en/journals/water-cycle/

Enhancing hydrological modeling in large basin with intensive human water use through hierarchical parameterization and bias-integrated calibration

Kaikui Cai [a], Jincheng Li [a,b], Qingsong Jiang [a], Lian Hu [a], Jiaxing Fu [a], Man Zhang [a], Yifan Li [a], Yue Qin [a,d,*], Yong Liu [a,c,d,**]

[a] College of Environmental Sciences and Engineering, State Environmental Protection Key Laboratory of All Material Fluxes in River Ecosystems, Peking University, Beijing, China
[b] Water Security Research Group, Biodiversity and Natural Resources Program, International Institute for Applied Systems Analysis (IIASA), Schlossplatz 1, Laxenburg, A-2361, Austria
[c] Southwest United Graduate School, Kunming, China
[d] Institute of Tibetan Plateau, Peking University, Beijing, China

ARTICLE INFO

Keywords:
Hydrological modeling
Hierarchical parameterization
Regionalized calibration
Water use
Community water model
Pearl river basin

ABSTRACT

Human management of water resources has profoundly altered the water cycle, creating complex and difficult-to-simulate human-water interactions. Traditional hydrological models, which have commonly focused solely on natural processes, struggle to accurately represent these changes especially in large basins with intensive human water use, highlighting an urgent need for more effective modeling methods to improve this challenge. This study proposed a hierarchical parameterization and bias-integrated calibration method to enhance modeling in those basins, and identified the optimal configuration through a comparative analysis of calibration scenarios based on the hydrological modeling of the Pearl River Basin (PRB) using the Community Water Model (CWatM). The key findings include: (a) Hierarchical calibration significantly improved simulation performance compared to non-regionalized methods, with average modified Kling-Gupta Efficiency (KGE) and NSE (Nash-Sutcliffe Efficiency) values increasing by over 0.5, and the third level of Water Resource Zones (WRZ3) was identified as the optimal calibration scale. (b) Integrating irrigation simulation bias into a single-objective function enabled the simultaneous optimization of both streamflow and irrigation simulations, which reduced irrigation bias from 327 % to 51 % with only a minor decrease in streamflow accuracy (KGE from 0.81 to 0.75), and the effective irrigation weighting coefficient was found to align with the basin's overall irrigation-to-total-runoff ratio. (c) The CWatM was confirmed as suitable for regional applications, although its performance is sensitive to meteorological and inflow boundary data, and it's important to customize the model's parameters to accurately reflect specific regional characteristics. The reproducible technical pathway presented in this paper could facilitate more precise hydrological modeling in similar basins.

1. Introduction

Over recent decades, accelerating global population growth and intensified human activities have profoundly altered the water cycle and hydrological fluxes (e.g., discharge) at various scales, from local watersheds to the global extent [1,2]. Consequently, the explicit representation of human intervention has become essential for the realistic simulation of global and regional hydrological processes. For example, accurately accounting for human water demands from sectors such as agriculture, industry, and domestic use is crucial, as these demands can

Peer review under the responsibility of Editorial Board of Water Cycle.
* Corresponding author. College of Environmental Sciences and Engineering, State Environmental Protection Key Laboratory of All Material Fluxes in River Ecosystems, Peking University, Beijing, China.
** Corresponding author. College of Environmental Sciences and Engineering, State Environmental Protection Key Laboratory of All Material Fluxes in River Ecosystems, Peking University, Beijing, China.
E-mail addresses: qinyue@pku.edu.cn (Y. Qin), yongliu@pku.edu.cn (Y. Liu).

https://doi.org/10.1016/j.watcyc.2025.10.003
Received 1 July 2025; Received in revised form 14 October 2025; Accepted 14 October 2025
Available online 17 October 2025

significantly influence estimated hydrological storage and fluxes [3,4]. In intensively managed large river basins, such as those in populous regions like India and China, human activities have substantially altered natural hydrological processes [5–7]. Specifically, water use, particularly for irrigation, has become a primary driver within the hydrological cycle, exerting complex interaction mechanisms on both natural hydrological processes and anthropogenic dynamics [8,9]. However, many conventional hydrological models, while considering the impact of human activities to some extent, still prioritize rainfall-runoff simulation and apply a simplified, conceptualized approach to anthropogenic influences [10]. This limited representation often results in poor performance for hydrological simulations, which in turn hinders the provision of reliable insights critical for effective water resource management [11]. This highlights an urgent need to enhance hydrological modeling capabilities in large basins experiencing intensive human water use.

Calibration is an essential step in hydrological modeling that directly determines the successful application of a model [12]. However, for large basins with significant spatial heterogeneity in geomorphology and hydraulics, using identical parameters for the entire basin is problematic [13,14]. A single, unified parameter set cannot adequately represent the diverse hydrological responses across the whole basin [15,16], thus necessitating regional parameterization for accurate hydrological simulations [17]. To address this, regional parameterization has been widely adopted, which involves dividing the basin into multiple calibration units, such as Hydrologic Response Units (HRUs) in models like SWAT [18–20]. Additionally, a novel hierarchical upstream-downstream calibration strategy has been developed, whose scheme divides a basin into independent sub-basins based on hydrological station locations and flow direction, allowing for the progressive calibration of model parameters from upstream to downstream [21,22]. This strategy has been successfully applied in major basins, including the Mississippi River [23], Lancang-Mekong River [13,24], and Yangtze River [25]. Nevertheless, the delineation of sub-basins or calibration units is often subjective, typically based on pre-defined subbasin divisions. Water Resource Zones (WRZ) are fundamental units for hydrological and water resource research, which are defined based on hydrological zones while also considering the unique characteristics of water resources [26]. In China, these divisions usually have one to three levels, commonly referred to as WRZ1, WRZ2, and WRZ3 [27]. There is limited research, however, on which level of WRZ should be chosen as the hydrological calibration unit for large river basins. Consequently, a systematic evaluation of optimal calibration unit scales remains insufficient.

Some of the currently used large scale hydrological models, such as Community Water Model (CWatM) [28], LISFLOOD [29], PCR-GLOBWB [30] and H08 [4,31], have integrated irrigation–soil moisture dynamics as a core module to enhance their simulation of water use management [32]. These models dynamically link the irrigation process with the water balance of soil water storage, surface water, and evapotranspiration over irrigated areas [10,28,30], thereby recognize irrigation as a key hydrological component that influences runoff. Although numerous studies have incorporated multi-variable calibration, including hydrological components like soil moisture, terrestrial water storage (TWS), and evapotranspiration [24,33,34], there is a lack of research that explicitly integrates irrigation water into the calibration process. To improve a model's ability to synergistically simulate multiple hydrological components, there is an urgent need to develop a coupled irrigation and streamflow calibration method. Furthermore, input data uncertainty is widely recognized as the dominant source of hydrological modeling uncertainty [35], with meteorological data uncertainty being particularly significant [36,37]. However, research assessing the specific impact of this input data uncertainty remains relatively scarce when global models are applied for regional modeling using a hierarchical parameterization strategy.

Based on the aforementioned literature review, three key research gaps persist: a) the influence of calibration unit scales on the simulation performance on hierarchical calibration strategies, b) the integration of irrigation water into the calibration framework, and c) the regional applicability of large-scale models and the impact of input data uncertainty. To investigate these questions, the CWatM was selected as the modeling framework due to its advanced grid-based structure, which systematically considers direct feedbacks between human water use and other terrestrial water fluxes [28]. It also possesses unique global and regional spatial representations and has been successfully applied at both scales [7,38–40]. The Pearl River Basin (PRB), which includes one WRZ1, seven WRZ2, and fifteen WRZ3 units, was selected as the study area, a region characterized by high population density and intensive water use [34]. Utilizing a scenario comparison method based on the established hydrological and water resources model, this study set up four groups of 11 calibration scenarios. The first group was designed to identify the optimal parameter set, while the second was designed to evaluate the impact of calibration unit scales on hierarchical parameterization. The third group aimed to optimize coupled irrigation and streamflow simulations, and the fourth was used to evaluate the effect of different meteorological datasets and inflow data on a hierarchical calibration strategy.

This paper attempted to enhance hydrological modeling in large, human-dominated basins by using a hierarchical parameterization and bias-integrated calibration approach, and validated the adaptability of a large-scale model for regional applications. The paper is structured as follows. Section 1 provides the introduction and background. Section 2 describes the study area, data preparation, a brief overview of CWatM, and the setup of the calibration scenarios. Section 3 presents the results and discusses the relevant research questions. Finally, Section 4 provides the conclusions.

2. Materials and methods

2.1. Modeling framework

To investigate how hierarchical parameterization and bias-integrated calibration can enhance hydrological modeling in basins with intensive human water use, a comprehensive modeling framework was proposed (Fig. 1). The framework began with the hydrological modeling of the PRB using CWatM, including the reconstruction of high-resolution water use data for model input. This was followed by a sensitivity analysis to identify sensitive parameters. Subsequently, four sets of calibration experiments were designed. The first group further screened parameters. The second group, which aimed to assess how calibration unit scales affect performance using the hierarchical parameterization approach, involved discretizing the basin into four distinct unit types: WRZ1-3 and site-specific catchments. The third group optimized coupled irrigation and streamflow simulations using a composite objective function. The fourth group assessed the impact of input data uncertainty (including meteorological and inflow data) on modeling robustness. Finally, the accuracy of the hydrological simulations was evaluated using the optimal calibration units and irrigation weighting coefficients identified from the previous steps.

2.2. Study area and data sources

2.2.1. Study area

The Pearl River Basin (PRB) (Fig. 2), situated in southern China (102°E–116°E and 21°N–27°N), is a vast hydrological system with a total drainage area of 4.5×10^5 km^2 [41]. Its river network is characterized by a maximum Strahler stream order of 7 (Fig. S1). The basin is conventionally divided into WRZ2, including Nan-bei Pan Jiang (NPJ), Hongliu Jiang (HJ), Yu Jiang (YJ), Xi Jiang (XJ), Bei Jiang (BJ), Dong Jiang (DJ), and the Pearl River Delta (PRD) basins [42]. Topographically, the basin exhibits significant elevation gradients, transitioning from high altitudes of up to 2852 m in the northwest to sea level in the southeast. The climate of the basin is a tropical and subtropical monsoon

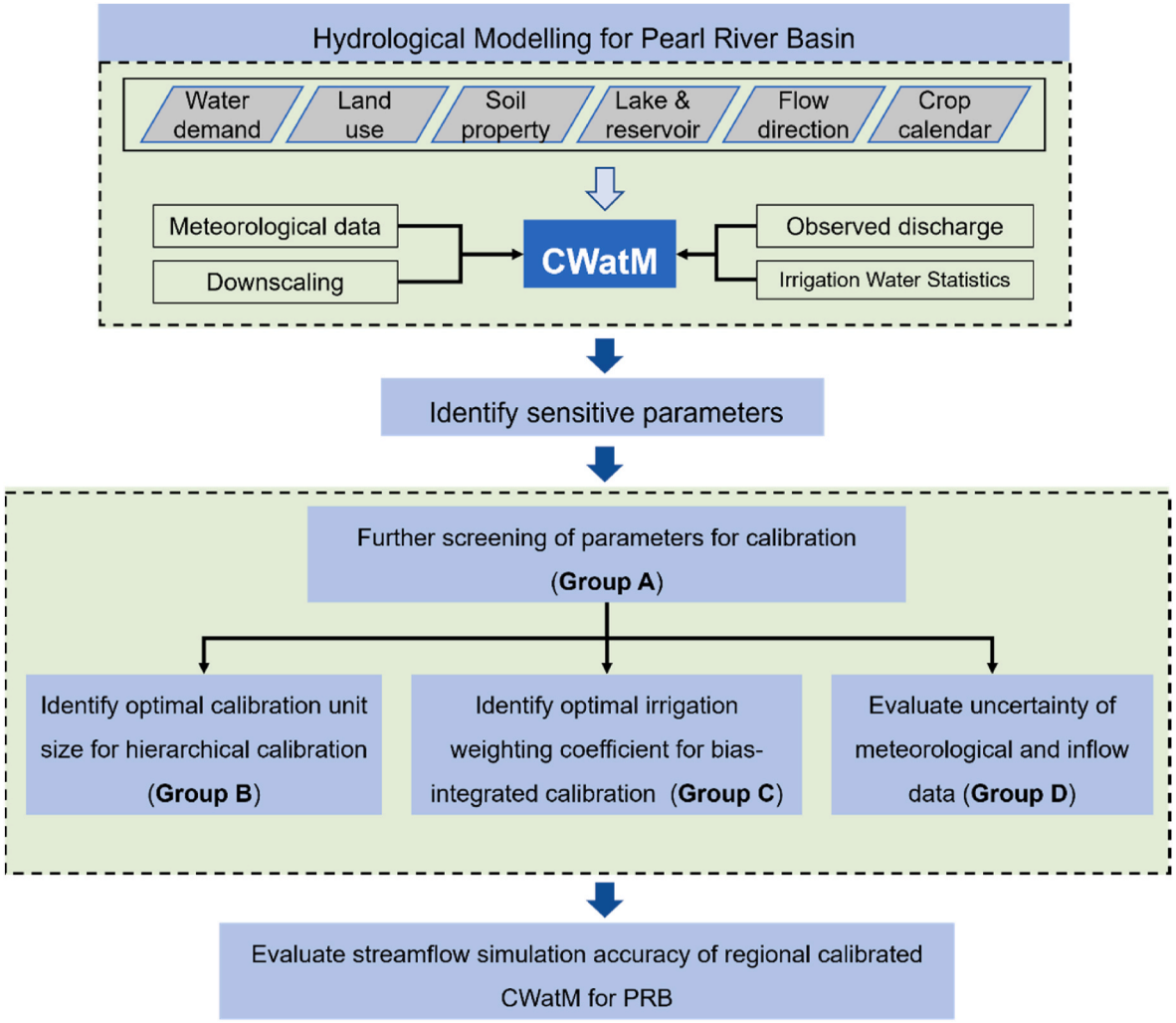

Fig. 1. Methodology framework. (Groups A-D represent the four sets of calibration experiments).

regime, characterized by concentrated precipitation primarily during the wet season (April to September), with annual precipitation ranging from 1200 to 2200 mm and average annual temperatures typically falling between 14 °C and 22 °C [43,44].

As illustrated in Fig. S2, the land cover composition within the study area is dominated by forests (61 %) and grasslands (12 %). Croplands account for 20 % of the area, with an even distribution between paddy and non-paddy irrigation. Minor land use types, including sealed areas (4 %) and water bodies (1.8 %), constitute a smaller proportion of the total area. The PRB ranked second among China's ten major river basins in total water resources, with an average of 3.4×10^{11} m^3/a from 2000 to 2020 (Fig. 2c). The midstream section holds the largest proportion of these resources at 58.7 %, followed by the downstream at 31.6 % and the upstream at 9.7 % (Table S1). The water withdrawal within the basin amounts to 6.05×10^{11} m^3/a, with the midstream and downstream sections accounting for the majority of the withdrawal at 45.6 % and 46.6 %, respectively, while the upstream section accounts for a much smaller portion (7.7 %) (Fig. 2c–Table S1). Agricultural irrigation represents the largest component of water withdrawal (55.3 %), followed by the industrial (26.6 %) and domestic (17.4 %) sectors. Although the PRB comprises only 5 % of China's land area, it supported 16.3 % of the national population and contributed 15.6 % to the national GDP in 2020, and its water withdrawal constitutes 13.3 % of the national total [45].

2.2.2. Basic data

For the modeling, a substantial amount of spatially distributed climatic and physiographic data were required. The local drainage direction map was adapted from HydroSHEDS (https://www.hydrosheds.org), and land use data were derived from the Global Resources Data Cloud (http://www.gis5g.com). Irrigation efficiency and water demand data were obtained from the Water Resources Bulletin (WRB) of PRB (https://www.pearlwater.gov.cn/zwgkcs/lygb/szygb/). Meteorological data were the critical input for hydrological models, and the specific climate indicators needed depend on the method used to calculate potential evapotranspiration [28]. In this study, the CWatM default Penman–Monteith method was employed, which required inputs for precipitation, average, maximum, and minimum 2 m temperatures, near-surface pressure, humidity, 10m wind speed, and long- and short-wave downward surface radiation fluxes. Additionally, temperature data were also used to determine if precipitation is snow or rain. The climate data were obtained from CMFD v2.0 (https://cstr.cn/18406.11.Atmos.tpdc.302088). To quantify the impact of meteorological uncertainty on model calibration, the GSWP3-W5E5 (ISIMIP3a) (https://www.isimip.org/gettingstarted/input-data-bias-adjustment/details/110/) data were also employed. All forcing data were spatially downscaled to 0.1° resolution using bilinear interpolation and then bias-corrected [46]. More details on the spatio-temporal resolution and sources of all input data can be found in Table S3.

Moreover, twenty-nine representative hydrological stations across the PRB were selected for model parameter sensitivity analysis and

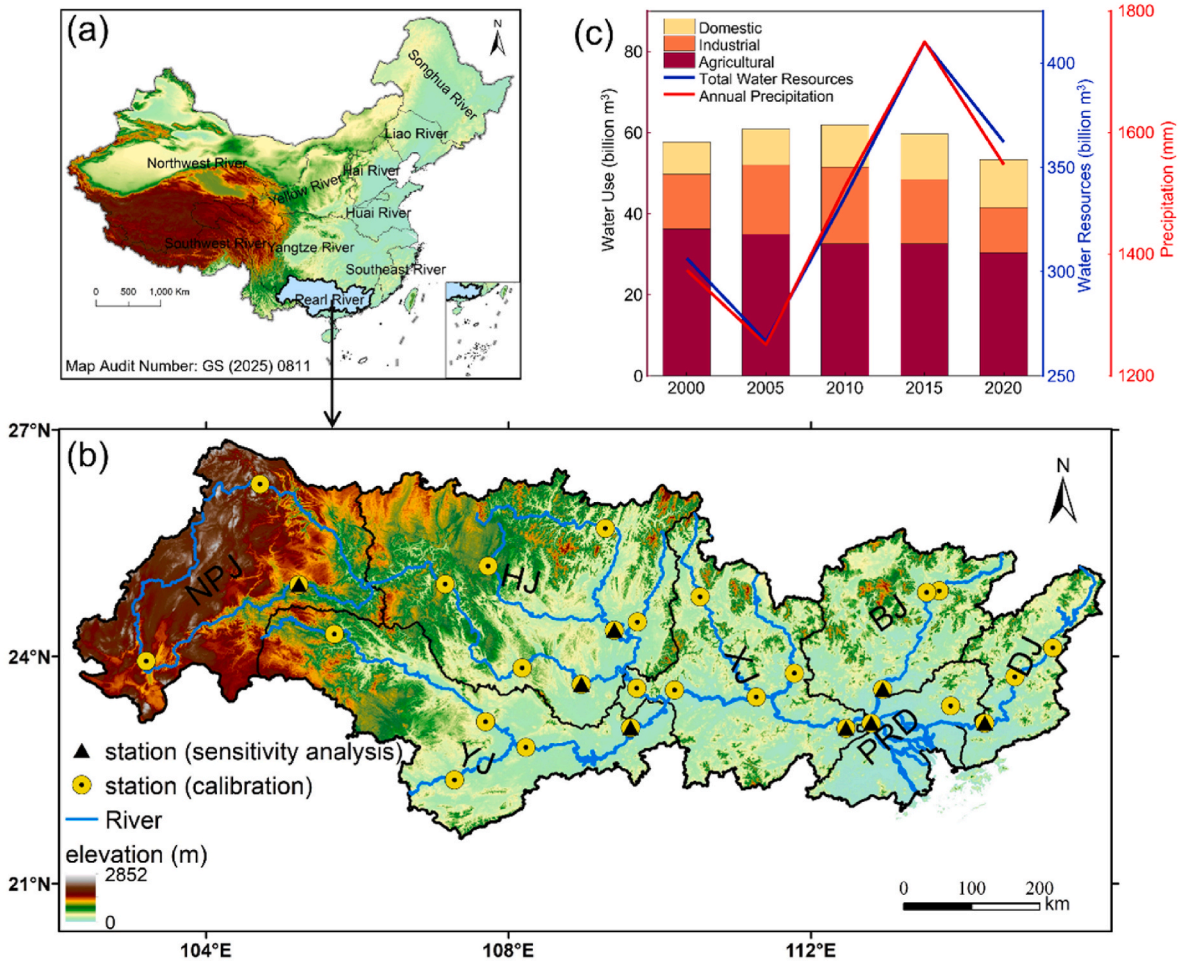

Fig. 2. The study area and water resource utilization. (a) and (b) show the study area and its location. Yellow circles represent hydrological stations for calibration, while black triangles indicate stations for sensitivity analysis. See Table S2 in the supporting information for details. (c) Shows the annual average population, precipitation, total water resources, and water withdrawal for the PRB.

estimation (Fig. 2b). It was determined with three key criteria: stations were located on mainstream (third-to fourth-order Strahler streams), exhibited low reservoir disturbance (by removing stations within 50 km downstream of large dams, verified using Google Earth Pro), and possessed data adequacy (requiring at least 10 years of daily streamflow records from 2006 to 2019 for both calibration and validation). The daily streamflow data were compiled from the National Hydrological Yearbooks of the PRB.

2.2.3. *Water demand data*

For the period 2000–2019, gridded (0.1°) water demand data, encompassing both water withdrawal and consumption for three sectors, were reconstructed using sector-specific methodologies [47,48] (Fig. 3). For the domestic sector, gridded withdrawal was estimated by downscaling WRZ2 level water withdrawal data from the WRB of PRB to 0.1° resolution, leveraging the spatial distribution from gridded population data WorldPop (https://hub.worldpop.org/). Domestic water consumption was subsequently calculated by multiplying this withdrawal by the corresponding WRZ2 level domestic water consumption rate. Similarly, for the industrial sector, WRZ2-level industrial water withdrawal from the WRB of PRB was downscaled to 0.1° using the China Industrial Water Withdrawal (CIWW) dataset's spatial distribution [49], with the industrial water consumption rate calculated similarly to that of the domestic sector. Livestock water demand was determined by multiplying livestock density (heads per 0.1° grid) by the temperature-dependent Livestock-specific Water Demand Intensity (LWDI) [47]. Livestock density for nine categories was derived by downscaling prefecture-level livestock numbers, sourced from the CNKI Chinese Economic and Social Big Data Research Platform (https://data.cnki.net/), to 0.1° using Gridded Livestock of the World distributions (GLW3 for 2000–2010; GLW4 for 2010–2019) [50,51]. LWDI for each category was dynamically adjusted based on three daily CMFD temperature ranges (≤15 °C, 15–35 °C, and ≥35 °C) [52,53], and the species-specific LWDI values are detailed in Table S4. Notably, livestock water consumption was assumed 100 % consumptive (no return flow) [3,28].

2.3. *Hydrological model and setup for PRB*

The CWatM developed by International Institute for the Applied Systems Analysis (IIASA) is a comprehensive hydrological and water resources model. Built upon established models such as PCR-GLOBWB and LISFLOOD, CWatM is capable of simulating water resource availability, human water demand, and the crucial role of water infrastructure (e.g., reservoirs) in water management [47,54]. Its reservoir routine is similar to that of LISFLOOD, effectively simulating dams as points within the channel network. Notably, CWatM internally embeds irrigation as a hydrological flux, while water demands for the livestock, industrial, and domestic sectors are provided as external model inputs [28]. A key advantage of CWatM is its modular design (Fig. S3) and fully open-source Python code, which enable easy integration and co-development with models from other sectors, such as energy and

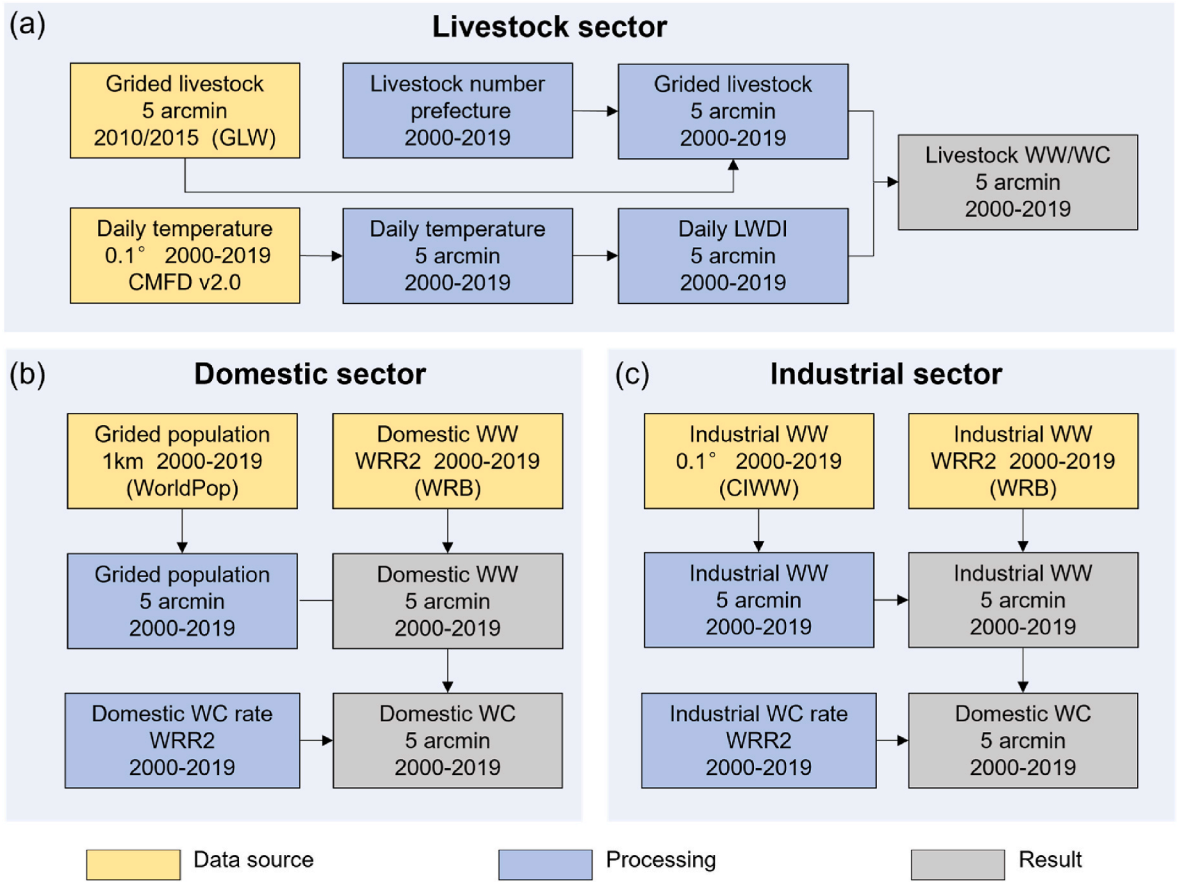

Fig. 3. Methods for calculating water demand in the livestock, domestic and industrial sectors. (WW: water withdrawal; WC: water consumption).

agriculture.

This study employed version 1.08 of CWatM for its modeling efforts, utilizing a spatial resolution of 0.1°. CWatM's operational requires Python along with essential packages such as NumPy, SciPy, netCDF4, and pandas. Given the substantial computational demands inherent in the calibration experiments, the model was executed on a high-performance computing platform equipped with a 32-core processor. All input data were pre-processed into the required 0.1° netCDF format. While primarily sourced from the global dataset provided by IIASA [28], several critical input datasets were updated. Specifically, water demand data for the industrial, domestic, and livestock sectors were reconstructed and served as crucial model inputs, as detailed in Section 2.2.3. Additionally, the map of irrigation efficiency at the sub-basin level within the PRB was updated, a step crucial for accurately simulating irrigation water consumption.

2.4. Parameter sensitivity analysis

Due to its thousands of spatially distributed parameters, a parameter sensitivity analysis of the CWatM is a necessary and crucial step to limit equifinality and reduce computational costs [23,28]. Based on our understanding of the CWatM model's mechanisms and a review of relevant calibration literature [7,28,39,55], this paper selected twenty parameters for the sensitivity analysis (Table S5). Parameter sensitivity varies across different regions, especially within large river basins [17,56,57]. To identify parameters that are important for characterizing different regional features, a sensitivity analysis was conducted at eight hydrological stations located at the outlets of WRZ2 (Fig. 2b). Moreover, this study designed a set of comparative scenarios to identify the most influential parameters for subsequent calibration. This study employed the variance-based Sobol method, as implemented by the SALib Python library [58]. Employing the Saltelli sampling method, 5376 parameter sets were generated per station, resulting in a comprehensive total of 43,008 sets. Simulations were conducted in parallel in batches of 32 runs. Each batch took approximately 12 min to complete the simulation period (01/01/2006-12/31/2008). Consequently, the total computational time for all parameter sets was approximately 269 h (about 11.2 days).

2.5. Calibration setting and performance matrix

The model was calibrated using the NSGA-II genetic algorithm, as implemented with the DEAP Python package. The algorithm was configured with an initial population (μ) of 256, a recombination pool size (λ) of 32, and a maximum of 30 generations. The parameters for calibration were selected based on the results of the sensitivity analysis. The number of parameters used for each calibration scenario can be found in Table 1. Simulations spanned a 19-year period (2001–2019), which was segmented into a 5-year spin-up (2001–2005), a 10-year calibration (2006–2015), and a 4-year validation (2016–2019). To accelerate the calibration process, we used parallel computing across 32 CPU cores with Python's multiprocessing library. This reduced the computational time per CDA unit to approximately 11 h when calibrating 7 parameters and 15 h when calibrating 11 parameters. Computational time for each calibration scenario varied depending on the number of CDA units, with specific details presented in Table 1. Model performance was evaluated using four metrics: the modified Kling-Gupta Efficiency (KGE) [59], the correlation coefficient (R), Nash-Sutcliffe Efficiency (NSE) [60], and Percent Bias (PBIAS) (Eqs. (1)–(4)). These metrics quantify the goodness-of-fit between simulated (Q_s) and observed (Q_o) streamflow.

Table 1
Configurations of the 11 calibration scenarios. (N is the number of hydrological stations in each CDA).

Group	Design for	Scenarios	Level	Number of CDAs	Number of parameters	time (h)	Objective Function	Inflow	Climate
A	Sensitive parameter identification	11_param	WRZ3	15	11	15 × 15	$1/N\sum_i^N KGE(Q)_i$	Simulation	CMFD 2.0
		7_param	WRZ3	15	7	15 × 11			
B	Performance of the hierarchical calibration	WRZ1_level	WRZ1	1	7	1 × 11	$1/N\sum_i^N KGE(Q)_i$	– Simulation	CMFD 2.0
		WRZ2_level	WRZ2	7		7 × 11			
		WRZ3_level	WRZ3	15		15 × 11			
		Site_level	site-specific	29		29 × 11	$KGE(Q)_i$		
C	Optimizing Coupled irrigation and Streamflow Simulations	wgt_0	WRZ3	15	7	15 × 11	$\frac{1}{N}\sum KGE(Q)_i$	Simulation	CMFD 2.0
		wgt_005					$(1-0.05)\times 1/N\sum\left(KGE(Q)_i + 0.05\times\left(1-\lvert Pbias_{Irrigation}\rvert\right)_i\right)$		
		wgt_01					$(1-0.1)\times 1/N\sum\left(KGE(Q)_i + 0.1\times\left(1-\lvert Pbias_{Irrigation}\rvert\right)_i\right)$		
		wgt_02					$(1-0.2)\times 1/N\sum\left(KGE(Q)_i + 0.2\times\left(1-\lvert Pbias_{Irrigation}\rvert\right)_i\right)$		
		wgt_03					$(1-0.3)\times 1/N\sum\left(KGE(Q)_i + 0.3\times\left(1-\lvert Pbias_{Irrigation}\rvert\right)_i\right)$		
D	Impact of input data uncertainty on model calibration	CMFD GS_W5	site-specific	29	7	29 × 11	$KGE(Q)_i$	–	CMFD 2.0 GSWP3-W5E5
		Inflow_sim	WRZ3	15	7	15 × 11	$\frac{1}{N}\sum_i^N KGE(Q)_i$	Simulation	CMFD 2.0
		Inflow_obs						Observation	CMFD 2.0

$$KGE = 1 - \sqrt{(R-1)^2 + (\overline{Q_m}/\overline{Q_O} - 1)^2 + \left((\sigma_{Q_m}/\overline{Q_m})/(\sigma_{Q_m}/\overline{Q_O}) - 1\right)^2} \quad (1)$$

$$R = cov(Q_m, Q_O)/(\sigma_{Q_m}\cdot\sigma_{Q_m}) \quad (2)$$

$$NSE = 1 - \left[\sum_{i=1}^{n}(Q_{oi} - Q_{mi})^2 \Big/ \sum_{i=1}^{n}(Q_{oi} - \overline{Q_o})^2\right] \quad (3)$$

$$\text{PBIAS} = \sum_{i=1}^{n}(Q_{mi} - Q_{oi}) \Big/ Q_{oi} \times 100\% \quad (4)$$

2.6. Experimental setup

In this study, four experimental groups comprising 11 calibration scenarios were designed to achieve four key objectives: identifying the most influential parameters for calibration, determining the optimal calibration scale for hierarchical parameterization, optimizing coupled irrigation and streamflow simulations, and evaluating the impact of different meteorological and inflow datasets on both model performance and the hierarchical calibration strategy. The specific configurations for each scenario are detailed in Table 1.

Group A: to identify the most representative parameters for whole-basin calibration, this group explored two selection methods. The first involved using the union of the five most sensitive parameters identified at each of the eight sites, resulting in a total of 11 parameters. However, this approach risked over-parameterization, potentially increasing computational demand and exacerbating issues of equifinality [61]. To mitigate this, a second, more constrained screening method was applied. This method selected only those parameters that were sensitive at over 50 % of the sites, yielding a unified set of 7 parameters. These two distinct sets of parameters were used to establish two calibration scenarios at the WRZ3 level, named ″11_param″ and ″7_param″, respectively.

Group B: to evaluate the effectiveness of hierarchical parameterization and determine the optimal scale for calibration, the basin was partitioned into four calibration unit configurations. Based on the water resource zones levels 1–3 boundaries and site-specific watershed boundaries, the basin was partitioned into a 1-unit (WRZ1 level), 7-unit (WRZ2 level), 15-unit (WRZ3 level), and 29-unit (site-specific level) configuration (as shown in Fig. 4). These configurations correspond to the ″WRZ1_level″, ″WRZ2_level″, ″WRZ3_level″, and ″Site_level″ scenarios, respectively. To effectively represent the process of hierarchical parameterization within the modeling framework, we referenced the definition of Calibration Data Assimilation (CDA) units [23], in which parameters are uniformly adjusted. The partitioning of the study area and the establishment of confluence relationships were performed using PCRaster and Python 3.12 (see Text S1 and Fig. S4).

Group C: to optimize coupled irrigation and streamflow simulations, a bias-integrated calibration method for irrigation was developed, which employed a composite objective function (Eq. (5)). To determine the optimal weighting coefficient for irrigation, this group established five distinct scenarios— ″wgt_0″, ″wgt_005″, ″wgt_01″, ″wgt_02″, and ″wgt_03″— corresponding to irrigation weighting coefficients of 0, 0.05, 0.1, 0.2, and 0.3, respectively. This range of values allowed us to explore the full spectrum of calibration objectives, from prioritizing pure streamflow optimization (″wgt_0″) to placing a strong emphasis on irrigation water reliability (″wgt_03″).

$$OF = (1 - wgt)\times KGE_Q + wgt\times\left(1 - \left|Pbias_{Irrigation}\right|\right) \quad (5)$$

where, OF represents the single objective function of the genetic algorithm, KGE_Q is the daily streamflow KGE value, $|Pbias_{Irrigation}|$ denotes the absolute value of the annual simulated irrigation water bias, and wgt is the assigned weight for irrigation.

Group D: to evaluate the comprehensive impact of both meteorological input data and upstream inflow boundary data on model performance, a dedicated experimental design was developed. First, the model's sensitivity to meteorological forcing data was assessed by setting up two distinct scenarios: the "CMFD" scenario, which utilized CMFD v2.0 data, and the ″GS_W5″ scenario, which used global-scale GSWP3-W5E5 data. Furthermore, to investigate the propagation effects of upstream errors during hierarchical parameterization, two inflow scenarios were established: the ″Inflow_sim″ scenario, which used optimal simulated values from upstream CDAs, and the ″Inflow_obs″ scenario, which employed observed daily streamflow data as boundary conditions.

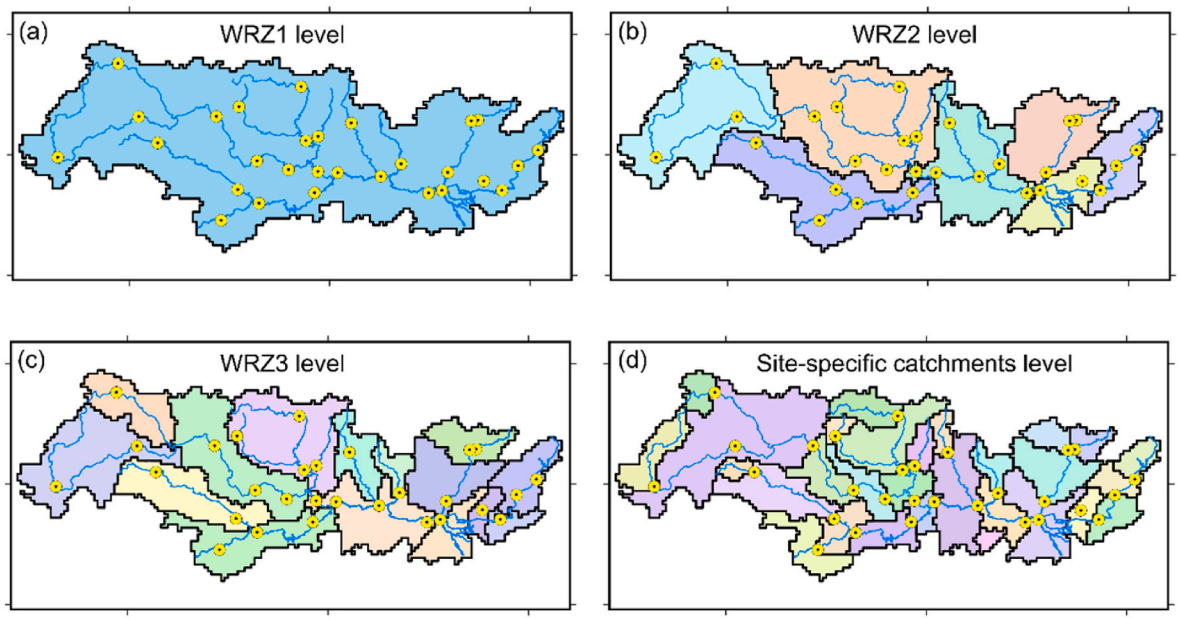

Fig. 4. (a) Calibration unit configurations. (a)–(d) respectively show four CDA units' configurations: 1-unit (WRZ1 level), 7-unit (WRZ2 level), 15-unit (WRZ3 level), and 29-unit (site-specific catchments level). Yellow circles represent hydrological stations for calibration.

3. Results and discussions

3.1. Identification of sensitive parameters through multi-site and multi-metric analysis

The parameter sensitivity analysis was conducted across eight distinct sites, utilizing four evaluation metrics: KGE, NSE, MAE, and RMSE (Fig. 5). The results showed strong consistency across the diverse evaluation metrics in identifying the most sensitive model parameters. Across all eight locations, a common set of highly sensitive parameters was consistently identified, including: crop_correct (a factor adjusting crop evapotranspiration), preferentialFlowConstant (an empirical shape parameter for the preferential flow relation), and factor_interflow (a factor adjusting the amount of interflow that percolates to groundwater). Beyond these common sensitivities, the choice of evaluation metric did influence the identification of sensitive parameters. For instance, soildepth_factor (a factor for the overall soil depth of soil layers 1 and 2) was consistently among the top five sensitive parameters only when evaluated using MAE. Furthermore, certain parameters exhibited marked sensitivity only at specific sites. For instance, the parameter normalStorageLimit, representing the normal storage volume of a reservoir, showed particular sensitivity in the NPJ and HJ basins. This localized sensitivity might be attributable to the substantial reservoir storage, which collectively hold 22 % and 35 % of the entire basin's total capacity, thereby significantly influencing their local hydrological processes. This finding also highlighted the spatial variability of sensitive parameters in a large basin, which aligned with existing research [17,56,57], and underscored the necessity of synthesizing sensitive parameters from multiple sites to derive basin-wide sensitive parameters.

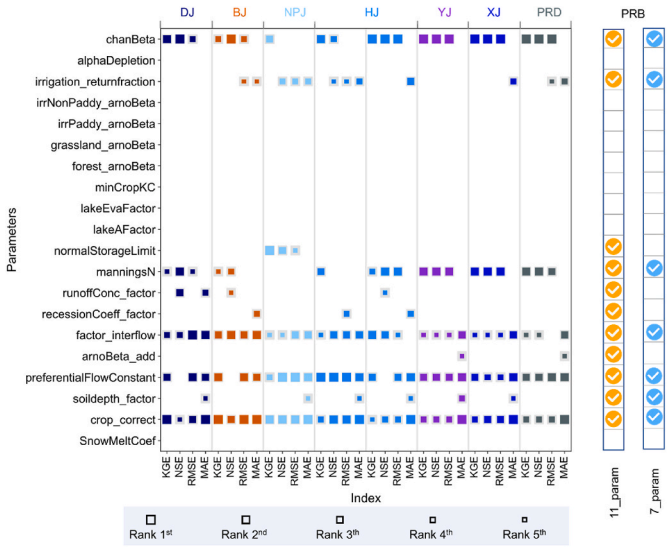

Fig. 5. The most influential parameters at the 8 stations based on the sensitivity of daily streamflow. The size of each box represents the magnitude of the total-order indices of a parameter. Two distinct methods for selecting basin-wide calibration parameters are illustrated on the right: the union of all site-specific sensitive parameters (orange check circles) and a more constrained set of parameters sensitive at over 50 % of the sites (blue check circles).

A comparison of the daily streamflow simulation results from the two sensitive parameter selection methods at 29 hydrological stations at WRZ3 levels revealed that the 11-parameter set showed a slightly better

Table 2
Benchmark Statistics for daily streamflow simulation from two sensitive parameter selection methods at 29 hydrological stations at WRZ3 levels. (calibration/validation).

Types	Matrix	Median	Mean	Maximum	Minimum
11_param (11 parameters)	KGE	0.86/0.79	0.83/0.72	0.96/0.94	0.58/0.34
	NSE	0.73/0.66	0.68/0.52	0.92/0.86	0.19/-0.28
	R	0.89/0.86	0.85/0.81	0.96/0.95	0.66/0.42
	Pbias (%)	0.1/2.7	−0.4/1.6	19.2/29.9	−22.2/21.5
7_param (7 parameters)	KGE	0.84/0.80	0.81/0.75	0.95/0.93	0.54/0.35
	NSE	0.71/0.65	0.67/0.54	0.90/0.88	0.11/-0.52
	R	0.88/0.85	0.85/0.81	0.95/0.95	0.62/0.41
	Pbias (%)	0.2/3.2	−0.6/2.8	24.7/35.8	−24.7/23.6

performance during the calibration period but no significant improvement or even a slight decrease during the validation period (Table 2). At the same time, the Mann-Whitney *U* test showed that there was no statistically significant difference between the two groups ($p < 0.05$, Table S6). This result suggested that calibrating parameters sensitive at over 50 % of the sites is a reasonable and effective approach to avoid the issue of equifinality. Since both methods produced comparable simulation results (Table 2), the more parsimonious set of 7 parameters was compiled for comprehensive basin-wide calibration (Table 3). This demonstrates the need for a multi-site and multi-metric analysis for a distributed and gridded model to capture the varied hydrological processes driven by spatially heterogeneous geohydrological properties. Future research could focus on developing a criterion for determining the optimal number of parameters for calibration, considering the spatial heterogeneity of parameters, particularly in large basins.

3.2. Impact of CDA unit scales on the performances of hierarchical parameterization

An ideal CDA unit delineation that fully captures watershed characteristics is essential for accurate basin modeling. This study investigated the impact of CDA unit scales on regional calibration performance across four distinct CDA unit scales (Fig. 4). When the entire basin was treated as a single CDA unit (the WRZ1-level scenario) and a single parameter set was employed, the average KGE and NSE values for the 29 stations were notably low (0.19 and −0.04, respectively). Conversely, significant improvements were observed when the basin was discretized into smaller units of 7, 15, and 29 CDAs (corresponding to the WRZ2-level, WRZ3-level, and Site-level scenarios, respectively). This resulted in improved mean KGE and NSE values by over 0.5, and PBIAS shifted from approximately −70 % to within ±5 % (Fig. 6). These results strongly indicated that a single parameter set was insufficient to adequately capture the complex hydrological processes of the PRB, underscoring the crucial role of spatial parameter discretization for accurate modeling within the basin [15].

Fig. 6 clearly demonstrated that using the finer scale of the CDA units or increasing the number of CDA units significantly improved performance across all four metrics. At the smallest CDA unit scale, the average daily streamflow simulation achieved a high KGE of 0.83 during the calibration period. However, this did not imply that the smallest scales were universally optimal, as they could introduce excessive spatial heterogeneity in parameter values, potentially leading to over-parameterization issues [62], and significantly increased computational demands (Table 1). In comparison to the Site-level scenario, the WRZ3-level scenario achieved a substantial reduction in the number of CDAs by nearly 50 %, which concurrently halved computational costs (from 319 to 165 h). Crucially, this significant gain in efficiency was realized while maintaining comparable performance across all four metrics (e.g., the mean KGE decreased by only 0.02). There is no statistically significant difference between the WRZ3-level and the WRZ2-level or Site-level, while a statistically significant difference was found between the WRZ3-level and WRZ1-level parameter sets (Mann-Whitney *U* test, $p < 0.05$, Table S6). Furthermore, a comparison of parameter set values across different CDA sizes (Fig. S5) revealed that some parameter sets within the Site-level scenario exhibited highly similar values. This suggested the potential for over-parameterization at finer scales, indicating that certain corresponding CDAs could potentially utilize the same parameter set. Therefore, considering simulation accuracy, computational efficiency, and parameterization effectiveness, the third-order (WRZ3) level was identified as the optimal scale for hierarchical parameterization in the PRB.

Table 3
CWatM pre-calibration parameters. (ET: evapotranspiration, SL: soil, GW: groundwater, RT: routing, RL: reservoir and lakes, WD: waterdemand.)

Parameter	Module	Description	range
crop_correct	ET	adjustment to crop evapotranspiration	[0.8, 1.8]
soildepth_factor	SL	a factor for the overall soil depth of soil layers 1 and 2	[0.8, 1.8]
preferentialFlowConstant	SL	empirical shape parameter of the preferential flow relation	[0.5, 8]
manningsN	RT	a factor roughness in channel routing	[0.1, 10]
chanBeta	RT	kinematic wave parameter	[0.5, 0.7]
factor_interflow	GW	a factor to adjust the amount which percolates from interflow to groundwater	[0.33, 3]
irrigation_returnfraction	WD	the fraction of non-efficient water	[0, 1]

While previous studies have evaluated the impacts of sub-watershed or HRU delineation schemes on simulation results using the SWAT model [19,20,63], research on the delineation of CDA units for gridded models was scarce. This result not only demonstrated the improved simulation performance of hierarchical calibration compared to a single basin-wide parameter set, but also identified the optimal CDA unit scale for modeling in PRB. The computational cost of hierarchical calibration was high (one CDA unit took about 11 h on a 32-core high-performance computing platform, Table 1). To reduce computational costs and improve efficiency, one approach is to use a parallelization scheme [64] and set a condition to terminate the calibration process upon parameter convergence, which would significantly reduce the required computation time. Another approach is to leverage parameter regionalization methods for ungauged basins, such as using clustering analysis or Random Forest algorithms to classify calibration units by similarity or establish a relationship between basin attributes and model parameters [65–67]. This would provide reference parameters values for uncalibrated units, reduce the parameter search space, and decrease the number of genetic algorithm iterations. In summary, the delineation of CDA units should be based on watershed characteristics, data availability, and available computational resources. This finding would provide a valuable reference for obtaining comprehensive, calibrated parameters for modeling in other basins.

3.3. Optimizing irrigation and streamflow simulations

The optimization of balanced annual irrigation and daily streamflow simulations within a single-objective calibration framework at WRZ3 levels was investigated (Fig. 7). It was observed that even a minimal irrigation weight (e.g., 0.05) significantly reduced irrigation simulation bias from 327 % to 120 %, accompanied by only a marginal decrease in streamflow accuracy (from 0.81 to 0.79 mean KGE). Conversely, an excessively high weighting (e.g., 0.3) led to an unacceptable degradation of streamflow accuracy (mean KGE dropped to 0.39), despite achieving high-precision irrigation estimates (bias <10 %). These results suggested that a moderate irrigation weighting striked an optimal balance between streamflow and agricultural water simulation. However, the determination of optimal irrigation weighting factors remains challenging due to the lack of universal criteria. A Mann-Whitney *U* test indicated that the "wgt_01" group was statistically significantly different from the "wgt_0", "wgt_02", and "wgt_03" groups ($p < 0.05$, Table S6). Further evaluation indicated that a weighting coefficient of wgt ≈0.1 achieved an optimal performance balance (Fig. 7a). Specifically, streamflow accuracy decreased moderately (from 0.81 to 0.75 mean KGE), while irrigation simulation bias substantially decreased from 327 % to 51 %. Notably, this empirical weighting aligned closely with the observed irrigation-to-discharge ratio (0.096) derived from the PRB's long-term water budget analysis, where mean annual irrigation water withdrawal (28.4 billion m^3) represents 9.6 % of the total river discharge (298.1 billion m^3). This finding provided a novel criterion that the regional irrigation-to-discharge ratio can serve as an informed basis for selecting the weighting factor in coupled single-objective calibration

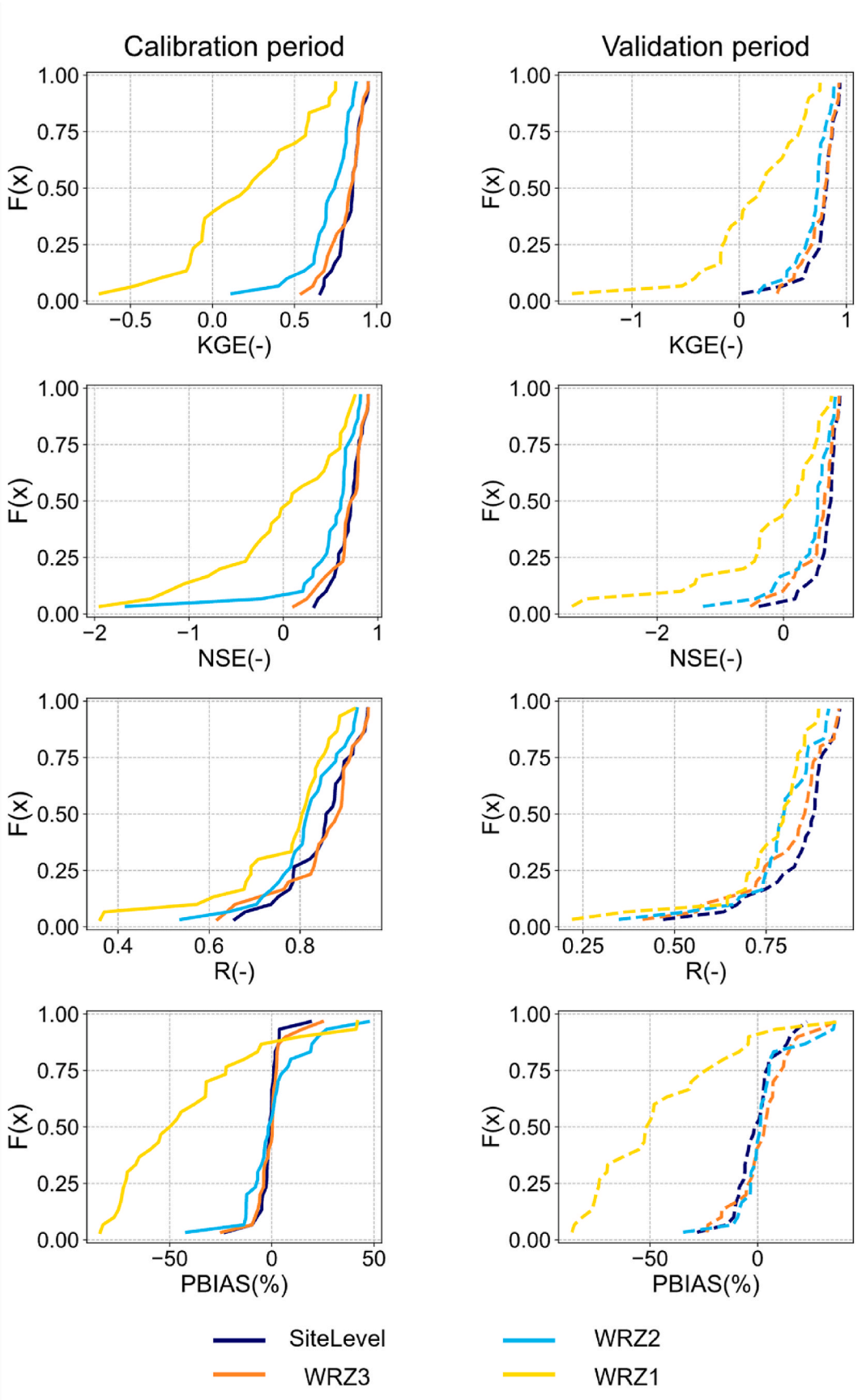

Fig. 6. Cumulative distributions of KGE, NSE, R and PBIAS coefficient over the 29 stations in calibration (left column) and validation (right column) period for 4 different scales of CDAs.

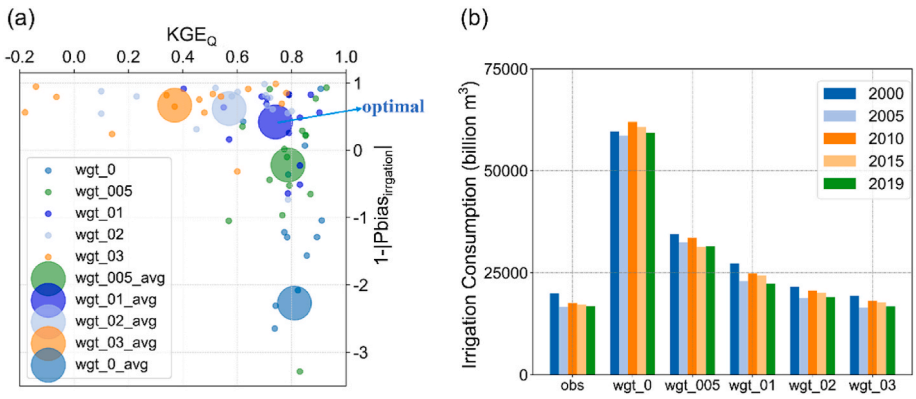

Fig. 7. Optimizing Coupled irrigation and Streamflow Simulations. (a) Daily streamflow and annual irrigation water simulation performance during the calibration period at WRZ3 levels. (b) Simulated annual irrigation water consumption in the PRB from 2000 to 2019. Results for both panels are presented under different irrigation weighting values (0, 0.05, 0.1, 0.2, 0.3).

frameworks.

The importance of accurately representing human water demand during model parameterization has been highlighted [10]. However, existing studies have predominantly focused on runoff calibration, with limited research dedicated to the calibration of water use. Given the dynamic interplay between water use and hydrological processes, it is crucial to incorporate irrigation water use into the calibration framework. While some studies have included human water use within model validation schemes, such as those in the Yangtze River Basin [9], Yellow River Basin [68], and Huaihe River Basin [10], these efforts were primarily for verifying model reliability rather than active calibration, but they failed to achieve the joint optimization of runoff and irrigation. Our study addressed this gap by proposing a novel bias-integrated calibration framework that jointly optimizes both irrigation and streamflow. In this framework, the choice of the irrigation weighting coefficient is particularly important. While the regional irrigation-to-discharge ratio in the PRB approximates the optimal weighting coefficient, this ratio is likely not be universally applicable to other basins. However, the ratio can be used as a reference initial value for the irrigation weighting coefficient, which can then be further refined to identify the optimal value, thereby effectively reducing computational costs.

3.4. Impact of input data uncertainty on model calibration

3.4.1. Impact of meteorological input data uncertainty on calibration

This study demonstrated that the use of high-spatial-resolution (10 km) CMFD v2.0 meteorological data significantly improved streamflow simulation performance compared to lower-resolution (0.5°) GSWP3-

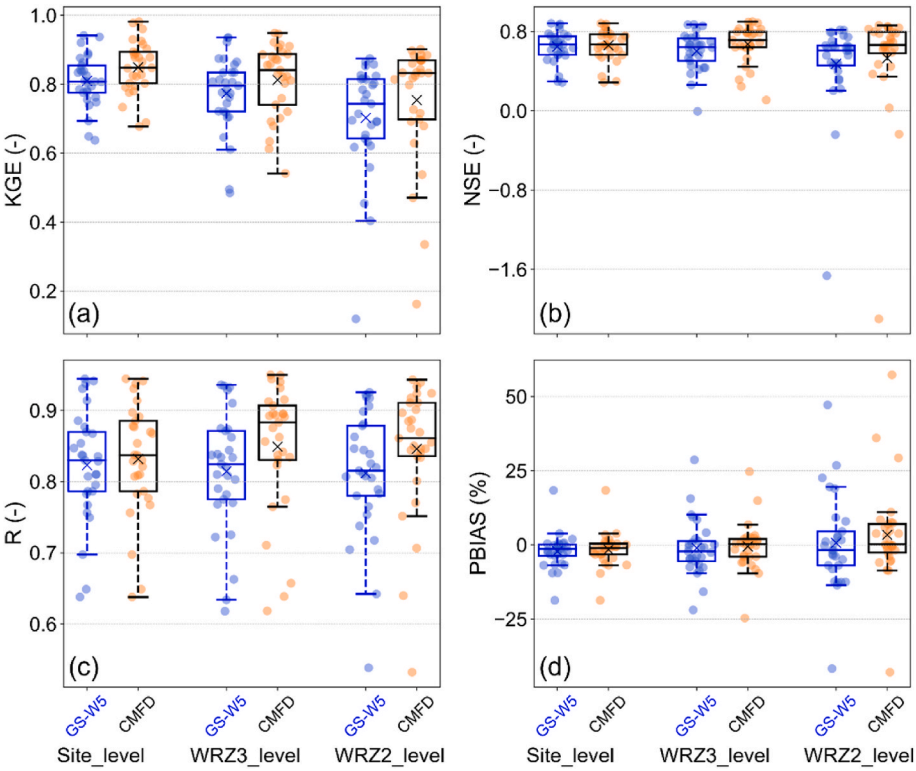

Fig. 8. Impact of meteorological forcing datasets (GSWP3-W5E5 (ISIMIP3a) vs. CMFD v2.0) at three different sizes of CDAs during the calibration period. (a)–(d) show the KGE, NSE, R, and PBIAS values, respectively, for daily streamflow simulations at 29 stations. The "x" markers represent the ensemble means, while the scattered data points illustrate the simulation performance for each of the 29 stations.

W5E5 inputs (Fig. 8). During the calibration period, although a Mann-Whitney *U* test showed no statistically significant difference between the two datasets ($p > 0.05$, Table S6), over 60 % of stations exhibited enhanced KGE, NSE, R, and PBIAS across three different CDA unit scales. Specifically, a detailed analysis of 29 stations revealed notable improvements at the WRZ3 level: mean KGE increased from 0.77 to 0.81, mean NSE improved from 0.60 to 0.67, and mean absolute PBIAS decreased by 25 %. Similar improvements were observed during the validation period (Fig. S6). These results underscore the potential unsuitability of global meteorological datasets for regional hydrological modeling, as they could lead to biased hydrological parameter estimates and propagate these biases into water resource simulations [69]. This finding corroborated prior research, which has shown that high-resolution meteorological data provide more reliable results than global datasets, emphasizing the critical need for high-quality meteorological forcing data for hydrological models [70].

The availability of high-quality meteorological datasets for China has increased in recent years. In addition to the CMFD dataset used in this study, other available datasets include the China Daily Meteorological Dataset (CDMet) [71], the CMA Land Data Assimilation System (CLDAS) [71], and station data from the China Meteorological Data Service Center (CMDC) (https://data.cma.cn/). Given the expanding availability, it is recommended that researchers comprehensively compare and evaluate various meteorological datasets before modeling to select the most appropriate data for their specific region. Furthermore, hydrological models are highly sensitive to biases in meteorological forcing data [72], and this study used observed near-surface meteorological data from approximately 700 weather stations in China as a reference for bias correction. In conclusion, regional hydrological modeling should prioritize meteorological data with high spatiotemporal

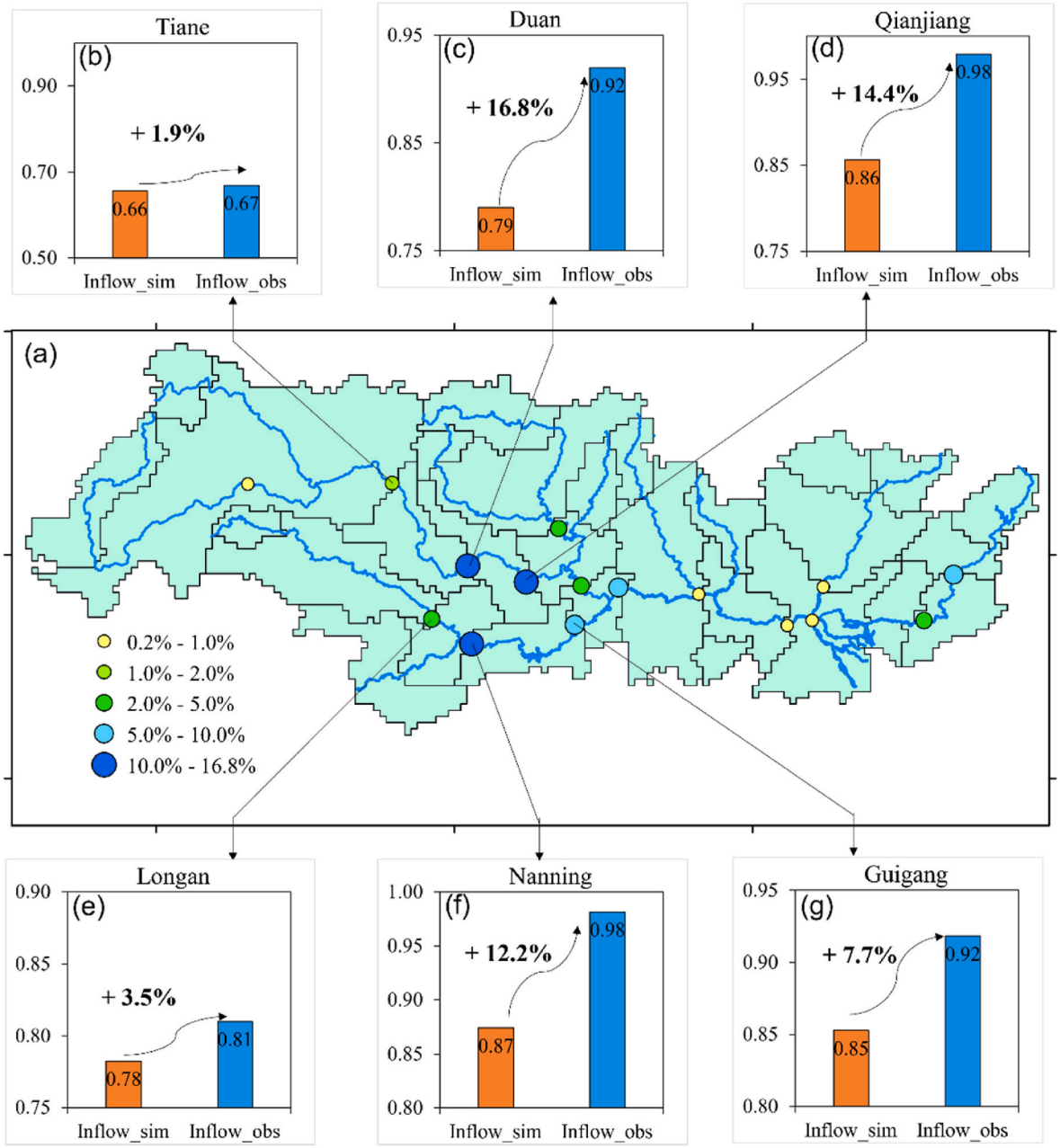

Fig. 9. Impact of simulated versus observed inflow on downstream CDA performance in a cascaded hierarchical calibration strategy at site-specific catchments level. (a) Illustrates the improvement in daily streamflow KGE when observed inflow is used compared to simulated inflow, while (b)–(g) show the improvement for six specific stations.

resolution and integrate bias correction using reliable reference data, such as station observation data.

3.4.2. Impact of inflow uncertainty on hierarchical parameterization

The uncertainty introduced by inflow in hierarchical parameterization was assessed using two distinct inflow scenarios, namely "Inflow_sim" (simulated inflow) and "Inflow_obs" (observed inflow). The analysis focused on sixteen CDA units that receive inflow at the site-specific catchments level. Although a Mann-Whitney *U* test showed no statistically significant difference between the two datasets ($p > 0.05$, Table S6), it was found that utilizing observed daily streamflow data as boundary inputs significantly improved performance in all sixteen CDA units compared to simulations based on simulated inflow (Fig. 9). This resulted in an average increase of 0.04 in the KGE across all sixteen CDA units. Specifically, six CDA units exhibited a KGE improvement exceeding 5 % for daily streamflow. The most significant improvement was observed at Duan Station on the Hongshui River, where the KGE rose from 0.79 to 0.92, representing a 16.8 % increase. Qianjiang Station showed the second highest improvement, with its KGE increasing from 0.86 to 0.98, a 14.4 % rise. This substantial improvement could be attributed to the correction of prior simulation inaccuracies at the upstream Tiane Station (KGE = 0.66) (Fig. 9b–d). When simulated values from Tiane Station were used as inflow, the inherent simulation error directly compromised the performance at Duan Station, subsequently affecting the downstream Qianjiang Station. A similar pattern was observed with the inaccurate simulation at Longan Station (KGE = 0.78) on the Yujiang River, which directly impacted the downstream Nanning and Guigang Stations (Fig. 9e–g).

The analysis highlighted that the performance of downstream CDA units was critically dependent on the accuracy of upstream simulation performance, as model biases were sequentially propagated through nested CDA units. It revealed a cascading effect where simulation biases in the hierarchical calibration method propagate from upstream to downstream. This illustrated a limitation of the hierarchical calibration method, as the strategy involved estimating parameters following a step-by-step process, and once upstream area was calibrated and its parameters were fixed, the resulting simulations were then used to calibrate the downstream areas [13,22]. Overall, this calibration strategy offerd the significant advantage of calibrating all grid parameters for an entire basin, making it a particularly useful method for large basins with abundant hydrological stations. However, it is essential to consider the upstream-to-downstream uncertainty propagation caused by inflow data.

3.5. Streamflow simulation accuracy of regional calibrated CWatM

This study assessed the accuracy of streamflow simulations from the regionally calibrated CWatM, and evaluated its performance for regional use by comparing it against several global data. The discharge from the regionally calibrated CWatM was simulated using the calibration findings from Sections 3.1-3.3. Specifically, it utilized seven parameters calibrated at the WRZ3 level, with upstream and downstream areas calibrated hierarchically, and the irrigation weight coefficient was set to 0.1. This was then compared to global discharge datasets from uncalibrated global hydrological models (CWatM, WaterGAP2-2e, H08) sourced from ISIMIP3a [62], as well as from calibrated models (PCR-GLOBWB and LISFLOOD). Firstly, compared to data from the uncalibrated global models, the regionally calibrated CWatM demonstrated superior daily streamflow simulations at three hydrological stations along the PRB's main tributaries (Fig. 10). The improvements were significant, the daily flow's KGE value consistently exceeded 0.85 (compared to 0.02–0.74 for uncalibrated models), and biases were constrained within ±1 %. The calibrated model more effectively captured seasonal flow variability and accurately replicated peak flow

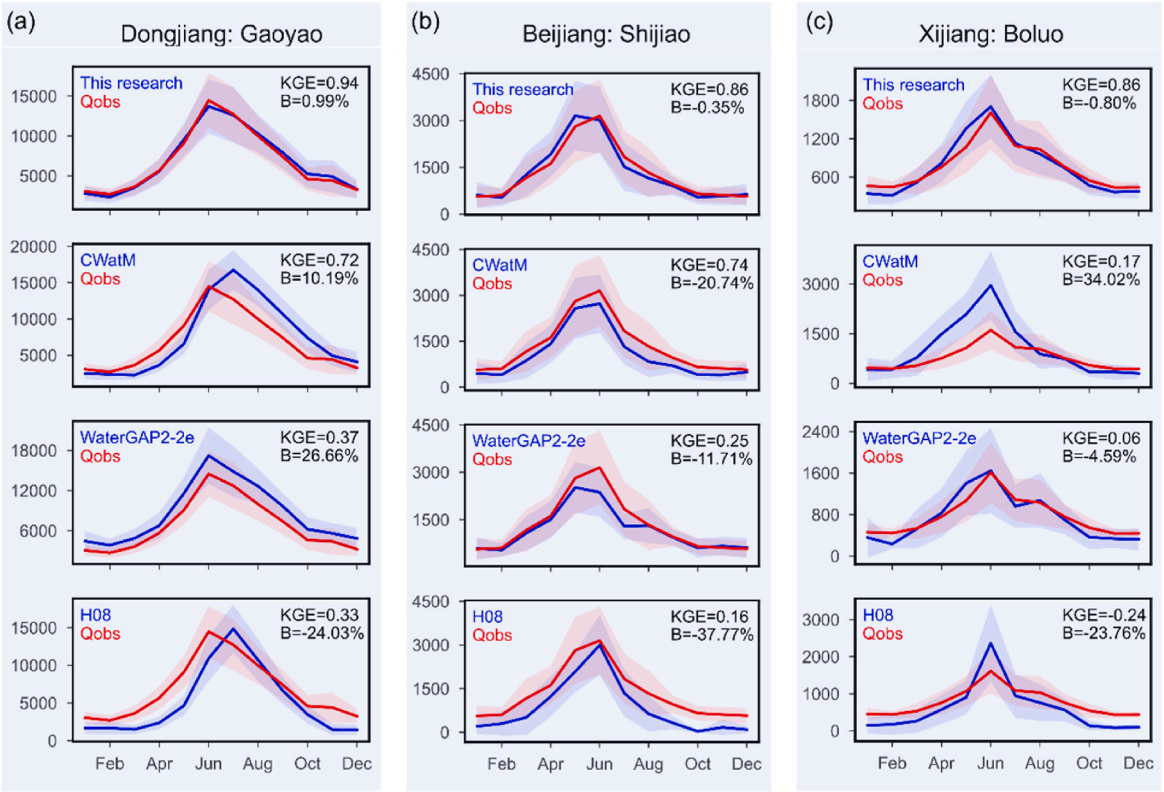

Fig. 10. Comparison of Daily Streamflow Simulation Performance between the regional calibrated CWatM and three uncalibrated global models (CWatM, WaterGAP2-2e, H08).
(a)–(c) show the performance at the three hydrological stations located on the three main stems, respectively.

magnitude and timing when compared to observed data. These results underscore the critical need for basin-specific modeling to ensure reliable streamflow simulations [73]. This indicates that a straightforward application of a global model with standard settings to a specific regional case study is often inadequate, necessitating careful consideration of how to configure global models for regional-scale investigations [2,25]. Increasing their spatial resolution was an advanced and effective approach to enhance the utility of global models for specific regional or catchment-scale applications [74]. For instance, applying global models at a finer resolution could lead to improved water level estimates by enhancing the precision of reach lengths and DEMs [75]. In summary, ongoing efforts are crucial to improve how global hydrological models can be applied regionally. This will ultimately advance both global models and regional simulations, leading to greater consistency between regional and global modeling approaches.

Secondly, to further assess the spatial consistency of the simulated monthly streamflow, it was compared with two datasets from calibrated global models of similar resolutions: DynWat (derived from PCR-GLOBWB) [76] and GloFAS-ERA5 (based on LISFLOOD) [77] (Fig. 11). The simulations showed strong agreement with DynWat, evidenced by a high R^2 value of 0.82. In contrast, the agreement with GloFAS-ERA5 was considerably weaker, with an R^2 of 0.41. This difference in alignment could be attributed to the structural similarities between CWatM and PCR-GLOBWB. Additionally, both DynWat and GloFAS-ERA5 were found to consistently overestimate streamflow in areas significantly impacted by human activities, such as the PRD basin and densely agricultural regions (Fig. S2). These overestimations might be due to the incorporation of more spatially accurate sectoral water use data into this study's modeling framework. The suboptimal performance of global datasets within the PRB was primarily attributed to limited regional parameterization in global models. For example, while GloFAS-ERA5 uses 1287 hydrological stations globally for calibration, fewer than 1 % are in China, with only a single station in the PRB [78,79]. A similar issue affects the spatial distribution of stations used for calibrating DynWat, with only one station situated within the PRB [80]. The performance of globally-calibrated models is suboptimal at the basin scale, failing to satisfy the requirements for regional hydrological simulations. Conversely, this study's approach to parameterizing the CWatM according to the regional characteristics of the PRB could provide a technical reference for parameterization in other basins.

4. Conclusion

This study demonstrated that a hierarchical parameterization and bias-integrated calibration approach is an effective method for enhancing hydrological modeling in large basins with intensive human water use. The findings offer a reliable technical pathway for achieving more precise simulations, which could lead to better-informed water resource management. The specific conclusions are as follows:

(a) Hierarchical calibration significantly improved simulation performance compared to non-regionalized methods, with average KGE and NSE values increasing by over 0.5. And the Water Resource Zones of level 3 (WRZ3) was identified as the optimal scales for calibration units.
(b) Integrating irrigation simulation bias into the single-objective function could enable the simultaneous optimization of both streamflow and irrigation simulations. And the effective irrigation weighting coefficient aligns with the basin's overall irrigation-to-total-runoff ratio.

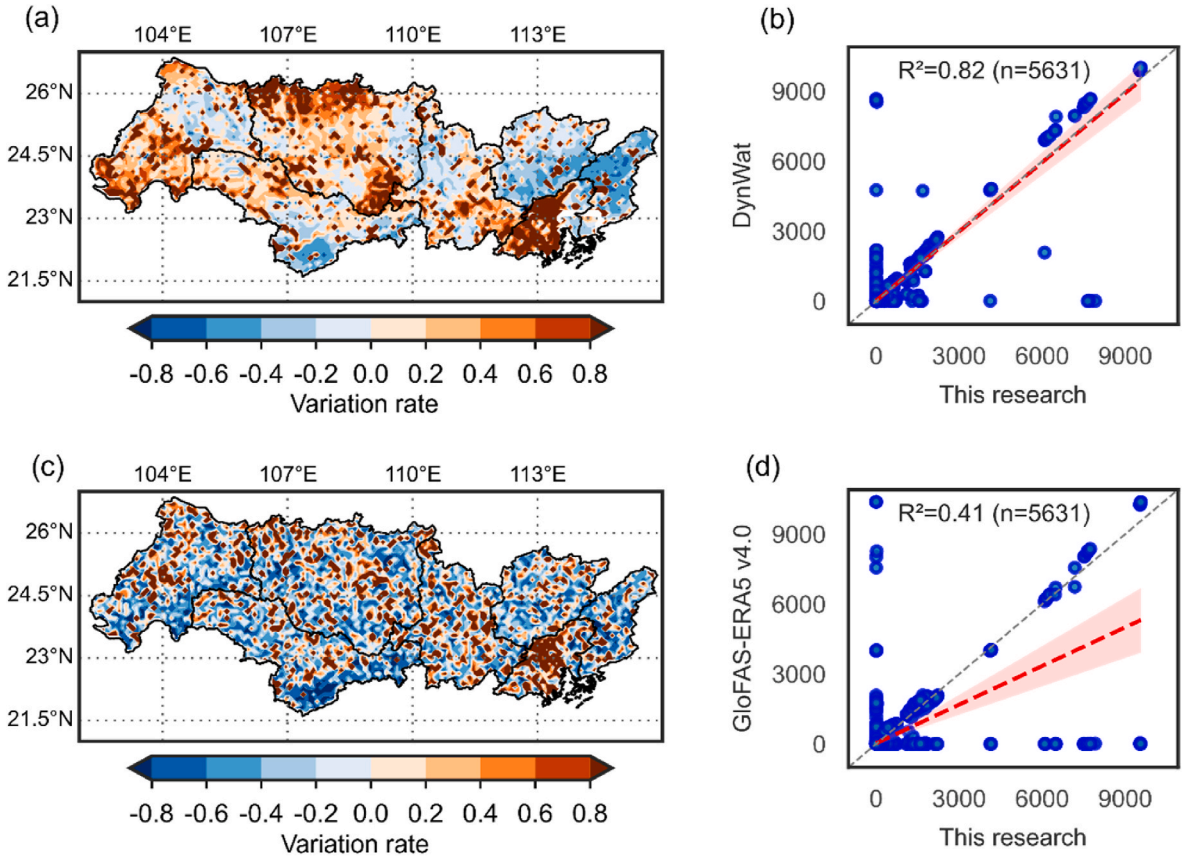

Fig. 11. Comparison of monthly Streamflow Simulation between the regional calibrated CWatM and two calibrated global models. (a,c) The spatial distribution of variation rate in multi-year mean streamflow between DynWat/GloFAS-ERA5 v4.0 and this study, respectively; (b,d) Temporal correlation of monthly streamflow across all grid cells for DynWat/GloFAS-ERA5 v4.0 versus this study, respectively.

(c) The CWatM is suitable for regional applications, but regional parameterization according to the regional characteristics for specific basins is vital. Additionally, the model's performance is sensitive to input data quality.

CRediT authorship contribution statement

Kaikui Cai: Writing – review & editing, Writing – original draft, Visualization, Resources, Methodology, Conceptualization. **Jincheng Li:** Investigation, Formal analysis, Data curation. **Qingsong Jiang:** Formal analysis, Data curation. **Lian Hu:** Formal analysis, Conceptualization. **Jiaxing Fu:** Formal analysis, Data curation. **Man Zhang:** Formal analysis, Data curation. **Yifan Li:** Formal analysis, Data curation. **Yue Qin:** Writing – review & editing, Methodology, Conceptualization. **Yong Liu:** Writing – review & editing, Supervision, Methodology.

Code and data availability data-availability

The forcing datasets for this study were obtained from the China Meteorological Forcing Dataset (CMFD) v2.0 (https://doi.org/10.11888/Atmos.tpdc.302088) and GSWP3-W5E5 (ISIMIP3a) (https://www.isimip.org/gettingstarted/input-data-bias-adjustment/details/110/). Water withdrawal and consumption data for various sectors were sourced from the Pearl River Resources Commission of the Ministry of Water Resources (https://www.pearlwater.gov.cn/zwgkcs/lygb/szygb/). The hydrological model employed, CWatM (Version 1.08), and the calibration method based on NSGA-II are publicly accessible on GitHub (https://github.com/iiasa/CWatM).

Funding

The paper was supported by the National Key R&D Program of China, China (2024YFD1701300), the Yunnan Provincial Science and Technology Project, China (202302AO370015), and National Natural Science Foundation of China (42471021 and 42277482).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors would like to thank the computational support of High-performance Computing Platform of Peking University, China.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.watcyc.2025.10.003.

References

[1] B.W. Abbott, K. Bishop, J.P. Zarnetske, C. Minaudo, F.S. Chapin Iii, S. Krause, D. M. Hannah, L. Conner, D. Ellison, S.E. Godsey, Human domination of the global water cycle absent from depictions and perceptions, Nat. Geosci. 12 (7) (2019) 533–540.
[2] Y. Wada, M.F. Bierkens, A. De Roo, P.A. Dirmeyer, J.S. Famiglietti, N. Hanasaki, M. Konar, J. Liu, H. Müller Schmied, T. Oki, Human–water interface in hydrological modelling: current status and future directions, Hydrol. Earth Syst. Sci. 21 (8) (2017) 4169–4193.
[3] Y. Wada, M. Flörke, N. Hanasaki, S. Eisner, G. Fischer, S. Tramberend, Y. Satoh, M. Van Vliet, P. Yillia, C. Ringler, Modeling global water use for the 21st century: the water futures and solutions (WFaS) initiative and its approaches, Geosci. Model Dev. (GMD) 9 (1) (2016) 175–222.
[4] N. Hanasaki, S. Fujimori, T. Yamamoto, S. Yoshikawa, Y. Masaki, Y. Hijioka, M. Kainuma, Y. Kanamori, T. Masui, K. Takahashi, A global water scarcity assessment under shared Socio-economic pathways–part 1: water use, Hydrol. Earth Syst. Sci. 17 (7) (2013) 2375–2391.
[5] T.I.E. Veldkamp, F. Zhao, P.J. Ward, H. De Moel, J.C. Aerts, H.M. Schmied, F. T. Portmann, Y. Masaki, Y. Pokhrel, X. Liu, Human impact parameterizations in global hydrological models improve estimates of monthly discharges and hydrological extremes: a multi-model validation study, Environ. Res. Lett. 13 (5) (2018) 055008.
[6] Z. Li, H. Sahotra, S. Ahmad, W. Wang, Z. Yang, P. Wu, E. Khan, L. Zhuo, A distributed machine learning model for blue and green water resources with transferable applications in similar climatic zones, Water Resour. Res. 61 (5) (2025) e2024WR039169.
[7] D. Shah, H.L. Shah, H.M. Dave, V. Mishra, Contrasting influence of human activities on agricultural and hydrological droughts in India, Sci. Total Environ. 774 (2021) 144959.
[8] D. Yang, Y. Yang, J. Xia, Hydrological cycle and water resources in a changing world: a review, Geogr. Sustain. 2 (2) (2021) 115–122.
[9] F. Wu, X. Yang, Z. Cui, L. Ren, S. Jiang, Y. Liu, S. Yuan, The impact of human activities on blue-green water resources and quantification of water resource scarcity in the Yangtze River Basin, Sci. Total Environ. 909 (2024) 168550.
[10] H. Cheng, W. Wang, P.R. Van Oel, J. Lu, G. Wang, H. Wang, Impacts of different human activities on hydrological drought in the Huaihe River Basin based on scenario comparison, J. Hydrol.: Reg. Stud. 37 (2021) 100909.
[11] Z. Zhou, Y. Jia, Y. Qiu, J. Liu, H. Wang, C.-Y. Xu, J. Li, L. Liu, Simulation of dualistic hydrological processes affected by intensive human activities based on distributed hydrological model, J. Water Resour. Plann. Manage. 144 (12) (2018) 04018077.
[12] P. Athira, Calibration of hydrological models considering process interdependence: a case study of SWAT model, Environ. Model. Software 144 (2021) 105131.
[13] J. Wang, X. Yun, Y. Pokhrel, D. Yamazaki, Q. Zhao, A. Chen, Q. Tang, Modeling daily floods in the lancang-mekong river basin using an improved hydrological-hydrodynamic model, Water Resour. Res. 57 (8) (2021) e2021WR029734.
[14] D. Yamazaki, S. Kanae, H. Kim, T. Oki, A physically based description of floodplain inundation dynamics in a global river routing model, Water Resour. Res. 47 (4) (2011).
[15] L. Feyen, J.A. Vrugt, B.Ó. Nualláin, J. Van Der Knijff, A. De Roo, Parameter optimisation and uncertainty assessment for large-scale streamflow simulation with the LISFLOOD model, J. Hydrol. 332 (3–4) (2007) 276–289.
[16] C. Brinkerhoff, C. Gleason, D. Feng, P. Lin, Constraining remote river discharge estimation using reach-scale geomorphology, Water Resour. Res. 56 (11) (2020) e2020WR027949.
[17] S. Wang, Z. Zhang, G. Sun, P. Strauss, J. Guo, Y. Tang, A. Yao, Multi-site calibration, validation, and sensitivity analysis of the MIKE SHE Model for a large watershed in northern China, Hydrol. Earth Syst. Sci. 16 (12) (2012) 4621–4632.
[18] H. Xie, L. Longuevergne, C. Ringler, B.R. Scanlon, Calibration and evaluation of a semi-distributed watershed model of Sub-Saharan Africa using GRACE data, Hydrol. Earth Syst. Sci. 16 (9) (2012) 3083–3099.
[19] L.-C. Chiang, Y. Yuan, The NHDPlus dataset, watershed subdivision and SWAT model performance, Hydrol. Sci. J. 60 (10) (2015) 1690–1708.
[20] B. Lin, X. Chen, H. Yao, Threshold of sub-watersheds for SWAT to simulate hillslope sediment generation and its spatial variations, Ecol. Indic. 111 (2020) 106040.
[21] H. Wu, J.S. Kimball, N. Mantua, J. Stanford, Automated upscaling of river networks for macroscale hydrological modeling, Water Resour. Res. 47 (3) (2011).
[22] H. Wu, J.S. Kimball, M.M. Elsner, N. Mantua, R.F. Adler, J. Stanford, Projected climate change impacts on the hydrology and temperature of Pacific Northwest rivers, Water Resour. Res. 48 (11) (2012).
[23] P. Döll, H.M.M. Hasan, K. Schulze, H. Gerdener, L. Börger, S. Shadkam, S. Ackermann, S.-M. Hosseini-Moghari, H. Müller Schmied, A. Güntner, Leveraging multi-variable observations to reduce and quantify the output uncertainty of a global hydrological model: evaluation of three ensemble-based approaches for the Mississippi River basin, Hydrol. Earth Syst. Sci. 28 (10) (2024) 2259–2295.
[24] H. Huang, J. Liu, L. Guillaumot, A. Chen, I.E. De Graaf, D. Chen, Contrasting impacts of irrigation and deforestation on Lancang-Mekong River Basin hydrology, Commun. Earth Environ. 6 (1) (2025) 107.
[25] F. Zhao, N. Nie, Y. Liu, C. Yi, L. Guillaumot, Y. Wada, P. Burek, M. Smilovic, K. Frieler, M. Büchner, Benefits of calibrating a global hydrological model for regional analyses of flood and drought projections: a case study of the Yangtze River Basin, Water Resour. Res. 61 (3) (2025) e2024WR037153.
[26] D. Yan, K. Wang, T. Qin, B. Weng, H. Wang, W. Bi, X. Li, M. Li, Z. Lv, F. Liu, A data set of global river networks and corresponding water resources zones divisions, Sci. Data 6 (1) (2019) 219.
[27] X. Zhi-Cheng, T. Xian-Qi, W. Min, L. Shang-Wu, D. Bing, Y. Xia, C. Zhao-Hui, Evaluation of River System connectivity on scales of water resource zones in China, J. Changjiang River Sci. Res. Institute 41 (10) (2024) 40.
[28] P. Burek, Y. Satoh, T. Kahil, T. Tang, P. Greve, M. Smilovic, L. Guillaumot, F. Zhao, Y. Wada, Development of the Community Water Model (CWatM v1. 04)-a high-resolution hydrological model for global and regional assessment of integrated water resources management, Geosci. Model Dev. (GMD) 13 (7) (2020) 3267–3298.
[29] J. Van Der Knijff, J. Younis, A. De Roo, LISFLOOD: a GIS-based distributed model for river basin scale water balance and flood simulation, Int. J. Geogr. Inf. Sci. 24 (2) (2010) 189–212.
[30] E.H. Sutanudjaja, R. Van Beek, N. Wanders, Y. Wada, J.H. Bosmans, N. Drost, R. J. Van Der Ent, I.E. De Graaf, J.M. Hoch, K. De Jong, PCR-GLOBWB 2: a 5 arcmin global hydrological and water resources model, Geosci. Model Dev. (GMD) 11 (6) (2018) 2429–2453.

[31] N. Hanasaki, S. Fujimori, T. Yamamoto, S. Yoshikawa, Y. Masaki, Y. Hijioka, M. Kainuma, Y. Kanamori, T. Masui, K. Takahashi, A global water scarcity assessment under Shared Socio-economic Pathways–Part 2: water availability and scarcity, Hydrol. Earth Syst. Sci. 17 (7) (2013) 2393–2413.
[32] C.-E. Telteu, H. Müller Schmied, W. Thiery, G. Leng, P. Burek, X. Liu, J.E. S. Boulange, L.S. Andersen, M. Grillakis, S.N. Gosling, Understanding each other's models: an introduction and a standard representation of 16 global water models to support intercomparison, improvement, and communication, Geosci. Model Dev. (GMD) 14 (6) (2021) 3843–3878.
[33] H.M.M. Hasan, P. Döll, S.-M. Hosseini-Moghari, F. Papa, A. Güntner, The benefits and trade-offs of multi-variable calibration of the WaterGAP global hydrological model (WGHM) in the Ganges and Brahmaputra basins, Hydrol. Earth Syst. Sci. 29 (2) (2025) 567–596.
[34] S. Shah, Z. Duan, X. Song, R. Li, H. Mao, J. Liu, T. Ma, M. Wang, Evaluating the added value of multi-variable calibration of SWAT with remotely sensed evapotranspiration data for improving hydrological modeling, J. Hydrol. 603 (2021) 127046.
[35] H.K. Mcmillan, I.K. Westerberg, T. Krueger, Hydrological data uncertainty and its implications, Wiley Interdiscip. Rev.: Water. 5 (6) (2018) e1319.
[36] J. Joseph, S. Ghosh, A. Pathak, A. Sahai, Hydrologic impacts of climate change: comparisons between hydrological parameter uncertainty and climate model uncertainty, J. Hydrol. 566 (2018) 1–22.
[37] A. Bárdossy, C. Kilsby, S. Birkinshaw, N. Wang, F. Anwar, Is precipitation responsible for the most hydrological model uncertainty? Front, Water 4 (2022) 836554.
[38] D. Long, W. Yang, B.R. Scanlon, J. Zhao, D. Liu, P. Burek, Y. Pan, L. You, Y. Wada, South-to-North Water Diversion stabilizing Beijing's groundwater levels, Nat. Commun. 11 (1) (2020) 3665.
[39] J. Kim, K.-H. Ahn, Understanding the influence of hydrologic parameter uncertainty on Community Water Model predictions: a diagnostic assessment through extensive ensemble simulations, Environ. Res. Lett. 20 (6) (2025) 064001.
[40] Y. Wang, Y. Zhou, X. Zhang, K.J. Franz, G. Jia, Water regulation mitigates but does not eliminate water scarcity under rapid economic growth in the Haihe River basin, Resour. Conserv. Recycl. 215 (2025) 108098.
[41] C. Wu, S. Yang, Y.-P. Lei, Quantifying the anthropogenic and climatic impacts on water discharge and sediment load in the Pearl River (Zhujiang), China (1954–2009), J. Hydrol. 452 (2012) 190–204.
[42] H. Wang, T. Zhu, J. Wang, Y. Liu, Microplastic pollution in Pearl River networks: characteristic, potential sources, and migration pathways, Water Res. 276 (2025) 123261.
[43] S. Chen, L. Ran, J. Zhong, B. Liu, X. Yang, P. Yang, M. Tian, Q. Yang, S.L. Li, Z. Yan, Magnitude of and hydroclimatic controls on CO2 and CH4 emissions in the subtropical monsoon Pearl River Basin, J. Geophys. Res.: Biogeosci. 129 (5) (2024) e2023JG007967.
[44] J. Yu, Y. Tian, X. Wang, T. Sun, M. Lancia, C.B. Andrews, C. Zheng, Integrated modeling of flow, soil erosion, and nutrient dynamics in a regional watershed: assessing natural and human-induced impacts, Water Resour. Res. 60 (9) (2024) e2024WR037531.
[45] B. Fu, G. Huang, H.O. Zhang, J. Zhang, Strengthening ecological protection and green development in Pearl River Basin to support construction of Guangdong-Hong Kong-Macao Greater Bay Area, Bull. Chin. Acad. Sci. 38 (10) (2023) 1440–1446.
[46] Y. Tong, Z. Han, X. Gao, Bias correction in climate extremes over China for high-resolution climate change RegCM4 simulations using QM and QDM methods, Environ. Res. 27 (2022) 383–396.
[47] Y. Wada, L.P. Van Beek, M.F. Bierkens, Modelling global water stress of the recent past: on the relative importance of trends in water demand and climate variability, Hydrol. Earth Syst. Sci. 15 (12) (2011) 3785–3808.
[48] Y. Zhang, Y. Yin, M. Yin, X. Zhang, A high-resolution gridded dataset for China's monthly sectoral water use, Sci. Data 12 (1) (2025) 1157.
[49] C. Hou, Y. Li, S. Sang, X. Zhao, Y. Liu, Y. Liu, F. Zhao, High-resolution mapping of monthly industrial water withdrawal in China from 1965 to 2020, Earth Syst. Sci. Data Discuss. 2023 (2023) 1–25.
[50] M. Gilbert, G. Nicolas, G. Cinardi, T.P. Van Boeckel, S.O. Vanwambeke, G. Wint, T. P. Robinson, Global distribution data for cattle, buffaloes, horses, sheep, goats, pigs, chickens and ducks in 2010, Sci. Data 5 (1) (2018) 1–11.
[51] M. Gilbert, G. Cinardi, D. Da Re, W.G.R. Wint, D. Wisser, T.P. Robinson, Gridded Livestock of the World - 2015 (GLW 4), 2022.
[52] H. Steinfeld, Livestock's Long Shadow: Environmental Issues and Options, 2006.
[53] M.A. Meehan, G.L. Stokka, M.S. Mostrom, Livestock Water Requirements [M], NDSU Extension Service, 2015.
[54] Y. Wada, D. Wisser, M.F. Bierkens, Global modeling of withdrawal, allocation and consumptive use of surface water and groundwater resources, Earth Syst. Dyn. 5 (1) (2014) 15–40.
[55] W. Cheng, Q. Feng, H. Xi, X. Yin, L. Cheng, C. Sindikubwabo, B. Zhang, Y. Chen, X. Zhao, Modeling and assessing the impacts of climate change on groundwater recharge in endorheic basins of Northwest China, Sci. Total Environ. 918 (2024) 170829.
[56] L. Wu, X. Liu, Z. Yang, Y. Yu, X. Ma, 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) (2022) e14616.
[57] S. Das, A. Ng, B. Perera, Sensitivity analysis of SWAT model in the Yarra River catchment, in: Proceedings of the 20th International Congress on Modelling and Simulation, 2013. Adelaide, F.
[58] A. Saltelli, P. Annoni, I. Azzini, F. Campolongo, M. Ratto, S. Tarantola, Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index, Comput. Phys. Commun. 181 (2) (2010) 259–270.
[59] H. Kling, M. Fuchs, M. Paulin, Runoff conditions in the upper Danube basin under an ensemble of climate change scenarios, J. Hydrol. 424 (2012) 264–277.
[60] J.E. Nash, J.V. Sutcliffe, River flow forecasting through conceptual models part I-A discussion of principles, J. Hydrol. 10 (3) (1970) 282–290.
[61] Y. Her, C. Seong, Responses of hydrological model equifinality, uncertainty, and performance to multi-objective parameter calibration, J. Hydroinform. 20 (4) (2018) 864–885.
[62] Y. Nan, F. Tian, Glaciers determine the sensitivity of hydrological processes to perturbed climate in a large mountainous basin on the Tibetan Plateau, Hydrol. Earth Syst. Sci. 28 (3) (2024) 669–689.
[63] A. Ozdemir, U.M. Leloglu, K.C. Abbaspour, Hierarchical approach to hydrological model calibration, Environ. Earth Sci. 76 (8) (2017) 318.
[64] A.A. Oubeidillah, S.-C. Kao, M. Ashfaq, B.S. Naz, G. Tootle, A large-scale, high-resolution hydrological model parameter data set for climate change impact assessment for the conterminous US, Hydrol. Earth Syst. Sci. 18 (1) (2014) 67–84.
[65] Y. Guo, Y. Zhang, L. Zhang, Z. Wang, Regionalization of hydrological modeling for predicting streamflow in ungauged catchments: a comprehensive review, Wiley Interdiscip. Rev.: Water. 8 (1) (2021) e1487.
[66] J. Gou, C. Miao, Z. Xu, Q. Duan, Parameter uncertainty analysis for large-scale hydrological model: challenges and comprehensive study framework, Adv. Water Sci. 33 (2) (2022) 327–335.
[67] H.E. Beck, M. Pan, P. Lin, J. Seibert, A.I. Van Dijk, E.F. Wood, Global fully distributed parameter regionalization based on observed streamflow from 4,229 headwater catchments, J. Geophys. Res. Atmos. 125 (17) (2020) e2019JD031485.
[68] C. Jiang, E.J. Parteli, Q. Xia, X. Yin, Y. Shao, A regional hydrological model for arid and semi-arid river basins with consideration of irrigation, Environ. Model. Software 157 (2022) 105531.
[69] G. Tang, M.P. Clark, W.J. Knoben, H. Liu, S. Gharari, L. Arnal, A.W. Wood, A. J. Newman, J. Freer, S.M. Papalexiou, Uncertainty hotspots in global hydrologic modeling: the impact of precipitation and temperature forcings, Bull. Am. Meteorol. Soc. 106 (1) (2025) E146–E166.
[70] Z.E. Asong, M.E. Elshamy, D. Princz, H.S. Wheater, J.W. Pomeroy, A. Pietroniro, A. Cannon, High-resolution meteorological forcing data for hydrological modelling and climate change impact analysis in the Mackenzie River Basin, Earth Syst. Sci. Data 12 (1) (2020) 629–645.
[71] J. Zhang, B. Liu, S. Ren, W. Han, Y. Ding, S. Peng, A 4 km daily gridded meteorological dataset for China from 2000 to 2020, Sci. Data 11 (1) (2024) 1230.
[72] D.M. Jose, G.S. Dwarakish, Uncertainties in predicting impacts of climate change on hydrology in basin scale: a review, Arabian J. Geosci. 13 (19) (2020) 1037.
[73] J. Towner, H.L. Cloke, E. Zsoter, Z. Flamig, J.M. Hoch, J. Bazo, E. Coughlan De Perez, E.M. Stephens, Assessing the performance of global hydrological models for capturing peak river flows in the Amazon basin, Hydrol. Earth Syst. Sci. 23 (7) (2019) 3057–3080.
[74] F.F. Hattermann, V. Krysanova, S.N. Gosling, R. Dankers, P. Daggupati, C. Donnelly, M. Flörke, S. Huang, Y. Motovilov, S. Buda, Cross-scale intercomparison of climate change impacts simulated by regional and global hydrological models in eleven large river basins, Clim. Change 141 (2017) 561–576.
[75] A. Fleischmann, R. Paiva, W. Collischonn, Can regional to continental river hydrodynamic models be locally relevant? A cross-scale comparison, J. Hydrol. X. 3 (2019) 100027.
[76] N.V.V. Wanders, M.T. H, Y. Wada, M.F.P. Bierkens, L.P.H.R. Van Beek, Global Monthly Discharge Dataset, Derived from Dynamical 1-D water-energy Routing Model (Dynwat) at 10 Km Spatial Resolution, 2018.
[77] S. Harrigan, E. Zsoter, L. Alfieri, C. Prudhomme, P. Salamon, F. Wetterhall, C. Barnard, H. Cloke, F. Pappenberger, GloFAS-ERA5 operational global river discharge reanalysis 1979–present, Earth Syst. Sci. Data Discuss. 2020 (2020) 1–23.
[78] F.A. Hirpa, P. Salamon, H.E. Beck, V. Lorini, L. Alfieri, E. Zsoter, S.J. Dadson, Calibration of the Global Flood Awareness System (GloFAS) using daily streamflow data, J. Hydrol. 566 (2018) 595–606.
[79] L. Alfieri, V. Lorini, F.A. Hirpa, S. Harrigan, E. Zsoter, C. Prudhomme, P. Salamon, A global streamflow reanalysis for 1980-2018, J. Hydrol. X. 6 (2020) 100049.
[80] N. Wanders, M.T. Van Vliet, Y. Wada, M.F. Bierkens, L.P. Van Beek, High-resolution global water temperature modeling, Water Resour. Res. 55 (4) (2019) 2760–2778.